

[Page de garde]



**HEC MONTRÉAL**

**Stimulation haptique et sous-titres en contexte de multimédia  
hédonique: L'utilisateur peut-il être sur-stimulé?**

**par**  
**Félix Giroux**

**Sciences de la gestion  
(Option Expérience Utilisateur dans un Contexte d'Affaires)**

*Mémoire présenté en vue de l'obtention  
du grade de maîtrise ès sciences en gestion  
(M. Sc.)*

Août 2020  
© Félix Giroux, 2020



## Résumé

Ce mémoire par articles s'intéresse à l'influence de la stimulation haptique sur le multitâche en contexte de multimédia hédonique. Les multimédias traditionnels, incluant ceux affichant des sous-titres, sont de plus en plus adaptés à des expériences multisensorielles rehaussées de stimulations haptiques (ex. siège vibro-cinétique) dans le but d'améliorer la qualité d'expérience. Cependant, la capacité d'un auditoire à effectuer une activité multitâche avec la lecture de sous-titres lors de telles expériences est un sujet de recherche inexploré. Considérant que les sous-titres sont une partie intégrante de certains multimédias et parfois nécessaires à leur compréhension, il est important de comprendre si la stimulation haptique peut nuire à une activité multitâche avec ces sous-titres, et donc compromettre la qualité de l'expérience. Dans un premier temps, ce mémoire vise à comprendre les réactions émotionnelles, cognitives et comportementales face à la présence de sous-titres intralinguistiques présentés dans la langue natale d'un auditoire d'opéra cinématographique. Ensuite, nous cherchons à comprendre l'effet de la stimulation haptique sur la qualité d'expérience résultant d'une activité multitâche avec les sous-titres à l'étude. Afin d'atteindre ces objectifs de recherche, une étude expérimentale a été menée pour répondre aux questions suivantes: *Quel est l'effet de la présence de sous-titres sur la qualité d'une expérience d'opéra cinématographique? Dans quelle mesure la stimulation haptique peut dégrader la qualité de l'expérience résultant d'une activité multitâche avec des sous-titres dans un contexte de multimédia?* Dans un premier article, nos résultats suggèrent que l'auditoire utilise et apprécie les sous-titres qui lui permettent une meilleure compréhension des dialogues du multimédia et ce, sans affecter ses réactions émotionnelles au contenu visuel, ni son effort cognitif nécessaire à la tâche. Les résultats du second article montrent que la stimulation haptique n'influence pas l'attention visuelle naturellement portée aux sous-titres. Cependant, la stimulation haptique ne semble pas avoir eu d'effet modérateur positif sur les réactions émotionnelles de l'auditoire ainsi que l'appréciation subjective de l'expérience. Ce résultat suggère une potentielle sur-stimulation empêchant l'auditoire de traiter et réagir à l'ensemble des informations qui lui sont communiquées lors d'une expérience de multimédia. Il est

nécessaire de poursuivre la recherche afin de mieux comprendre l'impact de stimulations haptiques dans des activités multitâches traditionnelles avec des multimédias.

**Mots clés :** Multitâche, Sous-titres, Multimédia, Hédonique, Haptique, Qualité d'Expérience, Émotion, Expression Faciale, Charge Mentale, Oculométrie

## Table des matières

Résumé .....	iii
Table des matières .....	v
Liste des tableaux et des figures .....	vii
Liste des abréviations .....	viii
Avant-propos .....	ix
Remerciements .....	x
Introduction .....	1
Questions de recherche .....	3
Objectifs et contributions .....	4
Construits clés du mémoire .....	6
1. Qualité d'Expérience (QoE) .....	6
2. Comportement multitâche avec des sous-titres .....	8
3. L'Haptique dans les multimédias .....	9
Informations au sujet des articles .....	11
Présentation sommaire du premier article .....	12
Présentation sommaire du second article .....	13
Contributions et responsabilités individuelles .....	15
Structure du mémoire .....	18
Chapitre 1: Premier Article Hedonic multitasking: The effects of Instrumental subtitles during video watching .....	19
Abstract .....	19
1. Introduction .....	19
2. Methods .....	21
2.1. Subjects and Stimuli .....	21
2.2. Measurements and Processing .....	21
2.3. Statistical Analyses .....	22
3. Results .....	22
4. Discussion .....	24
References .....	26

Chapitre 2: Deuxième Article Measuring the Effect of Haptic Stimulation on audience's Quality of Experience during Hedonic Multimedia Multitasking context .....	29
Abstract .....	29
1. Introduction .....	30
2. Related work .....	34
2.1. Quality of Experience in multimedia and mulsemedia .....	34
2.2. Multitasking with subtitles in multimedia.....	36
2.3. Haptics in Multimedia.....	38
3. Methodology .....	40
3.1. Sample.....	41
3.2. Experimental Design.....	41
3.3. Audiovisual stimulus.....	41
3.4. HFVK Stimulus.....	43
3.5. Procedure.....	43
3.6. Measures.....	43
3.7. Statistical Analysis .....	46
4. Results .....	46
5. Discussion and Conclusion .....	51
References .....	56
Chaptire 3: Conclusion .....	65
Défis méthodologiques du projet .....	66
Rappel des questions de recherche et résultats.....	67
Contributions du mémoire.....	70
Implications pratiques .....	73
Limites et recherches futures.....	73
Bibliographie.....	75
Annexes.....	83
Annexe 1: Méthode de revue de littérature .....	83
Annexe 2: Articles de journaux liés au projet de recherche.....	85
Annexe 3: Support d'oculomètres et caméra .....	87

# Liste des tableaux et des figures

## Liste des figures

### Premier article

Figure 1 - One subject's aggregated fixations associated with automatic facial expression-based sadness. Fixations are displayed in 2D plans. Mean sadness frequency in each condition's zone of interest is indicated by the darkness of the colored circles. 24

### Deuxième article

Figure 2: Experimental settings and audiovisual stimulus 42

Figure 3: Scatterplot and histogram and showing the threshold determination process according to the Y-axis fixation location value between the two fixation's population. 44

Figure 4: One subject's aggregated fixations associated with automatic facial expression-based sadness. Fixations are displayed in 2D plans. The mean sadness frequency in each condition's zone of interest is indicated by the darkness of the colored circles. 49

## Liste des tables

### Introduction

Table 1: Contribution dans le projet de recherche et rédaction d'article 15

### Deuxième article

Table 2: Interview questions 46

Table 3: Overall descriptive statistic's results 47



## **Liste des abréviations**

AU: Action Unit

AFE`Automatic Facial Expression

EEG : Electroencéphalography / Électrpencéphalographie

EMG : Electromyography / Électromyographie

FACS: Facial Action Coding System

HCI: Human Computer Interaction

HFVK: High Fidelity Vibro Kinetic

IHM: Interaction Humain-Machine

NI: Native Intralingual

PCPD: Percentage Change in Pupil Diameter

QoE: Quality of Experience / Qualité d'Expérience

SAM: Self-Assessemnt Manikin

TV: Television

UX: User Experience / Expérience utilisateur



## **Avant-propos**

Ce mémoire en Expérience utilisateur en contexte d'affaires a été soumis avec l'autorisation de la direction administrative du programme de la Maîtrise ès Science en Gestion.

Le projet de recherche lié à ce mémoire a été approuvé par le comité d'éthique de la recherche (CER) de HEC Montréal. Deux articles issus du projet sont compris dans le mémoire avec le consentement des coauteurs.

Le premier article est un acte de conférence accepté et présenté en juin dernier, portant sur une partie des données du projet de recherche (Giroux et al., 2020). Cet article explore les effets neurophysiologiques d'une activité multitâche avec des sous-titres instrumentaux dans un contexte de multimédia hédonique.

Le second article utilise le jeu de données complet du projet de recherche. Cet article en préparation pour soumission à la revue scientifique *ACM Transactions on Multimedia Computing, Communications and Applications* mesure l'impact de stimulations haptiques sur une activité multitâche avec des sous-titres instrumentaux dans un contexte de multimédia hédonique.

Finalement, le projet de recherche a fait l'objet de plusieurs articles de journaux, ainsi que d'une entrevue radio (voir *Annexe 2: Articles de journaux liés au projet de recherche*)



## **Remerciements**

J'aimerais d'abord remercier mes codirecteurs, Pierre-Majorique Léger et Sylvain Sénécal pour leur implication tout au long de mon mémoire.

Je tiens aussi à remercier Jared Boasen qui a eu un rôle de mentor important tout au long de la maîtrise, en plus de me transmettre sa passion pour la recherche.

Merci à l'équipe de D-BOX Technologies, et spécialement à Jean-François Ménard ainsi que Michel Paquette pour leur implication dans les projets de recherche toujours plus stimulants.

Merci au Conseil de recherche en sciences naturelles et génie (CRSNG) ainsi qu'à Prompt de m'avoir soutenu financièrement durant la maîtrise.

Merci à l'équipe d'opérations du Tech3Lab, incluant les assistant(e)s de recherche ayant contribué à la mise en place du projet ainsi qu'à la collecte de données.

Merci aux premiers collègues que j'ai rencontrés tels que Éric, Félix, Robin, Paul, Mario, Florian, Charles, Ariane, Caroline, Gabrielle, Hasna, Audrey, Marianne, Arianne, et tous les autres qui ont contribué à un juste équilibre de vie étudiante saine.

Finalement, je tiens surtout à remercier mes parents et mon frère pour leur support et leurs encouragements durant ces deux dernières années.



## Introduction

Les multimédias sont omniprésents dans notre monde moderne. Durant le deuxième trimestre de l'année 2018, la population adulte américaine passait en moyenne 10 heures 24 minutes exposée à différentes formes de médias, dont 8 heures 13 minutes passées devant des multimédias (Nielsen, 2018). Étant communiqués à la télévision, ou sur une tablette numérique ou un téléphone mobile, les multimédias sont caractérisés par leur utilisation de plusieurs médias (ex. texte, son, vidéo) sollicitant les sens de la vision et de l'ouïe (Akthar et al., 2019). Dans la mesure où une diversité d'informations numériques est de plus en plus accessible via les technologies d'information, les auditoires de multimédias hédoniques adoptent de plus en plus des comportements multitâches. En effet, 45% des répondants à un sondage indiquent qu'ils écoutent la télévision en utilisant simultanément un deuxième écran (ex. appareil mobile) (Nielsen, 2018).

Traditionnellement, un comportement multitâche se manifeste lorsqu'on tente d'accomplir plus d'une tâche simultanément. Alors que ce comportement a souvent pour objectif d'être plus rapide et efficace, un coût est associé à l'alternance d'attention entre les différentes tâches (Gazzaley & Rosen, 2016). En effet, lorsque le cerveau humain s'engage dans une activité multitâche, il n'effectue pas les tâches simultanément, mais doit plutôt alterner rapidement entre le traitement d'une tâche à l'autre en raison de sa capacité attentionnelle limitée (Baddeley et al., 1969; Kahneman, 1973; Lang, 2000). Ainsi, la division d'attention nécessaire entre les tâches a généralement été associée à une diminution de performance dans les tâches impliquées contrairement à si elles sont réalisées individuellement sans interruption ou distraction (Gazzaley & Rosen, 2016). Avec le développement et l'accès grandissant aux technologies d'informations et communications, les comportements et les formes d'activités multitâches se multiplient.

Par exemple, face à l'omniprésence des multimédias dans notre quotidien, l'humain moderne s'expose de plus en plus à des activités multitâches hédonique impliquant différentes sources de médias (Jeong & Fishbein, 2007). Une activité multitâche avec des médias se produit lorsqu'un utilisateur est simultanément exposé à plus d'une forme de média (ex. médias sociaux, film, etc.) pouvant être communiquée par le même ou plus

d'un médium (Bardhi et al., 2010). Les consommateurs de multimédias hédoniques s'engagent dans des activités multitâches avec d'autres sources d'information ou sensations, dans le but de rendre l'activité plus agréable ou afin d'obtenir des informations supplémentaires (Jeong & Fishbein, 2007; Hwang et al., 2014). Par contre, la littérature sur les multimédias indique généralement qu'un comportement multitâche peut avoir un effet négatif sur la compréhension, la mémorisation et la performance liée au traitement d'un média principal, et que ces effets sont plus importants lorsque la ou les tâches secondaires ne sont pas pertinentes au média principal (Jeong & Hwang, 2016). Ainsi, lorsqu'une tâche secondaire n'est pas complémentaire ou nécessaire à la réalisation d'une tâche principale, une activité multitâche peut potentiellement dégrader une expérience avec un multimédia, même lorsque les tâches sont traitées par des modalités sensorielles différentes.

Cependant, des études antérieures suggèrent qu'il est important de distinguer une activité multitâche lorsqu'une tâche secondaire est non-pertinente, par rapport à lorsque qu'elle est pertinente à l'activité principale (Wang et al., 2015; Hwang & Jeong, 2018). Une tâche est pertinente lorsqu'elle est liée ou sert à accomplir le même objectif qu'une seconde tâche effectuée avec un autre média (Wang et al., 2015). Par exemple, les résultats d'étude récente ont montré que la lecture d'un article non pertinent à un visionnement d'une vidéo de nouvelles pouvait réduire la compréhension de la vidéo par rapport à une condition sans activité multitâche (Hwang & Jeong, 2018). Cependant, dans cette étude, une activité multitâche impliquant la lecture d'un article correspondant au même sujet que la vidéo n'a pas affecté la compréhension de cette dernière, sans toutefois l'améliorer (Hwang & Jeong, 2018)

Diviser notre attention visuelle entre l'image et les sous-titres d'un multimédia à la télévision est une activité multitâche courante pouvant être pertinente à une expérience hédonique, même si elle résulte naturellement en une réduction du contenu visuel pouvant être traité. Toutefois, en contexte hédonique, la littérature indique que les consommateurs de multimédia sont capables de diviser leur attention visuelle entre les images et les sous-titres de façon efficace, soit sans affecter la compréhension de l'image (Perego et al., 2010). D'autres études suggèrent que les sous-titres peuvent même réduire la charge

cognitive de l'auditoire d'un multimédia (D'Ydewalle & De Bruycker, 2007; Kruger et al., 2013, Kruger et al., 2014), mesurée à l'aide de pupillométrie et d'électroencéphalographie, ce qui peut s'expliquer par une amélioration de la compréhension des dialogues et de la trame narrative (Lavaur and Bairstow, 2011; Lee et al., 2013). Ainsi, lors d'un visionnement d'un multimédia sous-titré, l'auditoire peut s'engager dans une activité multitâche active contribuant à améliorer la qualité de l'expérience.

Cependant, de plus en plus de contenu de multimédia traditionnel comprenant des sous-titres est adapté à des environnements mulsemcia (i.e. pour *multiple sensorial multimedia*) impliquant des effets sensoriels au-delà de la vision et de l'ouïe (ex. stimulations haptiques) (Timmerer et al., 2014). Malgré les bienfaits de l'utilisation d'effets haptiques dans des expériences hédoniques comme une écoute de musique (Giroux et al., 2019), une étude récente suggère que l'ajout de stimulations haptiques à une activité multitâche hédonique avec des informations numériques textuelles peut dégrader l'expérience globale (Tchanou et al., 2019). Les résultats de cette dernière étude suggèrent que lors d'une expérience d'écoute de musique synchronisée avec des stimulations haptiques, la lecture simultanée d'informations textuelles sur un appareil mobile peut diminuer l'attention et les réactions émotionnelles de l'auditoire par rapport à la musique. Ainsi, cette étude suggère qu'afin de bénéficier d'une expérience multisensorielle optimale d'écoute de musique, un auditoire devrait éviter d'adopter un comportement multitâche avec une lecture d'informations textuelles. Toutefois, l'effet de la stimulation haptique sur une activité multitâche avec des sous-titres, pouvant être nécessaires à une expérience de multimédia hédonique, est un sujet de recherche inexploré.

## Questions de recherche

Dans ce mémoire, nous mesurons les réactions psychophysiologiques et comportementales de l'auditoire, en lien avec une expérience d'écoute d'opéra cinématique accompagnée de sous-titres instrumentaux, ainsi que de stimulations haptiques livrées par un siège vibro-cinétique. Plus précisément, nous cherchons à comprendre si l'ajout de stimulations haptiques dans une expérience de multimédia peut

distraire une activité multitâche complexe effectuée avec les sous-titres, et ainsi dégrader la qualité d'une expérience de multimédia hédonique. Nous tenterons de répondre à cette question en étudiant une activité multitâche avec des sous-titres en contexte d'opéra cinématique, dont le stimulus auditif est supporté par des stimulations haptiques livrées par un siège.

Dans le premier article, l'effet de la stimulation haptique a été exclu de la question de recherche, afin de mieux comprendre les réactions psycho physiologiques et comportementales d'un auditoire dans une activité multitâche avec les sous-titres du multimédia à l'étude. Ce premier article permet de répondre à la question de recherche suivante: *Quel est l'effet de la présence de sous-titres sur la qualité d'une expérience d'opéra cinématique?* Plus précisément, nous voulons voir si la présence de sous-titres instrumentaux influence l'attention visuelle portée aux sous-titres, les réactions émotionnelles, ainsi que la charge cognitive liées au traitement du stimulus audiovisuel, et l'appréciation subjective de l'expérience d'écoute d'opéra cinématique.

Dans le second article, nous voulions tester si l'effet de la stimulation haptique pouvait affecter le comportement naturel de lecture des sous-titres. En d'autres mots, nous voulions voir si la stimulation haptique pouvait agir comme distracteur à une activité multitâche avec les sous-titres. Ainsi, ce deuxième article permet de répondre à la question de recherche suivante: *Dans quelle mesure la stimulation haptique peut dégrader la qualité de l'expérience résultant d'une activité multitâche avec des sous-titres dans un contexte de multimédia?* Plus précisément, nous voulons voir si la livraison de stimulations haptiques influence l'attention visuelle portée aux sous-titres, ainsi que les réactions émotionnelles et l'appréciation subjective de l'expérience d'écoute d'opéra cinématique sous-titré.

## Objectifs et contributions

L'objectif de ce mémoire est de tester si l'effet de stimulations haptiques peut nuire à la qualité d'expérience résultant d'une activité multitâche avec les sous-titres d'un multimédia hédonique. La stimulation haptique consiste plus précisément en une combinaison d'effets vibro-tactiles et cinétiques (HFVK) livrés par des actuateurs intégrés

dans un siège de cinéma. Dans notre étude, la stimulation HFVK est synchronisée avec les éléments auditifs non-verbaux du multimédia. Quant au comportement multitâche avec les sous-titres, il réfère plus spécifiquement à l'attention visuelle de l'auditoire, mesurée par le temps de fixation portée à la zone des sous-titres à l'écran. Ces derniers étaient plus précisément des sous-titres intralinguistiques (i.e. même langue que les dialogues) présentés dans la langue natale de l'auditoire. Finalement, nous évaluons la qualité de l'expérience multisensorielle en mesurant l'appréciation subjective, ainsi que les réactions émotionnelles à l'aide d'une méthode de détection automatique d'expressions faciales captées durant le visionnement du multimédia.

Le contexte d'opéra cinématique était pertinent pour traiter cet objectif de recherche, en raison de l'importance du rôle des sous-titres dans ce type de multimédia. En effet, les sous-titres sont généralement pertinents pour les auditoires d'opéra, puisque les dialogues ne sont pas toujours faciles à comprendre, même lorsqu'ils sont chantés dans la langue natale de l'auditoire (Desblache, 2008; Orero & Matamala, 2007). Aussi, l'opéra en général est reconnu pour la capacité de son art total à générer des émotions chez l'auditoire par différents mécanismes artistiques (Scherer et al., 2019). Ainsi, le contexte d'étude est donc aussi pertinent afin d'évaluer les réactions émotionnelles induites par le multimédia.

Nous croyons que ce mémoire apporte des contributions importantes aux niveaux théorique, pratique, et méthodologique. Nos contributions théoriques avancent la recherche sur les multimédias, et plus précisément sur les sous-titres, ainsi que l'évaluation de la QoE. Nous apportons aussi une contribution pratique en informant les concepteurs d'expériences mulsemcia sur la potentielle sur-stimulation sensorielle lorsqu'un auditoire est exposée à une activité multitâche avec des sous-titres instrumentaux. Finalement, nous croyons apporter une contribution méthodologique importante à la recherche sur les méthodes d'évaluation de la QoE dans les multimédias, et plus généralement à la recherche en Interaction-Humain-Machine (IHM) utilisant des systèmes de détection automatique d'expressions faciales. Plus spécifiquement, nous montrons comment les mesures simultanées de données oculométriques peuvent ajouter une information contextuelle aux mesures d'émotions faciales par rapport à la tâche visuelle associée (ex. tâche de lecture de texte v.s tâche de visionnement d'image). Ainsi,

selon la littérature, cette pratique pourrait permettre de mieux interpréter les résultats de systèmes de détection automatique d'expressions faciales.

## Construits clés du mémoire

### Qualité d'Expérience (QoE)

Avec la production croissante de multimédias, la nécessité d'évaluer la qualité du matériel produit a gagné beaucoup d'importance dans l'industrie des télécommunications. Traditionnellement, la qualité des multimédias est évaluée en termes de performance et de caractéristiques du système (ex. délais de communication d'information, résolution d'image) (Varela et al., 2014). Cependant, durant les dernières décennies, la recherche en multimédia et en Interaction Humain-Machine (IHM) a accordé de plus en plus d'importance aux besoins des utilisateurs dans l'évaluation de la qualité des systèmes et de l'expérience globale (Wechsung & De Moor, 2014). Ainsi la Qualité d'Expérience (QoE) se définit par le degré d'appréciation des utilisateurs d'applications et services en fonction de l'utilisabilité et le plaisir que ces derniers procurent, ainsi que la personnalité et l'état mental de l'utilisateur (Brunnström et al., 2003). Selon Reiter et al. (2014), la QoE peut être influencée par des facteurs techniques, humains et contextuels. Les facteurs techniques liés à la qualité d'un système ou service incluent les caractéristiques du contenu audiovisuel, du média (ex. fréquence d'image), du réseau de communication (ex. délais), ou du médium de communication (ex. résolution d'écran). Les facteurs humains représentent les caractéristiques des utilisateurs, qui peuvent être soit constantes durant l'interaction avec un système (ex. son statut socioculturel, ses connaissances, ses aptitudes pertinentes à une tâche), ou plus dynamiques (ex. ses émotions, sa motivation, son attention).

La recherche en multimédia accorde de plus en plus d'importance aux facteurs humains dans l'évaluation de la QoE (Wechsung & De Moor, 2014). Par exemple, les émotions générées dans une interaction avec un système peuvent avoir une influence importante sur la perception de la QoE pour un utilisateur de multimédia en contexte de divertissement (Schleicher & Antons, 2014). Cependant, les facteurs humains comme les émotions sont souvent complexes à évaluer puisqu'ils peuvent se manifester de manière consciente ou

inconsciente. Par exemple, le traitement d'un stimulus qui induit des émotions peut déclencher des réponses physiologiques conscientes qu'un humain peut ressentir, ou inconscientes (ex. micromouvements faciaux) ne pouvant être détectées que par des mesures objectives captées avec des outils neurophysiologiques (Walla, 2018).

Avec les avancées en recherche en UX, l'utilisation de méthodes objectives prend de plus en plus d'importance dans l'évaluation de la QoE en recherche sur le multimédia (Mandryk et al., 2006; Engelke et al., 2017). Les outils neuropsychologiques modernes permettent de mesurer le système nerveux central et autonome (ex., Électroencéphalographie ou EEG, Pupillométrie), ainsi que les mouvements oculaires (ex. Oculométrie) et musculaires (ex. Électromyographie ou EMG) d'utilisateurs en interagissant avec des systèmes. Cependant, la plupart de ces outils peuvent être intrusifs dans une expérience multimédia en raison de l'utilisation de capteurs (Akhtar et al., 2019). Des méthodes non-intrusives telles que l'analyse d'expressions faciales à partir de vidéos semblent avoir été négligées par les chercheurs du domaine des multimédias (Arndt et al., 2016; Moon & Lee, 2017; Engelke et al., 2017). Pourtant, ces méthodes de détection automatique d'expressions faciales sans contact avec l'utilisateur sont reconnues pour leur pertinence, afin d'analyser les émotions humaines dans la recherche sur les IHM (Pantic & Rothkrantz, 2000; Egger et al., 2019). Aussi, ces méthodes ont le potentiel de pouvoir être utilisées à l'extérieur d'expériences contrôlées tel qu'au domicile des utilisateurs, permettant ainsi une analyse d'émotions dans un contexte écologiquement valide. Bien que les concepts de QoE et de UX ne doivent pas être confondus, les chercheurs du domaine des multimédias reconnaissent la valeur de la littérature sur le UX pour avancer la recherche et la pratique liée à l'évaluation du QoE (Wechsung & De Moor, 2014).

Finalement, la QoE peut être influencée par des facteurs contextuels déterminés par l'environnement de l'interaction entre l'utilisateur et le système. Le contexte est défini en termes de caractéristiques contextuelles physiques (ex. emplacement d'expérience), temporelles (ex. durée d'expérience), sociales, économiques, techniques ou informationnelles, ou liées à la nature de l'expérience (Reiter et al., 2014). Par exemple, la QoE peut être influencée le type de tâche effectuée, ou les potentielles interruptions causées par du multitâche avec des activités en parallèle. Les résultats d'une étude ont

montré qu'une activité complexe nécessitant du multitâche entre un multimédia et une tâche additionnelle peut influencer négativement la qualité perçue du multimédia (Reiter & Weitzel, 2007). Les résultats ont montré que lorsqu'un utilisateur interagit activement avec un jeu à l'ordinateur, ce dernier perçoit moins bien la qualité d'un stimulus auditif musical, comparativement à lorsqu'il regarde passivement le jeu (Reiter & Weitzel, 2007). Ainsi, il est important de comprendre comment des facteurs contextuels introduits par le multitâche en contexte de multimédia peut en influencer la QoE.

### **Comportement multitâche avec des sous-titres**

La lecture de sous-titres en écoutant un multimédia peut être considérée comme étant une forme de multitâche pertinente dans la mesure où elle peut améliorer la compréhension des dialogues ainsi que de la cohérence narrative (Lauvaur & Bairstow 2011; Lee et al. 2013). Le sous-titrage est une technique de traduction grandement utilisé dans l'industrie des télécommunications afin d'accroître l'accessibilité de multimédias internationaux (Duarte & Fonseca, 2019). Traditionnellement, les sous-titres interlinguistiques ou standards sont utilisés afin de pallier à la barrière linguistique d'un multimédia en traduisant, dans la langue de l'auditoire, les dialogues de langue étrangère. Les sous-titres intralinguistiques peuvent aussi être une source redondante d'information textuelle afin de pallier à la barrière sensorielle d'un multimédia pour un auditoire malentendant (Gernsbacher, 2015), ou lorsque les dialogues présentés sont difficilement compréhensibles par l'auditoire comme à l'opéra (Desblache, 2008; Orero & Matamala, 2007). En effet, la littérature indique que la redondance audio-visuelle, soit lorsque deux informations identiques sont communiquées par les modalités sensorielles différentes (Hsia, 1971), peut faciliter le traitement et la mémorisation de l'information (Walma van der Molen, & Klijn, 2004; Walma van der Molen, & van der Voort, 2000).

Aussi, des recherches antérieures sur l'effet des sous-titres indiquent que peu importe leur type et leur nécessité pour l'auditoire, les sous-titres instrumentaux apparaissant à l'écran peuvent attirer automatiquement l'attention visuelle (d'Ydewalle et al., 1991; Bisson et al., 2014). En effet, à l'aide de mesures oculométriques, ces dernières études ont trouvé aucune différence de durée de fixation dans la zone des sous-titres entre des sous-titres

interlinguistiques et intralinguistiques, et ainsi suggèrent un comportement automatique de lecture de sous-titres (d'Ydewalle et al., 1991; Bisson et al., 2014). Ce comportement peut être expliqué par la saillance des sous-titres, qui est déterminée par leur contraste de couleur (ex. blanc sur noir) et leur aspect dynamique qui attire inconsciemment l'attention visuelle (Nothdurft, 1993; Engelke et al., 2011)

Cependant, en contexte de multimédia hédonique, la littérature suggère que la division d'attention visuelle entre l'image et les sous-titres se fait généralement de manière efficace, de sorte que la compréhension de l'image n'est pas affectée (Perego, 2010). À l'aide de mesures neurophysiologiques, telles que la dilatation de la pupille ainsi que des signaux électroencéphalographiques, certaines études ont même trouvé que les sous-titres peuvent réduire la charge cognitive associée au traitement d'un multimédia (Kruger et al., 2013; Kruger et al., 2014). Alors, la littérature indique qu'un comportement multitâche avec des sous-titres peut améliorer une expérience de multimédia hédonique. Cependant, la recherche devrait explorer davantage les divers types de sous-titres dans différents contextes, afin de mieux comprendre comment l'auditoire consomme les multimédias produits.

Par exemple, les sous-titres intralinguistiques présentés dans la langue natale de l'auditoire ont eu peu d'attention des chercheurs dans le domaine des multimédias et communications (Kruger et al., 2015). Pourtant, ces derniers sous-titres peuvent être très pertinents dans des contextes comme l'opéra cinématique, puisque les dialogues ne sont pas toujours faciles à comprendre, même chantés dans la langue natale de l'auditoire (Desblache, 2008). Aussi, l'arrivée de nouveaux effets sensoriels (ex. stimulations haptiques) dans les expériences traditionnelles de multimédias audiovisuels sous-titrés présenteront de nouvelles formes de multitâche sur lesquelles les chercheurs devront se pencher.

### **L'Haptique dans les multimédias**

Avec la diminution du potentiel d'amélioration marginale de la qualité d'information audiovisuelle, telle que la résolution d'image et la clarté du son, l'industrie du divertissement exploite de plus en plus d'autres sens, tel que le toucher comme troisième

canal sensoriel avec la vision et l'ouïe (Timmerer et al., 2014). Les technologies haptiques utilisent la sensation du toucher comme canal sensoriel additionnel servant à communiquer de l'information sous forme spatiale, temporelle, et physique d'un environnement, pouvant même être intégré avec des stimuli audiovisuels (El Sadik et al., 2011). Les stimulations haptiques permettent de communiquer des informations sous forme vibro-tactile ou cinétique, dans le but de solliciter respectivement la perception de sensations cutanées localisées sur la surface de la peau, et la perception des mouvements des parties du corps (El Saddik et al., 2011; Chaudhari et al., 2014).

Une technologie ayant reçu beaucoup d'attention de chercheurs en IHM durant les dernières années utilise des stimulations vibro-tactiles et cinétiques livrées avec haute-fidélité par un siège en synchronisme avec des effets audiovisuels d'un multimédia (Gardé et al., 2018; Pauna et al., 2017; Giroux et al., 2019; Tchanou et al., 2019; Boasen et al., 2020). La technologie High Fidelity Vibro Kinetic (HFVK) développée par l'entreprise DBOX Technologies (Longueuil, QC), peut contribuer à améliorer le sentiment d'immersion de l'auditoire durant une scène de film d'action (Pauna et al., 2017), ou en expérience de réalité virtuelle (Gardé et al., 2018). Les résultats d'une étude récente ont montré que la stimulation HFVK synchronisée avec des éléments de musique populaire (ex. rythme, mélodie) pouvait augmenter l'appréciation subjective, ainsi que les réactions émotionnelles positives liées à une expérience purement auditive (Giroux et al., 2019). Plus spécifiquement, cette dernière étude a trouvé que la stimulation HFVK pouvait augmenter la fréquence d'émotions positives détectées par un système de détection automatique d'expressions faciales.

Cependant, les résultats d'une autre étude ont montré que l'effet de la stimulation HFVK sur la fréquence d'émotions positives détectées dans le visage des participants était annulé par l'interaction avec des informations textuelles non pertinentes présentées sur un appareil mobile (Tchanou et al., 2019). De plus, les participants de cette dernière étude ont révélé qu'ils avaient de la difficulté à se concentrer sur les mouvements et vibrations générés par le siège et sur une tâche de lecture d'information présentée simultanément sur un appareil mobile, puisque l'un était une distraction à l'autre. Ces résultats semblent indiquer que l'auditoire a une capacité limitée de d'information pouvant être perçues ou

traitées dans une expérience hédonique multisensorielle impliquant du multitâche avec des informations textuelles.

Dans cette optique, il serait important d'explorer la stimulation haptique comme étant un distracteur potentiel à une activité multitâche avec des sous-titres, considérant que ces derniers sont pertinents et peuvent même être nécessaires à la compréhension d'un multimédia. En effet, malgré les bienfaits de l'ajout de stimulation haptique dans une expérience de multimédia, les créateurs d'expériences mulsemcia doivent garder en tête que l'auditoire a une capacité limitée de traitement d'informations multisensorielles (Baddeley et al., 1969; Kahneman, 1973; Lang, 2000). En effet, alors que l'objectif d'ajouts d'effets sensoriels est d'améliorer la qualité d'expériences de multimédias hédoniques, les créateurs de ces contenus doivent s'assurer que le comportement multitâche naturel de l'auditoire avec les sous-titres n'est pas affecté. Cette ligne de recherche est importante, afin de pouvoir identifier les compromis qu'un auditoire peut avoir à faire parmi la quantité et la variété d'informations multisensorielles accessibles.

## **Informations au sujet des articles**

Le premier article constituant ce mémoire a été présenté à la conférence scientifique NeuroIS Retreat 2020 en juin dernier. Le deuxième article constituant le mémoire a pour objectif d'être soumis à la revue scientifique ACM Transactions on Multimedia Computing, Communications and Applications, qui s'intéresse à l'application de nouvelles technologies (ex. haptique) dans les multimédias, en plus de méthodes d'évaluation de la QoE des utilisateurs. Les deux articles qui font partie du mémoire ont utilisé les données de la même expérience en laboratoire.

Il est à noter que l'auteur de ce mémoire a aussi écrit un article de conférence en lien avec la stimulation haptique en contexte d'écoute de musique populaire (Giroux et al., 2019). Cet article a été accepté à la conférence IEEE Transactions on Haptics, et présenté au Japon en juillet 2019 par l'auteur de ce mémoire. Ce dernier a aussi été co-auteur d'un article ayant été accepté et présenté à la conférence AIS SIG-HCI à San Francisco en décembre 2018. Cet article porte plus spécifiquement sur le multitâche en contexte d'écoute de musique rehaussée de stimulations haptiques. Les données utilisées pour ces

deux derniers articles proviennent d'un projet de recherche, pour lequel l'auteur du mémoire a activement contribué (ex. design expérimental, collecte, analyse, présentation au partenaire) dans le cadre d'un stage de recherche d'été (2018) de premier cycle (CRSNG). Finalement, les résultats de ces articles ont servi comme fondation empirique pour le développement des questions de recherche explorées dans les articles présentés dans ce mémoire.

### **Présentation sommaire du premier article**

Diviser son attention visuelle entre un texte et une vidéo est un comportement multitâche très commun dans notre monde technologique moderne. Les résultats de recherches montrent principalement les effets néfastes de ce comportement sur la performance liée à une tâche (Courtemanche et al., 2018, ou même sur les réactions émotionnelles d'utilisateurs d'une expérience hédonique (Tchanou et al., 2019). Dans cette étude, nous voulions explorer un comportement multitâche généralement reconnu pour améliorer une activité globale, soit la lecture de sous-titres dans un multimédia (Perego, 2010; d'Ydewalle & Gielen, 1992; d'Ydewalle & De Bruycker, 2007). Plus spécifiquement, nous cherchions à comprendre l'effet neurophysiologique de la présence de sous-titres intralinguistiques diffusés dans la langue natale de l'auditoire d'opéra cinématique. Nous avons posé l'hypothèse que la présence de sous-titres allait attirer l'attention visuelle vers la zone des sous-titres de l'écran. Conséquemment, nous avons suggéré que la présence des sous-titres allait globalement réduire l'effort cognitif associé au traitement du multimédia en améliorant la compréhension de l'auditoire. Cependant, nous avons aussi fait l'hypothèse que les réactions émotionnelles à l'image du multimédia allaient être réduites.

Pour cette étude, 17 participants ont été recrutés à partir du panel d'HEC Montréal afin de participer à une expérience en laboratoire au Tech3Lab. Les participants ont écouté 16 extraits vidéo d'un opéra cinématique d'une durée variant d'environ deux à trois minutes chacun. Les extraits couvrant une performance de Carmen de Bizet étaient joués dans l'ordre de la trame narrative de la performance. Les extraits étaient aussi randomisés selon les quatre conditions expérimentales utilisées dans ce mémoire, soit avec ou sans sous-

titres, et avec ou sans stimulation haptique. Les conditions avec et sans sous-titres ont uniquement été utilisées pour le premier article, alors que le deuxième article comprend également conditions avec stimulation haptique.

Dans cette expérience en laboratoire, nous avons évalué l'attention visuelle divisée entre l'image et la zone de sous-titres de l'écran de télévision à l'aide de données de localisation de fixations mesurée avec le système d'oculomètre Smart Eye Pro 6.2 (Gothenburg, Sweden). L'effort cognitif a été en partie évalué à l'aide de mesures de dilatation de la pupille, aussi obtenue avec l'oculomètre. Nous avons aussi mesuré l'effort cognitif à l'aide de la bande de fréquence électroencéphalographique Theta obtenue au moyen d'un casque muni de 32 électrodes (BrainProducts). Les expressions faciales des participants durant le visionnement des stimuli ont été captées à l'aide d'une webcam ainsi que le logiciel Media Recorder 2 de Noldus (Wageningen, Pays-Bas), et analysées à l'aide du logiciel de détection d'expressions faciales FaceReader 7.0 de Noldus (Wageningen, Pays-Bas). Finalement, nous avons évalué les réactions émotionnelles subjectives, ainsi que l'opinion des participants sur l'expérience à l'aide d'échelles et entrevues semi-dirigées.

Les résultats du premier article montrent que l'utilisation des sous-titres dans notre contexte d'étude a permis d'améliorer la compréhension de l'auditoire sans affecter l'effort cognitif à la tâche, ainsi que les réactions émotionnelles liées au traitement de l'image du multimédia. Ainsi le premier article introduit un contexte de multitâche avec des sous-titres instrumentaux qui permet d'améliorer la qualité d'une expérience de multimédia hédonique.

### **Présentation sommaire du second article**

Avec la production croissante de multimédia, il est important d'en évaluer la qualité de l'expérience des utilisateurs (QoE). La QoE peut être influencée par les caractéristiques du système multimédia, de l'utilisateur, et de son environnement (Reiter et al., 2014). Par exemple, faire du multitâche durant une activité de multimédia hédonique peut affecter négativement la QoE perçue associée au multimédia (Reiter & Weitzel, 2007).

Toutefois, il est important de distinguer une activité multitâche pertinente d'une activité multitâche non-pertinente en contexte de multimédia. Par exemple, la lecture de sous-titres en visionnant un film peut être considérée comme une activité multitâche pertinente pouvant contribuer à améliorer la qualité d'une expérience multimédia. En effet, les auditoires démontrent généralement un comportement de lecture efficace et même automatique (d'Ydewalle et al., 1991; Bisson et al., 2014) de sous-titres sans affecter la compréhension des images (Perego et al., 2010) d'un multimédia hédonique.

Cependant, avec l'adaptation des multimédias traditionnels aux expériences multisensorielles, il est important de comprendre l'impact de l'ajout de nouveaux types d'information (p. ex., stimulations haptiques) sur le multitâche avec des sous-titres. Les résultats d'une étude récente suggèrent que le multitâche avec la lecture d'informations textuelles sur un appareil mobile lors d'une écoute de musique rehaussée de stimulations haptiques pouvait empêcher l'auditoire de pleinement profiter de l'expérience multisensorielle (Tchanou et al., 2019). En effet, en raison la capacité limitée de l'humain à percevoir les informations sensorielles qui lui sont présentées (Baddeley et al., 1969; Kahneman, 1973; Lang, 2000), il est possible que l'auditoire ne puissent traiter simultanément l'ensemble des informations communiquées, même lorsqu'elles sollicitent des modalités sensorielles différentes.

Alors, dans la mesure où les sous-titres sont une partie intégrante d'un multimédia et peuvent contribuer à la QoE, il est important d'évaluer si la stimulation haptique peut nuire au multitâche avec ces sous-titres. Le second article cherche donc à comprendre l'effet de la stimulation haptique synchronisée avec les éléments auditifs non-verbaux d'un multimédia hédonique sur le multitâche avec des sous-titres intralinguistiques présentés dans la langue natale de l'auditoire.

Nous avons posé l'hypothèse que l'attention visuelle ou le temps de fixation porté aux sous-titres n'allait pas être influencé par la stimulation haptique lors du visionnement. Nous avons aussi posé l'hypothèse que la stimulation haptique allait accentuer les réactions émotionnelles, ainsi que l'appréciation subjective de l'expérience d'opéra cinématique sous-titré. Les résultats de ce second article nous montrent que la stimulation

haptique ne semble pas affecter le comportement naturel de lecture des sous-titres. Toutefois, les résultats suggèrent que la présentation de sous-titres instrumentaux en plus d'effets haptiques peut potentiellement causer une sur-stimulation, pouvant empêcher l'auditoire de réagir pleinement aux émotions induites par le multimédia.

### **Contributions et responsabilités individuelles**

La contribution et les responsabilités personnelles liées au projet de recherche du mémoire sont rapportées dans la table suivante (voir Table 1). Dans cette table, pour chaque étape du projet, la ou les tâches effectuées sont présentés avec la contribution (en pourcentage) de l'auteur de ce mémoire.

*Table 1: Contribution dans le projet de recherche et rédaction d'article*

Revue de littérature	Identifier des aires de recherche inexploré étant pertinents à l'expérience multisensorielle en contexte d'opéra cinématique assisté d'un siège à retour haptique ( <i>voir Annexe 1</i> ) – 100%
Définition de la question de recherche	<p>Identifier la question de recherche – 60%</p> <ul style="list-style-type: none"> <li>• L'auteur de ce mémoire a suggéré d'intégrer la manipulation de la présence de sous-titres dans l'expérience. De cette idée, une question de recherche autant pertinente pour le partenaire industriel DBOX que pour la recherche en expérience utilisateur a été développée à l'aide des codirecteurs.</li> </ul>
Conception du design expérimental	<p>Compléter le formulaire de demande au CER ainsi que les formulaires de demandes de modification de projet (F8) – 100%</p> <ul style="list-style-type: none"> <li>• Un membre de l'équipe d'opération du Tech3Lab s'est assuré que les formulaires étaient adéquats</li> </ul> <p>Élaborer et rédiger le protocole d'expérimentation – 100%</p> <p>Développer des scripts Python afin de permettre l'automatisation de la présentation de stimuli – 50%</p>

	<ul style="list-style-type: none"> <li>• Un développeur du partenaire industriel DBOX a transmis à l'équipe de recherche un gabarit de code Python, qui a été développé par l'auteur de ce mémoire ainsi qu'un membre de l'équipe de recherche</li> </ul> <p>Assurer la mise en fonction du système oculométrique Smart Eye Pro – 90%</p> <ul style="list-style-type: none"> <li>• Éric, un étudiant du Tech3Lab ayant aussi eu à travailler avec le système oculométrique SEP pour son projet, a contribué à l'apprentissage du fonctionnement du système.</li> </ul> <p>Organiser et préparer la salle de collecte de données – 50%</p> <ul style="list-style-type: none"> <li>• L'équipe de recherche a contribué à différentes étapes de l'organisation de la salle (p. ex. installation des panneaux coupe-son, installation d'équipements EEG, installation des ordinateurs)</li> <li>• L'auteur de ce mémoire s'est occupé de différentes étapes de l'organisation de la salle (p. ex. installation du matériel audiovisuel, construction et installation d'une rampe pour soutenir les oculomètres et la caméra, installation de rideaux)</li> </ul>
Recrutement des participants	<p>Élaborer et rédiger le questionnaire de recrutement – 100%</p> <ul style="list-style-type: none"> <li>• Recruter les participants et céder les collectes de données via des échanges de courriel</li> </ul>
Prétests et collecte de données	<p>Prendre en charge des opérations lors des collectes de données – 33%</p> <ul style="list-style-type: none"> <li>• Deux assistant(e)s de recherche contribuaient à la collecte de données (p. ex. installation des outils neurophysiologiques, lecture de verbatims, enregistrement des outils)</li> <li>• L'auteur de ce mémoire a été présent pour tous les collectes de données utilisées pour ce mémoire. Il s'est principalement occupé de la calibration de</li> </ul>

	<p>l'oculomètre, la présentation du stimulus, de l'entrevue semi-guidée avec les participants</p>
Extraction et transformation des données	<p>Extraire et mettre en forme des données psychométriques, oculométriques, et physiologiques – 70%</p> <ul style="list-style-type: none"> <li>Paul, un étudiant du Tech3Lab, a contribué au développement d'un code matlab afin d'extraire les données de fixations de fichier de données brutes.</li> </ul>
Analyse des données	<p>Préparer l'analyse des données oculométriques et physiologiques</p> <ul style="list-style-type: none"> <li>Combiner les données oculométriques et physiologiques afin d'obtenir les émotions faciales temporellement liées aux fixations des participants.</li> </ul> <p>Réaliser les analyses statistiques du mémoire – 90%</p> <ul style="list-style-type: none"> <li>Interprétation des résultats à l'aide du logiciel statistique SPSS.</li> <li>Aide d'un chercheur postdoctorant et coauteur des articles concernant l'utilisation du logiciel statistique SPSS (IBM) ainsi que les tests statistiques appropriés.</li> </ul>
Rédaction des articles	<p>Contribuer à la rédaction des articles présentés dans ce mémoire – 90%</p> <ul style="list-style-type: none"> <li>Les articles ont été peaufiné à l'aide de commentaires et suggestions des coauteurs.</li> </ul>
Préparation et présentation des résultats au partenaire industriel impliqué dans le projet	<p>Contribution à la rédaction du rapport client - 40%</p> <ul style="list-style-type: none"> <li>L'équipe de recherche ainsi que les directeurs ont contribué au développement du rapport</li> </ul> <p>Présenter les résultats au partenaire industriel - 20%</p> <ul style="list-style-type: none"> <li>L'auteur de ce mémoire a été présent pour la plupart des présentations de résultat liés à ce mémoire afin de répondre à des questions et pour compléter la</li> </ul>

	présentation avec des détails spécifiques au déroulement de la collecte ainsi qu'aux entrevues (qui ont été réalisées par l'auteur du mémoire)
--	------------------------------------------------------------------------------------------------------------------------------------------------

## Structure du mémoire

Le mémoire sera structuré ainsi. Le premier chapitre porte sur le premier article du mémoire qui a été présenté à la conférence NeuroIS en juin 2020. Cet article vise à répondre à la question de recherche suivante : Quel est l'effet de la présence de sous-titres sur la qualité d'une expérience d'opéra cinématique? Le deuxième chapitre du mémoire présente le deuxième article qui a pour objectif d'être soumis à la revue scientifique ACM Transactions on Multimedia Computing, Communications, and Applications. Ce dernier tente de répondre à la question de recherche suivante : Dans quelle mesure la stimulation haptique peut dégrader la qualité de l'expérience résultant du multitâche avec des sous-titres dans un contexte de multimédia? Finalement, le troisième chapitre du mémoire conclut avec un retour sur les questions de recherche en discutant les défis méthodologiques liés au projet de recherche, les résultats des articles, en plus de souligner les contributions et les limites du mémoire.

# **Chapitre 1: Premier Article**

## **Hedonic multitasking: The effects of instrumental subtitles during video watching<sup>1</sup>**

Félix Giroux, Jared Boasen, Sylvain Sénécal, Pierre-Majorique Léger

### **Abstract**

Hedonic videos are often accompanied by instrumental text/subtitles. However, the neurophysiological effects of task-switching between instrumental text processing and audiovisual processing during hedonic video watching remains unexplored. We investigated these effects using changes in pupil diameter and theta-band spectral power as indices of cognitive load, and self-reported and automated facial expression-based measures of emotion. We found that the presence of instrumental subtitles subjectively improved comprehension without negatively impacting cognitive and emotional response. These findings contribute to the literature by showing that in this specific context, multitasking does enhance the user experience. Further research is needed to explore the extent to which instrumental text is useful, and how its effects change in accordance with audio presence during passive hedonic experiences across various contexts.

**Keywords:** Multitasking · Hedonic · Subtitles · Facial Expression · Emotion · Cognitive Load · Eye Tracking

### **1. Introduction**

Switching between tasks involving text processing and those involving non-textual audiovisual processing is a form of multitasking which is commonplace in our modern technological world, and of great interest to information systems (IS) research [1].

---

<sup>1</sup> Giroux F., Boasen J., Sénécal S., Léger PM. (2020) Hedonic multitasking: The effects of instrumental subtitles during video watching. In: Davis F., Riedl R., vom Brocke J., Léger PM., Randolph A., Fisher T. (eds) Information Systems and Neuroscience. Lecture Notes in Information Systems and Organisation. Springer, Cham.

Typically, tasks involving text processing will interfere with other tasks and increase cognitive processing load. This in turn has been associated with attenuated task performance [2], reduced emotional responsiveness to hedonic activities [3], increased pupil diameter [4], and even sustained increases in electroencephalographic (EEG) theta and alpha band spectral power [5].

However, there are occasions where text processing can instrumentally support another task. One example of this is reading intralingual subtitles during video watching. Although the presence of intralingual subtitles during hedonic video watching can reportedly erode comprehension of the imagery content, reading subtitles in this context has nevertheless been associated with increased comprehension of dialogue and local narrative coherence [5,6]. Furthermore, reading intralingual subtitles during academic video watching has also reportedly been associated with increased comprehension and reduced cognitive load as indexed via pupil diameter [7]. Although these studies show that multitasking in such context has benefits for the viewer, they do not simultaneously take into account emotional, cognitive and behavioral responses of users. Thus, our understanding regarding the effect of subtitles on cognitive load and emotional response remains limited. Considering the ubiquitous instrumental use of subtitles in online advertising videos, educational training, news, games, and movies, further IS research targeting subtitles is highly desirable.

The present study takes a step in this direction by investigating the effect of intralingual subtitles during hedonic video watching. Using eye-tracking, EEG, automated facial expression-based emotion, and self-reported emotion, we compared neuropsychophysiological responses between scenes with and without subtitles. Given the aforementioned evidence on subtitle reading during hedonic video watching, we predicted that subtitle presence would improve participant's subjective dialogue comprehension (H1) but would distract from image content and thereby reduce emotional responsiveness to it. Correspondingly, for subtitled compared to non-subtitled conditions, we hypothesized that the presence of instrumental subtitles would decrease participant's cognitive load indexed with smaller pupil diameter and lower EEG theta-band spectral power (H2), and decrease AFE-based emotions associated with image content processing

(H3). Our study contributes to IS research by advancing our understanding of how instrumentally supportive text information affects users during multisensory hedonic experiences.

## 2. Methods

### 2.1. Subjects and Stimuli

Subjects comprised 17 healthy young adults (M:8, F:9; mean age  $\pm$  std:  $23.7 \pm 3.5$  years). They sat in a cinema seat and watched 16 scenes from a 720p high definition cinematic recording of Georges Bizet's opera, Carmen. Opera was chosen as it is intrinsically hedonic, and the display of subtitles is traditional [8], making it a relevant real-world case study and therefore congruent with the future direction of IS research [9]. The scenes were randomly presented with or without subtitles (four trials each per subject) to control for variance in emotionality. The video was presented centrally on the screen (hereafter: image zone) and flanked on top and bottom by black bars. Subtitles appeared in the bottom bar (hereafter: subtitle zone). Subjects viewed the stimuli on a 70x120 cm high definition TV at a distance of 262 cm, with the audio presented in 2.0 stereo. Subtitles were intralingual and native to the subjects.

### 2.2. Measurements and Processing

Gaze fixation location and pupil diameter were recorded at a 60 Hz sampling rate with Smart Eye Pro 6.2 (Gothenburg, Sweden). Image and subtitle zone fixations were delineated for each subject individually according to the Y-axis value between the two zones that had the lowest frequency of fixations. No fixation outliers were excluded on either the x or the y axis. Cognitive load was indexed via pupil diameter. Because, pupil diameter varies across individuals [10], mean pupil diameter in each condition was normalized as a percent change from the mean pupil diameter across both conditions. Synchronous with eye-tracking, the facial expressions of subjects were continuously sampled at 30 fps with a high definition webcam using MediaRecorder (Noldus, Wageningen, Netherlands). Based on these recordings, the AFE detection software FaceReader 7.0 (Noldus, Wageningen, Netherlands) was used to assess the frequency (%)

of one positive (happiness) and four negative basic emotions (anger, disgust, sadness, and scared). Subjective emotional valence toward the stimuli was assessed after viewing each scene using a nine-point version of the Self-Assessment Manikin (SAM) pictogram scale ranging from the feeling of sadness (1) to happiness (9) [11]. Finally, in post-experiment interview, subjects were asked their preference for presence or absence of subtitles and the reason for their response. 32-channel electroencephalographs (EEG) and electrode positions were additionally recorded for all subjects throughout the experiment at 500 Hz (BrainProducts). EEG signal was cleaned with independent component analysis and band-pass filtered from 1-40 Hz. Source-level Hilbert transforms were applied to extract mean theta (5-7 Hz) activity in 62 cortical areas (Mindboggle Atlas) for each condition based on 10 second post-scene windows normalized as a percent deviation from the mean activity over 10 second pre-scene windows for each condition.

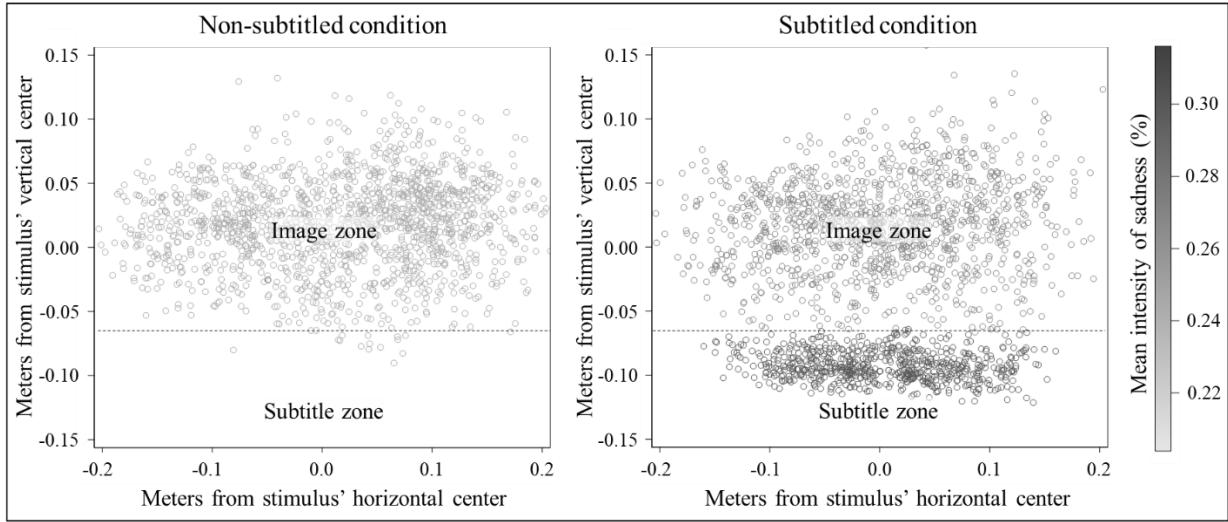
### **2.3. Statistical Analyses**

A three-way repeated measure analysis of variance (RM ANOVA) was performed to assess differences in the frequency of AFE-based emotions according to fixation zone and subtitle condition (2: conditions  $\times$  2: zones  $\times$  5: AFE-based emotion types). Paired t-tests were used to compare the difference in subjective emotional valence, and the difference in image zone PCPD, between subtitled and non-subtitled conditions. Note that we restricted our analyses of PCPD to the image zone as pupil diameter changes according to brightness of the point of focus [12], thereby making comparisons between the bright image zone and the dark subtitle zone inappropriate. Two-way RM ANOVA was performed to assess differences in cortical theta activity between conditions (2: conditions  $\times$  62: brain areas). All statistical analyses were performed using SPSS software version 22 (IBM, Armonk, NY, USA), with a threshold for statistical significance set at  $p \leq 0.05$ .

## **3. Results**

Three-way RM ANOVA revealed a marginally significant main effect of subtitle presence ( $F(1, 16) = 4.469, p = .051$ ), with higher AFE-based emotional frequency in the subtitled versus non-subtitled condition. There was also a significant main effect of AFE-based emotion type ( $F(4, 64) = 26.316, p < .001$ ), indicating differences in the occurrence

frequency of AFE-based emotions according to their type. There was no significant main effect of fixation zone ( $F(1, 16) = .220$ ,  $p = .225$ ), nor was there a significant interaction between subtitle presence and fixation zone ( $F(1, 16) = 1.929$ ,  $p = .184$ ). However, there was a significant interaction between subtitle presence and AFE-based emotion type ( $F(4, 64) = 4.878$ ,  $p = .002$ ), with simple main effects testing revealing significantly greater frequency sadness in the subtitled versus non-subtitled condition ( $p = .033$ ). Meanwhile, there was a significant three-way interaction between subtitle presence, fixation zone, and AFE-based emotion type ( $F(4, 64) = 4.366$ ,  $p = .003$ ). Simple-main effects testing revealed no difference in the frequency of any AFE-based emotion type in the image zone between subtitled and non-subtitled conditions. However, frequency of sadness was significantly higher in the subtitle zone compared to the image zone in subtitled conditions ( $p = .007$ ) (See Fig. 1). Paired t-test revealed no difference in subjective emotional valence between subtitled and non-subtitled conditions ( $t = .046$ ,  $p = .964$ ). In fact, post-experiment interviews revealed that 15 out of 17 subjects preferred subtitles as they improved narrative comprehension. That said, paired t-test revealed no significant difference in image zone PCPD between the subtitled and non-subtitled condition ( $t = -.022$ ,  $p = .983$ ). Meanwhile, two-way RM ANOVA on source-level theta activity revealed no main effect of condition ( $F(1, 17) = 1.991$ ,  $p = .176$ ) or brain area ( $F(61, 1037) = .741$ ,  $p = .931$ ), nor interaction between condition and brain area ( $F(61, 1037) = .826$ ,  $p = .827$ ).



*Figure 1: One subject's aggregated fixations associated with automatic facial expression-based sadness. Fixations are displayed in 2D plans. Mean sadness frequency in each condition's zone of interest is indicated by the darkness of the colored circles.*

## 4. Discussion

The present study explored task-switching between instrumental text processing and hedonic audiovisual processing by examining the effect of intralingual, native-language subtitle presence on neuropsychophysiological response to viewing cinematic opera. Gaze fixation measures clearly indicated that subjects looked at the subtitles (Fig. 1). In line with our hypothesis, subjects reported that the presence of subtitles improved their subjective comprehension of the opera (H1). However, the absence of correspondingly significantly smaller pupil diameter and lower EEG theta-band spectral power for subtitled compared to non-subtitled conditions indicated that whatever improvements in comprehension, reading subtitles did not translate to any marked changes in cognitive processing burden (H2). Meanwhile, a lack of significant differences in self-reported valence and AFE-based emotions associated with the image zone between conditions similarly indicated that the improvement in subjective comprehension due to subtitle presence did not come at the cost of emotional responsiveness to audiovisual content (H3). In short, the subtitles appeared to serve their instrumental purpose without significantly disrupting the hedonic experience.

Interestingly, our analysis of AFE-based emotion indicated a significantly higher frequency of sadness in the subtitled compared to the non-subtitled condition, which simple-main effects testing revealed to be driven specifically by a high detection rate of AFE-based sadness for fixations in the subtitle zone. Although the opera, Carmen, is a tragedy, increased sadness due to text information is suspicious, particularly when there was no difference in self-reported emotion between conditions. Riedl and Léger (2016) stressed the importance of considering the impact of non-emotional cognitive processes on facial expressions, and perhaps herein lies the explanation [13]. AFE software are based on the Facial Action Coding System (FACS), where emotion is determined according to specific muscle activity configurations or action units (AUs) [14,15]. The principal muscle activity responsible for the AU related to AFE-based measures of sadness is the corrugator muscle [16]. Reports have implicated greater corrugator muscle activity during reading in association with decreased text readability, and with increased mental effort during sustained information processing [17,18,19]. Meanwhile, Kaiser (2014) observed the mere act of reading itself to induce corrugator muscle activity [20]. Although our results did not indicate any cognitive differences between conditions, the observed AFE-based phenomenon, together with the above evidence strongly points to non-emotional drivers of facial muscle activity during reading.

The present study has limitations. For instance, it did not address more specific questions such as to what extent the text information is really necessary, and how do its effects change due to the presence/absence of audio during the hedonic video experience. These are important questions for hedonic systems, for as van der Heijden (2004) observed, subjective enjoyment and ease of use more strongly determine intent to use/consume than subjective usefulness [21]. Additionally, our neurophysiological measures were not sensitive enough to detect differences in correspondence with those regarding subjective comprehension. To better support these subjective results, future studies could benefit from an EEG approach based on eye-fixation related potentials, which could provide insight into cognitive processing load at the time of each eye fixation [22].

In conclusion, the present study provides first-time combined evidence that task-switching between instrumental subtitle processing and audiovisual processing during

hedonic video watching subjectively improves comprehension without negatively impacting cognitive and emotional response. Our findings also demonstrate the importance of multitasking studies relying on AFE-based emotion to differentiate measures according to task due to the potential influence of non-emotional cognitive drivers of facial muscle activity. This study highlights the importance and need for further neurophysiological investigations into instrumental text processing during hedonic video watching, not only in cinema, but also in other contexts such as educational training, online advertising, gaming, and even live contexts.

## References

1. Cameron, A. F., & Webster, J. Relational outcomes of multicomunicating: Integrating incivility and social exchange perspectives. *Organization Science*, 22(3), 754-771 (2011)
2. Courtemanche, F., Labonté-LeMoyne, E., Léger, P. M., Fredette, M., Senecal, S., Cameron, A. F., Faubert, J., & Bellavance, F. Texting while walking: An expensive switch cost. *Accident Analysis & Prevention*, 127, 1-8 (2019)
3. Tchanou, A.-Q., Giroux, F., Léger, P.-M., Senecal, S., Ménard, J.-F. Impact of Information Technology Multitasking on Hedonic Experience. Association for Information Systems. SIGHCI Proceedings (2018)
4. Katidioti, I., Borst, J. P., and Taatgen, N. A. What happens when we switch tasks: Pupil dilation in multitasking. *Journal of Experimental Psychology: Applied*, vol. 20, no. 4, pp. 380–396 (2014)
5. Lavaur J. M., and Bairstow, D. Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology*, vol. 46, no. 6, pp. 455-462 (2011)
6. Lee, M., Roskos, B., and Ewoldsen, D. R. The Impact of Subtitles on Comprehension of Narrative Film. *Media Psychology*, vol. 16, no. 4, pp. 412–440 (2013)

7. Kruger, J-L., Hefer, E., Gordon, M. Measuring the Impact of Subtitles on Cognitive Load: Eye Tracking and Dynamic Audiovisual Texts. Proceedings of Eye Tracking South Africa, Cape Town, 29-31 (2013)
8. Orero, P., Matamala, A. Accessible Opera: Overcoming Linguistic and Sensorial Barriers. Perspectives: Studies in Translatology, vol. 15, no. 4 (2008)
9. vom Brocke, J., Hevner, A., Léger, P. M., Walla, P., & Riedl, R. Advancing a neurois research agenda with four areas of societal contributions. European Journal of Information Systems, pp. 1-16 (2020)
10. Tsukahara, J. S. Harrison, T. L. and Engle, R. W. The relationship between baseline pupil size and intelligence. Cognitive Psychology, vol. 91, pp. 109–123 (2016)
11. Bradley, M. M. and Lang, P. J. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry, vol. 25, no. 1, pp. 49–59 (1994)
12. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. Eye tracking: A comprehensive guide to methods and measures. Oxford: Oxford University Press (2011)
13. Riedl, R., & Léger, P. M. Fundamentals of NeuroIS. Studies in Neuroscience, Psychology and Behavioral Economics. Springer, Berlin, Heidelberg (2016)
14. Ekman, P. Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), Nebraska symposium on motivation, Lincoln: University of Nebraska Press, pp. 207-283 (1972)
15. Ekman, P., & Friesen, W. V. Facial action coding system: A technique for the measurement of facial movement. Palo Alto, CA: Consulting Psychologists Press (1978)
16. Ekman, P., Friesen, W. V., & Hager, J. C. The facial action coding system CD-ROM. Salt Lake City, UT: Research Nexus (2002)

17. Larson, K., Hazlett, R.L., Chaparro, B.S., Picard, R.W. Measuring the Aesthetics of Reading. *People and Computers XX — Engage*, pp 41-56 (2004)
18. van Boxtel, A. and Jessurun, M. Amplitude and bilateral coherency of facial and jaw-elevator EMG activity as an index of effort during a two-choice serial reaction task. *Psychophysiology*, vol. 30, no. 6, pp. 589–604 (1993)
19. Waterink W. and van Boxtel, A. Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. *Biological Psychology*, vol. 37, no. 3, pp. 183–198 (1994)
20. Kaiser, M. The Prediction of an Emotional State through Physiological Measurements and its Influence on Performance. Thesis. Utrecht University (2014)
21. van Der Heijden, H. User Acceptance of Hedonic information systems. *MIS Quarterly*, vol. 28, no. 4, pp. 695–704 (2004)
22. Léger, P. M., Séncal, S., Courtemanche, F., de Guinea, A. O., Titah, R., Fredette, M., & Labonte-LeMoyne, É. Precision is in the eye of the beholder: Application of eye fixation-related potentials to information systems research. *Journal of the Association for Information Systems*, vol. 15, pp. 651–678 (2014)

## **Chapitre 2: Deuxième Article**

# **Measuring the Effect of Haptic Stimulation on audience's Quality of Experience during Hedonic Multimedia Multitasking context<sup>2</sup>**

### **Abstract**

Multitasking is a common contextual factor during hedonic multimedia experiences. When an additional task is irrelevant to the main media consumption activity, the quality of experience may be degraded. However, relevant multitasking, such as dividing our visual attention between subtitles and the image content of a multimedia, was shown to be beneficial for the audience's quality of experience by improving dialog's and narrative's comprehension. However, while subtitled audiovisual content is being increasingly adapted to mulsemcia (i.e., multiple sensory media) experiences involving new sensory effects (e.g., haptic stimulation), the extent to which these effects can influence traditional multimedia multitasking is unexplored. The purpose of this study is to investigate whether relevant multimedia multitasking with subtitles can be distracted by haptic stimulation. More specifically we tested the effect of High Fidelity Vibro-Kinetic (HFVK) stimulation on viewer's visual attention to Native Intralingual (NI) subtitles, as well as their emotional reactivity to the multimedia, respectively measured via gaze tracking and Automatic-Facial Expression (AFE) detection method. We invited 17 healthy participants to watch 16 sequences of a cinematic Carmen opera performance, with the sequences being presented randomly with or without NI subtitle, and with or without HFVK stimulation. Our results reveal that the HFVK stimulation did not influence the fixation time spent in the subtitle zone. Moreover, our results suggest that some participants may have unconsciously looked at the subtitles more than they perceived, which suggests automatic subtitle reading behavior. Finally, our results point out that both HFVK stimulation and NI subtitle presence have individually increased the audience's emotional reactivity and subjective appreciation of the experience.

---

<sup>2</sup> Giroux F., Boasen, J., Sénécal S., Léger P.M. (*Work in progress*) Measuring the Effect of Haptic Stimulation on audience's Quality of Experience during Hedonic Multimedia Multitasking context. *ACM Transactions on Multimedia Computing, Communications, and Applications*.

However, when both HFVK effects and NI subtitles are presented, it seems like the audience may experience sensory overstimulation which impedes it from fully processing and reacting to the hedonic stimulus.

**Keywords:** Quality of Experience (QoE) · Multimedia · Mulsemedia · Multitasking · Subtitles · Haptics · Facial Expressions · Eye Tracking

## 1. Introduction

With the increasing production of multimedia content, the need to assess the user's quality of experience (QoE) has become important in telecommunication research. The user's QoE in multimedia is known to be influenced by factors related to the system's technical features, the user's perception or mental and emotional state, and also the interaction context (Reiter et al, 2014). For instance, with the increasing access to digital information via information communication technologies, humans often engage in multitasking with other activities during hedonic multimedia experience (e.g., reading while watching a TV show). The literature shows that although audiences of hedonic multimedia engage in multitasking to seek pleasure and information via various forms of complex and novel information (Jeong & Fishbein, 2007; Hwang et al., 2014), this behavior may have a negative impact on tasks-related information processing and comprehension (Jeong & Hwang, 2016). A study has also shown that while actively interacting with a computer game, users could less perceive the quality of an auditory stimulus than when passively watching the game play itself (Reiter & Weitzel, 2007). Therefore, the literature suggests that when an additional task is not complementary or necessary to the main media activity, multitasking may degrade the processing and comprehension of a multimedia, and thus the audience's QoE, even when the tasks does not compete for the same sensory modalities.

However, it is important to distinguish irrelevant multitasking from relevant multitasking, in which two or more media tasks are related (Wang et al., 2015), since the latter may be beneficial for information processing (Hwang & Jeong, 2018). Dividing our visual attention between the image and the subtitles of multimedia content can be considered as

a form of relevant multitasking. Past studies using eye tracking measurement have suggested that saliency of subtitles may automatically draw the audience's visual attention to it (Bisson et al., 2014; d'Ydewalle et al., 1991). Consequently, since the human eye cannot attend both the subtitle reading and the image simultaneously, the processing of image content may be disrupted by the presence of subtitles that are not needed for comprehension. However, subtitle reading behavior has been shown to be non-detrimental in hedonic multimedia watching. Indeed, there is evidence that viewers are able to effectively switch their attention between the image content and subtitles without impacting image comprehension (Perego et al., 2010). Moreover, when multitasking with subtitles that can contribute to improve comprehension of dialogs and narrative coherence of a multimedia (Lauvaur & Bairstow 2011; Lee et al. 2013), the audience's QoE may be enhanced.

However, traditional multimedia, including those that display subtitles, are being increasingly adapted to mulsemcia experiences (i.e., multiple sensory media) involving the synchronisation of new sensory effects (e.g., haptic, olfactory) with audiovisual stimuli to enhance the audience's QoE (Walt & Timmerer, 2010; Timmerer et al., 2014; Yuan et al., 2014). Moreover, past research suggests that the addition of sensory effect (e.g., haptic stimulation) in hedonic media experience may introduce new forms of complex multitasking. Indeed, a recent study investigated the effect of multitasking with textual information on a mobile device during a music listening experience enhanced with haptic stimulation (Tchanou et al., 2019). Listening to music while reading or studying textual information is a common multitasking activity that was shown to have low impact on participants' recall (Middlebrooks et al., 2017) as compared to when they were actively completing a digit-detection task. However, in Tchanou et al., (2019), multitasking with text was shown to block the increasing effect of haptic stimulation on the participant's facial expressions of happiness. The subjects further revealed that simultaneously focusing their attention on both the music enhanced by haptic stimulations, as well as the textual information, was challenging since both stimuli were distracting the other (Tchanou et al., 2019). Therefore, the results of this study may suggest that multitasking with textual information while listening to music can be distracted by additional haptic

stimulation. Furthermore, it is suggested that the multisensory music listening experience would be optimized without multitasking with a non relevant textual information reading task on a mobile device (Tchanou et al., 2019).

However, research has not investigated multitasking with relevant activities (i.e. reading subtitles), during a hedonic multisensory multimedia experience. It would be important to address this gap in the literature in order to better understand impact of haptic stimulation on multitasking with subtitles, considering that the latter are relevant to a multimedia by improving the narrative's comprehension. Resulting insights regarding the mechanisms through which audience attend and react to multiple sensory stimuli involving haptics and instrumental text (e.g. subtitles) reading would benefit multisensory hedonic experience designers. Indeed, while the goal of adding haptic effects into traditional multimedia experiences is to improve the QoE, one should make sure that the QoE associated with narrative's comprehension through subtitle reading is not affected.

This study explores the following research question: Does haptic stimulation synchronized with a multimedia degrade the QoE resulting from multitasking with subtitles? To address this question, the present study investigates the influence of audio-targeted haptic stimulation on multitasking with instrumental subtitle reading during a cinematic experience. We assessed subtitle multitasking behavior via the eye fixation time spent in the subtitle zone as contrasted to the image content. The QoE was measured via emotional reactivity and subjective appreciation of an audiovisual stimuli, respectively assessed via the frequency of AFE-based discrete emotions detected during the opera watching, as well as post questionnaires and interviews.

First of all, we assumed that the participants would spend time fixating the subtitles either as an automatic behavior (d'Ydewalle et al., 1991; Bisson et al., 2014), or as an alternative source of information to the dialogs to improve dialogs comprehension and narrative coherence (Lauvaur & Bairstow 2011; Lee et al. 2013). As for the influence of haptic stimulations, literature has shown that tactile cue can increase the saliency of visual objects in a monitor when they occur simultaneously (Burg et al., 2009). However, the

haptic stimuli of this study were not synchronized with image features of the multimedia, but rather its non-verbal auditory stimuli. Therefore, we expected that the haptic stimulation would not distract the audience's visual attention to the subtitles. Consequently, we hypothesized that the haptic stimulation will not negatively moderate the relationship between instrumental subtitle presence and total fixation time spent in the subtitle zone of the audiovisual stimuli (H1).

Regarding the audience's emotional reactivity, we assumed that the enhanced lyrics' comprehension provided by the subtitles would increase the stimuli's induced emotions, as well as the multimedia activity appreciation. Indeed, literature on emotion-eliciting media suggest that subtitles can contribute to enhance narrative comprehension, thereby better conveying emotions to the audience (Rodica Baltes et al., 2011). This study found that learning the plot of a sad opera music before listening to it can significantly enhance self-perceived emotion of sadness. We can assume that since lyrics comprehension may lead to plot's comprehension (Scherer et al., 2019), subtitles can enhance the emotions induced by a hedonic multimedia such as cinematic opera. Literature further suggests that subtitle reading behavior should not affect image content comprehension (Perego et al., 2010). Therefore, we expected that the emotions induced by the image content would not be affected by subtitle reading.

We further predicted that without affecting the natural visual attention to the subtitles, the haptic stimulation would improve the QoE by increasing the emotional reactivity to and the subjective appreciation of the multimedia. Indeed, recent research has shown that haptic stimulation synchronized with auditory stimulus can enhance the frequency of automatic facial expressions (AFE) based discrete emotion of happiness induced by popular music, and the subjective appreciation of the experience (Giroux et al., 2019; Tchanou et al., 2019). Therefore, since popular music was associated with greater zygomatic facial muscle activity responsible for the positive emotion happiness (Lundqvist et al., 2009), it is suggested that haptic stimulation may intensify the emotions induced by a stimulus. While the prediction of the multimedia's specific emotions is not the focus of this study, we hypothesized that the effect of haptic stimulation will positively

moderate the relationship between subtitle presence and the frequency each AFE-based positive and negative discrete emotions (happiness, anger, sadness, fear, and disgust) (H2).

Finally, we believe that the subtitle presence will increase subjective appreciation of the multimedia hedonic experience by improving dialogs comprehension and narrative's coherence (Lauvaur & Bairstow 2011; Lee et al. 2013). We further hypothesized that the haptic stimulation would strengthen the relationship between subtitle presence and subjective appreciation of the experience (H3). The rest of the paper will be structured as follows. First, background related work will be provided, followed by our methodology. Then, the results will be presented and discussed. We will conclude with our study's implications for multimedia research and future directions.

## **2. Related work**

### **2.1. Quality of Experience in multimedia and mulsemcia**

The Quality of Experience (QoE) is defined as the degree of delight or annoyance of the user of an application or service, resulting from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of his or her personality and current state (Brunnström et al., 2013). QoE is a relatively new concept that has replaced the concept of Quality of Service (QoS), which mainly focused on system and service's characteristics' performance assessment (e.g., information communication delay, display resolution) without taking the user's needs into account (Varela et al., 2014). However, in the last decade of research in multimedia and HCI, the focus on the utilitarian approach of assessing information technologies has shifted towards a user-centric approach (Wechsung & De Moor, 2014), which gives more importance to humans and contextual influencing factors of QoE (Reiter et al., 2014).

Humans influencing factors are relatively stable characteristics (e.g., user's socio-cultural background, prior knowledge, relevant skills) or more dynamic characteristics (e.g.,

emotions, mood, motivation, or attention) that result from the interaction between the user and the system (Reiter et al., 2014). For instance, emotions elicited from the interaction with a system are known to have a strong influence on QoE in entertainment multimedia (Schleicher & Antons, 2014). Dynamic human factors influencing QoE have traditionally been assessed via the audience's perception of their own emotions or opinion towards the system (e.g. Mean Opinion Score) (Wechsung & De Moor, 2014) using self-reported measurement (e.g. Likert scale). However, scholars have stressed the reliability of assessing emotions using subjective measurement in emotion eliciting media experience (Scherer, 2004; Kayser, 2017). Indeed, subjective measures are unable to capture unconsciously experienced emotions, and can only grasp consciously felt event (Walla, 2018).

In the last decade, the development of objective methods to assess the User Experience (UX) in HCI research has led to trends towards neurophysiological measurement methods to assess QoE in multimedia and mulsemdia research (Wechsung & De Moor, 2014; Mandryk et al., 2006; Engelke et al., 2017). Neurophysiological measures encompass the central and autonomic nervous systems (e.g., EEG and pupillometry), as well as eye and muscles' movement (e.g., gaze tracking, EMG). For instance, Golland et al. (2018) investigated the dynamics of facial expressions during movie watching and found that EMG signals (i.e., zygomaticus and corrugator supercilii facial muscles) were temporally aligned with the movie's continuous subjective valence ratings developed in pre-tests. This suggests that multimedia's induced emotions may be reflected in viewers' facial expressions.

However, research on objective QoE has mostly focused on methods such as EMG that are characterized by the intrusiveness of their applied neurophysiological tools and sensors (Akhtar et al., 2019). Other non-obtrusive methods such as video based AFE detection analysis seems to have been generally overlooked by scholars in the field of multimedia (Arndt et al., 2016; Moon & Lee, 2017; Engelke et al., 2017). AFE detection tools allow to precisely model a human's face from live video recordings in order to infer the intensity of six discrete emotions based on face micro-movements. Thereby, this contact-less tool is known to be useful for emotion assessment in HCI research (Pantic &

Rothkrantz, 2000; Egger et al., 2019). Considering the low hardware needs, this method also has the potential to operate outside of controlled laboratory experiments in more ecologically valid settings such as users' own living room. Therefore, it is important to develop objective methods to assess human influencing factors of QoE, such as emotions.

Multimedia QoE can also be influenced by contextual factors given by the audience's environment. The latter is determined by characteristics related to the physical (e.g., location), temporal (e.g., duration), social, task (e.g., nature of experience), technical or informational (e.g., additional information), and economic context of the experience (Reiter et al., 2014). For instance, the QoE can be influenced by the task type, or interruptions caused by multitasking with parallel activities. Indeed, it was shown that a challenging task that requires multitasking between the system and an additional activity with same or different modalities can negatively affect the perceived quality of an experience (Reiter & Weitzel, 2007). Therefore, it is important to consider contextual factors, such as multitasking behavior, that are likely to occur or part of multimedia when assessing the QoE.

## **2.2. Multitasking with subtitles in multimedia**

With the ubiquitous presence of multimedia in our modern world, humans are constantly exposed to hedonic multitasking activities. Media multitasking occurs when the audience tries to simultaneously attend to more than one form of media (e.g., social media, movie) communicated via one or many mediums (Bardhi et al., 2010). Hedonic multimedia consumers will often engage in media multitasking to seek novel and various forms of complex information to make their activity more enjoyable (Jeong & Fishbein, 2007; Hwang et al., 2014). However, the literature has generally shown that media multitasking in academic or hedonic context has negative effects on cognitive outcomes (e.g., comprehension, recall, task performance), and that these effects are stronger when the tasks are unrelated or irrelevant to each other (Jeong & Hwang, 2016). Indeed, according to limited capacity models of attention (Baddeley et al., 1969; Kahneman, 1973; Lang, 2000) suggesting that humans have limited sensory information processing capacity due to limited available resources at a given

time. Thus, when the cognitive resources required to process a multimedia exceeds the audience's capacity, the information communicated may not be fully processed. Moreover, it is suggested that when the audience performs media multitasking, it rapidly switches its attention between different sources of information, which results in a switching attention cost that may impede processing (Gazzaley & Rosen, 2016).

However, past studies have shown the importance to distinguish irrelevant from relevant additional tasks when assessing media multitasking (Wang et al., 2015; Hwang & Jeong, 2018). The relevance of a task can be defined as the extent to which the two media tasks are related or serve a similar goal (Wang et al., 2015). Previous study has shown that, when compared to a non-multitasking newscast video watching, irrelevant multitasking (e.g. watching a video newscast while reading a text news article on a different topic) reduced the participant's performance on comprehension questions about the video, while relevant multitasking did not (e.g., watching a video newscast while reading a text news article on the same topic) (Hwang & Jeong, 2018). Indeed, according to Wang et al.'s (2015) dimensional framework of media multitasking, irrelevant multitasking may be detrimental for the media processing, while relevant multitasking can improve it.

For instance, dividing our visual attention between subtitles and image content of a multimedia is a form of relevant multitasking that may be beneficial for the audience's comprehension of dialogs and narrative coherence of multimedia (Lauvaur & Bairstow 2011; Lee et al. 2013). Displaying subtitles in multimedia is an audiovisual translation technique that has greatly improved the accessibility to the increasing worldwide traffic of audiovisual content (Duarte & Fonseca, 2019). Interlingual or standard subtitles traditionally allow the audience to read foreign dialogs in their native language, thereby removing the linguistic barrier to dialogs in movies or TV programs. Intralingual subtitles are displayed in the same language as the dialogs and can be used to reduce sensory barrier of hearing-impaired audiences, but also healthy viewers when the dialogs are difficult to understand (Gernsbacher, 2015). Indeed, the literature shows that audio-video or visual-verbal redundancy, which provide identical information via different sensory channels (Hsia, 1971) of a media can enhance

information processing and recall (Walma van der Molen, & Klijn, 2004; Walma van der Molen, & van der Voort, 2000).

Past studies measuring visual attention to subtitles with gaze measurement have found that, regardless of their type (e.g., interlingual, intralingual), instrumental subtitles can automatically draw the audience's visual attention even when they are not needed (d'Ydewalle et al., 1991; Bisson et al., 2014). For instance, d'Ydewalle et al., (1991) have shown that a native (English) group of viewers spent as much time as a foreign (Dutch) group of subjects looking at intralingual subtitles of a movie sequence presented in English. Similarly, Bisson et al., (2014) later showed that viewers allocated as much visual attention to standard (foreign language soundtrack and native language subtitles) and intralingual (foreign language soundtrack and foreign language subtitles) subtitles in films. Indeed, the last study found no differences in the fixation duration time located in the subtitle zone between standard and intralingual subtitles (Bisson et al., 2014), which suggests automatic subtitle reading behavior. Indeed, it is known that when exposed to salient visual cues, which can be determined by color contrast, orientation, or motion features of an object (Nothdurft, 1993), one's gaze can involuntarily be directed towards the object (Engelke et al., 2011).

However, multitasking with subtitles in hedonic multimedia has generally been shown to be effective, as the audience can divide its visual attention between the image and the subtitles without affecting image comprehension (Perego et al., 2010). Literature on subtitles generally points toward cognitive effectiveness of subtitles since the latter can contribute to reduce overall the cognitive load of a multimedia by improving spoken language comprehension (D'Ydewalle & De Bruycker, 2007; Kruger et al., 2013; Kruger et al., 2014 ).

### **2.3. Haptics in Multimedia**

Incorporation of haptics in multimedia entertainment experiences has become commonplace, particularly in cinema. Haptic technologies have been used to provide additional information via kinesthetic perception of body parts' position and movement, as well as tactile perception of localized cutaneous sensations delivered

on the body surface (Chaudhari et al., 2014). Thereby, haptic stimulation can serve as an additional media channel through which the human touch sensory system acquires spatial, temporal, and physical information about the environment, that is further integrated with other visual and auditory sensory input in multimedia (El Saddik et al., 2011).

In hedonic multimedia content, the synchronisation of sensory effects, such as haptic stimulation, with traditional audiovisual content can generally contribute to improve user's subjective QoE (Walt & Timmerer, 2010; Timmerer et al., 2014; Yuan et al., 2014). The QoE assessment of multimedia haptics systems generally include the perception of haptic stimulation's quality, as well as psychophysiological measures that grasp subjective and objective user's states (Danieau et al., 2013; Chaudhari et al., 2014; Mesfin et al., 2019).

One haptic technology that has recently received increasing attention by researchers uses both vibro-tactile and kinetic stimulation delivered with high fidelity and precision through a cinema seat. The High-Fidelity Vibro-Kinetic (HFVK) technology has been reported to increase immersiveness and subjective enjoyment of various audiovisual experiences such as movies (Pauna et al., 2017) and Virtual Reality experiences (Gardé et al., 2018). Traditionally, HFVK stimulation has been mainly supporting diegetic (i.e. visual) effects to enhance the sense of presence in multimedia, such as feeling the impact of a car crash in a movie scene. Recent research has also shown that HFVK stimulation targeting non-diegetic (i.e., non-visual) aspects of musical features (e.g., rhythm, melody) of pop music increased subjective enjoyment and emotional reactivity to the music listening experience (Giroux et al., 2019). More specifically, using AFE detection, the study reported that the musically aligned HFVK stimulation was associated with increased frequency of facial expressions of happiness.

However, in a follow-up study, Tchanou et al. (2019) have shown the degrading effect of multimedia multitasking with a mobile device during a multisensory music listening experience. Indeed, the additional task of browsing textual information,

which was non relevant to the music listening activity enhanced with haptic stimulation, reduced the frequency of AFE-based positive emotion detected (Tchanou et al., 2019). The study suggests that the processing of the enhanced music's emotionality was distracted by the mobile device, thereby impacting the subject's emotional reactivity to the stimuli. Moreover, the subjects also reported that they had difficulty focusing their attention on both the enhanced music and the information on the mobile device at the same time (Tchanou et al., 2019). This can be explained by the fact that humans have a limited attention which restricts the perceivable sensory information that can be attended, and thus reacted to, in an environment at once (Baddeley et al., 1969; Kahneman, 1973; Lang, 2000). Therefore, in the previous hedonic multitasking situation (Tchanou et al., 2019), it is suggested that the audience have a greater experience in both the information browsing task and the multisensory music listening activity by performing them individually.

However, the effect of haptic stimulation on the QoE related to a multimedia experience in which multitasking with textual information is integrated in the activity (e.g., reading subtitles in a movie) is unexplored. This line of research is important to understand the audience's sensory information processing limits in hedonic multisensory experiences, since there might be trade-offs in the multimedia content being processed. Indeed, while the goal of adding sensory effects into traditional multimedia activities (e.g., listening to music, watching a movie) is to improve the QoE (Walt & Timmerer, 2010; Timmerer et al., 2014; Yuan et al., 2014), content creators should make sure that the new information processing does not instead degrade it.

### **3. Methodology**

To answer our research question, we used a case of multitasking with native intralingual (NI) subtitles (i.e., presented in the audience's native tongue) in cinematic opera watching enhanced by haptic stimulation synchronized with non-verbal auditory stimulus of the multimedia. More specifically, the haptic stimulation was generated by High Fidelity Vibro-Kinetic (HFVK) technology via a cinema seat.

### **3.1. Sample**

Our sample comprised seventeen healthy students, (Males: 8, Females: 9; Mean age ± Std:  $23.7 \pm 3.5$  years), who were recruited via our institution's panel. Exclusion criteria were: non-french speakers, 17 years old or younger, hearing disabilities, epilepsy, cardiac pacemaker, skin sensitivity or skin allergy, neurological or other health related disorders. These criteria were confirmed via written informed consent prior to the commencement of the experiment. The study was approved by the ethics committee of our institution.

### **3.2. Experimental Design**

We designed a within-subject experimental study with two factors: with or without NI subtitles, and with or without HFVK stimulation. Our experiment consisted of watching 16 sequences from a 720p high definition cinematic recording of Georges Bizet's opera, during which NI subtitle presence and HFVK stimulation were randomized. The sequences were presented in chronological order according to the original recording. Prior to each sequence, a brief text was displayed and narrated by one of the researchers in order to provide context to the sequence and preserve the narrative of the opera. After viewing each sequence, subjects answered two multiple-choice questions regarding their emotional state and level of appreciation. A short interview was also conducted at the end of the experiment. See subsequent section on measures for further details on the questions and interview.

### **3.3. Audiovisual stimulus**

We chose cinematic opera for our audiovisual stimulus since Opera is considered as a “total artwork” in the music industry and known as the most emotional type of music (Cohen, 2016). Indeed, emotions can be elicited through different mechanisms involving the opera music, the staging or acting, as well as the lyrics and plot itself (Scherer et al., 2019). Therefore, cinematic opera appeared to be an interesting emotion-eliciting stimulus to assess the audience's emotional reactions. Furthermore, instrumental subtitles are traditionally shown in opera (Burton, 2009), making our audiovisual stimulus relevant for multitasking activity. More specifically, NI subtitles were originally displayed to

accommodate the hearing-impaired community (Gernsbacher, 2015). Even though they may seem to be redundant and unnecessary information for viewers with a healthy audition, they can in fact be useful when failing to entirely process the spoken dialogs of a multimedia. In opera, NI subtitles can be essential for the dialogue and narrative comprehension, since opera's international signers' vocals are not always easy to understand even when they are communicated in the audience's native language (Desblache, 2008; Orero & Matamala, 2007).

The auditory stimulus comprised four streams of information: the image, the non-verbal sound (orchestra), the verbal sound (dialog and singing), and subtitles. Verbal sound and subtitles were both presented in the audience's native language. The video was displayed on a 70x120 cm high definition TV positioned at 262 cm from the subjects, with the audio presented in 2.0 stereo on two Mission 761 speakers via a Pioneer VSX-324 AV receiver. The mean loudness of the speakers was set at approximately 65 dB, which was measured using a REED R8080 sonometer. The screen comprised an image zone vertically delimited by top and bottom black bars (see Figure 2). In the bottom black bar (hereafter: subtitle zone), the subtitles were displayed intermittently in white font.



*Figure 2: Experimental settings and audiovisual stimulus*

### **3.4. HFVK stimulus**

Our HFVK stimuli were artistically designed and programmed by D-Box Technologies Inc. (Longueuil, Canada), and further delivered by a motion seat. DBOX's motion

designers use a digital audio workstation to program and activate independent integrated actuators and vibrators synchronously with audiovisual stimulus (Roy et al., 2003; Boulais et al., 2011). As mentioned earlier, our HFVK stimulus was designed and delivered in high temporal synchrony with the non-verbal auditory elements of our audiovisual stimulus, such as the different elements of the orchestra.

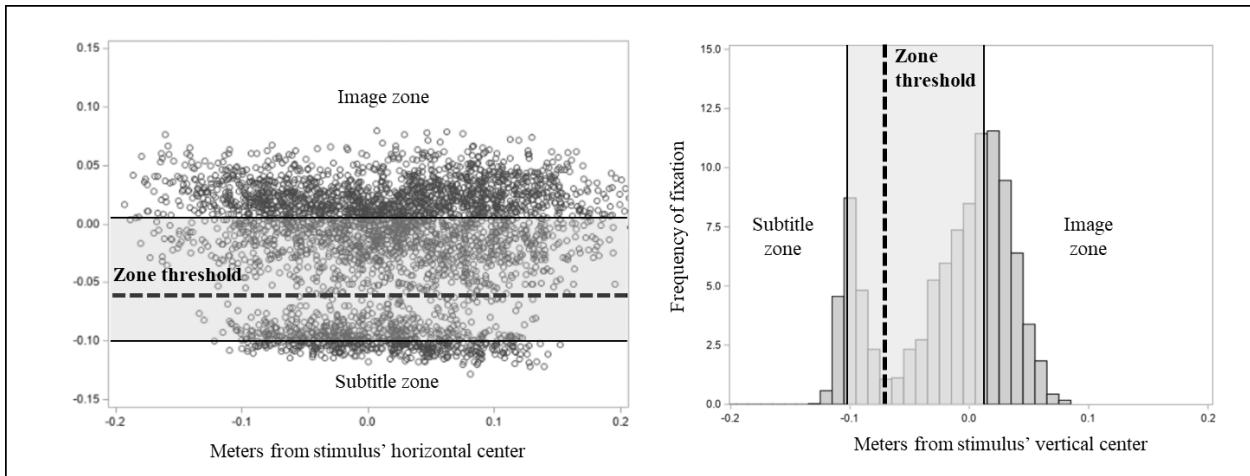
### **3.5. Procedure**

After welcoming the participants to the laboratory and completing informed consent procedures, physiological measurement devices were installed. They were then invited to sit in the HFVK seat. The HFVK seat was reclined such that their feet were lifted off the ground while maintaining as much as possible an upright posture with their upper bodies. Subjects were asked to keep their head still throughout the experiment to ensure stable face recordings by our fixed-camera for AFE detection software analysis. Lighting for this recording was also optimized using a table lamp to provide sufficient illumination of the subject's face in the darkened cinematic viewing environment. Prior to beginning the experiment, subjects read a one-page explanation about the study, which informed them that they would watch different sequences of the opera, and that the HFVK seat may or may not move, and subtitles may or may not appear depending on the randomly assigned experimental condition. After viewing all 16 sequences, we conducted a short interview (5 minutes) to assess the participant's perception about the multisensory experience, before compensating them with a \$40 gift card. Total time for the experiment was 120 minutes.

### **3.6. Measures**

Gaze tracking can be used to capture eye movements and underlying visual attentional mechanisms in a multimedia interface (Engelke et al., 2011; Engelke et al., 2017). Therefore, we assessed visual attention to the subtitles with eye tracking measurement. We recorded measures of gaze fixation location at a 60 Hz sampling rate with Smart Eye Pro 6.2 (Gothenburg, Sweden) eye tracking system. Similar as Perego et al. (2010), we delineated the image-related fixations' from the subtitle-related fixations by defining a threshold on the Y-axis of the stimulus between the two clear data populations for each

subject separately. The threshold was defined by finding, among all fixations captured during the opera watching, the Y-axis coordinate value rounded to three decimals with the lowest frequency of captured data located between the two peaks of frequency related to the subtitles and image zones (see Figure 3). When more than one Y-axis fixation value had the lowest frequency in the recorded data, we took the average value as the zone threshold. Then, fixation time in the subtitle zone was standardized according to the fixation time in the image zone during a sequence, such that (*Standardized fixation time in subtitle in sequence = fixation time in subtitle in sequence / fixation time in image in sequence*)



*Figure 3: Scatterplot and histogram and showing the threshold determination process according to the Y-axis fixation location value between the two fixation's population.*

Eye tracking data was simultaneously recorded with the video of the subjects' face, which was continuously sampled at 30 frames per second (fps) using a HD webcam and MediaRecorder (Noldus, Wageningen, Netherlands). The facial recordings of the participants when they were watching the sequences were analyzed with the AFE detection software FaceReader 7.0 (Noldus, Wageningen, Netherlands) at a 10 Hz sampling rate. FaceReader systems (Noldus, Wageningen, Netherlands) were traditionally validated by applying artificial neural network on a sample of static images of humans mimicking facial expressions of emotions (Skiendziel et al., 2019) according to the Facial Action Coding System (FACS) (Ekman and Friesen, 1978). FaceReader outputs the frequency (%) of a total of six discrete emotions, including one positive

(happiness), four negative (anger, disgust, sadness, and scared), and one neutral (surprise) discrete emotions based on face micro-muscle configurations. The AFE based emotion frequency data was upsampled, using MatLab (MathWorks, Massachusetts, USA), in order to match the eye tracker's sampling rate, permitting estimation of AFE based emotion for each gaze fixation event. Then, each of the five positive and negative AFE based emotions' frequency were individually aggregated by experimental condition, according to the two fixation zones.

Since we did not priorly assess the different opera scenes' induced emotions, we used subjective measurement of emotional valence to contrast with the AFE-based emotions detected. For instance, recent research in music psychology has assessed music induced-emotions via AFE detection methods and self-reported ratings of each of the six discrete emotions (Weth et al., 2015). Other studies have also used the 1-item SAM scale's emotional valence dimension (Bradley and Lang, 1994) as it is faster to administer than each discrete emotion when time can be a limitation in an experimental study. Indeed, the nine-point SAM scale version of the emotional valence dimension allows to assess one's own subjective feeling according to a pictorial scale ranging from a the negative feeling of sadness (1), to a neutral state (5), to the positive feeling of happiness (9). To remain consistent with the SAM scale emotional valence dimension, which evaluates the emotional states in terms of negative or positive values, we excluded surprise from the six discrete emotions provided by AFE detection systems as it is neither a negative or positive discrete emotion.

The subjective appreciation level of experience was measured via a 7-point scale, ranging from "I did not enjoy the experience at all" (1) to "I extremely enjoyed the experience" (7). Both the emotional valence dimension of the SAM scale and the subjective level of appreciation scale were administered after each sequence using Qualtrics software on a touch-screen tablet.

Finally, during a final retrospective interview, we asked the participants a series of five questions to gather additional subjective information to better interpret the user's QoE (see Table 2). The participants were also asked to elaborate on their answers.

*Table 2: Interview questions*

Interview questions
<b>Q1:</b> Generally, were you able to understand the lyrics of the opera without the subtitles being present?
<b>Q2:</b> Generally, did the subtitles allow you to better understand the opera narrative?
<b>Q3:</b> Generally, did you prefer viewing the opera with or without subtitles?
<b>Q4:</b> Generally, did you prefer viewing the opera with or without vibrations and movements generated by the chair?
<b>Q5:</b> Generally, do you think that the vibrations and movements generated by the chair have influenced your visual attention to the TV screen?

### 3.7. Statistical Analysis

We performed a four-way repeated measures analysis of variance (RM ANOVA) to assess the differences in the mean frequency of AFE-based emotions according to the HFVK stimulation, NI subtitle presence, the stimulus zone in which the fixation was associated, and the AFE-based emotion type (2: HFVK × 2: NI subtitle presence × 2: fixation zone × 5: AFE-based emotion type). We also performed two two-way RM ANOVAs to assess the differences in the subjective emotional valence and subjective level of appreciation according to the NI subtitle presence and the HFVK stimulation. Simple main effects have been reported for any significant two, three, or four-way interactions in the aforementioned ANOVAs. Finally, we performed a paired t-test to compare the total fixation time in the subtitle zone between the conditions with and without HFVK stimulation, but only with NI subtitle presence. We also aggregated the overall qualitative results from the final interviews, which are presented by questions. All statistical tests were done using SPSS software version 22 (IBM, Armonk, NY, USA), with the threshold for statistical significance set at  $p \leq 0.05$ .

## 4. Results

The following table (see Table 3) shows the overall mean and standard deviation for each variable of interest, by experimental condition (if applicable), and by gaze's fixation zone (if applicable).

*Table 3: Overall descriptive statistic's results*

Variables	Experimental conditions															
	HFVK + Subtitles + Audiovisual				HFVK + Audiovisual				Subtitles + Audiovisual				Audiovisual only			
	Image zone		Subtitle zone		Image zone		Subtitle zone*		Image zone		Subtitle zone		Image zone		Subtitle zone*	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Happiness	,047	,037	,046	,041	,053	,040	,062	,048	,050	,035	,050	,038	,052	,041	,050	,050
Sadness	,267	,139	,271	,149	,276	,150	,287	,174	,266	,124	,292	,141	,227	,118	,211	,148
Anger	,083	,045	,078	,045	,087	,046	,077	,050	,093	,051	,085	,050	,089	,052	,090	,051
Fear	,061	,046	,057	,044	,060	,046	,067	,054	,063	,047	,059	,048	,062	,044	,060	,051
Disgusted	,035	,064	,035	,068	,036	,063	,038	,066	,036	,061	,036	,065	,038	,075	,035	,070
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Proportion of fixation time in subtitle zone	,206		,017						,201		,021					
SAM scale emotional valence (1-9)	5.65		1.63		5.58		1.70		5.75		1.40		5.74		1.19	
Subjective appreciation level scale (1-7)	5.40		1.30		5.14		1.45		5.17		1.24		4.93		1.09	

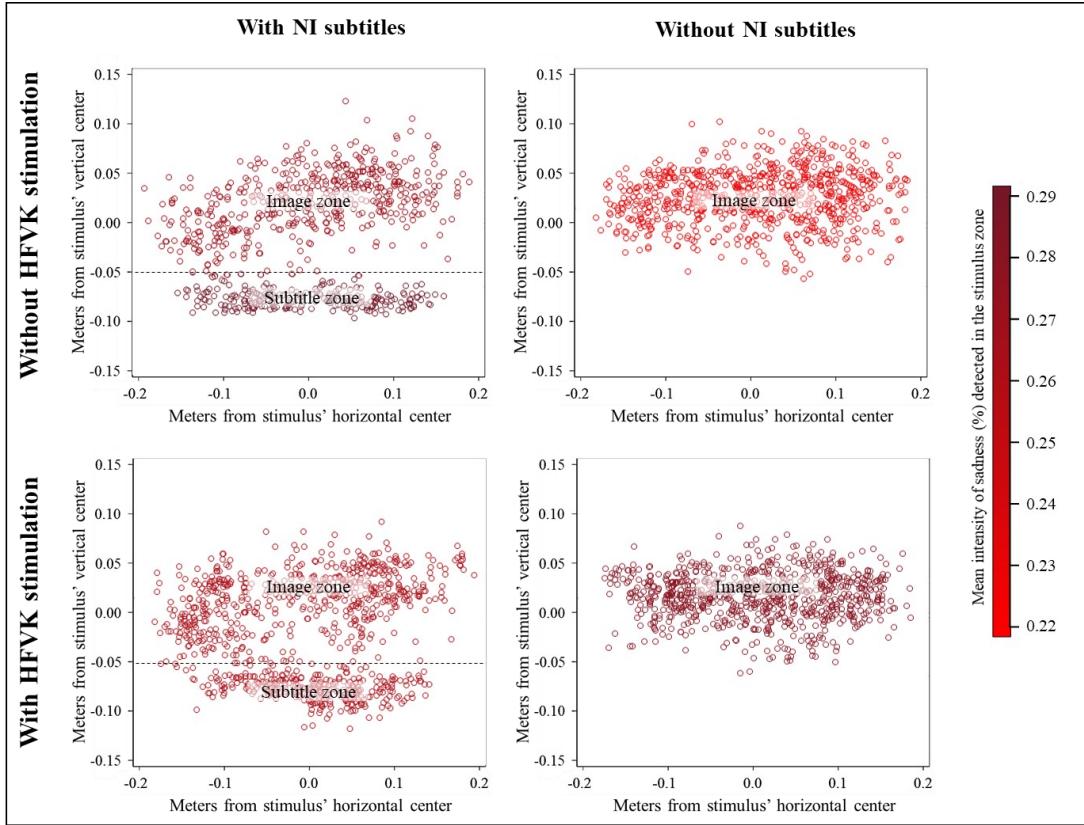
\* Low amount of fixation due to absence of subtitles in these conditions

Our four-way RM ANOVA's results revealed no main effect of HFVK stimulation ( $F(1,16) = 1.104$ ,  $p < .309$ ), nor NI subtitle presence ( $F(1,16) = .590$ ,  $p < .454$  ), nor gaze fixation zone ( $F(1,16) = .041$ ,  $p < .841$ ) on the mean frequency of AFE-based emotions. However, we found a significant main effect of AFE-based emotion type, which indicates that the frequency of AFE-based emotion detection varied among the emotion types ( $F(4,64) = 27.869$ ,  $p < .001$ ). The emotion type detected with the highest mean frequency was sadness ( $m = .263$ ,  $sd = .032$ ), followed by anger ( $m = .086$ ,  $sd = .011$ ), fear ( $m = .061$ ,  $sd = .011$ ), happiness ( $m = .052$ ,  $sd = .009$ ) and disgust ( $m = .037$ ,  $sd = .016$ ). The results indicate that the stimulus generally induced facial muscle configurations associated with negative emotions, particularly sadness.

The four-way RM ANOVA also revealed a significant interaction between NI subtitle presence and AFE-based emotion type ( $F(1,16) = 3.764$ ,  $p < .008$ ), and a significant three-way interaction between NI subtitle presence, AFE-based emotion type, and gaze fixation

zone ( $F(4,64) = 3.011$ ,  $p < .024$ ). Simple main effects testing showed that, when comparing the frequency of each AFE-based emotions between the subtitle zone and the image zone only when subtitles were present, greater frequency of the AFE-based sadness ( $p < .022$ ), but lower frequency of anger ( $p < .027$ ), and fear ( $p < .023$ ) were detected in the subtitle zone irrespective of HFVK stimulation.

With respect to HFVK, the four-way ANOVA did not reveal a significant interaction between HFVK stimulation and AFE-based emotion type ( $F(4,64) = 2.168$ ,  $p < .083$ ), but did show a significant interaction between HFVK stimulation and NI subtitle presence on the mean frequency of AFE-based emotions ( $F(1,16) = 8.311$ ,  $p < .011$ ), which further revealed a significant three-way interaction between HFVK stimulation, NI subtitle presence, and AFE-based emotion type ( $F(4,64) = 3.687$ ,  $p < .009$ ). Simple main effects comparisons between subtitle conditions indicated that, of the different AFE-based emotion types, sadness was detected with significantly higher frequency when subtitles were present ( $m = .282$ ,  $sd = .038$ ), compared to when subtitles were absent ( $m = .220$ ,  $sd = .032$ ), but only when no HFVK stimulation was delivered ( $p < .033$ ). Delivery of HFVK stimulation abolished this effect ( $p < .341$ ). Figure 5 shows the contrast between the frequency of sadness AFE based emotion detected in stimulus' screen according to the condition.



*Figure 4: One subject's aggregated fixations associated with automatic facial expression-based sadness. Fixations are displayed in 2D plans. The mean sadness frequency in each condition's zone of interest is indicated by the darkness of the colored circles.*

The four-way RM ANOVA results showed no significant interaction between subtitle presence and gaze fixation zone ( $F(1,16) = .146, p < .707$ ), nor between gaze fixation zone and AFE-based emotion type ( $F(4,64) = 1.188, p < .325$ ). Moreover, no significant interactions were observed between HFVK stimulation and gaze fixation zone ( $F(1,16) = .425, p < .524$ ); HFVK stimulation, gaze fixation zone, and AFE-based emotion type ( $F(4,64) = .263, p < .901$ ); HFVK stimulation, subtitle presence, and gaze fixation zone ( $F(1,16) = 1.819, p < .196$ ); and HFVK stimulation, subtitle presence, gaze fixation zone, and AFE-based emotion type ( $F(4,64) = 1.855, p < .129$ ).

As for subjective emotional valence, our second RM ANOVA revealed no main effect of HFVK stimulation ( $F(1,16) = 0.108, p < .746$ ), nor main effect of subtitle presence ( $F(1,16) = 0.050, p < .826$ ), nor interaction between HFVK stimulation and subtitle presence ( $F(1,16) = 0.089, p < .770$ ) on subjective emotional valence. Nevertheless, there

was a trend towards lower emotional valence scores in subtitled ( $m = 5.65$ ,  $sd = 1.63$  vs.  $m = 5.75$ ,  $sd = 1.40$ ) and non-subtitled conditions ( $m = 5.58$ ,  $sd = 1.70$  vs.  $m = 5.74$ ,  $sd = 1.19$ ) for HFVK compared to non-HFVK conditions, respectively.

Regarding the subjective appreciation level, our third RM ANOVA revealed a significant main effect of subtitle presence ( $F(1,16) = 5.292$ ,  $p < .035$ ) with greater appreciation of the cinematic opera experience with NI subtitles regardless of HFVK stimulation delivery. The cinematic opera experience was most appreciated when both NI subtitles and HFVK stimulation were present ( $m = 5.40$ ,  $sd = 1.30$ ), followed by without HFVK stimulation but with NI subtitles ( $m = 5.17$ ,  $sd = 1.24$ ), and with HFVK and without NI subtitles ( $m = 5.14$ ,  $sd = 1.45$ ), and finally without HFVK stimulation or NI subtitles ( $m = 4.93$ ,  $s = 1.09$ ). No main effect of HFVK stimulation ( $F(1,16) = 0.371$ ,  $p < .551$ ) nor interaction between NI subtitles and HFVK stimulation ( $F(1,16) = 0.031$ ,  $p < .863$ ) on subjective appreciation level was observed.

These results are in line with our interview reports. Indeed, when asked the question “Generally, were you able to understand the lyrics of the opera without the subtitles being present? (Q1)”, none of the participants answered that they could fully understand every of them. In fact, sixteen out of seventeen participants revealed the subtitles helped them better understand the story, when asked the question “Generally, did the subtitles allow you to better understand the opera narrative? (Q2)”. Our third question “ Generally, did you prefer viewing the opera with or without subtitles? (Q3)” revealed that fifteen out of seventeen participants preferred having subtitles to better understand the lyrics and the story. Although our subjective appreciation level scale results showed no effect of HFVK stimulation, when asked the question “Generally, did you prefer viewing the opera with or without vibrations and movements generated by the chair? (Q4)” forteen out of seventeen participants stated that they preferred watching the audiovisual content with the HFVK stimulation.

Regarding our last question “Generally, do you think that the vibrations and movements generated by the chair have influenced your visual attention to the TV screen? (Q5)”, eight of our participants revealed that the HFVK stimulation made them focus more on

the image content features such as the scene action. For instance, two participants stated: “Yes, it (the HFVK stimulation) drew my visual attention towards the dramaticality, the action, the image.” P11 or “The movements reproduced what the orchestra was playing, so they necessarily draw our eyes to the image. P17”

Interestingly, our one-way ANOVA on the proportion of fixation time in the subtitle zone when subtitles were present revealed no main effect of HFVK stimulation ( $F(1,16) = 0.06$ ,  $p < .809$ ), which indicates that mean proportion of fixation time located in the subtitle zone with HFVK stimulation ( $m = .206$   $sd = .017$ ) was not different than without it ( $m = .201$ ,  $sd = .021$ ). Moreover, our gaze tracking data showed that the proportion of fixation time in the subtitles of the group of eight participants that reported having allocated more visual attention the image content due to the HFVK stimulation ( $m = .265$ ,  $sd = .081$ ) was much similar to the rest of the participants ( $m = .261$ ,  $sd = .067$ ). This indicates that these participants may have unconsciously or naturally allocated visual attention to the subtitles despite the haptic stimulation.

## 5. Discussion and Conclusion

This study investigated the moderating effect of haptic stimulation on multimedia multitasking with instrumental text processing in a hedonic cinematic viewing experience. More specifically, we have shown that HFVK stimulation synchronized with non-verbal auditory stimulus of the multimedia does not influence the fixation time naturally allocated to the subtitles in cinematic opera watching. However, this is the first study that considers the overstimulation of the audience during a hedonic cinematic viewing experience involving haptic stimulations and subtitles in multimedia.

First, as we expected, our eye tracking results clearly indicate that the subjects visually attended the subtitle zone of the TV screen. Indeed, our final interviews revealed that the subtitles were generally necessary for the subject’s comprehension of lyrics. Our results further show that subjects spent the same amount of time fixating the subtitle zone of the screen regardless of HFVK stimulation, which suggests that the latter did not interfere with the natural subtitle reading behavior in this context, thus supporting H1. Interestingly, when we asked participants about their perception of the effect of

mouvements on their visual attention to the screen, most of them spontaneously revealed that their visual attention was drawn to the action in the image when the HFVK stimulation was delivered. However, our gaze tracking measurement suggests that participants' perception of their own visual attention might have been influenced by the HFVK stimulation. Indeed, our results suggest that participants unconsciously used the subtitles regardless of the HFVK stimulation, which is in line with automatic subtitle reading behavior in the literature (d'Ydewalle et al., 1991; Bisson et al., 2014)

Regarding our second hypothesis, the frequency of AFE-based emotions, and more specifically the emotion of sadness detected during the sequence was found to be increased with NI subtitle presence. It is possible that since Carmen is a tragedy, the increased comprehension of the dialogs and narratives provided by the NI subtitles increased the audiovisual stimuli's induced sadness. However, the increased frequency of sadness emotion during the sequence seemed to be due to the subtitle processing. Indeed, the frequency of AFE-based emotions detected when the subjects were looking at the subtitle zone was greater than when looking at the image zone. While the literature suggests a negative linear relationship between corrugator supercilii face muscle activity and affective words (Larsen et al, 2003), it could be that reading the lyrics of Carmen evoked facial expressions of sadness. The fact that more sadness is detected when reading could also be explained by past studies, which have shown greater activity in the corrugator supercilii facial muscle responsible for eyebrow frown when reading a text with poor typography as compared to optimized page settings (Larson et al., 2014). Similarly, it is possible the subtitles readability may have been impacted by the far distance between the participants and the T.V, which made them furrow their eyebrows or frown. Indeed, the discrete emotion of sadness can be determined via specific configurations of facial Action Units (AU) that contains AU4 (lowered brows or eyebrow "frown") according to the Facial Action Coding System (FACS) (Ekman and Friesen, 1978) on which AFE detection softwares such as FaceReader (Noldus, Netherlands) are based on. However, this assumption is weakened by the fact that lower frequency of anger and fear were detected in the subtitle zone, as compared to the image zone, independently of HFVK stimulation. Indeed, both anger and fear discrete emotions are also predicted with compounds of facial Action Units including AU4. In other words, they also involve

corrugator supercilii activation (AU4), and thereby rule out the possibility that corrugator supercilii face muscle activity may be triggered by the act of reading text. Therefore, we suggest that the increased frequency of sadness emotion was indeed emotionally driven by improved comprehension of the lyrics and tragedy.

As for the effect of HFVK stimulation, our results show that only when the latter was not delivered, the NI subtitle presence had an increasing effect of sadness frequency, suggesting that the delivery of HFVK abolished this effect. Interestingly, although not significant, our results regarding the non-subtitled conditions indicate that the HFVK stimulation may generate facial expressions of sadness, which would be consistent with prior study on sad music using measures of corrugator supercilii via EMG (Livingstone et al., 2009; Witvliet and Vrana, 2007). However, our results show that the effect of HFVK stimulation did not seem to add up with the effect of NI subtitles on the frequency of sadness detected, and might even have reduced this effect.

Regarding our subjective measurement, although the HFVK stimulation seems to have led to lower scores of subjective emotional valence, thus pointing towards the scale's negative pole (feeling of sadness), our results were not strong enough to show any significant effect of HFVK stimulation, nor NI subtitles. As mentioned by Scherer (2004), self-reported methods in music may only grasp conscious feelings related to music, which are often difficult to describe and prone to response bias. Therefore, it is possible that the participants' ratings were biased by their interpretation of how they should feel according to the hedonic experience, thus increasing positive emotions values. Indeed, Kayser (2017) recently suggested that since emotions were difficult to measure subjectively, facial expressions of discrete emotions may offer a better alternative to self-reported methods to assess music induced-emotions. Therefore, our objective measurement indicates that the effect of HFVK stimulation did not have a positive effect on emotional reactivity to subtitled cinematic opera, thus rejecting H2.

So far, H1 and H2 revealed that the HFVK stimulation does not influence the natural subtitle reading behavior, however the latter may reduce the enhancing effect of HFVK on the audience's emotional reactivity. Our results indicate that the NI subtitles

increased subjective appreciation of the experience regardless of HFVK stimulation. This was supported by our final interview revealing that most of the participants preferred having the subtitles, as all of them could not understand all the spoken dialogs. Although the participants generally revealed that they preferred having the HFVK stimulation than not having it, the HFVK stimulation did not significantly increase the subjective appreciation level assessed by scale. This can be explained by the fact that the NI subtitle presence had greater influence on the level of appreciation of the cinematic opera experience and might have potentially masked the effect of HFVK stimulation on the audience's perception. Indeed, since our experimental design emphasized on the story narratives by providing short introductory text on the events occurring between the presented sequences, the participants may have been encouraged to follow the narratives. Thereby, the NI subtitles may have been a necessity to the cinematic opera experience, assuming that their goal was to achieve comprehension, while the HFVK stimulation may have been a nonessential added value. Therefore, we reject H3.

Overall, our study is the first to consider overstimulation in a hedonic multimedia context. Indeed, our findings indicate that participants may have had a limited capacity in emotionally reacting to both the subtitles and the haptic-enhanced opera music. Therefore, given the growing evidence for automatic subtitle reading behavior (Bisson et al., 2014; d'Ydewalle et al., 1991) in multimedia audience, it could be interesting to assess the need to display all the subtitles, including those translating the sound effects (e.g. "cough", "explosion"), or redundant subtitles as they may unnecessarily disrupt the integration of haptic stimulation with other sensory information. Therefore, our study contributes to research on multimedia and more specifically multimedia haptics by highlighting potential trade-offs in the sensory information processed by the audience in mulsemedia experiences involving multitasking with textual information.

Our study also contributes to the evaluation of QoE in multimedia research, and more broadly to UX methods in HCI research using AFE detection methods. Popular commercially available AFE detection systems have recently faced criticism regarding their effectiveness in assessing spontaneous facial expressions in natural context (Dupré et al., 2020). Consequently, it is suggested that scholars integrate relevant contextual

information which may allow them to better interpret the ongoing facial micro-movements captured during a task. Our study computer-based experimental context integrates the visual task performed in the interface (e.g., text reading or image processing) as contextual information to assess fixation-related AFE-based emotions. We believe that this method can be a good practice, considering that non-emotional drivers of facial movements that are specific to visual tasks, such as the readability of a text (Larson et al., 2014), could be misrecognized as discrete emotions by AFE detection systems. Thereby, we contribute to the development of objective methods to assess emotions from spontaneous facial expressions captured in computer-based contexts involving text reading activities.

Finally, we hope that our study will encourage the use of AFE detection systems as non-obtrusive and objective tools to assess emotional reactivity to music, and to multimedia in general as part of the Quality of Experience assessment (Reiter et al., 2014). Indeed, as emotions can be challenging to measure subjectively (Kayser, 2017), AFE detection methods seem to be an interesting alternative to assess stimulus induced emotions in multimedia experience.

This study has limits. First, since the HFVK stimulation design is an artistic and thus subjective process, scholars must be cautious when generalizing the findings of the HFVK technology to other haptic technologies. Also, the HFVK effects were synchronized with the nonverbal auditory cues of our multimedia. Therefore, it would be important to investigate the effect haptic stimulation synchronized with visual cues on subtitle reading behavior, especially since it is known that simultaneously occurring tactile cue and visual object can attract visual attention to this object (Burg et al., 2009). Second, the chosen cinematic opera sequences were not tested and validated for their induced emotions prior to the experiment. Research should investigate the property of HFVK stimulation in enhancing the specific emotions that are induced by scientifically validated emotion eliciting multimedia. Finally, although our participants reported improved comprehension with the subtitles, we have not measured it directly. Therefore, it would be important to explore the extent to which HFVK stimulation can influence the audience's comprehension of a multimedia.

## References

- Akhtar, Z., Siddique, K., Rattani, A., Lutfi, S. L., & Falk, T. H. (2019). Why is multimedia quality of experience assessment a challenging problem? *IEEE Access*, 7, 117897–117915. <https://doi.org/10.1109/ACCESS.2019.2936470>
- Arndt, S., Brunnstrdnr, K., Cheng, E., Engelke, U., Moller, S., & Antons, J. N. (2016). Review on using physiology in quality of experience. *Human Vision and Electronic Imaging 2016, HVEI 2016*, 197–205. <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-125>
- Baddeley, A. D., Scott, D., Drynan, R., & Smith, J. C. (1969). Short-term memory and the limited capacity hypothesis. *British Journal of Psychology*. <https://doi.org/10.1111/j.2044-8295.1969.tb01175.x>
- Balteş, F. R., Avram, J., Miclea, M., & Miu, A. C. (2011). Emotions induced by operatic music: Psychophysiological effects of music, plot, and acting. A scientist's tribute to Maria Callas. *Brain and Cognition*, 76(1), 146–157. <https://doi.org/10.1016/j.bandc.2011.01.012>
- Bardhi, F., Rohm, A., & Sultan, F. (2010). Tuning in and tuning out: media multitasking among young consumers. *Journal of Consumer Behaviour*, 9(July–August), 316–332. <https://doi.org/10.1002/cb>
- Bisson, M.-J., Van Heuven, W., Conklin, K., & Tunney, R. J. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, 35(2), 399–418. <https://doi.org/10.1017/S0142716412000434>
- Boulais, S., Lizotte, J. M., Trottier, S., & Gagnon, S. (2011). US 7,934,773 B2. Canada.
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Brocke, J., Hevner, A., Léger, P. M., & Walla, P. (2020). Advancing a NeuroIS research

- agenda with four areas of societal contributions. European Journal of Information Systems, 29(1), 9–24. <https://doi.org/10.1080/0960085X.2019.1708218>
- Brunnström, K., Beker, S. A., De, K., Dooms, A., Egger, S., Garcia, M.-N., ... Larabi, M.-C. (2013). Qualinet White Paper on Definitions of Quality of Experience. European Network on Quality of Experience in in Multimedia Systems and Services (COST Action IC 1003), (March), 26.
- Burg, E. Van Der, Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2009). Poke and pop : Tactile – visual synchrony increases visual saliency. Neuroscience Letters, 450, 60–64. <https://doi.org/10.1016/j.neulet.2008.11.002>
- Burton, J. (2009). Burton, Jonathan (2009) The Art and Craft of Opera Surtitling, in Audiovisual translation: Language transfer on screen. Basingstoke, Hertfordshire: Palgrave Macmillan. <https://doi.org/10.1057/9780230234581>
- Chaudhari, R., Altinsoy, E., & Steinbach, E. (2014). Haptics. In Quality of Experience (pp. 261–276). <https://doi.org/10.1007/978-3-319-02681-7>
- Cohen, A. J. (2016). Music in performance arts: Film, theatre and dance. In The Oxford Handbook of Music Psychology.  
<https://doi.org/10.1093/oxfordhb/9780199298457.013.0041>
- D'Ydewalle, G., Praet, C., Verfaillie, K., & van Rensbergen, J. (1991). Watching subtitled television: Automatic reading behavior. Communication Research, 18, 650–665.
- D'Ydewalle, G., & Bruycker, W. De. (2007). Eye Movements of Children and Adults While Reading Television Subtitles. European Psychologist, 12(3), 196–205. <https://doi.org/10.1027/1016-9040.12.3.196>
- Danieau, F., Lécuyer, A., Guillotel, P., Fleureau, J., Mollet, N., & Christie, M. (2013). Enhancing audiovisual experience with haptic feedback: A survey on HAV. IEEE Transactions on Haptics, 6(2), 193–205. <https://doi.org/10.1109/TOH.2012.70>

- Desblache, L. (2008). Music to my ears, but words to my eyes?: Text, opera and their audiences. *Linguistica Antverpiensia*, (6), 155–170.
- Duarte, C., & Fonseca, M. J. (2019). Multimedia Accessibility. In Web Accessibility (pp. 461–475). [https://doi.org/10.1007/978-1-4471-7440-0\\_25](https://doi.org/10.1007/978-1-4471-7440-0_25)
- Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *PLoS ONE*, 15(4), 1–17. <https://doi.org/10.1371/journal.pone.0231968>
- Egger, M., Ley, M., & Hanke, S. (2019). Emotion Recognition from Physiological Signal Analysis: A Review. *Electronic Notes in Theoretical Computer Science*, 343, 35–55. <https://doi.org/10.1016/j.entcs.2019.04.009>
- Ekman, P. (1999). Basic Emotions. In *Handbook of Cognition and Emotion*. <https://doi.org/10.1002/0470013494.ch3>
- Ekman, P., & Friesen, W. V. (1978). The Facial Action Coding System. San Francisco, CA: Consulting Psychologists Press. <https://doi.org/10.1037/h0030377>
- El Saddik, A., Orozco, M., Eid, M., & Cha, J. (2011). Haptics Technologies: Bringing Touch to Multimedia. *Haptics Technologies*. <https://doi.org/10.1007/978-3-642-22658-8>
- Engelke, U., Darcy, D. P., Mulliken, G. H., Bosse, S., Martini, M. G., Arndt, S., ... Brunnström, K. (2017). Psychophysiology-Based QoE Assessment: A Survey. *IEEE Journal on Selected Topics in Signal Processing*, 11(1), 6–21. <https://doi.org/10.1109/JSTSP.2016.2609843>
- Engelke, U., Kaprykowsky, H., Zepernick, H., & Ndjiki-nya, P. (2011). Visual attention in quality assessment: Theory, advances, and challenges. *IEEE SIGNAL PROCESSING MAGAZINE*, (November), 50–59.
- Gardé, A., Léger, P., Sénechal, S., Fredette, M., Chen, S. L., Labonté-Lemoyne, É., & Ménard, J. (2018). Virtual Reality: Impact of Vibro-Kinetic Technology on

- Immersion and Psychophysiological State in Passive Seated Vehicular Movement. In Haptics: Science, Technology, and Applications. EuroHaptics 2018. Lecture Notes in Computer Science (pp. 264–275). Springer International Publishing. <https://doi.org/10.1007/978-3-319-93399-3>
- Gazzaley, A., & Rosen, L. D. (2016). *The distracted mind : ancient brains in a high-tech world*. MIT Press, Cambridge, MA.
- Gernsbacher, M. A. (2015). Video Captions Benefit Everyone. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 195–202. <https://doi.org/10.1177/2372732215602130>
- Giroux, F., Boasen, J., Sénécal, S., Fredette, M., Tchanou, A. Q., Ménard, F., ... Léger, P. (2019). Haptic Stimulation with High Fidelity Vibro-Kinetic Technology Psychophysiological Enhances Seated Active Music Listening Experience. In IEEE World Haptics Conference (WHC) Tokyo, Japan.
- Golland, Y., Hakim, A., Aloni, T., Schaefer, S., & Levit-binnun, N. (2018). Affect dynamics of facial EMG during continuous emotional experiences. *Biological Psychology*, 139(October), 47–58. <https://doi.org/10.1016/j.biopsych.2018.10.003>
- Hsia, H. J. (1971). The information processing capacity of modality and channel performance. *Educational Communication and Technology Journal*, 19(1), 51–75. <https://doi.org/10.1007/BF02768431>
- Hwang, Y., & Jeong, S.-H. (2018). Multitasking and Information Gain : Effects of Relevance between Tasks. *Journal of Media Economics & Culture*, 16(2), 50–77. <https://doi.org/10.21328/jmec.2018.5.16.2.50>
- Hwang, Y., Kim, H., & Jeong, S. H. (2014). Why do media users multitask?: Motives for general, medium-specific, and content-specific types of multitasking. *Computers in Human Behavior*, 36, 542–548. <https://doi.org/10.1016/j.chb.2014.04.040>
- Jeong, S. H., & Fishbein, M. (2007). Predictors of multitasking with media: Media

- factors and audience factors. *Media Psychology*, 10(3), 364–384.  
<https://doi.org/10.1080/15213260701532948>
- Jeong, S. H., & Hwang, Y. (2016). Media Multitasking Effects on Cognitive vs. Attitudinal Outcomes: A Meta-Analysis. *Human Communication Research*, 42(4), 599–618. <https://doi.org/10.1111/hcre.12089>
- Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.
- Kayser, D. (2017). Using facial expressions of emotion as a means for studying music-induced emotions. *Psychomusicology: Music, Mind, and Brain*, 27(3), 219–222.  
<https://doi.org/10.1037/pmu0000187>
- Kruger, J. L., Hefer, E., & Matthew, G. (2014). Attention distribution and cognitive load in a subtitled academic lecture: L1 vs. L2. *Journal of Eye Movement Research*, 7(5). <https://doi.org/10.16910/jemr.7.5.4>
- Kruger, J., & Matthew, G. (2013). Measuring the Impact of Subtitles on Cognitive Load: Eye Tracking and Dynamic Audiovisual Texts. *Proceedings of Eye Tracking South Africa*, (August), 29–31.
- Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>
- Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, 40(5), 776–785. <https://doi.org/10.1111/1469-8986.00078>
- Larson, K., Hazlett, R. L., Chaparro, B. S., & Picard, R. W. (2007). Measuring the Aesthetics of Reading. *People and Computers XX — Engage*, 41–56.  
[https://doi.org/10.1007/978-1-84628-664-3\\_4](https://doi.org/10.1007/978-1-84628-664-3_4)
- Lavaur, J. M., & Bairstow, D. (2011). Languages on the screen: Is film comprehension

- related to the viewers' fluency level and to the language in the subtitles? International Journal of Psychology, 46(6), 455–462.  
<https://doi.org/10.1080/00207594.2011.565343>
- Lee, M., Roskos, B., & Ewoldsen, D. R. (2013). The Impact of Subtitles on Comprehension of Narrative Film. *Media Psychology*, 16(4), 412–440.  
<https://doi.org/10.1080/15213269.2013.826119>
- Livingstone, S., & Thompson, William Forde Russo, F. A. (2009). Facial Expressions and Emotional Singing: A Study of Perception and Production with Motion Capture and Electromyography. *Facial Expressions in Song*, 26(5), 475–488.
- Lundqvist, L.-O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2009). Emotional responses to music: experience, expression, and physiology. *Psychology of Music*, 37(1), 61–90. <https://doi.org/10.1177/0305735607086048>
- Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behavior And Information Technology*, 25(2), 141–158.  
<https://doi.org/doi.org/10.1080/01449290500331156>
- Mesfin, G., Hussain, N., Covaci, A., & Ghinea, G. (2019). Using eye tracking and heart-rate activity to examine crossmodal correspondences QoE in Mulsemedia. *ACM Transactions on Multimedia Computing, Communications and Applications*, 15(2).  
<https://doi.org/10.1145/3303080>
- Middlebrooks, C. D., Kerr, T., & Castel, A. D. (2017). Selectively Distracted: Divided Attention and Memory for Important Information. *Psychological Science*, 28(8), 1103–1115. <https://doi.org/10.1177/0956797617702502>
- Moon, S. E., & Lee, J. S. (2017). Implicit Analysis of Perceptual Multimedia Experience Based on Physiological Response: A Review. *IEEE Transactions on Multimedia*, 19(2), 340–353. <https://doi.org/10.1109/TMM.2016.2614880>
- Nothdurft, H. C. (1993). Saliency effects across dimensions in visual search. *Vision*

Research, 33(5–6), 839–844. [https://doi.org/10.1016/0042-6989\(93\)90202-8](https://doi.org/10.1016/0042-6989(93)90202-8)

Orero, P., & Matamala, A. (2007). Accessible Opera : Overcoming Linguistic and Sensorial Barriers. Perspectives: Studies in Translatology, 15(4), 262–277. <https://doi.org/10.1080/13670050802326766>

Pantic, M., & Rothkrantz, L. Ü. M. (2000). Automatic analysis of facial expressions: The state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1424–1445. <https://doi.org/10.1109/34.895976>

Pauna, H., Léger, P. M., Sénécal, S., Fredette, M., Courtemanche, F., Chen, S. L., ... Ménard, J. F. (2017). The psychophysiological effect of a vibro-kinetic movie experience: The case of the D-BOX movie seat. In Information Systems and Neuroscience: Lecture Notes in Information Systems and Organisation (Vol. 25, pp. 1–7). [https://doi.org/10.1007/978-3-319-67431-5\\_1](https://doi.org/10.1007/978-3-319-67431-5_1)

Perego, E., Missier, F. Del, Porta, M., & Mosconi, M. (2010). The Cognitive Effectiveness of Subtitle Processing. Media Psychology, 13(3), 243–272. <https://doi.org/10.1080/15213269.2010.502873>

Reiter, U., Brunnström, K., De Moor, K., Larabi, Mohamed-Chaker Pereira, M., Pinheiro, A., You, J., & Zgank, A. (2014). Factors Influencing Quality of Experience. In Quality of Experience (pp. 55–72). <https://doi.org/10.1007/978-3-319-02681-7>

Reiter, U., & Weitzel, M. (2007). Influence of interaction on perceived quality in audiovisual applications: evaluation of cross-modal influence. Proc. 13th International Conference on Auditory Display, 380–385. Retrieved from <http://dev.icad.org/Proceedings/2007/ReiterWeitzel2007.pdf>

Riedl, René, Léger, P.-M. (2016). Fundamentals of NeuroIS: Information Systems and the Brain. Studies in Neuroscience, Psychology and Behavioral Economics (Vol. 1). Studies in Neuroscience, Psychology and Behavioral Economics. <https://doi.org/10.1136/bmj.1.3968.171-a>

- Roy, P., Bérubé, M., & Jacques, M. (2000). US 6,585,515 B1. Canada.
- Scherer, K. R. (2004). Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them? *Journal of New Music Research*, 33(3), 239–251. <https://doi.org/10.1080/0929821042000317822>
- Scherer, K. R., & Coutinho, E. (2019). Assessing Emotional Experiences of Opera Spectators in Situ. *Psychology of Aesthetics, Creativity, and the Arts*, 13(3), 244–258.
- Schleicher, R., & Antons, J.-N. (2014). Evoking Emotions and Evaluating Emotional Impact. In *Quality of Experience* (pp. 121–132). <https://doi.org/10.1007/978-3-319-02681-7>
- Skiendziel, T., Rosch, A. G., & Schultheiss, O. C. (2019). Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and Facial Action Coding System Scoring. *PLoS ONE*, 14(10), 1–18.
- Tchanou, Armel Quentin; Giroux, Félix; Léger, Pierre-Majorique; Senecal, Sylvain; and Ménard, J.-F. (2018). Impact of Information Technology Multitasking on Hedonic Experience Impact of Information Technology Multitasking on Hedonic. *SIGHCI 2018 Proceedings*. 7.
- Timmerer, C., Waltl, M., Rainer, B., & Murray, N. (2014). Sensory Experience: Quality of Experience Beyond Audio-Visual. In *Quality of Experience* (pp. 351–365). <https://doi.org/10.1007/978-3-319-02681-7>
- Varela, M., Skorin-Kapov, L., & Ebrahimi, T. (2014). Quality of Service Versus Quality of Experience. In *Quality of Experience* (pp. 85–96). <https://doi.org/10.1007/978-3-319-02681-7>
- Walla, P. (2018). Affective Processing Guides Behavior and Emotions Communicate Feelings : Towards a Guideline for the NeuroIS Community. In *Information Systems and Neuroscience: Lecture Notes in Information Systems and Organisation* (pp. 141–150). <https://doi.org/10.1007/978-3-319-67431-5>

Walma van der Molen, J. H., & van der Voort, T. H. A. (2000). Children's and Adults' Recall of Television and Print News in Children's and Adult News Formats. *Communication Research*, 27(2), 132–160.  
<https://doi.org/10.1177/07399863870092005>

Walma van der Molen, J. H., & Klijn, M. E. (2004). Recall of Television Versus Print News: Retesting the Semantic Overlap Hypothesis. *Journal of Broadcasting & Electronic Media*, 48(1), 89–107. <https://doi.org/10.1207/s15506878jobem4801>

Waltl, M., Timmerer, C., & Hellwagner, H. (2010). Improving the quality of multimedia experience through sensory effects. In 2010 2nd International Workshop on Quality of Multimedia Experience, QoMEX 2010 - Proceedings (pp. 124–129).  
<https://doi.org/10.1109/QOMEX.2010.5517704>

Wang, Z., Irwin, M., Cooper, C., & Srivastava, J. (2015). Multidimensions of media multitasking and adaptive media selection. *Human Communication Research*, 41(1), 102–127. <https://doi.org/10.1111/hcre.12042>

Wechsung, I., & De Moor, K. (2014). Quality of Experience Versus User Experience. In *Quality of Experience* (pp. 35–54). <https://doi.org/10.1007/978-3-319-02681-7>

Weth, K., Raab, M. H., & Carbon, C. C. (2015). Investigating emotional responses to self-selected sad music via self-report and automated facial analysis. *Musicae Scientiae*, 19(4), 412–432. <https://doi.org/10.1177/1029864915606796>

Witvliet, C. V. O., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition and Emotion*, 21(1), 3–25. <https://doi.org/10.1080/02699930601000672>

Yuan, Z., Chen, S., Ghinea, G., & Muntean, G. M. (2014). User quality of experience of mulsemmedia applications. *ACM Transactions on Multimedia Computing, Communications and Applications*, 11(1). <https://doi.org/10.1145/2661329>

## Chapitre 3: Conclusion

Le projet de recherche lié à ce mémoire a été mené dans le cadre d'une collaboration avec l'Opéra de Montréal et D-BOX Technologies. L'objectif de ce projet d'envergure était de comparer l'expérience d'écoute d'opéra en salle à une expérience d'opéra cinématique simulée en laboratoire. Ce projet ayant attiré l'attention de plusieurs médias (c.-à-d., Le devoir<sup>3</sup>, Radio-Canada<sup>4</sup>, La Presse<sup>5</sup>), a aussi permis à l'auteur de ce mémoire de donner une entrevue sur le sujet à l'émission radiophonique Les Années Lumières. L'auteur du mémoire a aussi activement participé à la collecte de données ambitieuse qui a eu lieu à la Salle Wilfrid-Pelletier de la Place des Arts à Montréal lors de six représentations de l'Opéra Carmen. Les données utilisées dans ce mémoire se rapportent toutefois uniquement à la collecte de données en laboratoire.

Plus spécifiquement, la collecte en laboratoire s'est faite en deux volets dans le but d'atteindre des objectifs spécifiques. La première version utilisait une séquence d'opéra en continu dans le but de pouvoir la comparer avec une expérience d'opéra en salle. Aussi, cette première version ciblait des participants correspondant au groupe d'âge de la clientèle de l'Opéra de Montréal. La deuxième version de l'expérience en laboratoire, ayant été spécifiquement conçue pour ce mémoire, cible plutôt les milléniaux. L'objectif du partenaire industriel était d'explorer le potentiel de la technologie HFVK à rehausser un type de multimédia hors du commun avec une nouvelle clientèle potentielle. Cette collecte de données s'est déroulée du 5 mars au 18 avril 2019 au Tech3lab, à HEC Montréal. La phase de préparation de la collecte, incluant le développement du design expérimental et du stimulus, la mise en place de la salle et des outils de mesures, ainsi que les pré-tests se sont déroulés durant le mois de février 2019. Cette collecte de données a mené à plusieurs défis méthodologiques liés à la présentation du stimulus, ainsi que la

---

<sup>3</sup> <https://www.ledevoir.com/societe/science/553547/sciences-l-opera-toute-une-experience>

<sup>4</sup><https://ici.radio-canada.ca/premiere/emissions/les-annees-lumiere/segments/reportage/113894/opera-science-marketing-carmen-montreal-tech3lab-hec-dbox-musique-cerveau>

<sup>5</sup> [https://plus.lapresse.ca/screens/95a58daa-a49d-4869-93e4-6cc5314880f5\\_7C\\_0.html](https://plus.lapresse.ca/screens/95a58daa-a49d-4869-93e4-6cc5314880f5_7C_0.html)

mise en place de la salle et des outils de mesure. La section suivante présente les défis rencontrés.

## Défis méthodologiques du projet

Le choix du stimulus audiovisuel ainsi que la sélection de 16 extraits marquants de la performance d'opéra Carmen ont été fait par des experts de l'Opéra de Montréal. Ensuite, les designers d'effets haptiques de D-BOX Technologies ont artistiquement codé chacun de ces extraits en supportant plus spécifiquement la trame musicale instrumentale de l'orchestre de l'opéra. Le stimulus a par la suite été livré à l'équipe du Tech3Lab en format sous-titré. En raison de la place importante qu'occupent les sous-titres dans l'opéra, il apparaissait intéressant d'explorer, à l'aide de mesures neurophysiologiques, l'effet de la présence des sous-titres dans le multimédia. Alors, l'auteur de ce mémoire a utilisé le logiciel de modification d'image vidéo Adobe Premiere Pro (CS6, Adobe Systems Inc, San Jose, California), afin de créer des versions d'extraits sans les sous-titres en recadrant l'image.

De plus, il était intéressant d'explorer l'effet de la stimulation haptique sur la présence des sous-titres. Ainsi, nous avons conçu une expérience 2x2 avec ou sans sous-titres, et avec ou sans stimulation haptique. Afin d'automatiser le processus d'exécution des 16 extraits d'opéra aléatoirement assignés à l'une de nos quatre conditions expérimentales, ainsi que pour minimiser le risque d'erreurs, un script Python a été développé. Ce script permettait essentiellement de lancer des listes de lecture prédéveloppées en appelant les fichiers .mp4 d'extraits vidéo (avec ou sans sous-titres) avec leur fichier de code d'informations haptiques correspondant.

Afin de concevoir une expérience écologiquement valide, nous avons réglé le volume de notre système audio au volume moyen en décibel capté par un membre de notre équipe lors d'une représentation du Metropolitan Opera de New York au cinéma. Nous avons aussi recouvert de panneaux de carton d'œufs la vitre adjacente au participant et à la télévision afin d'offrir une surface acoustique optimale. Aussi, des pré-tests nous ont permis d'ajuster la luminosité de la salle, afin que la qualité des enregistrements vidéo du

visage des participants soit optimale, tel qu'indiqué dans le manuel d'utilisation du logiciel de détection d'expressions faciales utilisé (FaceReader, Noldus).

Le contexte d'IHM étudié est différent des expériences utilisateurs traditionnelles en position de travail à l'ordinateur. En effet, nous avons dû être astucieux quant au positionnement des oculomètres et decaméra qui devaient être le plus près possible du participant, sans toutefois être intrusifs.

En effet, comme il n'était pas naturel d'installer un bureau devant le participant pour y déposer les oculomètres et la caméra, et que le bureau supportant la télévision était trop loin pour permettre des mesures de qualité, l'auteur de ce mémoire a assemblé un bras rotatif. Le bras rotatif positionné au-dessus des participants leurs permettait de s'assoir dans le siège D-BOX et d'ajuster le dossier et support de leurs jambes sans être obstrué (voir Annexe 3)

Finalement, un des défis importants de la collecte et l'analyse subséquente de données était que le stimulus audiovisuel devait être lu à partir du lecteur média du partenaire. Ce dernier n'était pas compatible avec des logiciels de présentation de stimulus tel que E-prime qui nous aurait permis de facilement intégrer les mesures oculométriques au stimulus audiovisuel en superposant les données de fixation sur ce dernier. Plutôt, nous avons dû travailler à partir de données oculométriques brutes, soit avec des coordonnées en X et Y de la localisation des fixations sur un plan en 2D. Les pré-tests nous ont permis de déterminer la faisabilité de distinction entre les deux populations de fixation liées aux zones d'image et de sous-titre du stimulus (voir Figure 3, Chapitre 2)

## Rappel des questions de recherche et résultats

Les données collectées pour ce mémoire ont permis de répondre à ses questions de recherche à travers les deux articles présentés qui le constituent. La première question de recherche, soit *Quel est l'effet de la présence de sous-titres sur la qualité d'une expérience d'opéra cinématique?* a été plus spécifiquement répondu dans le premier article du mémoire.

Dans cet article, les résultats ont montré que le multitâche avec des sous-titres intralinguistiques présentés dans la langue natale de l'audience permettait d'améliorer la compréhension des dialogues. Nous avons aussi trouvé que la présence des sous-titres instrumentaux n'affecte pas l'effort cognitif à la tâche, mesuré à l'aide de pupillométrie et d'électroencéphalographie, ni la fréquence d'émotions faciales associées au traitement de l'image du multimédia. Ainsi, les résultats de cette étude pointent vers un comportement de multitâche efficace qui permet de bénéficier une expérience multimédia hédonique.

Dans cet article, qui n'utilisait qu'une partie des données de l'étude, un phénomène intéressant a été soulevé concernant l'effet de la lecture de sous-titres instrumentaux sur les expressions faciales des participants. En effet, puisque les résultats de valence émotionnelle subjective indiquaient aucune différence entre les conditions avec et sans sous-titres, nous avions supposé que la grande fréquence d'émotion de tristesse détectée par notre logiciel pouvait plutôt être naturellement expliquée par la lecture d'informations textuelles. Plus spécifiquement, nous avions soulevé la possibilité que la tâche de lecture des sous-titres ait causé des mouvements faciaux générés par le muscle Corrugator Supercilii, en raison de la lisibilité plus ou moins bonne du texte à l'écran (Larson et al, 2014). Cependant, puisque l'étude dans ce mémoire n'a pas été conçue pour évaluer l'effet de la lecture d'information textuelle sur les mouvements faciaux, le premier article conclut ce sujet en suggérant d'explorer davantage les causes naturelles de mouvements faciaux en situation de lecture pouvant faussement influencer la détection d'expressions faciales par les systèmes automatiques.

L'analyse complète du jeu de données utilisé dans le second article a permis d'éclairer ce phénomène, et de réfuter la supposition que la lecture des sous-titres ait naturellement activé le muscle Corrugator Supercilii (AU4). Dans le deuxième article, nous avons aussi trouvé que, indépendamment de la stimulation haptique, les émotions faciales de tristesse ont été significativement plus fréquentes lorsque les participants regardaient les sous-titres par rapport à l'image. Cependant, nos résultats ont aussi révélé que les émotions faciales de colère et de peur, qui sont aussi détectées à l'aide du Corrugator Supercilii selon le FACS (Ekman and Friesen, 1978), ont été significativement moins fréquentes lorsque les

participants lisait les sous-titres que lorsqu'ils regardaient l'image. Ainsi, nos résultats nous ont permis déduire que l'augmentation de fréquence de tristesse n'était peut-être pas liée à l'activation du Corrugator Supercilii (AU4) dû à la tâche de lecture, mais plutôt générés par des réactions émotionnelles dû à la compréhension des dialogues de l'opéra. Le deuxième article a aussi permis de répondre à la seconde question de recherche de ce mémoire, soit *Dans quelle mesure la stimulation haptique peut dégrader la qualité de l'expérience résultant du multitâche avec des sous-titres dans un contexte de multimédia?*

Dans le second article, nos résultats ont montré que l'ajout de stimulations haptiques synchronisées avec les éléments auditifs non-verbaux d'un multimédia n'influence pas le comportement naturel de lecture des sous-titres de l'audience. En effet, nos résultats suggèrent que les participants allouent autant d'attention visuelle aux sous-titres avec ou sans stimulation haptique, ce qui indique que cette dernière ne semble pas affecter le multitâche avec les sous-titres.

De plus, nos résultats de mesures subjectives indiquent que les participants ont généralement préféré écouter les extraits d'opéra avec sous-titres plutôt que sans sous-titres, et avec stimulation haptique plutôt que sans stimulation haptique. Cependant, puisque les sous-titres étaient nécessaires à la compréhension des dialogues, ils semblent avoir joué un rôle globalement plus important que la stimulation haptique dans l'évaluation subjective de l'appréciation de l'expérience.

Finalement, nos résultats montrent que la présence des sous-titres a eu un effet positif sur la fréquence d'émotions de tristesse durant l'écoute d'opéra. Aussi, la stimulation haptique semble aussi avoir eu un effet positif sur la fréquence d'émotions faciales de tristesse associées au traitement de l'image du multimédia en absence de sous-titres. Cependant les effets de présence de sous-titres et de stimulation haptique sur la fréquence d'émotions faciales de tristesse des participants ne semblent pas s'être additionnés. Nos résultats suggèrent plutôt que le multitâche avec les sous-titres en plus des stimulations haptiques ont empêché les participants de pleinement réagir aux émotions induites par les deux sources d'informations autant que lorsqu'elles sont traitées individuellement.

Cette suggestion est cohérente avec la littérature sur le multitâche en contexte d'expérience multisensorielle hédonique (Tchanou et al., 2019), indiquant que l'audience a une capacité limitée de traitement d'informations sensorielles dans son environnement. Lorsque cette capacité est atteinte, ou lorsque l'auditoire est sur-stimulée, elle doit diviser son attention entre les sources d'informations ne pouvant pas être entièrement traitées.

## Contributions du mémoire

Le premier article de ce mémoire contribue à la recherche sur le multitâche en explorant un contexte de multimédia traditionnel dans lequel diviser son attention peut potentiellement améliorer une activité. En utilisant des méthodes de détection d'expressions faciales, des mesures de dilatation de la pupille, ainsi des signaux électroencéphalographiques, nous avons montré qu'un auditoire de milléniaux peut diviser son attention entre l'image et les sous-titres de manière efficace sans affecter la charge cognitive ainsi que les réactions émotionnelles liées au traitement de l'image. Ainsi, les résultats de ce mémoire s'ajoutent à la littérature sur l'efficacité du multitâche avec des sous-titres (Perego et al., 2010; Kruger et al., 2013; Kruger et al., 2014).

Ce mémoire contribue aussi à la recherche sur les sous-titres, et plus spécifiquement sur le comportement de lecture automatique des sous-titres (Bisson et al., 2014; d'Ydewalle et al., 1991). En effet, comme des études antérieures en contexte de multimédia ont suggéré, l'œil de l'auditoire est attiré vers les sous-titres même lorsqu'il n'en a pas réellement besoin pour comprendre les dialogues (Bisson et al., 2014; d'Ydewalle et al., 1991). Bien que les sous-titres utilisés dans ce mémoire se sont révélés utiles à la compréhension des participants, les résultats d'entrevue suggèrent un comportement potentiel de lecture automatique des sous-titres. En effet, lorsque des stimulations haptiques étaient livrées par le siège, plusieurs participants ont perçu que leur attention visuelle était davantage portée vers l'action dans l'image plutôt qu'aux sous-titres. Par contre, les résultats objectifs mesurés à l'aide de données de temps de fixation montrent que les participants regardaient autant les sous-titres avec ou sans stimulation haptique. Ainsi, nos résultats indiquent que les participants ont inconsciemment porté leur attention visuelle aux sous-titres, ce qui peut s'expliquer par le comportement de lecture

automatique des sous-titres chez les consommateurs de multimédias (Bisson et al., 2014; d'Ydewalle et al., 1991).

Ce mémoire contribue aussi à la recherche sur les multimédias utilisant des mesures objectives, afin d'évaluer les réactions émotionnelles de l'auditoire pouvant influencer la QoE (Reiter et al., 2014). Les résultats de ce mémoire indiquent que le stimulus audiovisuel a généré une plus grande fréquence d'émotions faciales de tristesse, qui peut s'expliquer par l'aspect dramatique de l'opéra cinématique. Plus spécifiquement, les émotions de tristesse détectées semblent avoir été liée à la compréhension des dialogues par la lecture des sous-titres. Ces résultats sont cohérents avec une étude antérieure ayant démontré à l'aide d'EMG que l'activation de muscles faciaux (c.-à.-d., Corrugator et Zygomaticus) étaient temporellement alignés et cohérents avec la valence émotionnelle subjective d'un film (Golland et al., 2018). Cependant, nos mesures subjectives n'ont pas permis d'observer de telles variations au niveau de la valence émotionnelle perçue. En effet, les mesures subjectives d'émotions sont souvent complexes à évaluer et ne peuvent capter les émotions vécues inconsciemment dans un contexte de multimédia (Kayser, 2017; Scherer, 2004). Ainsi, ce mémoire encourage l'utilisation de détection automatique d'expressions faciales comme mesure objective des réactions émotionnelles de l'auditoire, afin d'évaluer la QoE vécue avec un multimédia. De plus, cette dernière méthode de mesures d'émotions faciales est grandement avantageuse par le fait qu'elle soit non-intrusive et sans contact, et qu'elle requière du matériel de collecte de données très accessible (p. ex., webcam), par rapport à l'EMG (Pantic & Rothkrantz, 2000; Egger et al., 2019). Il serait même envisageable de collecter des mesures d'émotions faciales liées à une expérience de multimédia en dehors de contextes expérimentaux en laboratoire.

Ce mémoire apporte une contribution méthodologique importante pour la recherche sur les multimédias et plus largement la recherche en IHM utilisant les systèmes de détection automatique d'expressions faciales. En raison de leurs avantages, tels que leur accessibilité, leur rapidité d'analyse, et le fait qu'ils soient non-intrusifs en contexte expérimental, les systèmes de détection automatique d'expressions faciales sont de plus en plus populaires auprès des chercheurs dans le domaine des systèmes d'information (Fisher, et al., 2019). Cependant ces systèmes, tel que FaceReader (Noldus) ont

récemment été critiqués pour leur capacité à détecter des émotions à partir d'expressions faciales spontanées en contexte naturel (Dupré et al., 2020). En effet, ces systèmes ont généralement été entraînés et validés à partir d'images statiques d'expressions faciales humaines reproduisant les six émotions discrètes spécifiques (Skiedziel et al., 2019). Dupré et al. (2020) ont donc suggéré que l'échantillon restreint des six émotions discrètes ne pouvait pas expliquer l'ensemble des mouvements faciaux, et que cet échantillon devrait être élargi avec des émotions non-discrètes ou des états mentaux. Cependant, ces expressions faciales non liées aux six émotions discrètes ont eu peu d'attention des chercheurs en IHM. Ainsi, la recherche continue grandement d'interpréter les expressions faciales d'utilisateurs à travers l'influente *Basic Emotion Theory* ainsi que les six émotions discrètes (Ekman, 1999, Ekman and Friesen, 1978). Conséquemment, Dupré et al. (2020) suggèrent d'utiliser des informations contextuelles afin de mieux pouvoir interpréter les mouvements faciaux détectés par le système.

Ce mémoire démontre comment les mesures oculométriques d'attention visuelle peuvent apporter une information contextuelle et temporelle importante dans la détection d'expressions faciales, telle que le type d'information visuellement traité par un utilisateur. En effet, la littérature indique qu'une tâche de lecture peut avoir un effet sur les mouvements faciaux, et particulièrement le Corrugator Supercilli étant responsable du froncement de sourcil, en raison de la lisibilité (Larson et al., 2014). Il est donc important de distinguer les expressions faciales selon la tâche visuelle, dans l'optique où un froncement de sourcil, codé comme étant « AU4 » ou Action Unit 4 selon le FACS (Ekman & Friesen, 1978) est aussi utilisé pour prédire les émotions négatives telles que la colère, la tristesse, et la peur. En triangulant la localisation de fixation dans un stimulus avec les émotions faciales des utilisateurs, il est possible de détecter un effet potentiel dû à la tâche de traitement d'information (e.g. lecture de texte par rapport au visionnement d'une image) qui pourrait introduire un artefact dans les données d'émotions faciales. Par exemple, on pourrait ainsi identifier, en comparant les émotions faciales détectées dans deux sites web différents, si le temps passé à lire du texte par rapport à naviguer dans l'interface, ou si la complexité de la lecture peut avoir eu un effet sur les résultats de détection automatique d'expressions faciales.

## **Implications pratiques**

Ce mémoire apporte aussi des implications pour les créateurs de contenu *multimedia*, et plus précisément dans le domaine des multimédias haptiques. Les créateurs de contenu devraient tenir compte de la possibilité d'un comportement de lecture de sous-titres non-nécessaires, considérant que le multitâche avec ces derniers peut empêcher l'auditoire de profiter pleinement des émotions induites par un stimulus multisensoriel. En effet, puisque les sous-titres présentés dans nos écrans ne sont pas tous nécessaires à la compréhension d'un multimédia, tel est souvent le cas pour les descriptions sonores (ex. "Applaudissement", "Toussotement", "Sirènes"), les sous-titres redondants qui se répètent dans un court intervalle, ou les sous-titres simples, il serait intéressant d'évaluer la pertinence de tous les afficher. En effet, pour illustrer ce propos, il serait dommage que le visionnement d'une séquence de film, pour laquelle des efforts importants ont été mis dans la création et la synchronisation de stimulations haptiques, soit visuellement distrait par l'apparition d'un sous-titre inutile à la compréhension de l'auditoire. Bien entendu, cette option ne serait pertinente que pour un auditoire n'ayant aucun problème de santé auditive.

## **Limites et recherches futures**

Ce mémoire comporte certaines limites en raison du choix du stimulus audiovisuel et des séquences étudiées qui a été fait par un partenaire externe. En effet, les extraits ont été sélectionnés en fonction de leur pertinence dans la trame narrative de l'opéra Carmen.

L'équipe de recherche n'avait pas non plus de contrôle sur le développement des stimulations haptiques, qui ont aussi été développées par un partenaire externe. De plus, le procédé de création artistique des mouvements et vibrations est très subjectif et surtout confidentiel. Ainsi, le niveau de détails du stimulus haptique de l'étude, ainsi que la généralisation de ses implications pour les systèmes haptiques restent limités.

Les stimulations haptiques étudiées dans ce mémoire étaient synchronisées avec le stimulus auditif non verbal d'un multimédia. Il serait intéressant d'évaluer l'effet de

stimulations haptiques supportant les objets visibles à l'écran, sur l'attention visuelle portée aux sous-titres.

Aussi, nous n'avons pas évalué le niveau de compréhension lié à l'attention visuelle portée aux sous-titres. En effet, nous ne pouvons pas savoir si l'effet des stimulations haptiques a influencé la compréhension de la trame narrative de l'opéra. Des recherches futures devraient explorer l'effet de stimulations haptiques sur la performance de lecture de sous-titres mesurée à l'aide de questionnaires de compréhension

Pour conclure, ce mémoire contribue à la recherche sur les systèmes de divertissement du futur en mettant l'emphase sur la capacité de l'auditoire à percevoir et traiter la totalité des informations sensorielles qui lui sont communiquées. Le développement rapide des technologies d'information et communications ainsi que des *mulsemedia* mènent à de nouvelles formes d'activité multi-tâche allant au-delà d'une lecture de texte en écoutant de la musique, et d'un visionnement de programme télévisé sous-titré. Ainsi, il est important de comprendre comment l'utilisateur interagit avec ces systèmes pour évaluer et maximiser la qualité d'expérience.

## Bibliographie

- Akhtar, Z., Siddique, K., Rattani, A., Lutfi, S. L., & Falk, T. H. (2019). Why is multimedia quality of experience assessment a challenging problem? *IEEE Access*, 7, 117897–117915. <https://doi.org/10.1109/ACCESS.2019.2936470>
- Arndt, S., Brunnstrdnr, K., Cheng, E., Engelke, U., Moller, S., & Antons, J. N. (2016). Review on using physiology in quality of experience. *Human Vision and Electronic Imaging 2016, HVEI 2016*, 197–205. <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-125>
- Baddeley, A. D., Scott, D., Drynan, R., & Smith, J. C. (1969). Short-term memory and the limited capacity hypothesis. *British Journal of Psychology*. <https://doi.org/10.1111/j.2044-8295.1969.tb01175.x>
- Bardhi, F., Rohm, A., & Sultan, F. (2010). Tuning in and tuning out: media multitasking among young consumers. *Journal of Consumer Behaviour*, 9(July–August), 316–332. <https://doi.org/10.1002/cb>
- Bisson, M.-J., Van Heuven, W., Conklin, K., & Tunney, R. J. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, 35(2), 399–418. <https://doi.org/10.1017/S0142716412000434>
- Brunnström, K., Beker, S. A., De, K., Dooms, A., Egger, S., Garcia, M.-N., ... Larabi, M.-C. (2013). Qualinet White Paper on Definitions of Quality of Experience. European Network on Quality of Experience in in Multimedia Systems and Services (COST Action IC 1003), (March), 26.
- Kaiser, M. (2014). The Prediction of an Emotional State through Physiological Measurements and its Influence on Performance. Thesis. Utrecht University.
- Chaudhari, R., Altinsoy, E., & Steinbach, E. (2014). Haptics. In *Quality of Experience* (pp. 261–276). <https://doi.org/10.1007/978-3-319-02681-7>

D'Ydewalle, G., Praet, C., Verfaillie, K., & van Rensbergen, J. (1991). Watching subtitled television: Automatic reading behavior. *Communication Research*, 18, 650–665.

D'Ydewalle, G., & Bruycker, W. De. (2007). Eye Movements of Children and Adults While Reading Television Subtitles. *European Psychologist*, 12(3), 196–205.  
<https://doi.org/10.1027/1016-9040.12.3.196>

Desblache, L. (2008). Music to my ears, but words to my eyes? Text, opera and their audiences. *Linguistica Antverpiensia*, (6), 155–170.

Duarte, C., & Fonseca, M. J. (2019). Multimedia Accessibility. In *Web Accessibility* (pp. 461–475). [https://doi.org/10.1007/978-1-4471-7440-0\\_25](https://doi.org/10.1007/978-1-4471-7440-0_25)

Dupré, D., Krumhuber, E. G., Küster, D., & McKeown, G. J. (2020). A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *PLoS ONE*, 15(4), 1–17. <https://doi.org/10.1371/journal.pone.0231968>

Egger, M., Ley, M., & Hanke, S. (2019). Emotion Recognition from Physiological Signal Analysis: A Review. *Electronic Notes in Theoretical Computer Science*, 343, 35–55. <https://doi.org/10.1016/j.entcs.2019.04.009>

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion.

Ekman, Paul. (1999). Basic Emotions. In *Handbook of Cognition and Emotion*.  
<https://doi.org/10.1002/0470013494.ch3>

Ekman, Paul, & Friesen, W. V. (1978). *The Facial Action Coding System*. San Francisco, CA: Consulting Psychologists Press. <https://doi.org/10.1037/h0030377>

Ekman, Paul, Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System - Investigator's Guide*. FACS. <https://doi.org/10.1016/j.msea.2004.04.064>

El Saddik, A., Orozco, M., Eid, M., & Cha, J. (2011). Haptics Technologies: Bringing Touch to Multimedia. *Haptics Technologies*. <https://doi.org/10.1007/978-3-642->

- Engelke, U., Darcy, D. P., Mulliken, G. H., Bosse, S., Martini, M. G., Arndt, S., ... Brunnström, K. (2017). Psychophysiology-Based QoE Assessment: A Survey. *IEEE Journal on Selected Topics in Signal Processing*, 11(1), 6–21.  
<https://doi.org/10.1109/JSTSP.2016.2609843>
- Engelke, U., Kaprykowsky, H., Zepernick, H., & Ndjiki-nya, P. (2011). Visual attention in quality assessment: Theory, advances, and challenges. *IEEE SIGNAL PROCESSING MAGAZINE*, (November), 50–59.
- Gardé, A., Léger, P., Sénécal, S., Fredette, M., Chen, S. L., Labonté-Lemoyne, É., & Ménard, J. (2018). Virtual Reality: Impact of Vibro-Kinetic Technology on Immersion and Psychophysiological State in Passive Seated Vehicular Movement. In *Haptics: Science, Technology, and Applications. EuroHaptics 2018. Lecture Notes in Computer Science* (pp. 264–275). Springer International Publishing.  
<https://doi.org/10.1007/978-3-319-93399-3>
- Gazzaley, A., & Rosen, L. D. (2016). *The distracted mind: ancient brains in a high-tech world*. MIT Press, Cambridge, MA.
- Gernsbacher, M. A. (2015). Video Captions Benefit Everyone. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 195–202.  
<https://doi.org/10.1177/2372732215602130>
- Giroux, F., Boasen, J., Sénécal, S., Fredette, M., Tchanou, A. Q., Ménard, F., ... Léger, P. (2019). Haptic Stimulation with High Fidelity Vibro-Kinetic Technology Psychophysiological Enhances Seated Active Music Listening Experience. In *IEEE World Haptics Conference (WHC) Tokyo, Japan*.
- Hsia, H. J. (1971). The information processing capacity of modality and channel performance. *Educational Communication and Technology Journal*, 19(1), 51–75.  
<https://doi.org/10.1007/BF02768431>
- Hwang, Y., & Jeong, S.-H. (2018). Multitasking and Information Gain : Effects of

Relevance between Tasks. *Journal of Media Economics & Culture*, 16(2), 50–77.  
<https://doi.org/10.21328/jmec.2018.5.16.2.50>

Hwang, Y., Kim, H., & Jeong, S. H. (2014). Why do media users multitask?: Motives for general, medium-specific, and content-specific types of multitasking. *Computers in Human Behavior*, 36, 542–548.  
<https://doi.org/10.1016/j.chb.2014.04.040>

Jeong, S. H., & Fishbein, M. (2007). Predictors of multitasking with media: Media factors and audience factors. *Media Psychology*, 10(3), 364–384.  
<https://doi.org/10.1080/15213260701532948>

Jeong, S. H., & Hwang, Y. (2016). Media Multitasking Effects on Cognitive vs. Attitudinal Outcomes: A Meta-Analysis. *Human Communication Research*, 42(4), 599–618. <https://doi.org/10.1111/hcre.12089>

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kayser, D. (2017). Using facial expressions of emotion as a means for studying music-induced emotions. *Psychomusicology: Music, Mind, and Brain*, 27(3), 219–222.  
<https://doi.org/10.1037/pmu0000187>

Kruger, J. L., Hefer, E., & Matthew, G. (2014). Attention distribution and cognitive load in a subtitled academic lecture: L1 vs. L2. *Journal of Eye Movement Research*, 7(5). <https://doi.org/10.16910/jemr.7.5.4>

Kruger, J., & Matthew, G. (2013). Measuring the Impact of Subtitles on Cognitive Load: Eye Tracking and Dynamic Audiovisual Texts. *Proceedings of Eye Tracking South Africa*, (August), 29–31.

Kruger, Jan-Louis & Szarkowska, Agnieszka & Krejtz, Izabela. (2015). Subtitles on the Moving Image: An Overview of Eye Tracking Studies. *Refractory: A journal of Entertainment Media*. 25.

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal*

of Communication, 50(1), 46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, 40(5), 776–785. <https://doi.org/10.1111/1469-8986.00078>

Larson, K., Hazlett, R. L., Chaparro, B. S., & Picard, R. W. (2007). Measuring the Aesthetics of Reading. *People and Computers XX — Engage*, 41–56. [https://doi.org/10.1007/978-1-84628-664-3\\_4](https://doi.org/10.1007/978-1-84628-664-3_4)

Lavaur, J. M., & Bairstow, D. (2011). Languages on the screen: Is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology*, 46(6), 455–462. <https://doi.org/10.1080/00207594.2011.565343>

Lee, M., Roskos, B., & Ewoldsen, D. R. (2013). The Impact of Subtitles on Comprehension of Narrative Film. *Media Psychology*, 16(4), 412–440. <https://doi.org/10.1080/15213269.2013.826119>

Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behavior And Information Technology*, 25(2), 141–158. <https://doi.org/doi.org/10.1080/01449290500331156>

Moon, S. E., & Lee, J. S. (2017). Implicit Analysis of Perceptual Multimedia Experience Based on Physiological Response: A Review. *IEEE Transactions on Multimedia*, 19(2), 340–353. <https://doi.org/10.1109/TMM.2016.2614880>

Nielsen. (2018). The Nielsen Total Audience Report | Q2 2018. Retrieved from <https://www.nielsen.com/us/en/insights/report/2018/q2-2018-total-audience-report/#>

Nothdurft, H. C. (1993). Saliency effects across dimensions in visual search. *Vision*

Research, 33(5–6), 839–844. [https://doi.org/10.1016/0042-6989\(93\)90202-8](https://doi.org/10.1016/0042-6989(93)90202-8)

Orero, P., & Matamala, A. (2007). Accessible Opera : Overcoming Linguistic and Sensorial Barriers. Perspectives: Studies in Translatology, 15(4), 262–277. <https://doi.org/10.1080/13670050802326766>

Pantic, M., & Rothkrantz, L. Ü. M. (2000). Automatic analysis of facial expressions: The state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1424–1445. <https://doi.org/10.1109/34.895976>

Pauna, H., Léger, P. M., Sénécal, S., Fredette, M., Courtemanche, F., Chen, S. L., ... Ménard, J. F. (2017). The psychophysiological effect of a vibro-kinetic movie experience: The case of the D-BOX movie seat. In Information Systems and Neuroscience: Lecture Notes in Information Systems and Organisation (Vol. 25, pp. 1–7). [https://doi.org/10.1007/978-3-319-67431-5\\_1](https://doi.org/10.1007/978-3-319-67431-5_1)

Perego, E., Missier, F. Del, Porta, M., & Mosconi, M. (2010). The Cognitive Effectiveness of Subtitle Processing. Media Psychology, 13(3), 243–272. <https://doi.org/10.1080/15213269.2010.502873>

Reiter, U., Brunnström, K., De Moor, K., Larabi, Mohamed-Chaker Pereira, M., Pinheiro, A., You, J., & Zgank, A. (2014). Factors Influencing Quality of Experience. In Quality of Experience (pp. 55–72). <https://doi.org/10.1007/978-3-319-02681-7>

Reiter, U., & Weitzel, M. (2007). Influence of interaction on perceived quality in audiovisual applications: evaluation of cross-modal influence. Proc. 13th International Conference on Auditory Display, 380–385. Retrieved from <http://dev.icad.org/Proceedings/2007/ReiterWeitzel2007.pdf>

Riedl, René, Léger, P.-M. (2016). Fundamentals of NeuroIS: Information Systems and the Brain. Studies in Neuroscience, Psychology and Behavioral Economics (Vol. 1). Studies in Neuroscience, Psychology and Behavioral Economics. <https://doi.org/10.1136/bmj.1.3968.171-a>

- Scherer, K. R., & Coutinho, E. (2019). Assessing Emotional Experiences of Opera Spectators in Situ. *Psychology of Aesthetics, Creativity, and the Arts*, 13(3), 244–258.
- Schleicher, R., & Antons, J.-N. (2014). Evoking Emotions and Evaluating Emotional Impact. In *Quality of Experience* (pp. 121–132). <https://doi.org/10.1007/978-3-319-02681-7>
- Tchanou, Armel Quentin; Giroux, Félix; Léger, Pierre-Majorique; Senecal, Sylvain; and Ménard, J.-F. (2018). Impact of Information Technology Multitasking on Hedonic Experience Impact of Information Technology Multitasking on Hedonic. *SIGHCI 2018 Proceedings*. 7.
- Timmerer, C., Waltl, M., Rainer, B., & Murray, N. (2014). Sensory Experience: Quality of Experience Beyond Audio-Visual. In *Quality of Experience* (pp. 351–365). <https://doi.org/10.1007/978-3-319-02681-7>
- Varela, M., Skorin-Kapov, L., & Ebrahimi, T. (2014). Quality of Service Versus Quality of Experience. In *Quality of Experience* (pp. 85–96). <https://doi.org/10.1007/978-3-319-02681-7>
- Walla, P. (2018). Affective Processing Guides Behavior and Emotions Communicate Feelings : Towards a Guideline for the NeuroIS Community. In *Information Systems and Neuroscience: Lecture Notes in Information Systems and Organisation* (pp. 141–150). <https://doi.org/10.1007/978-3-319-67431-5>
- Walma van der Molen, J. H., & van der Voort, T. H. A. (2000). Children's and Adults' Recall of Television and Print News in Children's and Adult News Formats. *Communication Research*, 27(2), 132–160. <https://doi.org/10.1177/07399863870092005>
- Walma van der Molen, J. H., & Klijn, M. E. (2004). Recall of Television Versus Print News: Retesting the Semantic Overlap Hypothesis. *Journal of Broadcasting & Electronic Media*, 48(1), 89–107. <https://doi.org/10.1207/s15506878jobem4801>

Wang, Z., Irwin, M., Cooper, C., & Srivastava, J. (2015). Multidimensions of media multitasking and adaptive media selection. *Human Communication Research*, 41(1), 102–127. <https://doi.org/10.1111/hcre.12042>

Wechsung, I., & De Moor, K. (2014). Quality of Experience Versus User Experience. In *Quality of Experience* (pp. 35–54). <https://doi.org/10.1007/978-3-319-02681-7>

## Annexes

### Annexe 1: Méthode de revue de littérature

Notre méthode de recherche de littérature utilisé pour ce mémoire a été plutôt sélective que compréhensive ou représentative. Elle a aussi été conduite par thème regroupant les concepts mobilisés afin de développer les questions de recherche de ce mémoire. Ces thèmes sont présentés avec que leur stratégie et logique de recherche, ainsi qu'un court sommaire des conclusions de recherche importantes.

Une recherche booléenne a été effectuée avec les mots-clés suivants : *Subtitle (and) intra (and/or) inter (and/or) eye tracking (and/or) opera*. Cette recherche avait pour objectif de trouver des évidences empiriques qui allait nous permettre de répondre aux questions suivantes: Que rapportent les études sur l'utilisation des sous-titres en général, et plus spécifiquement les sous-titres intralinguistiques définissant notre stimulus? Quelles méthodes et mesures oculométriques ont utilisés dans les dernières études s'intéressant à la lecture de sous-titres. Cette recherche a révélé que peu importe le type de sous-titres, les consommateurs de multimédias adoptent généralement un comportement de lecture automatique de ces derniers. De plus, sans toutefois avoir fait une recherche représentative, cette dernière a permis d'identifier le manque de littérature sur des types de sous-titres et multimédias inexplorés mais grandement pertinents, soit les sous-titres intralinguistiques présentés dans la langue natale de l'audience dans un contextes d'opéra cinématique.

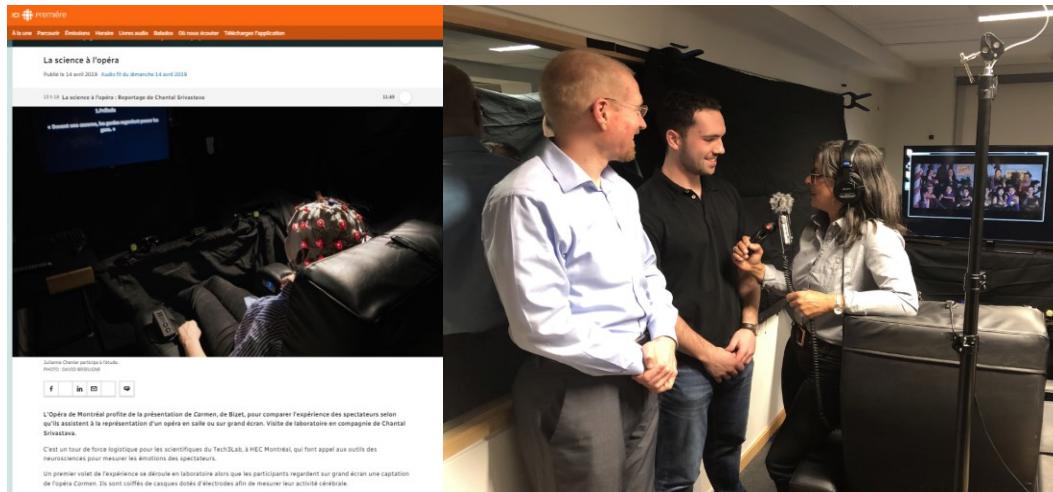
Une seconde recherche booléenne a été réalisée avec les mots clés suivants : *Multisensory stimulation (and) haptic (and/or) auditory (and/or) visual (and) multimedia*. Cette recherche avait pour but de d'identifier des articles pertinents pour répondre à la question suivante: Que rapporte la littérature sur les effets de l'intégration d'informations cohérentes simultanément traités par les modalités sensorielles pertinentes à une expérience de multimédia? Cette recherche a révélé les différents effets que pourrait créer l'intégration multisensorielle impliquant les sens de la vision, l'ouïe, et du touché dans un contexte de multimédia.

Une troisième recherche booléenne a utilisé les termes suivants : *Quality (and) User (and) experience (and) multimedia (and) measures (and) subjective (and/or) objective*. Les articles pertinents trouvés lors de cette recherche avaient pour objectif de répondre à la question suivante: Comment l'expérience utilisateur a été mesurée en contexte de multimédias? Cette recherche sélective a révélé que les émotions occupent un rôle important dans la qualité de l'expérience telle qu'elle est définie par les chercheurs du domaine (Reiter et al., 2014). Aussi, bien que cette recherche ne soit pas représentative, elle a permis d'identifier la sous-utilisation de méthodes de détections automatique d'expressions faciales afin mesurer les émotions induites par les multimédias.

Ces concepts ont été cherché manuellement à l'aide de Google Scholar afin d'identifier des articles clés sur le sujet. Ensuite, nous avons effectué une recherche dans les bibliographies ainsi que les citations de ces articles clés afin d'identifier d'autres articles potentiellement intéressants. Le processus de sélection des articles s'est fait selon la pertinence de ces derniers après une lecture du sommaire. Aucun critère d'exclusion précis a été utilisé, mais nous avons toutefois priorisé les publications de revues scientifiques plutôt que les actes de conférence. Les articles ont été accédés soit directement par Google Scholar ou en utilisant la base de données HEC Bibliothèque, ainsi que gérée et classifiée à l'aide du gestionnaire de référence Mendeley.

Notre échantillon d'évidences empiriques provient de différentes disciplines scientifiques, telles que les multimédias et télécommunications, la psychologie, ou les technologies de l'information. Nous avons aussi consulté des livres méthodologiques tel que le livre de Fundamentals of NeuroIS (Riedl et Léger, 2016) afin de nous informer sur les différentes méthodes et mesures d'outils neurophysiologiques.

## Annexe 2: Articles de journaux liés au projet de recherche



Article du journal Radio-Canada (14 avril 2019) <https://ici.radio-canada.ca/premiere/emissions/les-annees-lumiere/segments/reportage/113894/opera-science-marketing-carmen-montreal-tech3lab-hec-dbox-musique-cerveau>



Article du journal le Devoir (3 mai 2019)  
<https://www.ledevoir.com/societe/science/553547/sciences-l-opera-toute-une-experience>



## MARKETING CULTUREL **CARMEN EN LABORATOIRE**

Les airs de *Carmen* vous procurent-ils des frissons ou vous laissent-ils de marbre ? La belle bohémienne est-elle plus éclatante sur un écran ou sur scène ? Et que diriez-vous d'écouter *L'amour est un oiseau rebelle* dans un fauteuil D-Box ? Des chercheurs montréalais vont bientôt répondre à ces questions. Branchez les électrodes, le rideau monte !

MARIO GIRARD  
LA PRESSE

### Mesurer « le niveau de plaisir »

Au cours des prochaines semaines, l'Opéra de Montréal va mener une expérience unique en son genre. Lors des représentations de *Carmen* (du 4 au 13 mai), une trentaine de spectateurs regarderont le spectacle coiffés d'un casque équipé d'électrodes. Ce laboratoire, élaboré et mené par une équipe du Tech3lab de HEC Montréal, suscite une telle curiosité qu'un chercheur américain, Jared Boasen, spécialiste des neurosciences, est venu s'installer à Montréal afin de suivre chacune des étapes. Ce projet est rendu possible grâce à une bourse de 200 000 \$ offerte par la Fondation Getty. « En clair, nous voulons en savoir plus sur l'expérience que vit un spectateur en salle, explique Xavier Roy, directeur marketing de l'Opéra de Montréal. Nous allons analyser le niveau de plaisir des spectateurs. Comme plusieurs maisons d'opéra, nous voyons nos spectateurs nous défaillir au profit des projections proposées par le Met dans les salles de cinéma. Nous allons analyser et comparer plein d'autres choses. C'est un projet ambitieux et innovateur. »

Article du journal La Presse (28 janvier 2019) [https://plus.lapresse.ca/screens/95a58daa-a49d-4869-93e4-6cc5314880f5\\_7C\\_0.html](https://plus.lapresse.ca/screens/95a58daa-a49d-4869-93e4-6cc5314880f5_7C_0.html)

### Annexe 3: Support d'oculomètres et caméra





