

# Perception de la prosodie émotionnelle française par les croatophones

---

Vrbljanin, Gabrijela

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:131:181969>

Rights / Prava: [Attribution-NonCommercial-NoDerivs 3.0 Unported / Imenovanje-Nekomercijalno-Bez prerađivanja 3.0](#)

Download date / Datum preuzimanja: **2024-05-12**



Repository / Repozitorij:

[ODRAZ - open repository of the University of Zagreb Faculty of Humanities and Social Sciences](#)



Sveučilište u Zagrebu

Filozofski fakultet

Odsjek za romanistiku

**PERCEPCIJA EMOCIONALNE PROZODIJE FRANCUSKOG JEZIKA KOD  
HRVATSKIH GOVORNIKA**

Diplomski rad

Student: Gabrijela Vrbljanin

Mentor: Bogdanka Pavelin Lešić, red. prof.

Zagreb, rujan 2022.

Université de Zagreb  
Faculté de Philosophie et Lettres  
Département d'études romanes

**PERCEPTION DE LA PROSODIE ÉMOTIONNELLE FRANÇAISE PAR LES  
CROATOPHONES**

Mémoire de Master 2

Étudiante : Gabrijela Vrbljanin

Directrice de mémoire : Bogdanka Pavelin Lešić,  
professeure d'université

Zagreb, septembre 2022

## Table de matière

<b>1. Introduction</b> .....	1
<b>2. Cadre théorique</b> .....	2
2.1. Prosodie.....	2
2.1.1. Moyens prosodiques .....	3
2.1.2. Rôles de la prosodie.....	4
2.2. Émotions.....	6
2.2.1. Catégories émotionnelles.....	7
2.2.2. États émotionnels.....	9
2.2.3. Canaux de communication émotionnelle .....	10
2.2.4. Corpus émotionnels de parole .....	11
2.3. Expression vocale des émotions.....	12
2.3.1. Paramètres acoustiques.....	13
2.3.2. Caractéristiques vocales des expressions émotionnelles .....	15
2.4. Expression faciale et posturale des émotions.....	17
2.4.1. Mimiques ou expressions faciales .....	18
2.4.2. Gestes ou expressions posturales.....	19
2.5. Perception des émotions.....	20
2.5.1. Perception des émotions dans une langue seconde .....	22
<b>3. Étude</b> .....	25
3.1. Objectif et hypothèse.....	25
3.2. Participants .....	25
3.3. Instruments .....	26
3.4. Procédure.....	28
3.5. Résultats .....	28
3.6. Discussion .....	43
<b>4. Conclusion</b> .....	48
<b>Bibliographie</b> .....	50

## **Sažetak**

U ovom radu ispituje se točnost prepoznavanja emocionalne prozodije na materinskom i stranom jeziku. Sudionici su izvorni govornici hrvatskog jezika koji će prepoznavati emocije izražene u snimkama na hrvatskom i francuskom jeziku. Polovica govornika hrvatskog jezika nikada nije učila francuski jezik, dok druga polovica uči francuski jezik. Istraživanje je pokazalo da je stopa prepoznavanja emocija vrlo velika na oba jezika kod svih sudionika. Zaključak je da slušatelji mogu prepoznati emocije čak i ako ne razumiju jezik na kojem je emocija izražena. Potvrđeno je da poznavanje jezika donosi malu prednost. Stopa prepoznavanja emocija na materinskom jeziku malo je veća u audio snimkama, dok je na stranom jeziku veća u audiovizualnim snimkama. Istraživanje je otkrilo da sudionici koji se bave glazbom bolje prepoznaju emocije, na oba jezika, od sudionika koji se ne bave glazbom.

Ključne riječi: prozodija, prozodijska sredstva, emocije, prepoznavanje, materinski jezik, strani jezik

## **Résumé**

Dans ce mémoire nous abordons la réussite en reconnaissance de la prosodie émotionnelle dans la langue maternelle et dans la langue étrangère. Les participants sont des locuteurs natifs du croate qui reconnaîtront les émotions exprimées dans les enregistrements en croate et en français. La moitié des locuteurs croates sont non-francophones et l'autre moitié sont francophones. L'étude montre que le taux de reconnaissance des émotions est très élevé dans les deux langues par tous les participants. On conclut que les auditeurs peuvent reconnaître les émotions même s'ils ne comprennent pas la langue parlée. Il est confirmé que la connaissance de la langue apporte un petit avantage. Le taux de reconnaissance des émotions dans la langue maternelle est un peu plus élevé dans les enregistrements audio, tandis que dans la langue étrangère, il est plus élevé dans les enregistrements audiovisuels. La recherche révèle que les participants qui pratiquent la musique reconnaissent mieux les émotions, dans les deux langues, que les participants qui ne la pratiquent pas.

Mots clés : prosodie, moyens prosodiques, émotions, reconnaissance, langue maternelle, langue étrangère

## 1. Introduction

Les traits suprasegmentaux ont toujours été plus difficiles à définir que les traits segmentaux (Lehiste, 1970 : 1). En raison de sa spécificité et de sa complexité, la prosodie a été négligée pendant de nombreuses années. Aujourd'hui, nous sommes plus conscients de l'importance de la prosodie dans le fonctionnement de la parole. Par conséquent, on essaie de la traiter sur un pied d'égalité avec les autres niveaux de langue (Lacheret, 2011b). L'un des rôles fondamentaux de la prosodie réside dans la manifestation des émotions du locuteur (Peppé, 2009). La première modalité pour exprimer des émotions, c'est l'expression vocale ou la voix (Kerkeni, 2020). Les caractéristiques vocales les plus fréquemment analysées sont la fréquence fondamentale, l'intensité et la durée (Bänziger, 2004). L'autre modalité permettant la communication des émotions est l'expression faciale et posturale ou les mimiques faciales et les gestes (Kerkeni, 2020). Scherer (2003) a confirmé que la reconnaissance visuelle des émotions était supérieure à la reconnaissance auditive. Il est déterminé que l'information prosodique est suffisante pour identifier l'émotion (Bänziger, 2004). C'est pourquoi les auditeurs sont capables de reconnaître les émotions exprimées dans une langue étrangère (Bhatara & al., 2016). Cela indique qu'il existe des principes universaux qui relient certaines caractéristiques vocales (Scherer, 2003) et certaines mimiques faciales (Ekman, 1999) à certaines émotions. Certaines recherches démontrent que les émotions sont reconnues avec une grande précision dans différentes cultures et que les locuteurs natifs obtiennent généralement de meilleurs résultats que les locuteurs non natifs (Bhatara & al., 2016). Cependant, la présente recherche démontre que la différence statistique dans la reconnaissance des émotions entre le groupe de locuteurs natifs et celui de locuteurs non natifs n'est pas pertinente (cf. 3.5 Résultats et 4. Conclusion ci-après).

Ce mémoire est divisé en deux parties. La première partie comporte le cadre théorique. Au début, on traite de la prosodie, des moyens prosodiques et de leurs rôles. Après cela, on présente la problématique des émotions, leur définition et leur catégorisation. On continue avec l'expression vocale, faciale et posturale des émotions. Cette partie se termine par la perception des émotions en général en langue maternelle et par la perception des émotions en langue seconde. Les résultats des études précédentes ont été présentés. La deuxième partie décrit l'étude réalisée dans le cadre du présent mémoire. On observe le succès en reconnaissance de la prosodie émotionnelle dans la langue maternelle et dans la langue étrangère. Les participants sont des locuteurs natifs du croate qui sont censés reconnaître les émotions exprimées dans les

enregistrements en croate et en français. On observera dans quelle langue les émotions seront mieux reconnues. Les participants sont divisés en deux groupes - les locuteurs croates non-francophones et les locuteurs croates francophones. Il sera vérifié quel groupe reconnaîtra les émotions avec plus de succès. On verra si le succès en reconnaissance des émotions sera plus élevé dans les enregistrements audio ou dans les enregistrements vidéo véhiculant le son et l'image. Enfin, il sera observé comment les personnes qui chantent ou jouent d'un instrument reconnaîtront les émotions par rapport aux personnes qui n'exercent aucune activité musicale.

## 2. Cadre théorique

### 2.1. Prosodie

La prosodie vient du mot grec *prosōdía* et signifie « ce qui est chanté avec des syllabes (...) et c'est l'emphase de la syllabe en chantant ». En grec, c'était le nom d'une chanson chantée avec accompagnement instrumental (Simeon, 1969 : 211). Selon *Hrvatski jezični portal* (2022), d'un point de vue littéraire, la prosodie est synonyme de « versification » ou de « métrique ». Elle est définie comme science de la structure des vers, en particulier des pieds, des syllabes longues et courtes et des accents. D'un point de vue linguistique, la prosodie fait partie de la phonétique qui étudie les faits qui échappent à la double articulation. Selon Barić & al. (1995), la prosodie est une partie de la science du langage qui, au sein de la philologie, étudie les phénomènes prosodiques (accentuels) dans une langue (Barić & al., 1995). Cependant, la prosodie ou phonétique suprasegmentale ne se résume pas aux accents. Il s'agit de partie de la phonétique qui étudie ce qui dans la parole ne ressortit pas à la réalisation des phonèmes : rythme, débit, tons, intonations, accents, intensité, pauses etc. (Galisson & Coste, 1976 : 450) Les unités prosodiques font référence au côté sonore de l'expression des mots et appartiennent au niveau suprasegmental de la langue. D'autre part, les phonèmes appartiennent au niveau segmental de la langue (Barić & al., 1995). Selon Silić & Pranjković (2005), la prosodie est, au sens le plus large, la mélodie de la parole créée par des changements rythmiques de hauteur, d'intensité et de durée du ton de base de la parole, y compris son absence pendant les pauses. Ainsi, la prosodie peut être définie comme aspects suprasegmentaux de la parole qui sont les relations de ton (hauteur), d'intensité (volume) et de durée (Škarić, 2007).

Škarić (2007) fait la distinction entre la prosodie des mots et la prosodie des phrases. Silić & Pranjković (2005) notent que des traits distinctifs importants dans le discours sont, au niveau des mots, l'accent et la durée des syllabes, et au niveau des phrases, l'accent, les pauses et la mélodie (intonation). Martinović, Pletikos Olof, Vlašić Duić (2021) confirment que la prosodie des phrases est plus forte que la prosodie des mots. L'intonation de la phrase est primordiale et la prosodie des mots est modifiée.

La prosodie constitue un objet d'étude complexe par rapport à la composante segmentale de la parole. Elle a longtemps été exclue du domaine de la linguistique structurale. Il peut au premier abord paraître naturel d'assimiler la prosodie au domaine du *para-verbal* et de l'*extralinguistique*. Dans cette optique, le verbal ne peut désigner que le matériel segmental (phonèmes, morphèmes et syntagmes). Néanmoins, la prosodie appartient à part entière à la sphère linguistique parce qu'elle est indispensable pour une communication réussie (Lacheret, 2011a). On ne peut pas parler ni interpréter la parole sans recourir aux moyens prosodiques. Aujourd'hui, de plus en plus d'attention est portée à la communication affective et à l'importance de la dimension vocale suprasegmentale, plus particulièrement de la prosodie. Il faut montrer comment la prosodie affective, qui est « au cœur du verbal », participe à la construction du sens en contexte et fait partie intégrante du discours adressé à l'autre (Lacheret, 2011b).

### 2.1.1. Moyens prosodiques

En phonétique croate, les moyens prosodiques de la parole ont été mis en valeur pour la première fois par Petar Guberina (1938-39). Il les appelle *les valeurs de la langue parlée* et les définit comme moyens d'expression non-lexicaux qui réalisent leur valeur à la base du son et du mouvement et qui nous permettent de nous exprimer succinctement ou expressivement, ou à la fois succinctement et expressivement. Avec les moyens d'expression lexicaux, c'est-à-dire les mots et leurs combinaisons, les moyens suprasegmentaux de la parole déterminent la parole humaine. Lehiste (1970 : 2) utilise le terme *les caractéristiques suprasegmentales* et les définit comme traits dont le domaine s'étend à plus d'un segment, c'est-à-dire des traits dont l'arrangement, en motifs contrastés dans le temps, n'est pas limité à des segments individuels définis par leur qualité phonétique. Ivo Škarić (1991) introduit le terme *les moyens prosodiques de la parole* qui est le plus souvent utilisé dans la terminologie phonétique et linguistique moderne. Les moyens prosodiques englobent de nombreuses couches de discours : d'une

couche de mots et de phrases (couche textuelle) à une couche d'expressivité et de cris (couche vocale). Plus tard, Škarić (2007) les appellera *les caractéristiques trans-segmentaux et paralinguistiques*. Hess, Möbius, Pätzold (1993) utilisent également le terme *les moyens prosodiques*.

Guberina (1952 [1939]) divise les sept *valeurs de la langue parlée* en deux catégories. Les *valeurs acoustiques* sont l'intonation, l'intensité, le tempo/le débit de la phrase et la pause tandis que les *valeurs visuelles* sont les mimiques faciales, les gestes et le contexte réel. Lehiste (1970 : 4) dans les traits suprasegmentaux inclut le ton, l'intensité et la durée. Škarić (1991) énumère dix *moyens prosodiques* : le ton et l'intonation, le volume et l'accent, le timbre de la voix, la composition spectrale du son de la parole, les pauses, le débit de la parole, le rythme, les modulations de la parole, la manière de prononcer les phonèmes, les mimiques faciales et les gestes. Hess, Möbius, Pätzold (1993) mentionnent l'intonation, le rythme, le tempo, les pauses, la durée, l'intensité et la qualité de la voix comme moyens prosodiques.

Guberina (1938, selon Frankol & Pavelin Lešić, 2016) constate que le langage se réalise par les valeurs de la langue parlée. Les éléments constitutifs des valeurs de la langue parlée enrichissent l'expression à travers des variations stylistiques. Vuletić (2007) souligne que les valeurs acoustiques de la langue parlée sont solidaires, entrelacées avec d'autres faits de parole et ne peuvent pas apparaître sous une forme isolée. Il est impossible de prononcer les syllabes, les mots ou les phrases sans volume, sans changement de hauteur, sans vitesse ou sans pauses respiratoires (Vuletić, 2007). L'énoncé est avant tout réalisé comme une unité formelle globale de parole, comme un ensemble d'intonation et d'intensité dans lequel on insère des mots articulés. Les mots prennent sens au sein de l'énoncé. L'énoncé n'existe qu'en situation. Donc, les mots prennent sens en situation. Les valeurs de la langue parlée découlent de la situation. Elles font partie intégrante des pensées et des émotions (Pavelin Lešić, 2013 : 17). En lecture, les mots ne sont pas insérés dans l'ensemble sonore. C'est pourquoi il est difficile de recevoir un message écrit (Vuletić, 2007). Pour comprendre un texte écrit, il faut le mettre en situation/contexte et le recréer en énoncé, en lui attribuant les valeurs de la langue parlée (Pavelin Lešić, 2013 : 31).

### 2.1.2. Rôles de la prosodie

Peppé (2009) distingue quatre fonctions communicatives de la prosodie. Le rôle *indexical* fait référence aux modifications des paramètres physiques de la voix qui indiquent les caractéristiques du locuteur. Le rôle *affectif* implique les sentiments et les attitudes du locuteur face aux messages échangés. Le rôle *grammatical* indique des relations grammaticales (comme ponctuations à l'écrit). Le rôle *pragmatique* met l'accent sur certaines parties du message et attire l'attention sur les informations les plus importantes. Selon Lacheret (2011a), on trouve deux fonctions centrales de la prosodie en français : une fonction de structuration du message et une fonction pragmatique qui s'organise autour de contraintes communicatives plurielles (modalités énonciatives, expression des émotions). Guberina (1954 [1939]) constate que les valeurs de la langue parlée jouent un rôle majeur dans la syntaxe des phrases complexes. Les conjonctions et locutions conjonctives ne suffisent pas à analyser des propositions. La catégorisation des phrases complexes repose sur le type de rapport d'idées et non pas sur le type de conjonction ou locution conjonctive utilisé. On ne comprendra bien la valeur des conjonctions qu'en comparant la valeur logique et la valeur stylistique de deux phrases avec et sans conjonctions. Donc, les valeurs de la langue parlée ne portent pas uniquement sur la valeur stylistique de la construction, mais également sur la valeur logique. Hess, Möbius, Pätzold (1993) soutiennent la thèse de Guberina en soulignant que les moyens prosodiques jouent un rôle important dans le soutien de l'organisation syntaxique et sémantique des énoncés. De nombreuses études démontrent que la compréhensibilité et le caractère naturel des stimuli peuvent être considérablement améliorés si l'information prosodique est proposée. Le fait que la prosodie contribue à une communication linguistique réussie est mis en évidence par son rôle dans le développement de la parole et du langage de l'enfant. Avant de pouvoir utiliser les mots, l'enfant utilise les gestes, les mimiques faciales et les sons (Guberina, 1952). Les nourrissons montrent une préférence pour écouter le discours maternel rempli de prosodie suraccentuée (Thiessen & al., 2005). Le signal riche en prosodie facilite la segmentation de la parole en éléments linguistiquement significatifs, c'est-à-dire les syntagmes et les mots, et aide au développement de la syntaxe (Seidl, 2007). L'importance des moyens prosodiques est également mise en évidence par la méthode verbo-tonale de Petar Guberina. La méthode verbo-tonale découle de l'approche verbo-tonale et de la linguistique de la parole. La méthode verbo-tonale est utilisée dans la rééducation des sourds, dans la pathologie de la parole ainsi que dans l'enseignement des langues étrangères. La méthode verbo-tonale ne traite pas la parole au niveau des sons isolés et des mots, mais prête une grande attention à la structure globale de la parole. Le corps entier participe à la perception des sons et il faut l'observer dans son ensemble. La méthode verbo-tonale met l'accent sur l'importance du rythme, de l'intonation, du

mouvement corporel, du contexte et de la situation. On engage à travers les mouvements corporels, les stimulations musicales et les dramatisations le corps et l'être entier de l'apprenant dans le processus d'enseignement/apprentissage de la parole en langue maternelle ou bien en langue étrangère (Guberina, 1995 ; Frankol & Pavelin Lešić, 2016).

## 2.2. Émotions

Le mot « émotion », qui signifie *mettre en motion ou en mouvement vers l'extérieur*, remonte à l'année 1579. Son origine vient du mot français *se mouvoir* et du mot latin *emovere*, *emotum* (enlever, secouer). L'émotion est l'un des concepts les plus difficiles à définir en psychologie (Kerkeni, 2020). Il existe de nombreuses définitions, mais elles se ressemblent toutes. William James (1884, selon Kerkeni, 2020) a défini l'émotion comme une réaction physiologique. Selon les psychologues, c'est la réaction affective intense, la réponse de l'organisme à une situation donnée (Darwin, 1872, selon Lacheret, 2011b). En neurophysiologie, on considère les émotions comme un ensemble des réponses physiologiques, somatiques (battements accélérés du cœur, modification de la respiration, etc.) induites par des stimuli provenant de l'environnement (Lacheret, 2011b). Le dictionnaire (selon Kerkeni, 2020) la définit comme un état affectif intense, caractérisé par des troubles physiques et mentaux, ce qui est perçu comme le résultat d'un phénomène extérieur (appelé stimulus : une image, une odeur, un son...). Selon Ekman et Davidson (1994, selon Kerkeni, 2020), l'émotion est une réaction transitoire, provoquée par un stimulus spécifique et caractérisée par un ensemble de réponses cognitives, physiologiques et comportementales. Lecomte (selon Kerkeni, 2020) la définit aussi comme une réaction de l'organisme à un événement extérieur, et qui comporte des aspects cognitifs, physiologiques et comportementaux. Sous l'angle expressif, les émotions sont des réactions inconscientes à des stimuli extérieurs. Ces actions et mouvements se manifestent sur le visage, dans la voix et à travers les comportements. (Damasio 1999, selon Lacheret, 2011b). On peut donc considérer l'émotion comme une réaction physiologique de durée brève et transitoire à une situation réelle ou imaginaire (Bänziger, 2004 ; Kerkeni, 2020). Les émotions ont différentes conséquences physiques. Par exemple, la peur peut provoquer un cri ou des tremblements, la tristesse peut causer les larmes, la joie génère un sourire (Kerkeni, 2020).

Léon (1993, selon Lacheret, 2011b) distingue 3 types d'émotions : les émotions brutes peu ou pas contrôlées, les émotions mieux canalisées (sentiments) et les attitudes contrôlées et

codées linguistiquement au plan de l'expression (évidence, ordre, doute). Selon les définitions ci-dessus, les émotions dont on parle sont du premier type. Elles ne doivent pas être confondues avec les sentiments et les attitudes qui correspondent à une étape plus élaborée des processus émotionnels.

La plupart des scientifiques s'accordent à dire qu'une émotion a cinq composantes. La première composante est *l'expression*. Elle permet d'interagir avec l'environnement, de communiquer l'émotion et elle peut être faciale, vocale ou posturale. La deuxième composante est *la motivation* ou la tendance à l'action. La troisième composante est *la réaction corporelle*. Il peut s'agir d'un changement de la température du corps ou un changement du rythme cardiaque. La quatrième composante est *le sentiment* ou la prise de conscience de son état émotionnel. Il nous permet de verbaliser l'émotion ou de la réguler. La cinquième composante est *l'évaluation cognitive* qui détermine la nature et l'intensité de la réaction émotionnelle (Kerkeni, 2020). Bänziger (2004) ne mentionne que 3 compositeurs sur ces cinq. Il s'agit du sentiment subjectif (expérience émotionnelle, angl. *feeling*), de la réaction physiologique et de l'expression émotionnelle. Elle omet la motivation et l'évaluation.

### 2.2.1. Catégories émotionnelles

Il existe une très grande variété d'émotions avec leurs combinaisons. Tous les psychologues et philosophes ne sont pas d'accord sur la classification mais tout le monde admet qu'il existe des émotions principales et leurs nuances. Donc, les émotions peuvent être divisées en deux classes - les émotions primaires et les émotions secondaires. (Kerkeni, 2020)

Les émotions primaires ou *les émotions de base* sont universelles et innées. Elles sont aussi automatiques, inconscientes et ont un déclenchement rapide, telles des réflexes (Kerkeni, 2020). Les émotions primaires peuvent également être identifiées avec *les émotions fondamentales, les catégories émotionnelles, les types d'émotions, les dimensions émotionnelles* ou *les familles émotionnelles*. La liste des émotions primaires varie d'un auteur à l'autre. Les quatre émotions primaires les plus fréquemment utilisées sont la colère, la peur, la joie et la tristesse (Bänziger, 2004). On trouve la liste la plus longue chez Arnold (1960, selon Kerkeni, 2020). Il a énuméré 11 émotions primaires. : la colère, l'aversion, le courage, le dégoût, le désespoir, le désir, la peur, la haine, l'espoir, l'amour et la tristesse (Kerkeni, 2020). Gray (1982, selon Kerkeni, 2020) a proposé 4 émotions primaires : l'espoir, la panique, la peur

et la rage. Ekman, Friesen et Ellsworth (1972, selon Kerkeni, 2020) ont conçu une liste de 6 émotions primaires dont chacune correspond à une expression du visage. Les expressions sont les mêmes entre plusieurs personnes d'âge, de culture ou de sexe différents. Ces émotions sont : *la tristesse, la colère, la joie, le dégoût, la peur et la surprise*. Dans son ouvrage « L'expression des émotions chez l'Homme et les animaux », Charles Darwin a constaté que ces émotions étaient fondamentales et universelles à l'espèce humaine. En psychologie, ces 6 émotions sont communes à tous les auteurs (Bänziger, 2004 ; Kerkeni, 2020). Il y a un modèle expressif qui correspond à chaque catégorie émotionnelle. Les études ont montré que les émotions de base, sont facilement identifiables sur le plan des expressions vocales (Bänziger, 2004) et faciales (Ekman, 1976, selon Kerkeni, 2020). Les bébés de plus de 3 mois sont capables de discriminer les expressions de visage des 6 émotions de base et *l'expression neutre* (Ekman, 1976, selon Kerkeni, 2020). Ces sept émotions sont aujourd'hui mondialement connues (Kerkeni, 2020).

Les émotions secondaires ou *émotions complexes* découlent des émotions primaires et résultent d'un mélange de ces dernières. Par exemple, la colère peut donner de l'agressivité, de la haine ; à partir de la peur, de la culpabilité, de l'angoisse... Selon Ekman, les émotions secondaires sont des mélanges des émotions de base, il les nomme parfois *des émotions mixtes*. Pour lui la honte est une émotion mixte, parce que c'est un mélange de peur et de colère. Parmi ces émotions secondaires, on trouve aussi la jalousie, la culpabilité, l'embarras et l'envie... L'approche catégorielle (ou discrète) permet de représenter une infinité d'émotions obtenues en combinant les émotions primaires. Un exemple de modèle discret est celui de Plutchik qui est composé de 8 émotions de base faites de 4 paires opposées deux à deux (joie-tristesse, anticipation-surprise, colère-peur et dégoût-confiance) et de multiples variations. Contrairement aux émotions primaires, les émotions secondaires ne sont pas innées. Elles n'apparaissent qu'entre 1 an et 4 ans. Elles ne sont pas forcément automatiques, leur déclenchement n'est pas forcément rapide et leur action peut durer dans le temps (Kerkeni, 2020).

Dans l'approche dimensionnelle (ou continue), les émotions sont situées dans un espace de deux ou trois dimensions. Les deux dimensions principales sont la valence et l'activation. La valence permet de distinguer les émotions positives (par exemple, la satisfaction) des émotions négatives (par exemple, la peur). L'activation représente le niveau d'activation corporelle et distingue les émotions d'activation faible (par exemple, l'ennui) des émotions d'activation forte (par exemple, la joie) (Kerkeni, 2020). Les émotions positives ne sont pas

aussi bien étudiées que les émotions négatives. La plupart des émotions de base, celles qui sont le plus souvent examinées, sont négatives (Bhatara & al., 2016).

### 2.2.2. États émotionnels

Une catégorie émotionnelle implique plusieurs états émotionnels. Wundt (1897, selon Bänziger, 2004) a suggéré que les états émotionnels pouvaient être décrits, dans une perspective subjective, en fonction de leurs positions respectives à trois dimensions : *la valence*, *l'excitation* et *la tension*. Les dimensions des états émotionnels correspondent aux dimensions des catégories émotionnelles. La valence permet de distinguer des états positifs et des états négatifs ou des états agréables et des états désagréables. L'excitation représente le niveau d'activation corporelle et distingue des états d'activation faible ou des états perçus comme calmes et des états d'activation forte ou des états perçus comme excités. La tension distingue des états perçus comme relaxé et des états perçus comme tendu (Kerkeni, 2020). On a trois facteurs principaux de confusion des états émotionnels. Le premier facteur est la famille émotionnelle. Les états émotionnels appartenant à une même famille (par exemple colère chaude et colère froide) sont plus souvent confondus entre elles qu'avec des émotions appartenant à d'autres familles. Les deux autres facteurs de confusion sont deux dimensions principales - la valence et l'activation. Les états d'un même niveau d'activation seront plus souvent confondus entre eux (par exemple, joie très intense et tristesse très intense). De même, les expressions positives sont plus souvent confondues entre elles que ne le sont les expressions négatives (Banse & Scherer, 1996 selon Bänziger, 2004). Les études observant la différence entre les expressions correspondantes à des différents états d'une même famille émotionnelle (différents types de colère, de peur, de joie ou de tristesse) sont relativement rares. Cette variabilité des états émotionnels appartenant à la même catégorie entraîne des incohérences des résultats des différentes études (Bänziger, 2004).

Banse & Scherer (1996, selon Bänziger, 2004) et Burkhardt (2001, selon Bänziger, 2004) ont fait les distinctions des états émotionnels, à l'intérieur de chaque catégorie émotionnelle, basées sur le niveau d'activation. Juslin & Laukka (2001, selon Bänziger, 2004) fondent leur distinction sur l'intensité des émotions exprimées. Burkhardt (2001, selon Bänziger, 2004) distingue la colère froide et la colère chaude que Banse & Scherer (1996, selon Bänziger, 2004) appellent aussi l'irritation (colère froide) et la rage (colère chaude). Donc, la rage est la version active de la colère (Scherer, 2003). Juslin & Laukka (2001, selon Bänziger,

2004) font la différence entre l'intensité faible et l'intensité forte de colère. Frick (1986, selon Bänziger, 2004) fait la différence entre la colère frustrée et la colère agressive. Banse & Scherer (1996, selon Bänziger, 2004) ont introduit la distinction entre anxiété (ou peur anxieuse) et peur panique. Juslin & Laukka (2001, selon Bänziger, 2004) distinguent l'intensité faible et l'intensité forte de peur. Banse & Scherer (1996, selon Bänziger, 2004) distinguent joie calme et joie intense. La joie intense peut aussi être appelée l'exaltation (joie exaltée) ou la joie excitée (Katz, 1998, selon Bänziger, 2004). Juslin & Laukka (2001, selon Bänziger, 2004) distinguent l'intensité faible et l'intensité forte de joie. Banse & Scherer (1996, selon Bänziger, 2004) distinguent tristesse déprimée et désespoir (tristesse désespérée). La tristesse déprimée est chez Burkhardt (2001, selon Bänziger, 2004) appelée la tristesse calme. Juslin & Laukka (2001, selon Bänziger, 2004) distinguent l'intensité faible et l'intensité forte de tristesse ainsi que l'intensité faible et l'intensité forte de dégoût.

Dans chaque étude différentes catégories émotionnelles et différents états émotionnels ont été étudiés. Dans leurs recherches, Juslin & Laukka (2003) incluent 5 catégories émotionnelles : la colère, la peur, la joie, la tristesse et la tendresse. Caelen Hamont (2005, selon Kerkeni, 2020), dans son résumé de plusieurs études, mentionne aussi 5 émotions. À la place de la tendresse, il installe le dégoût. Scherer (2003) inclut 6 catégories émotionnelles : le stress, la rage, la peur panique, la tristesse, la joie intense et l'ennui. Il a précisé les catégories émotionnelles de joie, de colère et de peur qui ont été étudiées, dans ce cas, dans leurs versions intenses. Dans une autre étude, en plus des 6 émotions primaires, Scherer & Oshinsky (1977, selon Kerkeni, 2020) ont également mentionné l'ennui, le plaisir, l'activité émotionnelle et la puissance.

### 2.2.3. Canaux de communication émotionnelle

Il existe 3 modalités permettant de communiquer des émotions : *les signaux physiologiques*, *l'expression faciale et posturale* et *l'expression vocale*. L'expression vocale ou la voix est la forme la plus naturelle de communication humaine. Elle est considérée depuis longtemps comme une mesure de l'émotion et comme un miroir de la personnalité. Les expressions faciales et posturales, c'est-à-dire les mimiques et les gestes, expriment aussi des émotions. Les mouvements du corps, et surtout les mouvements du visage et le regard véhiculent de riches informations. Les émotions primaires sont véhiculées par des configurations faciales spécifiques, en termes de positions des sourcils, forme de la bouche...

Selon Ekman, ces expressions faciales sont universelles. Cependant, la mimique faciale a des limites dans son efficacité à communiquer des émotions. Certaines recherches montrent que les femmes sont meilleures que les hommes pour la reconnaissance des émotions à partir d'expressions faciales. En outre, certaines configurations faciales sont communes à plusieurs émotions et cela rend difficile l'interprétation. Par exemple, le visage manifeste la peur et la surprise par sourcils levés. Lors de l'expression des émotions, nous assistons à un certain nombre d'indices physiologiques. La manifestation des émotions est étroitement liée au système nerveux végétatif. Selon les émotions, nous pouvons observer des augmentations ou des diminutions de la fréquence cardiaque, de la sudation, du débit sanguin cutané (rougissement ou pâleur), variations de température et la conductivité électrique de la peau, variations du volume respiratoire, etc. Ces caractéristiques peuvent être mesurées et utilisées pour reconnaître et interpréter une émotion (Kerkeni, 2020).

#### 2.2.4. Corpus émotionnels de parole

On distingue trois catégories de corpus émotionnel utilisées dans le domaine de la reconnaissance automatique des émotions : *le corpus naturel*, *le corpus induit* et *le corpus simulé*. Le corpus naturel est le plus authentique. Il est constitué d'émotions spontanées et naturelles. Il est obtenu à partir des enregistrements réalisés dans les contextes non contrôlés, tel que les émissions de TV, les centres d'appels, les interviews de consommateurs... L'inconvénient : ces enregistrements sont très limités en quantité de données émotionnelles, ils sont généralement de courtes durées et de mauvaise qualité. Pour ce type de corpus, l'étiquetage émotionnel est particulièrement complexe parce que dans une même phrase il peut y avoir plusieurs émotions. À cause de tout cela, la reconnaissance des émotions n'est pas garantie. Le corpus induit est moins authentique. Les émotions de cette catégorie sont induites en utilisant plusieurs techniques : l'exposition du sujet à des tâches ou la présentation des images, des films, des jeux permettant d'induire des émotions. Les enregistrements sont de bonne qualité, mais les émotions induites sont souvent d'intensité faible. De plus, la reconnaissance des émotions est moins garantie parce que les mêmes protocoles d'induction n'induisent pas nécessairement les mêmes états émotionnels chez les individus. Le corpus « acté » (corpus avec des émotions actées / simulées ; Vaudable, 2012 / *corpus with the acted emotions* ; Vogt & André, 2005, selon Kerkeni, 2020) ou simulé consiste à enregistrer des acteurs professionnels ou semi professionnels qui simulent l'émotion en se basant sur des

scenarios. Ce corpus présente beaucoup d'avantages. Il permet de collecter, d'une manière simple, un grand nombre de données très prototypiques. Le contrôle et l'exploitation du corpus sont faciles. Les performances obtenues sont largement meilleures à celles obtenues avec des données réalistes. Bien que ce type de corpus suscite aussi des critiques, car les expressions sont hors contexte et ne correspondent pas exactement aux expressions véritables, il est le plus fréquemment utilisé dans les études menées sur la reconnaissance automatique des émotions (Kerkeni, 2020).

### 2.3. Expression vocale des émotions

On a déjà mentionné que l'expression vocale ou la voix est la modalité la plus importante pour la communication en général et pour la communication des émotions (Kerkeni, 2020). Le chemin de la voix est divisé en 3 parties : les poumons, le larynx et les cordes vocales, les résonateurs (la gorge, la bouche, les fosses nasales). Les poumons produisent un flux d'air pour permettre la vibration des cordes vocales : plus le flux est faible, moins les cordes vocales vibrent et moins la voix s'entend. La voix humaine est une émission de sons qui est caractérisée par trois paramètres acoustiques : l'intensité, la fréquence et le timbre. Chaque individu a sa propre voix qui n'est pas toujours la même : elle varie suivant la situation et le contexte (Kerkeni, 2020). Les modifications corporelles/physiologiques, qui accompagnent certains états émotionnels d'un locuteur, induisent les modifications de la voix (Picard, 1997, selon Kerkeni, 2020). Le fonctionnement des appareils respiratoire, phonatoire et articulatoire est modifié par les changements autonomes et somatiques associés à la réaction émotionnelle (Scherer, 2003). Donc, les facteurs physiologiques (tels que la fréquence cardiaque, la tension musculaire et le flux sanguin) affectent les paramètres vocaux (Kerkeni, 2020).

Le signal vocal est un signal réel, continu, d'énergie finie, et stationnaire sur une durée courte (25 à 30 ms) au cours du temps. Il y a deux grands types d'informations dans le signal vocal. L'information linguistique fait partie sémantique du discours. Elle est véhiculée par les énoncés faits selon les règles de prononciation. L'information paralinguistique fait référence aux messages implicites tels que l'état émotionnel du locuteur. Plusieurs études démontrent que cette information est suffisante pour reconnaître l'état émotionnel (Kerkeni, 2020). Les auditeurs peuvent identifier correctement les émotions sans difficulté à partir des caractéristiques vocales (Bänziger, 2004). C'est pourquoi un être humain est capable de

comprendre les émotions transmises par des énoncés dans une langue étrangère que celles dans sa langue maternelle (Kerkeni, 2020).

### 2.3.1. Paramètres acoustiques

Les études menées sur la reconnaissance automatique des émotions (RAE) récupèrent l'information paralinguistique et négligent toutes les autres informations (Kerkeni, 2020). Lorsque l'on parle d'analyse acoustique des caractéristiques vocales des expressions émotionnelles, les paramètres pris en compte varient en fonction des études. Les caractéristiques les plus fréquemment analysées correspondent aux paramètres dérivés du contour de la fréquence fondamentale, du contour de l'intensité et de la durée des expressions (Bänziger, 2004). Lacheret (2011b) explique que l'organisation prosodique du signal de la parole est organisée par ces mêmes paramètres. Elle y ajoute également de la qualité vocale. Les corrélats perceptifs de ces paramètres sont la mélodie, la sonie<sup>1</sup>, la longueur et le timbre. Kerkeni (2020) confirme aussi que chacun des phénomènes prosodiques qui modèlent l'émotion du locuteur (l'accent, le rythme, l'intonation, la mélodie de la voix) se manifeste par des variations au niveau de la fréquence fondamentale, de l'intensité et/ou du débit de la parole.

La fréquence fondamentale (F0) correspond au nombre d'ouvertures/fermetures par seconde des cordes vocales (Kerkeni, 2020) ou au nombre de répétitions de la période fondamentale du signal acoustique par seconde (Bänziger, 2004). Elle, exprimée en hertz (Hz) (Bänziger, 2004), permet de distinguer des sons graves (bas) et des sons aigus (hauts). La fréquence fondamentale correspond d'un point de vue perceptif à la hauteur de la voix. Le terme de *ton/tonalité* est aussi utilisé pour désigner la hauteur d'un son (Kerkeni, 2020). Selon Lehiste (1970 : 54), la hauteur (angl. *pitch*) fait référence au corrélat perceptif de la fréquence fondamentale, le ton (angl. *tone*) concerne le niveau du mot et l'intonation (angl. *intonation*) fonctionne au niveau de la phrase. C'est pourquoi Škarić (1991) énumère le ton et l'intonation comme un moyen prosodique unique. Guberina (1952) fait mention du ton comme d'un élément de l'intonation. Les contours de F0, qui peuvent être assimilés à l'intonation, représentent l'évolution de la F0 au cours d'une expression (Bänziger, 2004). Škarić (1991) constate que l'intonation ou la mélodie de la parole est formée par des changements successifs de ton. Barić & al. (1995) s'accordent et définissent l'intonation comme ligne mélodique d'une phrase, une

---

<sup>1</sup> Intensité de la sensation sonore en rapport avec la pression acoustique (Larousse).

ligne qui représente le mouvement du ton de la phrase. En fonction du ton, pour des raisons physiologiques, l'intensité et le timbre varient. L'intensité forte est aussi le signe d'un ton haut, le timbre sombre d'un ton très bas, etc. (Škarić, 1991).

L'intensité de la voix (l'énergie) est mesurée en tant que puissance de parole et permet de distinguer un son fort d'un son faible. Elle est exprimée en décibels (dB). Le corrélat perceptif de l'intensité est le volume ou la sonie (Bänziger, 2004). L'intensité dépend de l'amplitude de la vibration des cordes vocales : plus elle est élevée, plus le son est fort ; plus l'amplitude est faible, plus le son est faible. L'amplitude dépend de la variation de la pression de l'air sur les cordes vocales (Kerkeni, 2020). Guberina (1952) mentionne l'intensité comme une valeur acoustique de la langue parlée tandis que Škarić (1991) mentionne le volume comme un moyen prosodique. La sonie est de 35 dB dans les chuchotements, 65 dB dans une conversation normale et 100 dB dans les cris (Škarić 1991). Le contour d'intensité correspond à l'évolution de l'intensité acoustique au cours d'une expression (Bänziger, 2004).

La durée a plusieurs aspects. Certains d'entre eux sont le débit de la parole ou la durée proportionnelle des pauses par rapport à la durée totale des expressions (Bänziger, 2004). Donc, la longueur et le débit de la parole sont tous deux des corrélats perceptifs de la durée. (Lehiste, 1970 : 51). Selon Škarić (1991), le débit de la parole ou le tempo de la parole est un moyen prosodique qui s'exprime par le nombre d'unités de parole par unité de temps. C'est généralement le nombre de syllabes par seconde. Selon Guberina (1952), le tempo de la phrase est la valeur acoustique de la langue parlée. C'est la vitesse à laquelle les éléments d'une phrase et les phrases entières sont prononcés. Ainsi, il mesure la rapidité avec laquelle la parole est produite (Kerkeni, 2020). Le tempo peut être lent, moyen ou rapide (Vuletić, 2007). Le débit normal de la parole est de 4 à 7 syllabes par seconde, et le maximum peut aller jusqu'à 13 à 14 syllabes par seconde (Škarić, 1991).

Le timbre de la voix correspond aux sons sombres et sons clairs. L'air circule dans les résonateurs (la gorge, la bouche, les fosses nasales) et va prendre sa couleur, son timbre et ses harmoniques. C'est grâce au timbre qu'on distingue une note jouée à la guitare et au piano, qu'on reconnaît la voix de chaque personne (Kerkeni, 2020). Le timbre a la partie permanente et la partie variable. La partie permanente est unique à chacun. Elle résulte de propriétés organiques et ne varie pas. La partie variable est la capacité expressive du locuteur (Škarić, 1991). Le timbre de la voix est lié au spectre de la voix (Kerkeni, 2020). Le timbre est aussi appelé la qualité. La qualité est définie comme qualité du timbre de la voix qui est un corrélat

perceptif de la complexité des signaux sonores (acoustiques) de la caractéristique spectrale du son (Varošanec-Škarić, 2005). Selon Škarić (1991), la composition spectrale et le timbre de la voix sont des moyens prosodiques.

Chaque étude prend un nombre différent de paramètres acoustiques. En plus de ces quatre dimensions fondamentales et les plus courantes, de nombreuses autres dimensions vocales sont évaluées dans diverses études. La précision de l'articulation (la parole est très bien articulée à très mal articulée) est souvent mentionnée (Bezooijen, 1984 ; Juslin & Laukka, 2003 ; Bänziger, 2004 ; Caelen Hamont, 2005) ainsi que les variations de F0 (Scherer & Oshinsky, 1977 ; Juslin & Laukka, 2003 ; Caelen Hamont, 2005) et la qualité rauque (la voix est très rauque à ne pas rauque) (Bezooijen, 1984 ; Bänziger, 2004). Bezooijen (1984, selon Bänziger, 2004) mentionne le degré d'arrondissement et le degré d'extension des lèvres, le degré de tension du larynx, le degré de relâchement du larynx, le degré de « grincement » (angl. *creak*), de tremblement, de chuchotement, etc. Bänziger (2004) inclut l'intonation ou la mélodie (la voix est très monotone à très modulée), la fluidité (la parole est très hachée à très fluide), l'instabilité (la voix est très stable à très tremblante), la qualité perçante (la voix est très perçante à ne pas perçante), etc. Juslin & Laukka (2003) évaluent jusqu'à 14 dimensions. Certains d'entre eux sont la variabilité de l'intensité, l'énergie dans les hautes fréquences, les régularités microstructurales, la proportion de pauses, le formant 1 (hauteur), le formant 1 (largeur), *jitter*, *glottal waveform*, etc. On trouve aussi l'enveloppe (Scherer & Oshinsky, 1977, selon Kerkeni, 2020) et l'amplitude de F0 (Caelen Hamont, 2005, selon Kerkeni, 2020).

### 2.3.2. Caractéristiques vocales des expressions émotionnelles

Les résultats des études indiquent que les paramètres acoustiques reflètent essentiellement l'activation émotionnelle. Les émotions qui incluent une activation forte - telle que la joie, la colère et la peur - sont caractérisées par l'augmentation de la F0 et de l'intensité ainsi que l'accélération de la parole alors que les émotions qui incluent un degré d'activation faible - tels que la tristesse, la tendresse ou l'ennui - sont caractérisées par la diminution de la F0 et de l'intensité de même que le ralentissement de la parole (Bänziger, 2004). Vuletić (1986) démontre que les variations des constituants de *valeurs de la langue parlée* sont déployées symétriquement dans les évaluations de l'expression parlée d'une émotion particulière, et que les émotions opposées (la joie et la tristesse) sont exprimées par des constituants de valeurs

parlées opposées. Certaines des caractéristiques vocales de 4 émotions de base (colère, peur, joie, tristesse) et de plusieurs autres émotions seront présentées ci-dessous.

Selon Juslin & Laukka (2003), Scherer (2003) et de nombreux autres auteurs (Kerkeni, 2020), la colère est généralement caractérisée par la fréquence fondamentale haute, l'intensité forte et le débit de la parole rapide. De plus, Scherer & Oshinsky (1977, selon Kerkeni, 2020) constatent que la colère manifeste le débit le plus rapide de toutes les émotions. L'énergie dans les hautes fréquences est forte alors que la proportion de pauses est faible (Juslin & Laukka, 2003). Les résultats de Scherer (2003) sont les mêmes, avec une seule différence. Les contours de F0 (intonation) chez Juslin & Laukka (2003) sont montants tandis que chez Scherer (2003) sont descendants. Lacheret (2011b) confirme la mélodie descendante. Les résultats de Sasserath & al. (2018 selon Kerkeni, 2020) sont en accord avec ceux de Juslin & Laukka (2003) qui montrent une augmentation de la fréquence fondamentale. Chez Vuletić (1986), on trouve une contradiction intrigante. Sa recherche démontre que la manifestation de la colère se caractérise par l'intensité forte et le ton bas. Cependant, comme déjà mentionné ci-dessus, le ton haut est généralement associé à la colère.

Plusieurs auteurs (Picard, 1997 ; Juslin & Laukka, 2003 ; Scherer (2003) ; Kerkeni, 2020) ont confirmé que la peur se manifeste par la fréquence fondamentale haute (la voix aigüe), l'intensité forte et le débit de la parole rapide. Les contours de F0 sont montants tandis que l'énergie dans les hautes fréquences est forte. La proportion de pauses est faible (Juslin & Laukka, 2003). Chez la peur, on trouve la plus grande variabilité des résultats. Bien que la peur soit associée à l'intensité forte, la moitié des études de Juslin & Laukka ont rapporté l'intensité faible (ou moyenne). De plus, Vuletić (1986) a déterminé que l'expression de la peur se caractérise par le ton haut, l'intensité faible et le débit de la parole rapide. On a donc une grande divergence d'intensité. L'intensité est-elle forte, moyenne ou faible ? Scherer & Oshinsky (1977, selon Kerkeni, 2020) ont confirmé que la peur a les variations de l'intensité les plus fortes.

Selon Vuletić (1986), Juslin & Laukka (2003) et Scherer (2003), la joie est caractérisée par la fréquence fondamentale haute, l'intensité forte et le débit de la parole rapide. Les contours de F0 (contour intonatif global) sont montants (Lacheret, 2011b) alors que l'énergie dans les hautes fréquences est forte. La proportion de pauses est faible (Juslin & Laukka, 2003 ; Scherer, 2003). Néanmoins, plusieurs auteurs ont confirmé que la joie pouvait aussi s'exprimer par le débit de la parole lent (Murray & Arnott, 1993, Kerkeni, 2020). Contrairement à la peur,

la joie a les variations de l'intensité les plus faibles (Scherer & Oshinsky, 1977, selon Kerkeni, 2020).

Selon Picard (1997), Vuletić (1986), Juslin & Laukka (2003), Scherer (2003), Lacheret (2011b), Sasserath & al. (2018) et beaucoup d'autres auteurs (Kerkeni, 2020), la tristesse est caractérisée par la fréquence fondamentale basse (la voix grave), l'intensité faible et le débit de la parole lent. Sasserath & al. (2018, selon Kerkeni, 2020) confirme une diminution de la fréquence fondamentale. Cela correspond aux contours de F0 descendants chez Juslin & Laukka (2003) et Scherer (2003) ainsi qu'à la mélodie descendante chez Lacheret (2011b). L'énergie dans les hautes fréquences est faible tandis que la proportion de pauses est forte (Juslin & Laukka, 2003 ; Scherer, 2003). Scherer & Oshinsky (1977, selon Kerkeni, 2020) ajoute que la tristesse a le débit le plus lent de toutes les émotions. Les résultats pour la tristesse sont très cohérents.

Plusieurs auteurs ont confirmé que le dégoût se caractérise par la fréquence fondamentale basse, l'intensité faible et le débit de la parole lent (Kerkeni, 2020). Selon Scherer & Oshinsky (1977, selon Kerkeni, 2020), le dégoût a les variations de F0 les plus faibles. L'expression de la surprise se caractérise par le ton haut et le débit de la parole rapide (Vuletić, 1986). Selon Scherer & Oshinsky (1977, selon Kerkeni, 2020), la surprise a les variations de F0 les plus fortes. Selon Juslin & Laukka (2003), la tendresse est caractérisée par la fréquence fondamentale basse, l'intensité faible et le débit de la parole lent. Les contours de F0 sont descendants alors que l'énergie dans les hautes fréquences est faible. La proportion de pauses est forte. Selon Scherer (2003), le stress est caractérisé par la fréquence fondamentale haute et l'intensité forte. L'ennui se manifeste par la fréquence fondamentale basse (la voix grave), l'intensité faible et le débit de la parole lent (Picard, 1997, selon Kerkeni, 2020). Scherer & Oshinsky (1977, selon Kerkeni, 2020), confirment le ton bas tandis que Scherer (2003), confirme le débit de la parole lent. Selon Scherer & Oshinsky (1977, selon Kerkeni, 2020), l'activité émotionnelle a le ton le plus haut. L'expression du mépris se caractérise par le débit de la parole lent, le ton bas et l'intensité forte (Vuletić, 1986).

#### 2.4. Expression faciale et posturale des émotions

L'autre modalité permettant de communiquer des émotions sont les expressions faciales et posturales ou en d'autres termes les mimiques faciales et les gestes. Pour Guberina (1952),

les mimiques faciales et les gestes sont, avec le contexte réel, les valeurs visuelles de la langue parlée. Škarić (1991) a accordé un peu moins d'attention aux mimiques et aux gestes qu'aux autres moyens prosodiques. Selon Guberina (1952), les mimiques faciales et les gestes sont les éléments de mouvement corporel pertinents à l'expression linguistique. Ce sont les moyens que l'homme peut utiliser pour exprimer plus rapidement ou/et expressivement ce qu'il peut exprimer avec des mots articulés. L'intonation et l'intensité donnent à l'énoncé la forme de base de l'expression de la pensée. Škarić (1991) a défini les mimiques faciales en tant que mouvements faciaux et les gestes en tant que mouvements corporels (bras, tête, épaules). Ils encouragent et soutiennent d'autres moyens prosodiques vocaux et les prolongent en un signal visible. Ils sont harmonisés avec les sons de la voix et véhiculent le même sens. Vuletić (2007) confirme que les mimiques faciales et les gestes font partie de la communication verbale et expriment des états émotionnels ou des attitudes interpersonnelles (compréhension, accord, satisfaction, doute, etc.). Il ajoute que, puisque le visage est au centre de l'attention de l'interlocuteur, les mimiques faciales ont un rôle très important qui consiste à suivre et à maintenir la communication. Škarić (1991) souligne que les mimiques faciales et les gestes qui sont indépendants de la parole (tics, actions incidentes) et qui ont lieu en dehors du temps de parole ne sont pas prosodiques. Les manifestations faciales et gestuelles peuvent être conventionnelles (hocher la tête, salutations) et non conventionnelles (clin d'œil). Guberina (1952) dit que ces mouvements ne sont pas de simples imitations, mais sont au service de l'homme en tant qu'être social. La méthode verbo-tonale de Guberina met en relief l'importance du mouvement dans l'apprentissage des langues étrangères. L'une de ses hypothèses de base est que la direction et la position des organes phonatoires (la micromotricité) peuvent être représentés par des mouvements du corps entier (la macromotricité). La macromotricité est susceptible de faciliter la perception et la reproduction des unités intonatives ainsi que créer des conditions pour atteindre le rythme et l'intonation d'une langue (Frankol & Pavelin Lešić, 2016 ; Vuletić, 2007).

#### 2.4.1. Mimiques ou expressions faciales

Les expressions faciales des émotions primaires sont spécifiques et universelles (Ekman, 1999). Ekman (1976, selon Kerkeni, 2020) a développé un outil (*The Facial Action Coding System*) qui lui a permis d'identifier les photographies de visages exprimant les 6 émotions de base (tristesse, colère, joie, dégoût, peur, surprise) et l'expression neutre (aucune

émotion). Même les bébés de plus de 3 mois sont capables de discriminer ces expressions faciales quand on les leur montre. Certaines recherches confirment que le mépris et la honte disposent aussi des expressions faciales universelles, mais les preuves sont les plus importantes pour 6 émotions primaires. Toute mimique faciale d'émotion a plusieurs variations possibles. Par exemple, des versions de la colère et du dégoût avec la bouche fermée et la bouche ouverte, et des variations dans l'intensité des contractions musculaires pour chaque émotion. Certaines preuves suggèrent qu'il existe plus d'une expression possible pour véhiculer certaines émotions (Ekman, 1999).

L'autre outil de Ekman *The Facial Affect Scoring Technique* consiste à reconnaître une émotion à partir d'une photographie d'une partie du visage. On a trois zones du visage capables de mouvements indépendants : 1. les sourcils/le front ; 2. les yeux/les paupières et la racine du nez ; 3. la partie inférieure du visage : les joues, la bouche, la majeure partie du nez et le menton. Chaque photo est associée à l'une des six émotions de base. La tristesse et la peur sont le mieux reconnues dans la zone des yeux. La joie, la colère et le dégoût sont le mieux reconnus dans la zone inférieure du visage. La surprise se reconnaît aussi bien dans la zone des yeux que dans la zone inférieure du visage. Si l'on prend en compte tous les pourcentages, les émotions sont le mieux reconnues dans la zone des yeux et le pire dans la zone des sourcils (Ekman 1977).

#### 2.4.2. Gestes ou expressions posturales

En tant que modalité d'expression des émotions, les mimiques faciales ont été beaucoup plus étudiées que les gestes. Ekman a suggéré que le visage est le canal le plus efficace pour exprimer des émotions spécifiques. Cependant, certaines émotions peuvent être décodées avec précision à partir des gestes. La plupart des études sur la kinésique ont une approche holistique, c'est-à-dire qu'elles se concentrent sur la dynamique du corps entier. Le niveau général d'activité du mouvement et l'étendue spatiale semblent être des caractéristiques importantes pour reconnaître les émotions dans le mouvement corporel. Les mouvements de colère sont généralement larges, rapides et relativement saccadés tandis que les mouvements de peur et de tristesse sont moins énergiques, plus petits et plus lents (Lhommet & Marsella, 2015). Des postures et des mouvements fréquents pour exprimer émotions primaires seront présentés ci-dessous.

La colère se caractérise par la tête en arrière (pas de poitrine en arrière, pas de torsion abdominale), les bras levés vers l'avant et vers le haut, les épaules levées. Puis, le mouvement de la main (bras) latéralisé, les bras tendus vers l'avant, la plus grande amplitude de mouvement du coude, la plus grande vitesse d'extension du coude, le bras le plus élevé. Les caractéristiques de la peur sont la tête en arrière (pas de torsion abdominale), les bras levés vers l'avant, les épaules vers l'avant. Le mouvement caractéristique est les bras tendus sur les côtés. La joie se caractérise par la tête en arrière (pas de poitrine en avant), les bras levés au-dessus des épaules et droits au niveau des coudes, les épaules levées. Le mouvement des mains pour exprimer la joie se caractérise par la vitesse maximale des muscles fléchisseurs et extenseurs du coude et les bras tendus vers l'avant. Les caractéristiques de la tristesse sont la tête en avant, la poitrine en avant (pas de torsion abdominale), les bras sur le côté du torse, une posture effondrée. Puis, le temps de mouvement de la main le plus long, la plus petite amplitude du mouvement du coude, la vitesse du muscle extenseur du coude la plus faible. Le dégoût est reconnaissable par les épaules en avant, la tête en bas. La surprise se caractérise par la tête en arrière, la poitrine en arrière, la torsion abdominale, les bras levés avec les avant-bras droits (Walters and Walk, rapporté dans Coulson, 2004 ; Atkinson et al., 2004 ; Wallbott, 1998 ; Wallbott, 1998 ; Gross, Crane, Fredrickson, 2010, selon Lhommet & Marsella, 2015). La colère, la tristesse et la joie sont aussi bien reconnues lorsqu'elles sont exprimées avec une posture corporelle statique qu'avec l'expression faciale (Walters et Walk, rapportés dans Coulson, 2004, selon Lhommet & Marsella, 2015). Pourtant, certaines émotions ne sont pas reconnues (dégoût), et d'autres sont confuses (surprise et joie) à partir des postures (Coulson, 2004, selon Lhommet & Marsella, 2015).

## 2.5. Perception des émotions

La perception des émotions est la capacité de l'être humain à identifier et à reconnaître les émotions d'autrui (Kerkeni, 2020). La reconnaissance des émotions peut être auditive (à partir de la voix) et visuelle (à partir de l'expression faciale et posturale). Scherer (2003) a confirmé que le pourcentage de reconnaissance visuelle des émotions (75 %) est supérieur au pourcentage de reconnaissance auditive (55% à 65%). Cela peut indiquer que le stimulus vocal produit des modèles moins stables que les configurations des muscles faciaux. De plus, cela pourrait suggérer que les états émotionnels d'une même famille émotionnelle diffèrent davantage les uns des autres dans les expressions vocales que dans les expressions visuelles.

L'étude de Scherer (2003) démontre que les émotions les plus reconnues à partir de la voix, ce sont la colère et l'expression neutre avec une précision d'environ 70%. Elles sont suivies de tristesse (65%), de peur (50%) et de joie (43%). L'émotion la moins reconnue est celle du dégoût avec un score de 31%. Bänziger (2004) a confirmé que ces résultats concordent avec les résultats de la grande majorité des études. La tristesse et la colère sont les catégories émotionnelles les mieux reconnues. La peur et la joie sont moins bien reconnues. Les expressions de dégoût sont systématiquement très mal reconnues. Lee & al. (2004) ont aussi effectué des tests de reconnaissance auditive des émotions. Les participants ont reconnu les émotions avec une précision de 68.3%. La joie était l'émotion la plus difficile à distinguer (le dégoût n'était pas dans l'étude). Il y avait principalement des confusions entre l'expression neutre/la tristesse et la colère/la joie. Bänziger (2004) ajoute qu'il existe une forte corrélation positive entre l'intensité de peur et l'intensité de tristesse. Les expressions qui reçoivent des jugements élevés de peur reçoivent également des jugements élevés de tristesse. Cette corrélation suggère la possibilité d'une confusion entre la perception de la tristesse et celle de la peur. L'intensité de tristesse est corrélée négativement à l'intensité de colère et de joie.

Selon Scherer (2003), l'émotion la plus reconnue de l'expression faciale est celle de la joie avec une précision de 92%, puis la surprise de 83%, la tristesse de 77%, le dégoût de 74%. Les émotions les moins reconnues sont la peur et la colère autour de 70%. Scherer (2003) a souligné que les émotions moins bien reconnues sur le plan des manifestations vocales - notamment le dégoût et la joie - sont très bien reconnues sur le plan des manifestations faciales. Cette observation indique que certaines manifestations - telles que les expressions de la joie et du dégoût - sont préférentiellement communiquées par le canal facial/visuel alors que d'autres manifestations - telles que les expressions de la colère et de la tristesse - sont de préférence communiquées par le canal vocal/auditif.

Il existe des recherches qui ne se préoccupent pas de la reconnaissance des catégories émotionnelles, mais elles se focalisent sur les états émotionnels. Certains états émotionnels sont toutefois moins bien discriminés que d'autres. Les expressions d'anxiété sont surtout confondues avec les autres expressions faiblement intenses (la joie calme, l'irritation et la tristesse déprimée). Elles sont également confondues avec les expressions de peur panique et de désespoir (activation émotionnelle forte). Les expressions de tristesse déprimée et de rage sont bien discriminées à partir des paramètres acoustiques. Les expressions de joie exaltée sont confondues avec les autres états fortement intenses (la rage, la peur panique, le désespoir) et avec l'irritation. Les trois états faiblement intenses - l'anxiété (la peur anxieuse), la joie calme

et la tristesse déprimée - sont relativement bien distinguées par l'analyse mais partiellement confondues entre elles. Les trois états fortement intenses - la panique, le désespoir et l'exaltation - sont moins bien distinguées. On les confond entre eux mais on les confond également avec l'état d'anxiété (notamment quand il s'agit de l'état de désespoir et de panique), il arrive de confondre l'état d'exaltation avec celui de la rage et de l'irritation. La rage et l'irritation sont facilement confondues, mais elles sont quand même mieux discriminées par l'analyse que la panique, le désespoir et l'exaltation (Bänziger, 2004).

### 2.5.1. Perception des émotions dans une langue seconde

Les auditeurs peuvent souvent identifier correctement l'état émotionnel d'un locuteur, même s'ils ne comprennent pas la langue parlée, en s'appuyant sur les informations prosodiques et visuelles. De nombreuses études ont montré une grande précision dans la reconnaissance vocale des émotions dans les cultures très différentes (Bhatara & al., 2016). Il en va de même pour la reconnaissance visuelle des émotions. Ekman (1999) a confirmé que les mimiques faciales des émotions de base sont les mêmes partout dans le monde. Dans son étude, les expressions faciales des Américains ont été montrées aux Néo-Guinéens et vice versa. Les deux groupes ont reconnu avec succès presque toutes les émotions présentées. Cela prouve l'universalité de l'expression faciale des émotions. Des recherches de Scherer (2003) montrent que les expressions vocales et faciales des émotions des membres d'une culture seront reconnues par les membres d'autres cultures avec une précision élevée de 66 %. C'est la preuve qu'il existe des principes (biophysiques) universaux dans toutes les cultures qui relient certaines caractéristiques vocales à certaines émotions. Même les confusions dans la reconnaissance étaient similaires dans tous les pays. L'étude de Vuletić (1976) démontre que les étrangers peuvent comprendre le contenu exprimé par l'intonation. Il a effectué le test de compréhension de l'intonation. Les locuteurs croates et les locuteurs français ont écouté les phrases prononcées en croate. Les deux groupes ont obtenu des résultats élevés et les différences entre les groupes étaient minimales. Cependant, les locuteurs français ont obtenu le résultat un peu plus faible (Vuletić, 1976). Bien que les émotions soient reconnues dans différentes cultures, de nombreuses études démontrent un avantage au sein du groupe. Cela signifie que les locuteurs natifs obtiennent généralement de meilleurs résultats que les locuteurs non natifs (Bhatara & al., 2016). Par ailleurs, plus la différence entre la langue des participants et la langue dans laquelle les exemples auditifs ont été prononcés est grande, plus la précision

de la reconnaissance est faible. Cela suggère l'existence de schémas paralinguistiques spécifiques à la culture et/ou à la langue dans l'expression vocale des émotions (Scherer, 2003). Vuletić (1976) confirme qu'il y a une partie de l'intonation qui est caractéristique d'une langue particulière et une autre partie de l'intonation que tout le monde comprendra. Elfenbein & Ambady (2003, selon Bhatara & al., 2016) ont proposé une théorie du dialecte suggérant que la façon dont les émotions sont communiquées varie entre les groupes. Ainsi, il existe des aspects universels et des aspects culturellement spécifiques de la représentation des émotions qui modulent la perception des émotions. Néanmoins, plusieurs études ont comparé l'identification des émotions à partir de la prosodie émotionnelle chez les apprenants d'une langue et chez les locuteurs natifs de cette langue. Les résultats étaient variés. Dans les études comparant des groupes avec et sans expérience en langue seconde, il y avait également des facteurs de confusion et/ou des résultats contradictoires. La raison peut résider dans le fait que différentes émotions peuvent différer dans leur universalité d'expression (Bhatara & al., 2016). Frontera & Paone (2019) ont examiné la reconnaissance auditive de la prosodie émotionnelle dans la langue seconde. L'étude a impliqué des Russes qui apprennent l'italien et des Italiens qui ne connaissent pas la russe. Les stimuli étaient en russe et en italien. Les deux groupes ont identifié les émotions en italien avec un niveau de précision similaire (sauf la joie). Cependant, les locuteurs italiens ont atteint un niveau de précision nettement inférieur dans l'identification des émotions en russe. Les résultats suggèrent que la connaissance de la langue seconde pourrait contribuer à la reconnaissance des émotions de la langue seconde. En revanche, l'étude de Bhatara & al. (2016) a révélé qu'une meilleure connaissance de la langue seconde peut parfois être associée à une identification moins précise des émotions positives (mais non à l'interprétation des émotions négatives). Ils ont examiné l'effet de l'expérience langagière sur la perception auditive des émotions dans une langue seconde au sein d'un même groupe culturel. Des locuteurs natifs de français, aux différents niveaux de compétence en anglais, ont identifié des émotions à partir des expressions vocales en anglais. Les émotions positives et négatives ont été observées séparément. Les résultats ont confirmé que l'expérience langagière n'affecte pas la perception des émotions globalement, mais elle a un effet négatif sur la reconnaissance des émotions positives.

Elfenbein & Ambady (2002, selon Bhatara & al., 2016) ont confirmé que, dans toutes les études, la colère et la tristesse étaient les émotions les plus reconnues de manière interculturelle tandis que la joie était la moins reconnue dans les cultures. L'étude de Frontera & Paone (2019) confirme ce fait parce que les locuteurs russes ont identifié avec succès toutes

les émotions sauf la joie dans les stimuli italiens (leur langue seconde). De plus, l'étude de Bhatara & al. (2016) démontre qu'une meilleure connaissance de la langue seconde peut parfois être associée à une identification moins précise des émotions positives. Sauter & al. (2010, selon Bhatara & al., 2016) ont montré que les émotions positives, comme la joie, peuvent être culturellement plus spécifiques que les émotions négatives. Si la fonction des émotions positives est de faciliter la cohésion du groupe, cela pourrait expliquer pourquoi elles sont plus reconnaissables pour les membres du même groupe que pour les étrangers. De même, Ekman (selon Bhatara & al., 2016) suggère que les émotions positives peuvent être évolutivement moins importantes que les émotions négatives. Dans l'étude de Frontera & Paone (2019), dans les stimuli russes, les locuteurs italiens (qui ne connaissaient pas la russe) confondaient souvent la tristesse avec l'expression neutre et l'expression de la surprise ainsi que celle de la joie avec la peur et la surprise. Dans l'étude de Vuletić (1976) la différence la plus significative entre les locuteurs croates et les locuteurs français (qui ne connaissaient pas le croate) est présente dans la compréhension de la colère. Les locuteurs croates ont reconnu cette émotion bien mieux que les locuteurs français. Ce résultat est incompatible avec le fait mentionné que la colère est l'une des émotions les plus reconnues de manière interculturelle (Elfenbein & Ambady, 2002, selon Bhatara & al., 2016). Le décalage peut être expliquée par les caractéristiques phonétiques des langues croate et française. La langue française est beaucoup plus tendue que la langue croate. La plus grande tension de prononciation dans l'expression de la colère en langue croate s'identifie à la tension normale en langue française. La différence est également évidente dans la compréhension de la joie et de l'ironie. Il n'y a aucune différence dans la compréhension de l'affirmation neutre, de l'interrogation et de la tristesse (Vuletić, 1976). Contrairement à la reconnaissance auditive, dans la reconnaissance visuelle la joie est très bien identifiée. Dans l'étude de Ekman (1999) la joie était l'émotion la plus reconnue. Cela peut s'expliquer par l'affirmation de Scherer (2003) selon laquelle la joie est très bien reconnue sur le plan facial. La seule difficulté pour les participants était la confusion entre la peur et la surprise. Ces deux émotions ne se distinguent pas facilement l'une de l'autre dans toutes les cultures parce qu'elles se manifestent par presque la même expression faciale. La peur et la surprise se manifestent sur les visages par des sourcils levés (Ekman, 1999).

### 3. Étude

#### 3.1. Objectif et hypothèse

L'objectif de ce mémoire est d'observer le taux de réussite en reconnaissance de la prosodie émotionnelle en croate (langue maternelle) et en français (langue étrangère) par les locuteurs natifs croates. On essaiera de souligner l'importance des moyens prosodiques dans la communication et de montrer que la compréhension du contenu de la parole transcende les frontières linguistiques grâce à la prosodie.

Conformément à ce qui est décrit précédemment, on suppose que :

- 1) Les locuteurs croates reconnaîtront les émotions avec le taux de réussite plus ou moins égal dans les enregistrements en langue croate et dans les enregistrements en langue française. On s'attend à un petit avantage dans la reconnaissance des émotions dans la langue maternelle.
- 2) Les locuteurs croates non-francophones et les locuteurs croates francophones reconnaîtront les émotions avec un taux de réussite à peu près égal. Dans les enregistrements en français, on s'attend à un petit avantage en faveur des locuteurs croates francophones.
- 3) Le taux de reconnaissance des émotions sera plus élevé dans les enregistrements vidéo (audiovisuels) que dans les enregistrements audio (croates et français).
- 4) Les locuteurs croates qui pratiquent le chant ou jouent d'un instrument reconnaîtront les émotions, dans les enregistrements en croate et français, avec plus de succès que les locuteurs croates qui ne les pratiquent pas.

#### 3.2. Participants

L'étude implique 48 participants qui sont des locuteurs natifs de la langue croate : 35 femmes (73%) et 13 hommes (27%). Ils ont entre 19 et 28 ans. L'âge moyen est celui de 24 ans. Tous les participants sont étudiants. Au moment du déroulement de cette étude, la plupart des participants étaient en Master (75%) - Master 1 (33%), Master 2 (27%) ou en année supplémentaire en Master (15%). Les autres participants étaient en Licence (25%) - Licence 1 (2%), Licence 2 (2%), Licence 3 (8%) ou en année supplémentaire en Licence (13%). De tous

les participants, 13 participants chantent ou jouent d'un instrument tandis que 35 participants ne chantent ni ne jouent d'aucun instrument musical.

Les participants sont répartis en deux groupes. Le groupe A comprend les étudiants (de toute filière d'études universitaires sauf les études françaises) qui n'ont jamais appris le français, c'est-à-dire les locuteurs croates non-francophones. Le groupe B comprend les étudiants de français qui sont par conséquent locuteurs croates francophones. Chaque groupe est composé de 24 participants. Dans le groupe A il y avait 17 femmes et 7 hommes tandis que dans le groupe B il y avait 18 femmes et 6 hommes. Les étudiants non-francophones (groupe A) fréquentaient différentes facultés et différentes filières. Trois étudiants faisaient les études à la Faculté d'Économie - les filières économie et économie d'entreprise (2 participants). Huit étudiants faisaient les études à la Faculté d'Agriculture. Ils suivent des filières agroécologie, agronomie (2 participants), vinification, biotechnologie microbienne. La plupart des étudiants, treize étudiants, venaient de la Faculté de Philosophie et Lettres. Ils suivaient des filières ou des combinaisons de filières suivantes : archéologie et anthropologie, études croates et phonétique, allemand et bibliothéconomie, communication et sociologie, communication (3 participants), psychologie (2 participants), espagnol et portugais, études croates et espagnol, anglais et espagnol, études croates et italien. Les étudiants francophones (groupe B) apprennent le français depuis différentes périodes, entre 4 et 19 ans. La durée moyenne d'apprentissage du français est de 9 ans.

### 3.3. Instruments

Pour mener cette étude, nous avons créé un questionnaire (une enquête) à l'aide d'outil électronique *Google Forms*. Le questionnaire a été rédigé en croate. Il fallait 2 à 5 minutes pour le remplir. Dans la partie introductive de l'enquête les conditions de participation étaient présentées. On a présenté l'objectif de l'étude. Il y est souligné que le questionnaire était anonyme (identité des participants est inconnue), que toutes les données collectées seraient traitées au niveau du groupe et seraient utilisées exclusivement à des fins d'étude. Il y est expliqué que la participation était facultative et qu'il était possible de quitter à tout moment. Dans cette partie, il y avait aussi l'adresse électronique du chercheur afin que les participants puissent contacter le chercheur s'ils avaient des questions supplémentaires. Les participants ont donné leur consentement pour participer à l'étude en appuyant sur la touche *Suivant*. Ainsi, ils

passaient à l'enquête. À la fin, les participants devaient appuyer sur la touche *Envoyer* pour enregistrer leurs réponses.

Le questionnaire est composé de deux parties. La première partie contient 7 questions qui révèlent des informations générales sur les participants. Il y a 5 questions du type QCM<sup>2</sup> (il faut choisir une réponse parmi plusieurs options proposées) tandis que 2 questions exigent d'écrire des réponses courtes. La première question est : « Quel est votre sexe ? » Les réponses proposées sont « homme » et « femme ». La deuxième question est « Quel est votre âge ? » Les réponses proposées sont les numéros 18 à 28. La troisième question est « À quel groupe appartenez-vous ? » Les participants doivent choisir entre : « je suis étudiant(e) ; je n'ai jamais appris le français (A) » et « je suis étudiant(e) de français (B) ». La quatrième question s'adresse uniquement aux participants appartenant au groupe A. Les étudiants qui n'étudient pas le français doivent écrire une réponse courte à la question : « Qu'est-ce que vous étudiez ? » La cinquième question est « En quelle année d'études êtes-vous actuellement ? » Les réponses proposées sont « Licence 1 », « Licence 2 », « Licence 3 », « année supplémentaire en Licence », « Master 1 », « Master 2 » et « année supplémentaire en Master ». La sixième question est destinée aux participants du groupe B. On demande aux étudiants de français : « Depuis combien d'années apprenez-vous le français ? » Il faut répondre brièvement par un chiffre. La septième question est « Est-ce que vous pratiquez le chant (en chœur...) ou/et jouez d'un instrument ? » Les participants doivent choisir entre « oui » et « non ».

La deuxième partie du questionnaire se compose en 16 enregistrements - 8 enregistrements (audio et audiovisuels) en croate et 8 enregistrements (audio et audiovisuels) en français. Les enregistrements audio ne comprennent que des stimuli auditifs. Les enregistrements audiovisuels sont des vidéos. Ils comportent à la fois des stimuli auditifs et des stimuli visuels. Tous les enregistrements ont été tirés de séries et de films croates et français. Ainsi, le corpus utilisé est « acté » ou simulé. Chaque enregistrement dure 4 à 6 secondes. Sous chaque enregistrement, la question suivante est posée : « Quelle émotion est exprimée dans l'enregistrement X (numéro d'enregistrement) ? ». Puis, 4 réponses sont proposées : « colère », « joie », « peur » et « tristesse ». Les participants doivent écouter (et regarder) chaque enregistrement et choisir l'émotion qui y est exprimée. Les enregistrements sont divisés en 4 groupes. Au début, 4 enregistrements audio en croate sont présentés. Puis, on a 4 enregistrements audio en français. Ensuite, on présente 4 enregistrements audiovisuels en

---

<sup>2</sup> Le questionnaire à choix multiple

croate. À la fin, on confronte les étudiants avec 4 enregistrements audiovisuels en français. Les quatre états émotionnels (colère, joie, peur, tristesse), proposés dans les réponses, sont représentées à chaque groupe. Dans chaque groupe, les états émotionnels se succèdent dans un ordre différent. Par exemple, dans le premier groupe l'état de joie est exprimé dans le premier enregistrement tandis que dans le deuxième groupe l'état de joie est exprimé dans le quatrième enregistrement.

### 3.4. Procédure

Le questionnaire a été distribué aux étudiants par courriel et par réseaux sociaux. La recherche a été menée à distance. Nous avons attendu quelques jours pour que les étudiants, qui voulaient participer, remplissent l'enquête. Une fois qu'un nombre suffisant de participants avaient répondu au questionnaire, les données ont été traitées à l'aide d'un tableur *Excel*. Après avoir rempli l'enquête, les participants n'ont reçu aucune rétroaction. Mais chaque participant pouvait contacter individuellement le chercheur s'il souhaitait connaître les résultats de l'étude.

### 3.5. Résultats

Dans ce chapitre, nous allons présenter les résultats de notre enquête. Toutes les données obtenues sont présentées dans 16 tableaux ci-dessous. Les résultats sont divisés suivant 4 questions de base.

La première question est « Est-ce que les locuteurs croates reconnaîtront les émotions avec plus de succès dans les enregistrements en croate ou dans les enregistrements en français ? » Le *Tableau 1* montre les erreurs que les locuteurs croates (tous les participants) ont faites lors de la reconnaissance des émotions dans les enregistrements en croate. Les données pour les enregistrements audio et les données pour les enregistrements audiovisuels ont été observées séparément. La deuxième colonne (à gauche) du tableau répertorie l'état émotionnel présenté dans chaque enregistrement. Dans la colonne suivante se trouvent les états émotionnels que les participants ont confondus avec l'état émotionnel exprimé. Le numéro entre parenthèses indique combien de participants ont confondu cette émotion avec l'émotion exprimée. S'il n'y a pas de numéro, cela signifie qu'un seul participant a confondu cette émotion avec l'émotion exprimée. La troisième colonne contient le nombre d'erreurs pour chaque

enregistrement. Par exemple, le numéro 3 signifie que 3 participants (sur 48) ont choisi l'émotion erronée pour cet enregistrement. Dans la colonne précédente, nous voyons quelle est l'erreur/ quelles sont les erreurs. La dernière colonne indique le nombre total d'erreurs dans les enregistrements audio croates et le nombre total d'erreurs dans les enregistrements audiovisuels en croate. Le nombre d'erreurs pour tous les enregistrements en croate varie entre 0 et 3. Dans les enregistrements audio croates, l'émotion la plus reconnue est celle de la joie. Tous les participants l'ont reconnue. La peur est l'émotion la moins reconnue. Un participant l'a confondu avec la joie et l'autre avec la tristesse. La colère et la tristesse ont chacune une erreur. Dans les enregistrements audiovisuels en croate, la joie est l'émotion la moins reconnue. 2 participants l'ont confondu avec la colère et un participant avec la tristesse. 3 autres émotions (colère, peur, tristesse) ont chacune une erreur. Il n'est donc pas possible de dire quelle émotion est la plus reconnue. Le nombre total d'erreurs dans les enregistrements audio croate est de 4 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en croate est de 6.

Le *Tableau 2* montre les erreurs que les locuteurs croates (tous les participants) ont faites lors de la reconnaissance des émotions dans les enregistrements en français. Le *Tableau 2* a toutes les parties identiques au *Tableau 1*. Le nombre d'erreurs pour tous les enregistrements en français varie entre 0 et 8. Dans les enregistrements audio français, l'émotion la plus reconnue est encore celle de la joie. Tous les participants l'ont reconnue. La colère est très bien reconnue, avec une seule erreur. La tristesse comporte 5 erreurs, dont 4 sont en confusion avec la peur. La peur est l'émotion la moins reconnue. Un participant l'a confondu avec la joie et 7 participants l'ont confondu avec la tristesse. Dans les enregistrements audiovisuels en français, la colère est l'émotion la plus reconnue. Tous les participants l'ont reconnue. La peur est aussi l'émotion la moins reconnue. 5 participants l'ont confondu avec la tristesse. La joie a 2 erreurs alors que la tristesse a une erreur. Le nombre total d'erreurs dans les enregistrements audio français est de 14 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en français est de 8.

Le *Tableau 3* compare le taux de reconnaissance des émotions dans les enregistrements en croate et dans les enregistrements en français par les locuteurs croates (tous les participants). Sur 3 lignes on voit le nombre total d'erreurs, le nombre de réponses correctes (sur 384) et le taux de reconnaissance (le pourcentage de réussite en reconnaissance). Les données des enregistrements audio et des enregistrements audiovisuels sont additionnées et observées ensemble. Le nombre total d'erreurs dans les enregistrements en croate est de 10 tandis que dans les enregistrements en français est de 22. Dans les enregistrements en croate on a 374

réponses correctes tandis que dans les enregistrements en français il y en a 362. Le taux de reconnaissance des émotions dans les enregistrements en croate est de 97.39%. Le taux de reconnaissance des émotions dans les enregistrements en français est 94.27%. Donc, les locuteurs croates reconnaissent les émotions dans les enregistrements en croate avec d'environ 3 % plus de succès que dans les enregistrements en français.

Tableau 1 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en croate (audio et audiovisuels observés séparément) par les locuteurs croates.

Enregistrements en croate				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	joie	1	4
	joie	-	0	
	peur	joie tristesse	2	
	tristesse	colère	1	
Audiovisuels	colère	joie	1	6
	joie	colère (2) tristesse	3	
	peur	tristesse	1	
	tristesse	peur	1	

Tableau 2 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels observés séparément) par les locuteurs croates.

Enregistrements en français				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	joie	1	14
	joie	-	0	
	peur	joie tristesse (7)	8	
	tristesse	colère peur (4)	5	
Audiovisuels	colère	-	0	8
	joie	colère	2	
	peur	tristesse	5	
	tristesse	joie	1	

Tableau 3 : Le taux de reconnaissance des émotions dans les enregistrements en croate par rapport à celui de reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels inclus) par les locuteurs croates.

	Enregistrements en croate	Enregistrements en français
Nombre total d'erreurs	10	22
Nombre de réponses correctes	374/384	362/384
Taux de reconnaissance	97.39%	94.27%

La deuxième question est « Qui reconnaîtra avec le plus de succès les émotions - des locuteurs croates non-francophones ou des locuteurs croates francophones ? » Bien qu'il soit plus logique de vérifier la différence entre ces deux groupes uniquement dans la reconnaissance des émotions dans les enregistrements en français, nous vérifierons également les enregistrements en croate. Nous allons commencer par des enregistrements en croate. Le *Tableau 4* montre les erreurs que les locuteurs croates non-francophones (groupe A) ont faites lors de la reconnaissance des émotions dans les enregistrements en croate. Ce tableau ressemble

au *Tableau 1*. La seule différence est qu'il ne se réfère pas à tous les participants, mais au ceux du groupe A. Le nombre d'erreurs pour tous les enregistrements en croate varie entre 0 et 2. Dans les enregistrements audio en croate, l'émotion la plus reconnue (reconnue par tous) est celle de la joie. Les 3 autres émotions (colère, peur, tristesse) ont remporté chacune une erreur. Dans les enregistrements audiovisuels en croate, la joie est l'émotion la moins reconnue avec 2 erreurs. Les 3 autres émotions ont remporté chacune une erreur. Le nombre total d'erreurs faites par le groupe A dans les enregistrements audio en croate est de 3 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en croate est de 5.

Le *Tableau 5* ressemble au *Tableau 4*, sauf qu'il se réfère aux résultats du groupe B. Le *Tableau 5* montre les erreurs que les locuteurs croates francophones ont faites lors de la reconnaissance des émotions dans les enregistrements en croate. Dans les enregistrements audio en croate, les étudiants de français n'ont fait qu'une seule erreur. Un participant a confondu la peur avec la tristesse. Dans les enregistrements audiovisuels en croate, on trouve également une seule erreur. La joie est remplacée par la colère. Donc, le nombre total d'erreurs faites par les participants du groupe B dans les enregistrements audio en croate est de 1 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en croate est également de 1.

Le *Tableau 6* compare le taux de reconnaissance des émotions dans les enregistrements en croate par les locuteurs croates non-francophones et par les locuteurs croates francophones. Les données des enregistrements audio et des enregistrements audiovisuels sont observées ensemble. Le *Tableau 6* contient les mêmes éléments que le *Tableau 3*. La différence est que ce tableau ne compare pas les résultats de tous les participants dans 2 types d'enregistrements (en croate et en français), mais compare les résultats de 2 groupes de participants dans les mêmes enregistrements (en croate). Le nombre total d'erreurs faites par les locuteurs croates non-francophones est de 8 tandis que le nombre total d'erreurs faites par les locuteurs croates francophones est de 2. Le groupe A détient 184 (sur 192) réponses correctes tandis que le groupe B détient 190 réponses correctes. Le taux de reconnaissance des émotions dans les enregistrements en croate par le groupe A est 95.83%. Le taux de reconnaissance des émotions dans les enregistrements en croate par le groupe B est 98.95%. Donc, les locuteurs croates francophones reconnaissent les émotions dans les enregistrements en croate avec 3% plus de succès que les locuteurs croates non-francophones.

Tableau 4 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en croate (audio et audiovisuels observés séparément) par les locuteurs croates non-francophones (groupe A).

Enregistrements en croate - Non-francophones				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	joie	1	3
	joie	-	0	
	peur	joie	1	
	tristesse	colère	1	
Audiovisuels	colère	joie	1	5
	joie	colère tristesse	2	
	peur	tristesse	1	
	tristesse	peur	1	

Tableau 5 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en croate (audio et audiovisuels observés séparément) par les locuteurs croates francophones (groupe B).

Enregistrements en croate - Francophones				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	-	0	1
	joie	-	0	
	peur	tristesse	1	
	tristesse	-	0	
Audiovisuels	colère	-	0	1
	joie	colère	1	
	peur	-	0	
	tristesse	-	0	

Tableau 6 : Le taux de reconnaissance des émotions dans les enregistrements en croate (audio et audiovisuels inclus) par les locuteurs croates non-francophones (groupe A) et par les locuteurs croates francophones (groupe B).

Enregistrements en croate		
	Non-francophones	Francophones
Nombre total d'erreurs	8	2
Nombre de réponses correctes	184/192	190/192
Taux de reconnaissance	95.83%	98.95%

Comparons les résultats du groupe A et du groupe B mais cette fois-ci à partir des enregistrements en français. Le *Tableau 7* montre les erreurs que les locuteurs croates non-francophones ont faites lors de la reconnaissance des émotions dans les enregistrements en français. Le tableau est le même que le *Tableau 4*. Le nombre d'erreurs pour tous les enregistrements en croate varie entre 0 et 5. Dans les enregistrements audio en français, l'émotion la mieux reconnue (reconnue par tous) est celle de la joie tandis que l'émotion la moins reconnue est celle de la peur (5 erreurs). Dans les enregistrements audiovisuels en français, les émotions que tous les membres du groupe A ont reconnues sont la colère et la tristesse. La peur est l'émotion la moins reconnue (3 erreurs). Le nombre total d'erreurs faites par le groupe A dans les enregistrements audio en français est de 10 et dans les enregistrements audiovisuels en français est de 4.

Le *Tableau 8*, ainsi que le *Tableau 5*, se réfèrent au groupe B. Le *Tableau 8* montre les erreurs que les locuteurs croates francophones ont faites lors de la reconnaissance des émotions dans les enregistrements en français. Le nombre d'erreurs pour tous les enregistrements en français varie entre 0 et 3. Dans les enregistrements audio en français, tous les étudiants de français ont reconnu la colère et la joie. La peur n'a pas été reconnue par 3 participants. Dans les enregistrements audiovisuels en français, tous les membres du groupe B ont reconnu la colère. La peur est à nouveau l'émotion la moins reconnue avec 2 erreurs. Le nombre total d'erreurs faites par le groupe B dans les enregistrements audio en français est de 4 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en français est aussi de 4.

Le *Tableau 9* compare le taux de reconnaissance correcte des émotions dans les enregistrements en français par les locuteurs croates non-francophones et par les locuteurs croates francophones. Les données des enregistrements audio et des enregistrements

audiovisuels sont observées ensemble. Le nombre total d'erreurs faites par les locuteurs croates non-francophones est de 14 alors que le nombre total d'erreurs faites par les locuteurs croates francophones est de 8. Le groupe A détient 178 (sur 192) réponses correctes tandis que le groupe B détient 184 réponses correctes. La reconnaissance correcte des émotions dans les enregistrements en français par le groupe A est 92.70%. La reconnaissance correcte des émotions dans les enregistrements en français par le groupe B est 95.83%. Tout comme dans les enregistrements en croate, les locuteurs croates francophones reconnaissent les émotions dans les enregistrements en français avec 3% plus de succès que les locuteurs croates non-francophones.

Tableau 7 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels observés séparément) par les locuteurs croates non-francophones (groupe A).

Enregistrements en français - Non-francophones				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	joie	1	10
	joie	-	0	
	peur	tristesse	5	
	tristesse	colère peur (3)	4	
Audiovisuels	colère	-	0	4
	joie	colère	1	
	peur	tristesse	3	
	tristesse	-	0	

Tableau 8 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels observés séparément) par les locuteurs croates francophones (groupe B).

Enregistrements en français - Francophones				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	-	0	4
	joie	-	0	
	peur	joie tristesse (2)	3	
	tristesse	peur	1	
Audiovisuels	colère	-	0	4
	joie	colère	1	
	peur	tristesse	2	
	tristesse	joie	1	

Tableau 9 : Le taux de reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels inclus) par les locuteurs croates non-francophones (groupe A) et par les locuteurs croates francophones (groupe B).

Enregistrements en français		
	Non-francophones	Francophones
Nombre total d'erreurs	14	8
Nombre de réponses correctes	178/192	184/192
Taux de reconnaissance	92.70%	95.83%

La troisième question est : « Est-ce que la réussite en reconnaissance des émotions sera plus élevée dans les enregistrements audio ou dans les enregistrements audiovisuels ? » Le *Tableau 10* compare le taux de reconnaissance des émotions dans les enregistrements audio et dans les enregistrements audiovisuels par les locuteurs croates francophones et non-francophones (tous les participants). Dans la troisième ligne, on trouve le nombre d'erreurs faites par les participants lors de la reconnaissance des émotions dans les enregistrements en croate et en français séparément. Nous avons déjà observé ces données dans les tableaux

précédents. Les chiffres de la troisième ligne sont additionnés aux ceux présentés en quatrième ligne. À gauche, on trouve le nombre total d'erreurs faites par les participants lors de la reconnaissance des émotions dans les enregistrements audio en croate et en français. Ce nombre est 18. À droite, on voit que le nombre total d'erreurs faites par les participants lors de la reconnaissance des émotions dans les enregistrements audiovisuels en croate et en français est 14. Dans la cinquième ligne, il y figure le nombre de réponses correctes. Dans les enregistrements audio il y a 366 réponses correctes (sur 384) alors que dans les enregistrements audiovisuels il y en a 370. Dans la dernière ligne, on voit que le taux de reconnaissance des émotions, par les deux groupes, dans les enregistrements audio (en croate et en français) est de 95.31%, et dans les enregistrements audiovisuels est de 96.35%. Ainsi, la reconnaissance des émotions est d'environ 1 % plus élevée dans les enregistrements audiovisuels que dans les enregistrements audio.

Tableau 10 : Le taux de reconnaissance des émotions dans les enregistrements audio par rapport au taux de reconnaissance des émotions dans les enregistrements audiovisuels par les locuteurs croates.

	Enregistrements audio		Enregistrements audiovisuels	
	croates	français	croates	français
Nombre d'erreurs	4	14	6	8
Nombre total d'erreurs	18		14	
Nombre de réponses correctes	366/384		370/384	
Taux de reconnaissance	95.31%		96.35%	

La quatrième question est « Qui reconnaîtra avec le plus de succès les émotions - les locuteurs croates qui ne chantent ni ne jouent d'instrument ou les locuteurs croates qui chantent et/ou jouent d'un instrument ? » Nous ferons une comparaison entre ces deux groupes aussi bien dans les enregistrements en croate que dans les enregistrements en français. Il a déjà été mentionné que sur 48 participants, 35 ne chantent ni ne jouent d'instrument tandis que 13 chantent et/ou jouent d'un instrument. Sur les 35 participants qui ne chantent ni ne jouent d'instrument, 11 sont non-francophones tandis que 6 sont francophones. Sur les 13 participants qui chantent et/ou jouent d'un instrument, 4 sont non-francophones alors que 9 sont francophones. Le *Tableau 11* montre les erreurs que les locuteurs croates qui ne chantent ni ne

jouent d'instrument (non-musiciens) ont faites lors de la reconnaissance des émotions dans les enregistrements en croate. Le nombre d'erreurs pour tous les enregistrements en croate varie entre 0 et 3. Dans les enregistrements audio en croate, l'émotion la mieux reconnue (reconnue par tous) est celle de la joie. Les 3 autres émotions (colère, peur, tristesse) ont remporté chacune une erreur. Dans les enregistrements audiovisuels en croate, la joie est l'émotion la moins reconnue avec 3 erreurs. Les 3 autres émotions ont remporté chacune une erreur. Le nombre total d'erreurs faites par les locuteurs croates qui ne chantent ni ne jouent d'instrument dans les enregistrements audio croates est de 3 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en croate est de 6.

Le *Tableau 12* montre les erreurs que les locuteurs croates qui chantent et/ou jouent d'un instrument (musiciens) ont faites lors de la reconnaissance des émotions dans les enregistrements en croate. Dans les enregistrements audio en croate, ils n'ont fait qu'une seule erreur. Un participant a confondu la peur avec la tristesse. Dans les enregistrements audiovisuels en croate, on ne trouve aucune erreur. Le nombre total d'erreurs faites par les locuteurs croates qui chantent et/ou jouent d'un instrument dans les enregistrements audio en croate est de 1 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en croate est de 0.

Le *Tableau 13* compare le taux de reconnaissance des émotions dans les enregistrements en croate par les locuteurs croates qui ne chantent ni ne jouent d'instrument et par les locuteurs croates qui chantent et/ou jouent d'un instrument. Les données des enregistrements audio et des enregistrements audiovisuels sont observées ensemble. Le nombre total d'erreurs faites par les participants qui ne chantent ni ne jouent d'instrument est de 9 tandis que le nombre total d'erreurs faites par les participants qui chantent et/ou jouent d'un instrument est de 1. Les non-musiciens ont 271 réponses correctes (sur 280) tandis que les musiciens ont 103 réponses correctes (sur 104). Le taux de reconnaissance des émotions dans les enregistrements en croate par les non-musiciens est 96.78%. Le taux de reconnaissance des émotions dans les enregistrements en croate par les musiciens est 99.03%. Donc, les locuteurs croates qui chantent et/ou jouent d'un instrument reconnaissent les émotions dans les enregistrements en croate avec 2,5% plus de succès que les locuteurs croates qui ne chantent ni ne jouent d'instrument.

Tableau 11 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en croate (audio et audiovisuels observés séparément) par les locuteurs croates qui ne chantent ni ne jouent d'instrument (non-musiciens).

Enregistrements en croate - Non-musiciens				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	joie	1	3
	joie	-	0	
	peur	joie	1	
	tristesse	colère	1	
Audiovisuels	colère	joie	1	6
	joie	colère (2) tristesse	3	
	peur	tristesse	1	
	tristesse	peur	1	

Tableau 12 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en croate (audio et audiovisuels observés séparément) par les locuteurs croates qui chantent et/ou jouent d'un instrument (musiciens).

Enregistrements en croate - Musiciens				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	-	0	1
	joie	-	0	
	peur	tristesse	1	
	tristesse	-	0	
Audiovisuels	colère	-	0	0
	joie	-	0	
	peur	-	0	
	tristesse	-	0	

Tableau 13 : Le taux de reconnaissance des émotions dans les enregistrements en croate (audio et audiovisuels inclus) par les locuteurs croates qui ne chantent ni ne jouent d'instrument (non-musiciens) et par les locuteurs croates qui chantent et/ou jouent d'un instrument (musiciens).

Enregistrements en croate		
	Non-musiciens	Musiciens
Nombre total d'erreurs	9	1
Nombre de réponses correctes	271/280	103/104
Taux de reconnaissance	96.78%	99.03%

Nous allons maintenant refaire une comparaison de ces deux groupes de participants, mais cette fois dans les enregistrements en français. Le *Tableau 14* montre les erreurs que les locuteurs croates qui ne chantent ni ne jouent d'instrument (non-musiciens) ont faites lors de la reconnaissance des émotions dans les enregistrements en français. Le nombre d'erreurs pour tous les enregistrements en français varie entre 0 et 7. Dans les enregistrements audio en français, l'émotion la plus reconnue (reconnue par tous) est celle de la joie tandis que l'émotion la moins reconnue est celle de la peur (7 erreurs). Dans les enregistrements audiovisuels en français, l'émotion la plus reconnue (reconnue par tous) est celle de la colère alors que l'émotion la moins reconnue est encore une fois celle de la peur (5 erreurs). Le nombre total d'erreurs faites par les locuteurs croates qui ne chantent ni ne jouent d'instrument dans les enregistrements audio en français est de 13 et dans les enregistrements audiovisuels en français est de 8.

Le *Tableau 15* montre les erreurs que les locuteurs croates qui chantent et/ou jouent d'un instrument (musiciens) ont faites lors de la reconnaissance des émotions dans les enregistrements en français. Dans les enregistrements audio en français, un seul participant a fait l'erreur de confondre la peur avec la tristesse. Dans les enregistrements audiovisuels en français, il n'y a pas d'erreurs. Le nombre total d'erreurs faites par les locuteurs croates qui chantent et/ou jouent d'un instrument dans les enregistrements audio en français est de 1 tandis que le nombre total d'erreurs dans les enregistrements audiovisuels en français est de 0.

Le *Tableau 16* compare le taux de reconnaissance des émotions dans les enregistrements en français par les locuteurs croates qui ne chantent ni ne jouent d'instrument et par les locuteurs

croates qui chantent et/ou jouent d'un instrument. Les données des enregistrements audio et des enregistrements audiovisuels sont observées ensemble. Le nombre total d'erreurs faites par les participants qui ne chantent ni ne jouent d'instrument est de 21 alors que le nombre total d'erreurs faites par les participants qui chantent et/ou jouent d'un instrument est de 1. Les non-musiciens ont 259 réponses correctes (sur 280) tandis que les musiciens ont 103 réponses correctes (sur 104). Le taux de reconnaissance des émotions dans les enregistrements en français par les non-musiciens est 92.5%. Le taux de reconnaissance des émotions dans les enregistrements en français par les musiciens est 99.03%. Donc, comme dans les enregistrements en croate, les locuteurs croates qui chantent et/ou jouent d'un instrument reconnaissent les émotions dans les enregistrements en français avec 6,5% plus de succès que les locuteurs croates qui ne chantent ni ne jouent d'instrument.

Tableau 14 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels observés séparément) par les locuteurs croates qui ne chantent ni ne jouent d'instrument (non-musiciens).

Enregistrements en français - Non-musiciens				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	joie	1	13
	joie	-	0	
	peur	joie tristesse (6)	7	
	tristesse	colère peur (4)	5	
Audiovisuels	colère	-	0	8
	joie	colère	2	
	peur	tristesse	5	
	tristesse	joie	1	

Tableau 15 : Les erreurs dans la reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels observés séparément) par les locuteurs croates qui chantent et/ou jouent d'un instrument (musiciens).

Enregistrements en français - Musiciens				
	Émotion exprimée	Émotion confondue	Nombre d'erreurs	Nombre total d'erreurs
Audio	colère	-	0	1
	joie	-	0	
	peur	tristesse	1	
	tristesse	-	0	
Audiovisuels	colère	-	0	0
	joie	-	0	
	peur	-	0	
	tristesse	-	0	

Tableau 16 : Le taux de reconnaissance des émotions dans les enregistrements en français (audio et audiovisuels inclus) par les locuteurs croates qui ne chantent ni ne jouent d'instrument (non-musiciens) et par les locuteurs croates qui chantent et/ou jouent d'un instrument (musiciens).

Enregistrements en français		
	Non-musiciens	Musiciens
Nombre total d'erreurs	21	1
Nombre de réponses correctes	259/280	103/104
Taux de reconnaissance	92.5%	99.03%

### 3.6. Discussion

Cette recherche s'est intéressée au taux de reconnaissance de la prosodie émotionnelle dans la langue maternelle (croate) et dans la langue étrangère (française). 48 locuteurs natifs de la langue croate ont participé à cette étude. Tous les participants ont été étudiants entre 19 et 28 ans. Les participants ont été divisés en deux groupes de 24 élèves chacun. Le groupe A a été composé de locuteurs croates qui n'ont jamais appris le français (non-francophones). Le groupe B a été composé de locuteurs croates étudiant le français (francophones). Pour fins d'étude, nous avons rédigé un questionnaire en ligne composé de enregistrements en croate et enregistrements en français. Il y a eu des enregistrements audio et audiovisuels dans les deux langues. Les participants ont dû écouter (et regarder) chaque enregistrement et choisir l'émotion (entre 4 émotions proposées) qui y était exprimée. Avant de commencer la recherche, nous avons formulé 4 hypothèses.

La première hypothèse : les locuteurs croates reconnaîtront les émotions à proportion similaire dans les enregistrements en langue croate et dans les enregistrements en langue française. On s'attendait à un petit avantage dans la reconnaissance des émotions dans la langue maternelle. Cette hypothèse a été complètement confirmée. Le taux de reconnaissance des émotions dans les enregistrements en croate est de 97.39% et dans les enregistrements en français est 94.27%. Ainsi, le taux de reconnaissance des émotions est très élevé (supérieure à 90%) dans les deux langues, mais il est de 3 % plus élevé dans la langue maternelle.

Dans les enregistrements audio en croate, l'émotion la mieux reconnue a été la joie (0 erreurs). Ce résultat ne confirme pas recherches précédentes (Scherer, 2003 ; Bänziger, 2004 ; Lee & al., 2004) dans lesquelles la joie avait été l'une des émotions les plus difficiles à distinguer sur le plan des expressions vocales. La peur (2 erreurs) a été l'émotion la moins reconnue. Ce fait n'est pas surprenant car la peur est généralement l'une des émotions le moins bien reconnue sur le plan des expressions vocales (Scherer, 2003 ; Bänziger, 2004). Un participant a confondu la peur avec la joie et l'autre avec la tristesse. La confusion entre la peur et la tristesse est logique parce qu'il existe une forte corrélation positive entre l'intensité de peur et l'intensité de tristesse (Bänziger, 2004). Ces émotions sont généralement associées à l'intensité faible. La confusion entre la peur et la joie peut s'expliquer par le fait que la peur produit les variations de l'intensité les plus fortes (Scherer & Oshinsky, 1977, selon Kerkeni, 2020). Ainsi, la peur peut parfois avoir l'intensité aussi forte que la joie. Le ton haut et le débit de la parole rapide sont les caractéristiques des deux émotions. Un participant a confondu la colère (1 erreur) avec la joie.

Ces deux émotions se manifestent par le ton haut, l'intensité forte et le débit de la parole rapide. Un participant a confondu la tristesse (1 erreur) avec la colère. Cette confusion n'est pas courante car la tristesse et la colère ont toutes les caractéristiques vocales opposées. Dans les enregistrements audiovisuels en croate, la joie (3 erreurs) a été l'émotion la moins reconnue. Ce fait contredit à nouveau les recherches précédentes dans lesquelles la joie avait été très bien reconnue à partir des expressions faciales (Scherer, 2003 ; Bänziger, 2004). 2 participants l'ont confondue avec la colère et un participant avec la tristesse. La colère se manifeste par les caractéristiques vocales presque identiques à la joie. La tristesse se manifeste par les caractéristiques vocales complètement opposées à la joie. Mais par rapport aux expressions faciales, ces deux émotions sont assez différentes de la joie. Sur le visage, la joie se distingue de toutes les autres émotions par un sourire. 3 autres émotions ont eu chacune une erreur. La colère s'est confondue avec la joie, la peur avec la tristesse et la tristesse avec la peur.

Dans les enregistrements audio en français, l'émotion la plus reconnue a été encore celle de la joie (0 erreurs). Ceci est encore une fois incompatible avec les recherches précédentes qui avaient révélé que la joie était la moins reconnue dans les cultures différentes (Elfenbein & Ambady, 2002, selon Bhatara & al., 2016 ; Frontera & Paone, 2019 ; Bhatara & al., 2016 ; Sauter & al., 2010, selon Bhatara & al., 2016). Les recherches citées avaient montré que les manifestations des émotions positives, comme la joie, pouvaient être culturellement plus spécifiques que celles des émotions négatives. L'émotion la moins reconnue a été celle de la peur (8 erreurs). 7 participants ont confondu la peur avec la tristesse. Cette confusion ne surprend pas pour la raison déjà mentionnée. 1 participant a confondu la peur avec la joie. Cette confusion est également survenue dans l'étude de Frontera & Paone (2019). La manifestation de la tristesse (5 erreurs) est aussi assez mal reconnue. 4 participants l'ont confondue avec celle de la peur et un participant avec celle de la colère. En revanche, dans les recherches précédentes, la tristesse avait été l'une des émotions les plus reconnues de manière interculturelle (Elfenbein & Ambady, 2002, selon Bhatara & al., 2016 ; Vuletić, 1976). La colère a été confondue par un seul participant avec la joie. Ces résultats sont cohérents avec les résultats de recherche de Elfenbein & Ambady (2002, selon Bhatara & al., 2016) dans laquelle la colère (avec la tristesse) avait été l'émotion la plus reconnue de manière interculturelle. Au contraire, dans l'étude de Vuletić (1976), la colère avait été l'émotion la moins reconnue par les locuteurs non natifs. Dans les enregistrements audiovisuels en français, la colère (0 erreurs) a été l'émotion la plus reconnue. La peur a été l'émotion la moins reconnue. 5 participants l'ont confondue avec la tristesse. Cela suggère que la peur et la tristesse, en plus de par leurs

caractéristiques vocales similaires, sont également exprimées par les expressions faciales et posturales similaires. La joie a été 2 fois confondue avec la colère. Un participant a confondu la tristesse avec la joie. Il est difficile de confondre tristesse et joie, surtout dans l'enregistrement audiovisuel.

En observant les résultats dans les enregistrements en croate et en français, les enregistrements audio et audiovisuels, la confusion la plus courante est entre la peur et la tristesse. Elle est particulièrement évidente dans les enregistrements en français. Il y a 19 erreurs de ce type. Nous avons déjà dit qu'il existe une forte corrélation positive entre l'intensité de peur et l'intensité de tristesse qui explique cette confusion (Bänziger, 2004). Cependant, la peur peut être associée à l'intensité faible et à l'intensité forte (Scherer & Oshinsky, 1977, selon Kerkeni, 2020). De plus, la peur et la tristesse produisent d'autres caractéristiques vocales différentes. La peur fait produire le ton haut tandis que la tristesse fait produire le ton bas. La peur rend le débit de la parole rapide alors que la tristesse rend le débit de la parole lente. On pourrait conclure que la peur nous fait produire non seulement de fortes variations de l'intensité, mais aussi de fortes variations du ton et du tempo (débit). Il est possible que les gens associent parfois la peur au ton bas ou au débit de la parole lent. La deuxième confusion la plus courante, dans tous les enregistrements, est celle entre la colère et la joie. C'était aussi l'une des confusions principales dans la recherche de Lee & al. (2004). Dans notre étude, cette erreur s'est produite 9 fois. On a déjà mentionné que la colère et la joie partagent les mêmes 3 caractéristiques vocales principales. C'est pourquoi cette confusion est attendue.

La deuxième hypothèse : les locuteurs croates non-francophones et les locuteurs croates francophones reconnaîtront les émotions à un taux de réussite similaire. Dans les enregistrements en français, on s'attendait à un petit avantage en faveur des locuteurs croates francophones. Cette hypothèse a été confirmée. Le taux de reconnaissance des émotions dans les enregistrements en croate par le groupe A est 95.83% et par le groupe B est 98.95%. Le taux de reconnaissance des émotions dans les enregistrements en français par le groupe A est 92.70% et par le groupe B est 95.83%. Notre hypothèse de départ est confirmée, on constate que les deux groupes reconnaissent les émotions avec un pourcentage similaire (très élevé, supérieure à 90%) à la fois dans la langue maternelle et dans la langue étrangère. Les locuteurs croates francophones reconnaissent les émotions dans les enregistrements en français de 3% mieux que les locuteurs croates non-francophones. Le fait surprenant est que les locuteurs croates francophones reconnaissent les émotions dans les enregistrements en croate aussi de 3% mieux que les locuteurs croates non-francophones. Cela pourrait indiquer que la

connaissance d'une langue étrangère améliore également les compétences en langue maternelle. La question de la fiabilité des résultats doit certainement être posée. L'échantillon des participants n'est pas grand. D'autant plus, il est à noter que parmi les locuteurs croates non-francophones, il y avait 4 participants qui étudiaient les langues romanes. Deux participants étudiaient l'espagnol, un participant étudiait l'espagnol et le portugais et un participant étudiait l'italien. Il est possible que ces participants aient mieux reconnu les émotions dans les enregistrements en français, même s'ils ne connaissaient pas le français, parce qu'ils connaissaient des langues de la même famille que le français. Grâce à eux, le groupe A aurait pu avoir un meilleur résultat final.

La troisième hypothèse : le taux de reconnaissance des émotions sera plus élevé dans les enregistrements audiovisuels que dans les enregistrements audio. Cette hypothèse a été confirmée. Le taux de reconnaissance des émotions, par les deux groupes, dans les enregistrements audio (croates et français) est de 95.31%, et dans les enregistrements audiovisuels est 96.35%. Ainsi, le taux de reconnaissance des émotions est d'environ 1 % plus élevé dans les enregistrements audiovisuels que dans les enregistrements audio. Cependant, la différence de résultats entre les enregistrements audio et les enregistrements audiovisuels est très faible. Il a été attendu que la réussite en reconnaissance des émotions dans les enregistrements audiovisuels soit beaucoup plus élevée que dans les enregistrements audio. La reconnaissance dans les enregistrements audiovisuels aurait dû être beaucoup plus facile car elle avait impliqué à la fois des stimuli auditifs et des stimuli visuels. De plus, Scherer (2003) a confirmé que le pourcentage de reconnaissance visuelle des émotions était supérieur au pourcentage de reconnaissance auditive. Il est possible que certains participants eussent des problèmes pour ouvrir les vidéos, surtout s'ils remplissaient le questionnaire via les téléphones portables. La connexion Internet stable est nécessaire pour ouvrir les enregistrements audiovisuels car ils contiennent l'image avec le son. Dans ce cas, il est probable que des participants aient choisi au hasard n'importe quelle émotion. Dans les enregistrements audiovisuels en français (8 erreurs) le taux de reconnaissance des émotions est plus élevé que dans les enregistrements audio en français (14 erreurs). Cependant, dans les enregistrements audiovisuels en croate (6 erreurs) le taux de reconnaissance des émotions est plus faible que dans les enregistrements audio en croate (4 erreurs). Cette information inattendue pourrait à nouveau être attribuée au problème technique d'ouverture des enregistrements audiovisuels. Sinon, les participants auraient pu être confus en se concentrant sur plusieurs sens (ouïe et vue) en même temps. Mais cela ne concernait que la langue maternelle. Dans la langue étrangère

(dans l'interprétation des émotions dans les enregistrements en français), le résultat est meilleur lorsque les participants écoutent et regardent. Il est également possible qu'il y eût problème dans les enregistrements eux-mêmes. Comme on a déjà dit, les enregistrements ont été extraits des séries et des films. Ainsi, les émotions ont été mises en scène par différents acteurs, utilisant différents matériaux lexicaux. Certains exemples peut-être manifestent une certaine émotion de manière plus crédible que d'autres.

La quatrième hypothèse : les locuteurs croates qui pratiquent le chant ou jouent d'un instrument (musiciens) reconnaîtront les émotions, dans les enregistrements en croate et en français, avec plus de réussite que les locuteurs croates qui ne font pas de musique (non-musiciens). Cette hypothèse, qui a été confirmée, découle du fait que la formation musicale produit un effet positif sur la perception et sur la production des sonorités dans une langue étrangère. Des résultats de nombreuses études ont révélé une meilleure perception, discrimination et catégorisation des tons et des sons dans une langue seconde chez les musiciens par rapport aux non-musiciens (Besson & Chobert, 2013). Outre le fait de faciliter l'apprentissage d'une langue étrangère, pratiquer la musique influe sur le système auditif et améliore les compétences cognitives auditives dans la langue maternelle. Les stimulations musicales sont également utilisées dans la rééducation de l'audition et de la parole par la méthode verbo-tonale. Leur objectif est d'aider le sujet à développer une parole complexe et naturelle en corrigeant la production à l'aide des valeurs musicales (rythme et mélodie). De tout cela, nous avons conclu que les musiciens auraient une écoute « plus fine » et reconnaîtraient mieux les émotions dans les deux langues. Le taux de reconnaissance des émotions dans les enregistrements en croate par les non-musiciens est 96.78%. Le taux de reconnaissance des émotions dans les enregistrements en croate par les musiciens est 99.03%. Donc, les locuteurs croates qui chantent et/ou jouent d'un instrument reconnaissent les émotions dans les enregistrements en croate de 2,5% plus correctement que les locuteurs croates qui ne chantent ni ne jouent d'instrument. Le taux de reconnaissance des émotions dans les enregistrements en français par les non-musiciens est 92.5%. Le taux de reconnaissance des émotions dans les enregistrements en français par les musiciens est 99.03%. Ainsi, les locuteurs croates qui chantent et/ou jouent d'un instrument reconnaissent les émotions dans les enregistrements en français 6,5% plus précisément que les locuteurs croates qui ne chantent ni ne jouent d'instrument. On peut constater que les participants qui chantent et/ou jouent d'un instrument obtiennent de meilleurs résultats dans la reconnaissance des émotions dans les deux langues. Leur avantage est particulièrement visible dans les enregistrements en français. De plus, cette

différence de pourcentage des non-musiciens et des musiciens dans les enregistrements en français est la plus grande de toutes les autres différences de pourcentage dans cette étude. À partir de ces résultats, on pourrait conclure que pratiquer la musique améliore la reconnaissance des émotions dans la langue maternelle, et surtout dans la langue étrangère. On pourrait également conclure que pratiquer la musique améliore la reconnaissance des émotions dans une langue étrangère encore plus que la connaissance de cette langue. Ces résultats ne sont pas complètement fiables car l'étude n'a pas inclus le même nombre de participants qui pratiquaient la musique et ceux qui ne la pratiquaient pas. 35 participants ne chantaient ni ne jouaient d'un instrument tandis que 13 chantaient et/ou jouaient d'un instrument. De plus, parmi les non-musiciens il y avait plus de locuteurs croates non-francophones tandis que parmi les musiciens il y avait plus de locuteurs croates francophones. Cela aurait pu contribuer aux meilleurs résultats des musiciens dans l'interprétation des émotions dans les enregistrements en français. Dans la présente étude, l'accent a été mis sur le lien entre le niveau de connaissance d'une certaine langue et la reconnaissance des émotions dans cette langue. Il serait intéressant de mener des recherches plus approfondies sur la contribution de la musique à la reconnaissance des émotions dans une langue seconde.

#### **4. Conclusion**

Les hypothèses de départ ont été confirmées par cette étude. Conformément à la première hypothèse, la réussite en reconnaissance des émotions par les locuteurs croates est légèrement plus élevée en croate (langue maternelle) qu'en français (langue étrangère). La confusion la plus courante, dans les deux langues, est possible entre la reconnaissance de la peur et celle de la tristesse. Les participants ont fait cette erreur surtout dans les enregistrements en français. La deuxième confusion la plus courante est entre la reconnaissance de la colère et celle de la joie. Ces deux confusions étaient également fréquentes dans des recherches précédentes. Comme nous l'avons supposé dans la deuxième hypothèse, les locuteurs croates non-francophones et les locuteurs croates francophones reconnaissent les émotions, dans les deux langues, avec un taux de réussite à peu près égal. Comme prévu, dans les enregistrements en français, les francophones reconnaissaient les émotions avec un peu plus de succès que les non-francophones. Il est intéressant que, dans les enregistrements en croate, les locuteurs croates francophones aient également obtenu un résultat légèrement meilleur. Selon la troisième hypothèse, le taux de reconnaissance des émotions est un peu plus élevé dans les

enregistrements audiovisuels que dans les enregistrements audio. Néanmoins, on s'attendait à ce que l'avantage en faveur des enregistrements audiovisuels soit encore plus grand. En français, la reconnaissance des émotions est meilleure dans les enregistrements audiovisuels, mais en croate, elle est légèrement meilleure dans les enregistrements audio. Conformément à la quatrième hypothèse, les locuteurs croates qui chantent et/ou jouent d'un instrument obtiennent de meilleurs résultats dans la reconnaissance des émotions que les locuteurs croates qui ni chantent ni jouent d'instrument musical, dans les deux langues. L'avantage est encore plus grand dans les résultats des interprétations d'enregistrements en français. Les résultats montrent que pratiquer la musique pourrait améliorer la reconnaissance des émotions dans une langue étrangère plus que la connaissance de cette langue. Enfin, il est important de souligner que le taux de reconnaissance des émotions est généralement très élevé (supérieure à 90%) dans les deux langues par tous les participants.

Cette étude confirme que la compréhension des états émotionnels exclusivement par les moyens prosodiques est très élevée. Les expressions orales des émotions de base représentent un moyen expressif puissant et efficace, sans ambiguïté et presque universellement compréhensible (Vuletić, 1986). Donc, les auditeurs sont susceptibles de reconnaître les émotions même s'ils ne comprennent pas la langue parlée. La connaissance de la langue apporte un petit avantage. Ces résultats révèlent la grande importance des moyens prosodiques dans la communication. Le texte segmental ne suffit pas pour une compréhension complète. On a encore besoin de moyens prosodiques (Vuletić, 1976). Les moyens prosodiques introduisent la subjectivité et l'affectivité dans le discours et contribuent à la compréhension du message (Guberina, 1952). Il est important de souligner que le contenu exprimé par les moyens prosodiques a tendance de remporter sur le contenu exprimé par le matériel lexical. Autrement dit, si les mots sont en conflit avec les moyens prosodiques, les messages que nous envoyons par la voie prosodique prendront le dessus (Vuletić, 1976).

## Bibliographie

- Bänziger, T. (2004). *Communication vocale des émotions : perception de l'expression vocale et attributions émotionnelles*. Mémoire de master. Université de Genève.
- Barić, E., Lončarić, M., Malić, D., Pavešić, S., Peti, M., Zečević, V., Znika., M. (1995). *Hrvatska gramatika*. Zagreb: Školska knjiga.
- Besson, M., Chobert, J (2013). Musical Expertise and Second Language Learning. *Brain Science*, 3 (2), 923-940.
- Bhatara, A., Laukka, P., Boll-Avetisyan, N., Granjon, L., Elfenbein, H. A., Bänziger, T. (2016). Second Language Ability and Emotional Prosody Perception. *Plos One*.
- Ekman, P. (1977). Facial Expression. In A. Siegman & S. Feldstein (eds.), *Nonverbal Behavior and Communication*. New Jersey: Lawrence Erlbaum Association.
- Ekman, P. (1999). Facial Expressions. In T. Dalgleish & M. Power, *Handbook of Cognition and Emotion*. New York: John Wiley & Sons Ltd.
- Frankol, D., Pavelin Lešić, B. (2016). Mouvement, geste, parole : les valeurs de la langue parlée dans la correction phonétique. *La technologie aux limites de l'humain en didactique des langues*. Mons, Belgique: CIPA. 155-166.
- Frontera, M., Paone, E. (2019). *Emotional prosody perception in Italian as a second language*. ExLing 2019: Proceedings of 10th International Conference of Experimental Linguistics. Lisbon.
- Galisson, R., Coste, D. (1976). *Dictionnaire de didactique des langues*. Paris : Hachette.
- Guberina, P. (1938-39). Govorni jezik i pisani jezik. *Hrvatski jezik*, 1, 6-7, 114-124.
- Guberina, P. (1952). *Zvuk i pokret u jeziku: problemi ljudskog izraza*. Zagreb: Matica hrvatska.
- Guberina, P. (1954 [1939].). *Valeur logique et valeur stylistique des propositions complexes : théorie générale et application au français*, Zagreb : Epoha. [1<sup>ère</sup> édition sous le titre : *Valeur logique et valeur stylistique des propositions complexes en français et en croate*].
- Guberina, P. (1995). Filozofija verbotonalnog sistema. *Filologija*, 24-25, 157-164.
- Hess, W., Möbius, B., Pätzold, M. (1993). Analysis and synthesis of German F0 contours by means of Fujisaki's model. *Speech Communication*, 13, 53-61.
- Kerkeni, L. (2020). *Analyse acoustique de la voix pour la détection des émotions du locuteur*. Thèse de doctorat en cotutelle de l'Université de Sousse et Le Mans Université.
- Juslin, P. N., Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, 129 (5), 770-814.
- Lacheret, A. (2011a). La prosodie au cœur du verbal. *Rééducation orthophonique*, 246, 87-104.

- Lacheret, A. (2011b). Le corps en voix ou l'expression prosodique des émotions. *Evolutions Psychomotrices*, 23 (90), 25-37.
- Lee, C. M., Yildirim, S., Bulut, M., Kazemzadeh, A., Busso, C., Deng, Z., Lee, S., Narayanan, S. (2004). *Emotion recognition based on phoneme classes*. In Eighth International Conference on Spoken Language Processing.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Massachusetts, and London: The MIT Press.
- Lhommet, M., Marsella, S. C. (2015). Expressing emotion through posture and gesture. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford handbook of affective computing*. Oxford University Press, 273-285.
- Martinović, B., Pletikos Olof, E., Vlašić Duić, J. (2021). *Naglasak na naglasku*. Sveučilište Jurja Dobrile u Puli.
- Pavelin Lešić, B. (2013). *Vizualna obilježja govorenoga jezika*. Zagreb: FF press.
- Peppé, S. J. E. (2009). Why is prosody in speech-language pathology so difficult? *International Journal of Speech-Language Pathology*, 11, 258-271.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech communication*, 1-2, 227-256.
- Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, 57, 24-48.
- Silić, J., Pranjković, I. (2005). *Gramatika hrvatskoga jezika za gimnazije i visoka učilišta*. Zagreb: Školska knjiga.
- Simeon, R. (1969). *Enciklopedijski rječnik lingvističkih naziva : na 8 jezika : hrvatsko-srpski, latinski, ruski, njemački, engleski, francuski, talijanski, španjolski. 2, P-Ž*. Zagreb: Matica hrvatska.
- Škarić, I. (1991). Fonetika hrvatskoga književnog jezika. U: S. Babić, D. Brozović, M. Moguš, M. Pavešić, I. Škarić i S. Težak, *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika: nacrti za gramatiku*. Zagreb: HAZU, Globus, 71-378.
- Škarić, I. (2007). Fonetika hrvatskoga književnog jezika. U S. Babić, D. Brozović, I. Škarić i S. Težak, *Glasovi i oblici hrvatskoga književnog jezika*. Zagreb: Nakladni zavod Globus 17-157.
- Thiessen, E. D., Hill, E. A., Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, 7, 1, 53-71.
- Varošanec-Škarić, G. (2005). *Timbar*. Zagreb: FF press.
- Vaudable, C. (2012). *Analyse et reconnaissance des émotions lors de conversations de centres d'appels*. Thèse présentée pour obtenir le grade de Docteur de l'université Paris 11.
- Vuletić, B. (1976). *Fonetika književnosti*. Zagreb: Liber.
- Vuletić, B. (1986). Govorni izraz emocije. *Govor*, 3, 1, 33-38.
- Vuletić, B. (2007). *Lingvistika govora*. Zagreb: FF press.

Sitographie :

*Hrvatski jezični portal*. URL:

[https://hjp.znanje.hr/index.php?show=search\\_by\\_id&id=d15uXxI%3D](https://hjp.znanje.hr/index.php?show=search_by_id&id=d15uXxI%3D)

[La dernière date de la visite du site : le 12 mai 2022]

*Larousse*. URL: <https://www.larousse.fr/dictionnaires/francais/sonie/73462>

[La dernière date de la visite du site : le 19 juillet 2022]

*Poliklinika SUVAG*. URL: <http://www.suvag.hr/guberina/3/>

[La dernière date de la visite du site : le 9 juin 2022]