



International Journal of African Sciences
— IJAS —

DOI

10.58610/IJAS.2711

ISBN

978-2-38489-043-9

EAN

9782384890439

INTERNATIONAL JOURNAL OF AFRICAN SCIENCES

REVUE INTERNATIONALE DES SCIENCES AFRICAINES

Vol. 01 No. 02, Nov. 2023



EDITIONS
LUMUMBA

A New Momentum for African Publishing



International Journal of African Sciences

— IJAS —

Revue Internationale Des Sciences Africaines

Published by Editions Lumumba



A New Momentum for African Publishing



International Journal of African Sciences
— IJAS —

« Influence de la dégradation du contenu audiovisuel sur le spectateur mineur »

[Influence of the degradation of audiovisual content on the minor viewer]

LOKOSSOU Bonaventure
Université d'Abomey-Calavi/Bénin

International Journal of African Sciences, Vol. 01 No. 2, Nov. 2023, Pages 131 – 145

DOI : <https://doi.org/10.58610/IJAS.2711>



EDITIONS LUMUMBA

A New Momentum for African Publishing

International Journal of African Sciences – IJAS –

DOI : 10.58610/IJAS.2711

ISBN : 978-2-38489-043-9

EAN : 9782384890439

© Editions Lumumba, November 2023

Legal deposit: No. 7.20.2022.75, 2nd quarter

National Library of Congo - Ordinance No. 89-010 of January 11, 1989 -

Establishment permit No. BNC/DPHK/08/2022



Published by Editions Lumumba with the support of Afriscience, the International Journal of African Sciences (IJAS) is a bilingual interdisciplinary scientific journal (French and English) that aims to promote and disseminate African studies by providing a platform for exchange and dialogue among researchers, academics, and professionals from both Africa and around the world. The Lumumba Editions are registered with the International ISBN Agency through the Francophone Agency for International Book Numbering, under the publisher identifier: 978-2-38489. They have an international editorial committee composed of teachers, researchers, and specialists from Africa and elsewhere, ensuring a diversity of expertise and perspectives.

Registered in accordance with the Ordinance Law No. 89-010 of January 11, 1989, of the Congolese State (DR Congo), the Lumumba Editions operate under the establishment permit No. BNC/DPHK/08/2022. As a member of Crossref, one of the organizations based in the United States that participates in the global indexing of scientific content, the Lumumba Editions benefit from a DOI prefix (10.58610). This prefix allows assigning a DOI to each publication, whether it be books, articles, journals, or conference proceedings, thus providing a persistent link to the online location of the edited work. In addition to their network of distributors and international partners, thanks to the open DOI APIs, the publications of the Lumumba Editions are accessible to thousands of other Crossref members and hundreds of organizations worldwide, significantly increasing the visibility and international impact of the edited works.

Content

1- Formation des formateurs en TICE : Axes d'intervention socio-éducative pour les enseignants du District de Fianarantsoa et limites	10
RAFANOMEZANTSOA Jean Marie Vianney	10
<i>Université de Fianarantsoa/Madagascar</i>	10
2- Promotion de la culture de la paix dans les écoles secondaires : Un impératif stratégique pour l'avenir de la République Démocratique du Congo	33
Benjamin Kilambe	33
<i>Université de Lubumbashi/RD Congo</i>	33
3- Avis des parents et des enseignants sur la réforme du programme éducatif du Domaine d'Apprentissage des Sciences (DAS) dans la ville de Lubumbashi	54
Malamba Lubelo Jimmy	54
<i>Institut Supérieur de Statistiques/ RD Congo</i>	54
4- L'autorégulation des médias congolais à l'épreuve de l'ordre professionnel des journalistes	69
KITIKI KIWAKA Guy Blaise	69
<i>Université de Kinshasa /RD Congo</i>	69
5- Etat de lieu des installations électriques basse tension face aux éventuelles décharges dues aux coups de foudre	79
ILUNGA LWAMBA Trésor	79
NYEMBO WA SANGWA Ken	79
MULEMBA WA KABULO Degaulle	79
KISALE KYALWE Clodio	79
NGOIE-MWANA-MFUMU Laurent Fabrice	79
<i>Enseignants-chercheurs/RD Congo</i>	79
6- Le discours, véritable mécanisme de persuasion des églises de réveil au Cameroun	89
Ngaouri Landri	89
Bitye Cynthia Laure	89
<i>Université de Ngaoundéré/Cameroun</i>	89
7- Esquisse des pratiques liées à la délocalisation des communautés locales par les entreprises minières à Kolwezi : Une analyse criminologique	101
KABWE RWANFIZI Christian	101
UMBA KONGOLO Nathan	101
MPOO MOKUBA Alliance David	101
<i>Enseignants-chercheurs/RD Congo</i>	101
8- Cannabis et anxiété des exploitants miniers artisanaux à Kolwezi	112
José NGOY WA NGOY TWITE	112
<i>Université de Kolwezi/RD Congo</i>	112
9- Influence de la dégradation du contenu audiovisuel sur le spectateur mineur	131
LOKOSSOU Bonaventure	131
<i>Université d'Abomey-Calavi/Bénin</i>	131

10- Evaluation de la rentabilité de l'investissement en transport : cas de moto taxi, tronçon Lukalaba-Mbujimayi	146
Paul NKOMBUA MUTSHIMUNE	146
<i>Institut Supérieur Pédagogique de Lukalaba/ RD Congo</i>	146
Estelle LENGIE KASUMBA	146
<i>Université de KABINDA/RD Congo</i>	146
Timothée MULUMBA NTUMBA	146
<i>Institut Supérieur Pédagogique de Lukalaba/RD Congo</i>	146
11- Elaboration d'une politique optimale de production dans une entreprise industrielle. Cas de Chemical of Africa(Chemaf)	158
Mwamba Kongolo Coalice	158
<i>Université de Lubumbashi/RD Congo</i>	158
12- La main d'œuvre féminine face à la modification des prestations salariales dues au congé de la maternité en RDC. Incidences et voie de sortie honorable	172
MALUMALU MARTIN GAEL	172
<i>Université de Lubumbashi/RD Congo</i>	172
13- Le masque des groupes extrémistes violents dans les communes de Materi et de Coby au Bénin : vers une sociologie des fragilités sociopolitiques	181
Thelesphore Toliton DIKPO	181
Abou-Bakari IMOROU	181
Herbert SINA BIO	181
<i>Université d'Abomey Calavi/ Bénin</i>	181

Journal Team

Editorial Board :

- Editor-in-Chief : Kaniki Wa Cilombo Joseph-Robert, DR Congo
- Managing Editor : Valery Ngoy Ndala, DR Congo
- Lead Reviewer : Jean-Micky Kafuwa Musongo, DR Congo
- Editorial Secretary : William Agbaka, Benin
- Coordinator for Membership Process : Marie Mbatshiona Mundi, DR Congo
- Administrative and Financial Manager : Banon Zinsou Côme, Benin
- Partnerships Coordinator : Fidèle Bukasa Kabwe, DR Congo
- Statistics Manager : Kalunga Shakoli Dieu Merci, DR Congo
- Chief Translator : Dr. Fo-Koku D. Woameno, Togo
- Principal Proofreader : Jean-Micky Kafuwa Musongo, DR Congo
- Online Publishing Manager : Mrs. Blin Acouh Marie-José, Ivory Coast

Editorial Committee :

- Abdoulaye Ouedraogo, PhD, Ministry of Economy, Finance and Planning (MEFP), Burkina Faso
- Abi-Kaberou Gildas, PhD, UAC/ENS-Porto-Novo/Laboratory of Pedagogy and Didactics of Humanities (LaPEDIH), Benin
- Adaba Koffi Amessou, PhD, University of Lomé, Togo
- Akimou Tchagnaou, PhD, University of Zinder, Niger
- Ambombi Eyolo Azede Espoir, PhD, Chaire MBA, Afrikan campus, Republic of Congo
- Amoussou Franck, PhD, University of Abomey-Calavi, Benin
- Ballo Drissa, PhD, Teacher-Researcher, Mali
- Basile Mulwani Makelele, PhD, University of Lubumbashi, Democratic Republic of Congo
- Bekolo Engoudou Bruno, PhD, University of Douala, Cameroon
- Bitouga Bernard Aristide, PhD, University of Douala, Cameroon
- BOUMA Carine Nadège, PhD, University of Bamenda, Cameroon

- Christian Bumute, PhD, University of Likasi, DR Congo
- Essome Lele Gislain Arnaud, PhD, ECLLA Research Unit, Jean-Monnet University Saint-Étienne, Cameroon
- Félix N'dia Anon, PhD, Félix Houphouet-Boigny University, Côte d'Ivoire
- Fidélie Ntshikala Mbuya, PhD, University of Kolwezi, DR Congo
- Fo-Koku D. Woameno, PhD, University of Lomé, Togo
- Gaby Ilunga Mutombo, PhD, University of Lubumbashi, DR Congo
- Gano Nouhou, PhD, Cheikh Anta Diop University of Dakar, Senegal
- Gninneyo Sylvestre-Pierre NIYA, PhD, Ecole Normale Supérieure, Burkina Faso
- José Ngoy Wa Ngoy Twite, PhD, University of Kolwezi, DR Congo
- Kabiena Kuluila Pierre Valery Dieudonné, PhD, Official University of Mbuji-Mayi, DR Congo
- Kabuya-Kabeya Sthilobo Hilaire, PhD, Official University of Mbuji-Mayi, DR Congo
- Kaniki Wa Cilombo Joseph-Robert, PhD, Official University of Mbuji-Mayi, DR Congo
- Kongue Toyindou, PhD, University of Lomé, Togo
- Lago Blé Angelin, PhD, Jean-Lorougnon Guede University, Daloa, Côte d'Ivoire.
- Lokossou Bonaventure, PhD, University of Abomey-Calavi, Benin
- Matthias Cinyabuguma, PhD, World Bank, DR Congo
- Melingui Ayissi, PhD, University of Douala, Cameroon
- Metsena Ndjavoua, PhD, University of Maroua, Cameroon
- Mohamed Atteyoub H. dit Modibo SIDIBÉ, PhD, University of Social Sciences and Management of Bamako (USSGB), Mali
- Moussa Coulibaly, PhD, Assane Seck University of Ziguinchor, Senegal
- Moussa Dourfaye Abdoul-kadze, PhD, Health Service, Niger
- Nebie Boukary, PhD, University of Fada N'Gourma, Burkina Faso
- Ngala Ntumba Peter, PhD, Official University of Mbuji-Mayi, DR Congo
- Ngonu Ossango Pangrace, PhD, University of Yaoundé, Cameroon
- Nkongolo Mulami Kapweka Alphonse, PhD, Official University of Mbuji-Mayi, DR Congo

- Ntita Ntita Jean Christ, PhD, Official University of Mbuji-Mayi, DR Congo
- Nyebe Atangana Sandrine, PhD, Ministry of Secondary Education, Cameroon
- Patrice M'Bétien KONE, PhD, Félix Houphouët Boigny University, Côte d'Ivoire
- Philippe Kasongo Maloba Tshikala, PhD, University of Lubumbashi, DR Congo
- Raymond-Bernard Ahouandjinou, PhD, University of Abomey Calavi, Benin
- Serge Caleb Mbula Musasa Mwadianvita, PhD, National Pedagogical University (UPN) of Kinshasa, DR Congo
- Sidibé Ousmane, PhD, University of San Pedro, Côte d'Ivoire
- Yamsoumouna Ndimma Joël., PhD, Center for Studies and Research in International and Community Law (CEDIC), Chad

Editorial

Dear readers and researchers,

We are delighted to present to you the latest edition of the International Journal of African Sciences (IJAS) for the year 2023, bearing witness to our ongoing commitment to push the boundaries of knowledge on a global scale. In this edition, we embark on an intellectual exploration journey, transcending borders and disciplines to unveil a collection of innovative research contributions.

This edition stands as a beacon of excellence, highlighting the dedication and ingenuity of researchers from diverse backgrounds. Each article, from cutting-edge advancements to insightful analyses, encapsulates the essence of groundbreaking research. We express our sincere gratitude to the researchers whose rigorous efforts have enriched the academic landscape.

Alongside the evolving academic landscape, our commitment to providing a platform that fosters collaboration, innovation, and the dissemination of impactful research remains unwavering. In this spirit, we invite institutions, academics, and political and social leaders to join us in supporting the International Journal of African Sciences (IJAS). Together, let's forge new paths in the relentless pursuit of knowledge and contribute to the global dialogue.

Our gratitude extends to the diligent members of the review committee, whose expertise ensures the highest standards of academic rigor. Their meticulous evaluation has been essential in maintaining the quality and integrity of the journal.

Special mention is reserved for our dedicated editorial team, whose tireless efforts have transformed ideas into this captivating reality. Their commitment to excellence is a driving force behind the success of IJAS.

As we navigate through the ever-expanding frontiers of research, our goal is to cultivate an inclusive and dynamic space for academic exchange. We aspire to be a catalyst for ideas, a place where innovation converges with tradition, and where the global community comes together to shape the future of knowledge.

To our cherished readers and contributors, thank you for being an integral part of this intellectual odyssey. Your ongoing support fuels our commitment to excellence, and we look forward to exploring new horizons together in the upcoming issues.

Welcome to the intellectual journey that is the International Journal of African Sciences (IJAS).

Dr. Joseph-Robert KANIKI WA CILOMBO

Professor and Vice-Dean in charge of Research at the Faculty of Law of the Université Officielle de Mbuji-Mayi/ Congo DR

« Influence de la dégradation du contenu audiovisuel sur le spectateur mineur »

[Influence of the degradation of audiovisual content on the minor viewer]

LOKOSSOU Bonaventure¹

Université d'Abomey-Calavi/Bénin

Email : bonaventurelokossou2018@gmail.com

Resumé

Dans le contexte compétitif des services audiovisuels (AV), assurer une expérience de qualité (QOE) est crucial. Actuellement, la QOE se mesure principalement par la perception de la qualité audiovisuelle, évaluée par des participants selon des normes internationales. Cependant, cette approche ne capture pas tous les aspects de l'expérience, comme la fatigue induite par la qualité restituée. Cette étude explore une méthode alternative pour évaluer la qualité d'expérience en considérant des indicateurs oculaires et physiologiques. Deux protocoles ont été testés, révélant que la qualité audiovisuelle influence l'expérience physiologique, non entièrement reflétée par les notes de qualité. Bien que l'influence de la qualité audiovisuelle sur les mesures physiologiques et oculaires ne soit pas toujours évidente, certains indicateurs ont réagi aux dégradations cumulées avec d'autres facteurs. Cette approche, moins sujette aux biais subjectifs, offre une perspective plus complète de l'impact de la qualité audiovisuelle sur le spectateur, soulignant l'importance de considérer des facteurs tels que la fatigue et l'effort mental dans l'évaluation de la QOE.

Mots-clefs : Qualité d'expérience, qualité audiovisuelle, mesures subjectives, mesures physiologiques, mesures oculaires.

Abstract

In the competitive context of audiovisual (AV) services, ensuring a quality experience (QOE) is crucial. Currently, QOE is primarily measured through the perception of audiovisual quality, evaluated by participants according to international standards. However, this approach does not capture all aspects of the experience, such as fatigue induced by the rendered quality. This study explores an alternative method to assess the quality of experience by considering ocular and physiological indicators. Two protocols were tested, revealing that audiovisual quality influences physiological experience, not entirely reflected by quality ratings. Although the influence of audiovisual quality on physiological and ocular measures is not always evident, some indicators responded to degradations combined with other factors. This less susceptible-to-bias approach provides a more comprehensive perspective on the impact of audiovisual quality on the viewer, emphasizing the importance of considering factors such as fatigue and mental effort in QOE assessment.

Keywords : Quality of Experience, audiovisual quality, subjective measures, physiological measures, ocular measures.

¹ Docteur, enseignant à l'École Nationale des Sciences Techniques de l'Information et de la communication de l'Université d'Abomey-Calavi (UAC/ENSTIC) en sciences de communication audiovisuel, Spécialité : Techniques de Communication Audiovisuel et Image Appliquée.

0. Introduction

Dans le cadre de cette étude, Hands (2004), cité par le Dr. Julie Lassalle de l'Université de Bretagne, a souligné l'importance de l'influence du contenu sur l'évaluation perceptive de la qualité. Il a également souligné la nécessité de proposer différents types de contenus de test caractérisés, tels que le niveau de mouvements présents dans la vidéo ou la relation entre les médias audio et vidéo (par exemple, les paroles ou les commentaires). Les résultats de l'expérimentation A ont confirmé cette observation, mettant en évidence une influence du contenu tant sur les mesures subjectives que psychophysiologiques. L'objectif des expérimentations présentées dans ce chapitre est de proposer un ensemble de descripteurs permettant de caractériser de manière aussi complète que possible les contenus de test utilisés. L'impact du contenu sur la perception de la qualité et, plus globalement, sur la qualité de l'expérience du spectateur pourra donc être étudié à partir des descripteurs spécifiés. La caractérisation doit permettre, d'une part, de mieux comprendre comment le contenu, décrit par un certain nombre de critères, influence la perception de la qualité et, plus largement, la qualité de l'expérience du spectateur. D'autre part, elle doit faciliter l'interprétation des mesures psychophysiologiques. La norme UIT-T P.912 (UIT, 1999) fournit un certain nombre de critères pour décrire les séquences audiovisuelles de test. L'ensemble de ces critères est présenté dans le tableau 7.1 ci-dessous.

Vidéo 7.1. Catégories proposées par la norme UIT-T P.912 pour décrire les contenus audio et vidéo d'une séquence audiovisuelle.

• Expérimentation A : caractérisation et influence du contenu

La description et la classification proposées considèrent l'audio et la vidéo séparément, sans tenir compte du lien entre son et image. De manière générale, la méthode UIT-T P.912 ne tient pas compte des aspects sémantiques (modalité dominante), techniques (changement de plans ou de scènes, mouvement, *etc.*) ou hédoniques (valence et arousal) du contenu audiovisuel. Pourtant, différentes études ont attiré l'attention sur l'influence de ces facteurs comme celle de la dynamique et de la modalité dominante (Hands, 2004), du niveau d'intérêt (Palhais *et al.*, 2012) ou encore de la présence de mouvements et de changements de plans ou de scènes (Lang A. *et al.*, 2000 ; Simons *et al.*, 1999) cité par docteur Julie LASSALLE sur l'évaluation de la qualité perçue et sur les mesures psychophysiologiques. Comme l'a indiqué Hands (2004), une modalité donnée, audio ou vidéo, peut participer de manière plus importante à la note de qualité audiovisuelle en raison de son apport sémantique dominant. La présence de dégradations sur la modalité dominante serait alors d'autant plus gênante. La perception de qualité repose donc sur différents critères du contenu dont dépendra le jugement du spectateur. Les séquences audiovisuelles doivent être décrites de manière à pouvoir obtenir une interprétation plus précise de la note de qualité attribuée aux signaux audio et/ou vidéo et des influences de la qualité sur la *qualité d'expérience* étudiée à partir de mesures subjectives complémentaires (c.-à-d. autres que la seule note de qualité), physiologiques et oculaires. L'objectif final est de dégager les principaux critères participant à la perception de qualité et plus largement à la *qualité d'expérience* du spectateur. La caractérisation de contenus

de test s'est déroulée selon plusieurs phases. Premièrement, une base exhaustive de descripteurs de contenu a été élaborée avec l'aide d'un expert du domaine de l'audiovisuel (technicien professionnel de l'audiovisuel –société *Digipictoris*, Brest). Dans un second temps, chaque contenu de test a été découpé en unités signifiantes proches d'une analyse plan par plan. Chaque unité de chaque contenu a ensuite été caractérisée à partir du répertoire de descripteurs préalablement élaboré. Une dernière phase a consisté à identifier les descripteurs clés devant constituer le répertoire final. Le corpus de contenu de test a été enrichi de façon à proposer un plus grand nombre de contextes audiovisuels, à savoir les contenus :

- **Danse** : extrait du ballet *Balé de Rua* (14 min 21),
- **Documentaire** : documentaire entier sur Jean-Marc Mormeck (15 min 25),
- **Opéra** : extrait d'une adaptation de *Don Giovanni* (15 min 36),
- **Sport** : extrait de la finale de Roland Garros 2011 (15 min),
- **Théâtre** : extrait d'une adaptation des *Fourberies de Scapin* (10 min 29).

A la suite de cette caractérisation experte, des séquences de quelques secondes ont été extraites de chaque contenu et présentées à un panel de participants. Leur tâche était de caractériser à leur tour les séquences proposées sur la base, entre autres, de descripteurs utilisés par l'expert (expérimentation B). Cette étape devait remplir deux objectifs :

- **Expérimentation B : caractérisation et influence du contenu**

- **vérifier la pertinence d'un ensemble de descripteurs** considéré comme plus « perceptifs », afin d'être en mesure d'utiliser la caractérisation experte pour l'intégralité des contenus,
- **étudier la pertinence de descripteurs supplémentaires** davantage liés à la *qualité d'expérience* spectateur (plaisir ou intérêt par **exemple**).
- Enfin, les interactions entre contenus et qualité perçue ont été étudiées à la lumière de ces descripteurs (expérimentation B2).

1. Méthodologie

1.1. Expérimentation a : caractérisation des contenus

1.1.1. Selection des descripteurs

Un contenu peut être décrit à partir de différentes catégories de descripteurs, par exemple, des descripteurs techniques relatifs au choix de réalisation tels que le nombre de changements de plans, la dynamique de caméra (zooms, travellings, etc.) ou des descripteurs sémantiques tels que la modalité dominante, le niveau de compréhension ou encore la quantité d'information perçue. La caractérisation experte a été réalisée à l'aide de vingt-huit descripteurs pouvant être regroupés au sein de deux grandes catégories : les descripteurs techniques et les descripteurs sémantiques. Un exemple du support utilisé par l'expert pour décrire une séquence donnée est apporté par l'annexe 7-A.

Certaines nomenclatures existent pour décrire des contenus audiovisuels. Notamment la norme MPEG7 (ISO/IEC, 2004) propose une description standard de contenus multimédias dans le cadre d'applications de recherches étendues de documents archivés. Elle fournit notamment un ensemble de descripteurs dit de

bas-niveau d'abstraction tels que le mouvement de caméra (fixe, panoramique - rotation horizontale-, travelling –mouvement transversal horizontal-, zoom, *etc.*), la texture (niveau de détail), la température de couleur ou encore la dynamique, définie comme la notion intuitive de l'intensité ou du rythme de l'action d'une séquence vidéo. Des descripteurs de plan ont également été proposés par Amiar (1995) : durée, angle de vue (plongée, contre-plongée), mouvements de caméra, cadrage (gros plan, plan d'ensemble, *etc.*), profondeur de champ (flou, courte, grande, *etc.*). Le choix des **descripteurs techniques** s'est appuyé sur l'ensemble de ces spécifications.

Au total, treize descripteurs techniques ont été retenus : niveau de détail (faible-moderé-fort), température de couleur (chaude, jour, froide), luminosité (faible-moderée-forte) et caractéristiques de caméra (générale, mobilité, angle de prise de vue -horizontal et vertical-, cadrage, nombre de *cuts*, zoom, rotation caméra, profondeur de champ, angle de vue, pour plus de détails voir annexe 7-A). La sélection des **descripteurs sémantiques** a été réalisée sur la base des descripteurs proposés par la norme MPEG7 ainsi que ceux suggérés par Amiar (1995). Cet auteur propose notamment des paramètres scénaristiques (intérieur/extérieur, jour/nuit, visuel/dialogue, action tension/inaction-immobilité, nombre de personnages, intime/collectif/public), des caractéristiques audio (parole, bruit, musique) ou encore la qualification des relations image/son (son diégétique : son *in* ou hors-champ ; son extra-diégétique : son *off*, ces relations seront mieux définies ci-après).

Expérimentation A : caractérisation et influence du contenu

Par ailleurs, selon Zettl (1991, cité par Simons *et al.*), le mouvement (*motion*) d'un film ou d'un contenu télévisuel peut être décrit à la fois comme le mouvement d'un objet présent à l'image (balle de tennis par exemple), le mouvement des caméras (travelling, zoom, panoramique, inclinaison, *etc.*) et le mouvement de la séquence (changement de plans par utilisation de *cut* ou tout autre moyen de transition). Au total, quinze descripteurs sémantiques ont été retenus : modalité dominante (audio, vidéo, audiovisuelle : sur la base des résultats de l'expérimentation A et des constats de Hands, 2004), présence de mouvements, présence d'informations textuelles, dynamique de contenu (faible-moderée-forte), dynamique caméra (faible-moderée-forte), expression sonore (parole, musique, bruit), type de parole (dialogue-monologue, commentaires, chant), relations image/son (son *in*, *off* ou hors-champ) ainsi que l'ensemble des critères scénaristiques proposés par Amiar : intérieur/extérieur, jour/nuit, clair-sombre, visuel/dialogue, intime/collectif/public, nombre de personnages, action/inaction. Les vingt-huit descripteurs sémantiques et techniques ont été utilisés par l'expert pour caractériser la totalité des contenus du corpus. Pour permettre ce processus, chaque contenu a été segmenté par l'expert en différentes séquences de temps (proche d'une analyse plan par plan). Chacune de ces séquences peut être considérée, comme défini par Goliot-Lété et Vanoye (1993, p. 28, cité par Amiar), comme une unité de sens, c'est-à-dire une suite de scènes qui ne se déroulent pas forcément dans le même décor, mais qui forme un tout avec un sens lui étant propre. La caractérisation experte a permis de faire émerger neuf descripteurs sémantiques et techniques principaux. L'ensemble des descripteurs n'a en effet pas été retenu en raison de la redondance de certaines

informations ou de la granularité parfois trop fine de certains descripteurs (par exemple le type de cadrage, voir annexe 7-A). Les descripteurs techniques qui ont été sélectionnés sont :

- **la luminosité** (faible, modérée, forte),
- **la température de couleur** (chaude -orangée-, jour -lumière blanche-, froide –bleutée sect. 1.4, chap. I),
- **la dynamique caméra** (faible, modérée, forte : regroupe les différents mouvements de caméra - travellings, rotations, zooms, *etc.* – incluant *cuts*/changements de plan),
- **le niveau de détail** (faible, modéré, fort). Les descripteurs sémantiques qui ont été sélectionnés sont:
- **le rapport audiovisuel** ou **diégèse** (son *in*, *off* ou hors-champ),

Expérimentation B : caractérisation et influence du contenu

Dans ce document, la notion de *diégèse*¹⁵ fera référence à l'ensemble des sons pouvant être qualifié de son *in*, *off* ou hors-champ. Deux types de sons *in* (sons diégétiques c.-à-d. se déroulant dans le même espace-temps que l'action) peuvent être distingués : dans le champ ou son *in* (accompagne l'action et entendu par les personnages de la scène¹⁶) et hors-champ (hors de la scène -hors du champ de la caméra et donc du spectateur- mais entendu par les personnages¹⁷). Un son *off* (son extra-diégétique) se définit par un son en-dehors de l'espace-temps de l'action et qui n'est pas entendu par les personnages de la scène mais par le spectateur (voix *off* de narration¹⁸ ou musique *off*¹⁹). Dans les études suivantes, tous sons *in* (dans le champ et hors-champ) seront considérés comme diégétiques tandis que les sons *off* seront considérés comme extra-diégétiques.

- **l'expression sonore** (parole, musique, bruit),
- **le nombre de personnages** (faible ≤ 2 , modéré 2 à 5, fort ≥ 5),
- **la dynamique de contenu** (faible, modérée, forte)
- Le terme de contenu est accolé à la notion de dynamique dans l'intention d'établir une distinction nette avec le premier descripteur de dynamique relatif aux mouvements de caméra (descripteur technique). La *dynamique du contenu* réfère à l'action des personnages ou des objets.
- **la modalité dominante** (A, AV).

La *modalité dominante* peut se définir comme la modalité porteuse de l'information primordiale et sans laquelle la compréhension de la séquence serait mise à mal.

1.1.2. Retour sur expérimentation a: interprétations et explications

La caractérisation des contenus par l'expert permet déjà de confirmer ou tout au moins de consolider les éléments d'explications soumis lors de l'expérimentation A. Une influence du niveau de luminosité sur les mesures du diamètre pupillaire (DP) était supposée (visualisation des contenus Documentaire, Opéra et Sport, pour rappel : DP Documentaire > DP Sport et Opéra). La caractérisation experte des séquences tend à confirmer ce postulat. En effet, les contenus *Opéra* et *Sport* ont été caractérisés comme les plus lumineux (d'après la caractérisation experte réalisée sur les contenus entiers, le niveau de luminosité variait peu au sein d'un même

contenu). De ce fait, le diamètre pupillaire significativement plus petit pour ces deux contenus, par rapport au contenu

Documentaire, peut s'expliquer par un effet du niveau de luminosité (confirmation d'un réflexe photo-moteur). Un effet du niveau de dynamique a également été suggéré pour expliquer l'augmentation de l'activité électrodermale (AED) durant le contenu *Sport* par rapport aux contenus *Opéra* et *Documentaire*. La caractérisation experte des séquences tend également à consolider ce postulat : *Sport* a été caractérisé par un niveau de dynamique de contenu plus élevé que les contenus *Documentaire* et *Opéra*. Une autre différence notable entre ces contenus provient du niveau de dynamique de caméra. En effet, *Sport* présentait un niveau élevé pour ce descripteur tandis qu'à l'inverse, la dynamique caméra pour *Documentaire* et *Opéra* était faible ou modérée.

Ainsi, il semblerait que l'AED soit sensible aux aspects de dynamique relatifs à l'activité à la fois des personnages ou objets d'intérêt et des caméras.

Expérimentation B : caractérisation et influence du contenu

Comme l'ont indiqué Lang A. *et al.* (1999), Lang A. *et al.* (2000), Simon *et al.* (1999) ou encore

Yoon *et al.* (1998) cité par docteur Julie LASSALLE, l'AED est sensible (augmentation) au mouvement et aux changements de plan ou de scène. La caractérisation tend à aller dans le sens des constats de ces auteurs et à indiquer que l'AED pourrait être sensible au niveau de dynamique de manière générale (action des personnages, changement de plan, zooms, travelling, *etc.*). Ainsi, la caractérisation réalisée accrédite ou consolide les postulats émis pour expliquer les résultats de l'expérimentation A. Ce constat, souligne l'importance de caractériser les contenus de test pour une meilleure compréhension de l'influence de la qualité sur la perception du spectateur et les mesures psychophysiologiques recueillies. La caractérisation réalisée par l'expert puis par le panel de participants « naïfs » a donc permis d'apporter des éléments de réponses quant aux facteurs capables d'influencer la perception de qualité (niveau de dynamique par exemple) et la *qualité d'expérience* du spectateur subjectivement ou physiologiquement étudiée. Afin de mieux comprendre les interactions entre contenu et perception de qualité, le corpus de séquences de test précédemment utilisé et caractérisé a été présenté avec différentes conditions de qualité lors d'une seconde expérimentation (B2).

1.2. Expérimentation b : contenu et qualité

1.2.1. Objectifs

L'objectif de l'étude B1 est d'étudier l'influence du contenu à travers des séquences caractérisées lors de l'expérimentation B1 (c'est-à-dire décrites par un certain nombre de descripteurs) sur la perception de la qualité audiovisuelle. La connaissance des descripteurs les plus critiques du point de vue de l'évaluation de qualité doit permettre d'optimiser l'étape de caractérisation des séquences de test dans le cadre d'une méthode d'évaluation de la *qualité d'expérience* du spectateur.

1.2.2. Participants

Dans cette expérimentation, quarante participants (19 femmes, 21 hommes) entre 20 et 55 ans devaient évaluer le niveau de qualité audiovisuelle après chaque séquence de test visualisée. Les participants étaient rémunérés pour leur participation.

1.3. Matériel

a) Configuration générale

Vid. 7.11. Schéma de la configuration de la salle de test (250×300×320 cm) de l'expérimentation B1.

La place du participant est représentée par un point rouge, l'écran est figuré par un rectangle.

L'expérience a été conduite dans une salle insonorisée, les conditions de visualisation et d'écoute correspondaient aux recommandations de la norme UIT-T P.912 (voir annexe 7-G). Les paramètres de brillance et de contraste de l'écran respectaient la recommandation UIT-R BT.814-2 (UIT, 2007). Concernant l'affichage, un écran LCD de 46" (117 cm), full HD (1080p, 16/9) de type LG modèle KDL-46HX920 a été utilisé. La distance de visualisation, en accord avec la norme UIT-T P.912, a été fixée à 3,2 fois la hauteur de l'écran, soit 192 cm. Des haut-parleurs Genelec de modèle 8040A ont été réglés à une hauteur de 98 cm et placés de manière à être équidistants du centre de l'écran (115 cm) et de la tête du participant (208 cm). La vidéo 7.11 ci-dessus présente la configuration de la salle de test, celle-ci respectait les recommandations fournies par la norme UITR BS.1286 (UIT, 1997). En accord avec la recommandation UIT-T P.911, le volume sonore restitué par les HP se situait autour de 80 dB (mesuré à l'aide d'un sonomètre), le son ambiant de la salle de test était inférieur à 31 dB.

b) Configuration technique

Comme l'indique la vidéo 7.12 ci-après, le signal audiovisuel global était stocké et diffusé par un magnétoscope numérique (DVS - Digital Vidéo System - Pronto2K). Les signaux audio et vidéo étaient ensuite acheminés séparément vers le matériel de restitution audio (c.-à-d. HP externes *via* un amplificateur -SPL 2380-) ou vidéo (écran). Avant cela, chaque signal était converti au bon format d'entrée, c'est-à-dire d'un signal numérique (signal de sortie du DVS) à un signal analogique (signal d'entrée HP) pour l'audio (convertisseur Méridian 564) et une conversion SDI (signal de sortie du DVS) vers DVI (signal de sortie du convertisseur, Gefen HD-SDI to DVI scaler) pour permettre une entrée HDMI (signal d'entrée écran) pour la vidéo.

Vid. 7.12. Configuration technique de l'expérimentation B.

Le logiciel SEOVQ (Subjective Evaluation and Optimization of Video Quality, solution d'interface pour protocole d'évaluation d'images multimédias) a permis de déclencher, à partir de l'interface utilisée, la lecture des séquences vidéo (les séquences ne pouvaient vues et entendues qu'une seule et unique fois). Cette configuration permettait l'utilisation de *playlist*, avec un ordre aléatoire de présentation des séquences différent pour chaque participant.

1.4. Stimuli

Les vingt séquences de test utilisées lors de l'expérimentation B ont été présentées en format 2D full HD 1080p avec dix conditions de qualité différentes choisies pour être représentatives de dégradations susceptibles de survenir dans des conditions réelles de visualisation. L'ensemble des séquences audio a été normalisé en niveau sonore. Les conditions de qualité étaient les suivantes :

- **Condition1 (REF)** : séquence audiovisuelle de référence (sans dégradations) présentant un signal audio de 48 Kbps/16 bit et un signal vidéo .avi non compressé,
- **Condition 2 (D)** : séquence audiovisuelle présentant une désynchronisation entre l'image et le son avec un délai de 1500 ms appliqué sur le signal vidéo seulement (retard de l'image par rapport au son). Les caractéristiques des signaux audio et vidéo étaient identiques à celles de la séquence référence,
- **Condition 3 (V-DEB)** : séquence audiovisuelle présentant une réduction du débit vidéo. Sur la base des résultats de l'expérimentation A, un débit adaptatif a été appliqué dans cette étude. Le signal vidéo était compressé avec le codeur AVCx264 pour un débit de 93 à 1600 kbps selon le niveau de détail (complexité) de la séquence à coder.

Expérimentation B : caractérisation et influence du contenu

L'utilisation d'un débit adaptatif permet d'obtenir des dégradations plus homogènes du point de vue perceptif entre les différentes séquences (d'après les résultats de l'expérimentation A un seuil identique de débit appliqué à un contenu vidéo « simple » - plus facile à coder - dégraderait moins la qualité perçue qu'un contenu complexe),

- **Condition 4 (V-Gel)** : séquence audiovisuelle présentant un gel vidéo pour laquelle une série d'image "gelées" (figées) était introduites. Plus précisément, cinq périodes de "gel" d'une durée de 16 images chacune (soit 720 ms pour un total de 3,6 s dégradées) étaient introduites de manière aléatoire sur les dix secondes présentées,
- **Condition 5 (A-PP)** : séquence audiovisuelle présentant une perte de paquets d'information audio pour laquelle un taux de 12% de perte de paquets était introduit aléatoirement sur le flux audio, • **Condition 6 (A-DEB)** : séquence audiovisuelle présentant une réduction du débit audio (A-DEB) avec le signal audio compressé à 64Kbps/8Ko (codeur AC3),
- **Condition 7** : séquence audiovisuelle présentant la combinaison des dégradations APP
* V-Gel,
- **Condition 8** : séquence audiovisuelle présentant la combinaison des dégradations APP
* V-DEB,
- **Condition 9** : séquence audiovisuelle présentant la combinaison des dégradations **ADEB**
+ V- DEB,

- **Condition 10** : séquence audiovisuelle présentant la combinaison des dégradations ADEB + V- Gel.

1.5. Protocole

Chaque participant visualisait un total de deux cents séquences (20 séquences × 10 conditions de qualité) d'une durée de six à dix secondes, présentées dans un ordre aléatoire différent pour chaque individu. Entre chaque séquence, les participants disposaient d'une pause de cinq secondes pour évaluer la qualité AV globale sur une échelle en neuf points et cinq items (*Excellent, Bon, Satisfaisant, Médiocre, Mauvais*) conformément à la méthode ACR de la recommandation UIT-T P.912. Chaque séquence était visualisée une seule et unique fois. Le logiciel SEOVQ a été utilisé pour recueillir les jugements des participants. Une illustration de l'interface utilisée est apportée par la Figure 7.13 ci-dessous. Avant la passation, une fiche de consignes était distribuée au participant. Les consignes sont présentées dans l'annexe 7-H. La durée totale de la passation de test était d'environ une heure trente.

Vid. 7.13. Interface d'évaluation du logiciel SEOVQ permettant au participant de reporter son jugement sur l'échelle affichée. Une fois l'évaluation effectuée, le participant passait à la séquence suivante au moyen d'un *clic bouton* (a). Chaque séquence ne pouvait cependant être visualisée et entendue qu'une seule et unique fois. Les résultats pour chaque participant étaient ensuite enregistrés (b), après chaque passation, dans un fichier dédié.

1.6. Observables et hypothèses

Dans cette étude, seule la qualité audiovisuelle globale était évaluée par les participants, conformément à la norme UIT-T P.912. L'objectif de cette expérimentation était d'étudier l'influence du contenu sur la qualité audiovisuelle perçue. En conséquence, il était attendu que les notes de qualité audiovisuelle (MOSAV) recueillies ne dépendent pas seulement des dégradations de qualité mais aussi du contenu qualifié par les descripteurs de bas et hautniveau précédemment définis.

H0 : au-delà d'un effet de la dégradation, les scores MOSAV sont également influencés par la séquence de test (décrite à partir des descripteurs de B).

2. Résultats

Une ANOVA à mesures répétées considérant la variable indépendante « Séquences » et la variable dépendante « Qualité » à dix modalités a indiqué un effet significatif de ces deux variables sur les scores MOSAV obtenus, à savoir $F(19, 627) = 377,60$, $p < 0,001$ pour l'effet de la séquence et F

$(9, 297) = 15,61$, $p < 0,001$ pour l'effet de la qualité. Une interaction de la variable « Séquences » et de la variable « Qualité » a également été constatée : $F(171, 5643) = 12,33$, $p < 0,001$. Ces résultats confirment H0 en montrant que les séquences de test (décrites à partir des descripteurs de B) influencent la perception de qualité. Les notes

Passage à la séquence suivante Clôture du test et enregistrement données MOSAV

obtenues pour chaque dégradation sont présentées pour chaque contenu par les figures ci-dessous : Danse (vid . 7.14), Documentaire (vid . 7.15), Opéra (vid. 7.16), Théâtre (vid.7.17) et

Sport (vid. 7.18). Une présentation par type de dégradation est également apportée dans l'annexe 7-I. Cette présentation permet de mieux observer certains effets discutés lors de l'analyse des résultats.

Danse-1,Danse-2,Danse-3,Danse-4,Référence,A-DEB,A-PP,V-DEB,V-GEL,Desynchro,V-DEB*A-PP,V-DEB*A-DEB,V-GEL*A-PP,V-GEL*A-DEB
MOSAV

Vid. 7.14. MOSAV obtenues pour chaque condition de qualité pour les séquences du contenu Danse.

Doc-1,Doc-2,Doc-3,Doc-4,Référence,A-DEB,A-PP,V-DEB,V-GEL,Desynchro,V-DEB*A-PP,V-DEB*A-DEB,V-GEL*A-PP,V-GEL*A-DEB

Expérimentation B : caractérisation et influence du contenu

2.1. Effet de la désynchronisation

Les vidéos ci-dessus indiquent que la désynchronisation a été à l'origine d'une diminution des notes MOSAV sur l'ensemble des séquences extraites du contenu *Opéra* ainsi que sur les séquences Théâtre-1 et Doc-3 (diminution de la note d'environ deux points par rapport aux séquences des autres contenus). Ces séquences ont en commun un contexte audio à la fois verbal et diégétique. La totalité des autres séquences, pour lesquelles la désynchronisation ne semble pas avoir dégradé la qualité, correspondait soit à des scènes auditives caractérisées par les modes *Musique* ou *Bruit* du descripteur « Expression sonore », que ce dernier soit caractérisé comme diégétique (Danse, Théâtre-2 et -4) ou non (Théâtre-3, Doc.-1), soit à des contenus verbaux (mode Parole) mais présentant un son extra-diégétique (son *off*) comme c'est le cas pour le contenu *Sport* et les séquences Doc-4, Doc-2 (voir annexe 7-B). Ainsi, pour altérer la perception de qualité, la désynchronisation ne doit pas seulement survenir lors de séquences verbales mais lors de séquences verbales présentant aussi un son diégétique (*in* et hors-champ).

2.2. Effet des dégradations audio

Les Vidéos ci-avant montrent que globalement la dégradation A-PP a été perçue comme dégradant plus fortement la qualité que la dégradation A-DEB. Par ailleurs, le niveau de qualité perçue de certaines séquences semble avoir été plus fortement influencé par les dégradations audio. Il s'agit notamment des séquences : Opéra-1 (plus fortement dégradée par A-PP que les autres séquences du contenu), Théâtre-1 et Sport-2 (plus fortement dégradées par A-DEB que les autres séquences du contenu) et les séquences Doc-2, Doc-3 et Doc-4 du contenu *Documentaire* (plus fortement dégradées par A-DEB que l'ensemble des séquences du corpus, exceptée Sport-2). La connaissance de la modalité dominante permet de mieux comprendre ces différences. En effet, neuf séquences ont été caractérisées par une modalité dominante audio : Opéra (1, 2, 3 et 4), Théâtre 1, Documentaire (2, 3 et 4) et Sport-

2 (voir Annexe 7-B). Ainsi, il semble que la modalité dominante permette d'expliquer les différences observées. Le fait que les séquences du contenu *Opéra* n'aient pas été plus fortement dégradées par A-PP et A-DEB que les autres contenus (en-dehors d'*Opéra*¹) pourrait s'expliquer par le fait que les scènes auditives correspondaient à des paroles chantées et de langue étrangère. Il est alors envisageable que les dégradations audio dégradent l'intelligibilité du contenu sonore, celle-ci n'étant pas dégradée dans le cas d'un contenu de paroles étrangères (pas d'accès au sens). L'influence plus importante de la dégradation audio lors de contenus à dominante audio semble notamment vraie pour la dégradation A-DEB. Le fait que la dégradation

A-PP ait moins été soumise à l'influence du contenu pourrait s'expliquer par un effet moins subtil de cette dégradation qui viendrait lisser les effets de contenu (dégradation plus fortement perceptible et ce, indépendamment du contenu).

Expérimentation B : caractérisation et influence du contenu

2.3. Effet des dégradations vidéo

Globalement la dégradation V-DEB a été perçue comme dégradant plus fortement la qualité que V-GEL. Pour cette dernière dégradation, les notes de QAV attribuées aux séquences des contenus *Opéra*, *Documentaire* et *Théâtre* tendent à être plus élevées que celles attribuées aux séquences des contenus *Danse* et *Sport* (pour une meilleure observation de cette tendance voir annexe 7-I). Comme indiqué dans le paragraphe ci-dessus, la modalité dominante constitue un élément d'explication des variations observées : les séquences des contenus *Opéra* et *Documentaire* étaient principalement *Audio* tandis que celles des contenus *Danse*, *Sport* et *Théâtre* étaient principalement

Vidéo ou *Audiovisuelle*. Néanmoins, le descripteur de modalité ne permet pas d'expliquer les notes attribuées aux séquences du contenu *Théâtre*. Une influence du niveau de dynamique peut ici être supposée : toutes les séquences des contenus *Sport* et *Danse* présentaient une dynamique de contenu élevée (*Danse*, excepté *Danse-1*) et/ou une dynamique de caméra élevée (*Sport*), ce qui n'était pas le cas des séquences du contenu *Théâtre*. Ainsi, la présence de saccades (gel d'images) dégraderait plus fortement la qualité lorsque la séquence présente une dynamique forte (mouvements des personnages ou de caméra plus difficiles à suivre en présence de saccades). Une autre remarque concerne la séquence *Théâtre-3* ayant obtenu la note MOSAV la plus faible du corpus lors de la présentation de V-GEL. Cette séquence était la seule avec *Doc-1* à présenter à la fois une modalité dominante vidéo et un son *off* de musique. En d'autres termes, aucune information auditive complémentaire ne pouvait permettre au participant de compenser la dégradation vidéo. Le fait que *Doc-1* ait reçu des notes de QAV supérieures à *Théâtre-3* pourrait être expliqué par le nombre de personnages. En effet, *Doc-1* était caractérisée par un niveau faible (1 seul personnage) tandis que *Théâtre-3* présentait un niveau modéré (3 personnages) dont le suivi des déplacements serait rendu plus difficile par la présence de saccades vidéo (images gelées).

Ainsi, V-GEL aurait plus fortement dégradé *Théâtre-3* en raison de l'absence d'information auditive pertinente (musique *off*) et en présence de mouvements sur le média vidéo (déplacements des personnages). Ce constat tend à appuyer l'impact

plus important de V-GEL lorsque cette dégradation est appliquée lors de séquences dynamiques (mouvements des caméras ou des personnages) par rapport à des séquences peu dynamiques. La note MOSAV la plus faible du corpus de séquences de test a été attribuée à la séquence Danse-1 en présence de la dégradation V-DEB. Cette séquence s'était également distinguée lors de l'expérimentation B1 (notes plus faibles de QAV et QV par rapport aux autres séquences du contenu) bien qu'aucune dégradation n'était présente. Danse-1 correspond à une séquence pauvre du point de vue informationnel tant sur le média audio (mode *Musique* du descripteur *Expression Sonore*) que vidéo (niveau *Faible* des descripteurs *Dynamique de contenu* et *Dynamique de caméra*). Danse-1 était la seule séquence du corpus à présenter la combinaison « musique » (modalité dominante vidéo) et « dynamique faible » (contenu et caméra). Le fait que la diminution de qualité observée pour Danse-1 ne soit pas retrouvée en présence de V-GEL laisse supposer qu'une dégradation par diminution du débit vidéo (perte de résolution) dégrade plus fortement la qualité perçue, que la présence de saccades vidéo, pour ce type de séquence (modalité dominante vidéo -musique- et dynamique faible).

2.4. Effets des dégradations audio-vidéo

Expérimentation B : caractérisation et influence du contenu

De manière générale, les variations des notes MOSAV retranscrivent celles observées lors de l'analyse des dégradations vidéo seules. En effet, les dégradations AV incluant la dégradation V DEB ont été perçues comme dégradant plus fortement la qualité que celles incluant V-GEL. Cette observation reflète bien les résultats obtenus lorsque les dégradations vidéo étaient présentées seules. Les descripteurs *Couleur*, *Luminosité*, *Détail* ainsi que les descripteurs de haut-niveau n'ont pas permis d'apporter des éléments d'explications aux effets observés. Il semblerait que les influences potentielles de ces descripteurs aient été masquées par celles des autres critères.

• Conclusions B1

Globalement, les résultats de l'expérimentation B2 ont montré que **l'altération vidéo par réduction du débit (V-DEB) a été jugée comme la dégradation de qualité la plus forte** (diminuant le plus la note de qualité AV globale). Cela peut être constaté autant pour une présentation isolée de cette dégradation que pour une présentation combinée avec une dégradation audio (A-DEB ou A-PP). Concernant les dégradations audio, **la dégradation par perte de paquets (A-PP) a été jugée comme altérant plus fortement la qualité** que la dégradation par réduction du débit (A-DEB). Cet effet est également retrouvé lors d'une présentation combinée avec une dégradation vidéo (V-DEB ou V-GEL). Ce constat peut s'expliquer soit par des différences objectives entre les dégradations audio et vidéo (quantité de dégradations) soit par l'existence de régularités dans la perception de certaines dégradations. Les **dégradations liées au débit vidéo** (perte de résolution spatiale) **seraient perçues comme dégradant plus fortement la qualité** qu'une dégradation par rupture de la continuité du signal (saccades), indépendamment du type de contenu. Pour l'audio, à l'inverse, la rupture de la continuité (perte d'information) serait perçue comme plus gênante que

les dégradations liées au débit (distorsion). Il se peut que la rupture de continuité du flux audio entraîne une diminution de l'intelligibilité en raison de la perte de certaines informations notamment verbales (phonème, syllabe, *etc.*), pour la vidéo, la perte de netteté constituerait une perte d'information visuelle plus importante que la rupture de continuité. Ce résultat peut également être mis en regard des spécificités des systèmes auditifs et visuels : l'acuité temporelle de l'audition et donc sa sensibilité (à la rupture de la continuité) est plus élevée que celle de la vision tandis que l'acuité spatiale de la vision et donc sa sensibilité (à la perte de résolution) est plus élevée que celle de l'audition. Les résultats obtenus ont également permis d'observer **qu'une dégradation audio diminue davantage le niveau de qualité perçue lorsque la nature de la modalité dominante de la séquence était audio. Le même constat peut être apporté pour une dégradation vidéo et une modalité dominante vidéo.** Par ailleurs, lorsque la modalité dominante est vidéo, la présence de saccades sur le signal vidéo (V-GEL) tend à dégrader plus fortement des séquences de dynamique forte (dégradation de l'information pertinente) tandis que la perte de résolution vidéo (V-DEB) dégraderait plus fortement des séquences de dynamique faible (dégradation des rares informations visuelles disponibles).

Enfin, la dégradation **Désynchronisation a été perçue comme dégradant la qualité lorsque celle-ci survenait sur des séquences présentant un contenu verbal en lien avec l'action affichée à l'écran (son diégétique).** Lorsque ces deux conditions sont remplies, le phénomène de fusion des informations auditive et visuelle est alors altéré. Dans le cas contraire, les notes obtenues étaient proches de celles attribuées aux séquences présentées sans dégradations. Le cas échéant, la note de qualité ne permet pas de savoir si la désynchronisation était perçue ou non. Cette dernière observation semble confirmer l'importance d'ajouter une question spécifique à la perception de la désynchronisation lorsque cette dernière est étudiée. L'expérimentation B1 a permis de mieux définir la relation entre contenu, caractérisé à partir d'un ensemble de descripteurs, et perception de qualité. Les influences du contenu observées, par exemple sur la perception de désynchronisation, sont des informations importantes à considérer lors de la sélection des séquences de test.

3. Conclusion

Les expérimentations B et B1 présentées dans ce chapitre ont permis d'étudier l'influence du contenu, d'une part, sur la *qualité d'expérience* du spectateur (en matière d'intérêt, de plaisir, *etc.*, B) et d'autre part, sur la perception de qualité (B). L'expérience B1 a permis de disposer d'un corpus de contenus et de séquences de test audiovisuels caractérisés selon des descripteurs sémantiques (de bas et haut-niveau), techniques et hédoniques. B a montré que la caractérisation réalisée par l'expert (descripteurs de bas-niveau seulement) peut être considérée comme pertinente pour une description des contenus compréhensible par des participants naïfs en-dehors du descripteur de température de couleur. D'un point de vue méthodologique, cela autorise à ne mobiliser qu'un expert seul pour l'annotation des séquences ou contenus de test. L'expérience B1 a permis de mieux comprendre les interactions entre le contenu, défini par les descripteurs proposés, et la

perception de qualité. La modalité dominante, la dynamique, la relation audiovisuelle (diégétique) et l'expression sonore influencent notamment la qualité audiovisuelle perçue par le spectateur. Ce constat souligne l'importance de l'étape de sélection des séquences de test ainsi que la description préalable à réaliser. Par exemple, pour être étudiée, la désynchronisation devra être appliquée à des séquences présentant un segment audio à la fois verbal et diégétique. Une seconde campagne de test pourrait être conduite afin de tester le questionnaire utilisé dans l'expérimentation B1 pour un protocole similaire à l'expérimentation B1. Cela permettrait d'étudier plus précisément les effets des dégradations non plus sur la perception de qualité seule mais également sur la *qualité d'expérience* finale du spectateur. Il pourrait également être intéressant d'élargir le cadre de la norme UIT-T P.912 aux évaluations des qualités audio et vidéo pour permettre l'étude du poids de chaque qualité à la perception de qualité audiovisuelle. De plus, dans le cadre du corpus de séquences étudiées, les descripteurs *Modalité* et *Dynamique* étaient fortement liés (dynamique faible associée à une dominance audio). Ce constat va dans le sens de celui de Hands (2004) cité par docteur Julie LASSALLE, cependant, cette association pourrait être l'expression d'une limite liée au questionnaire où le terme de dynamique réfère uniquement à la modalité vidéo. Une étude complémentaire où il serait demandé aux participants d'évaluer à la fois la dynamique audio et vidéo pourrait être envisagée pour approfondir ce point.

4. Bibliographie

- Ahlstrom, U. et Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, 36(7), 623-636.
- Akselrod, S., Gordon, D., Ubel, F. A., Shannon, D. C., Berger, A. et Cohen, R. J. (1981). Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, 213(4504), 220-222.
- Allport, D. A. (1980). Patterns and actions: Cognitive mechanisms are content-specific. In G.
- Amiar., Y. (1995). L'analyse du film et de l'image fixe: approche méthodologique *Revue Recherche sur l'information scientifique et technique*, 5(2), 23-28.
- Andreassi, J. L. (2007). *Psychophysiology: Human behavior and physiological response* (5e éd.). Mahwah, New Jersey : Lawrence Erlbaum.
- Andreassi, J.L., Rapisardi, S.C. et Whalen, P.M. (1969). Autonomic responsivity and reaction time under fixed and variable signal schedules. *Psychophysiology*, 6, 58-69.
- Appel, M. L., Berger, R. D., Saul, J. P., Smith, J. M. et Cohen, R. J. (1989). Beat to beat variability in cardiovascular variables: noise or music? *Journal of the American College of Cardiology*, 14(5), 1139-1148.
- Arrighi, R., Alais, D. et Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision*, 6(3), 260-268.
- Backs, R. W. et Boucsein, W. (2000). *Engineering psychophysiology: issues and applications*. Mahwah, New Jersey: Lawrence Erlbaum Mahwah.
- Baddeley, A. D. (1986). *Working memory*. Oxford, R. -U. : Clarendon.
- Bahill, A. T. et Stark, L. (1975). Overlapping saccades and glissades are produced by fatigue in the saccadic eye movement system. *Experimental Neurology*, 48(1), 95-106.
- Bahrack, L. E. (1987). Infants' intermodal perception of two levels of temporal structure in natural events. *Infant Behavior and Development*, 10(4), 387-416.

- Barry, R. J. et Sokolov, E. N. (1993). Habituation of phasic and tonic components of the orienting reflex. *International Journal of Psychophysiology*, 15(1), 39-42.
- Barry, R. J., Clarke, A. R., McCarthy, R., Selikowitz, M. et Rushby, J. A. (2005). Arousal and activation in a continuous performance task: An exploration of state effects in normal children. *Journal of psychophysiology*, 19(2), 91-99.
- Bauer, L. O., Strock, B. D., Goldstein, R., Stern, J. A. et Walrath, L. C. (1985). Auditory discrimination and the eyeblink. *Psychophysiology*, 22(6), 636-641.
- Baumstimler, Y. et Parrot, J. (1971). Stimulus generalization and spontaneous blinking in man involved in a voluntary activity. *Journal of experimental psychology*, 88, 95-102.
- Beatty, J. et Lucero-Wagoner, B. (2000). The pupillary system. Dans J. T. Cacioppo, L. G.
- Bechara, A., Damasio, H., Damasio, A. R. et Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience*, 19(13), 5473-5481.
- Beerends, J. et De Caluwe, F. (1999). The influence of video quality on perceived audio quality and vice versa. *Journal of the Audio Engineering Society (JAES)*, 47(5), 355-362.
- Bermant, R. I. et Welch, R. B. (1976). Effect of degree of separation of visual-auditory stimulus and eye position upon spatial interaction of vision and audition. *Perceptual and motor skills*, 43(2), 487-493.
- Berntson, G. G., Cacioppo, J. T. et Fieldstone, A. (1996). Illusions, arithmetic, and the bidirectional modulation of vagal control of the heart. *Biological Psychology*, 44(1), 1-17.
- Bertelson, P. et Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditoryvisual spatial discordance. *Attention, Perception et Psychophysics*, 29(6), 578-584.
- Berthoz, A. et Petit, L. (1996). Les mouvements du regard: une affaire de saccades. *La Recherche*, 289, 58-65.
- Boonnithi, S. et Phongsuphap, S. (2011). Comparison of heart rate variability measures for mental stress detection. Dans *Proceedings of the IEEE Computing in Cardiology conference*, 38, 85-88.
- Claxton (Ed.), *Cognitive psychology: New directions* (p. 32-59). London, R. -U.: Routledge & Kegan Paul.
- Tassinary et Berntson, G.G. (dir.), *Handbook of psychophysiology* (p.142-162). New York, NY: Cambridge University Press.

INTERNATIONAL JOURNAL OF AFRICAN SCIENCES



REVUE INTERNATIONALE DES SCIENCES AFRICAINES

Editions Lumumba



A New Momentum for African Publishing

10, De la Moto Street, Gambela, Lubumbashi City - Democratic Republic of Congo

Tel: 00243 90 433 70 19
contact@editionslumumba.com
www.editionslumumba.com

Printed in the Democratic Republic of the Congo