

Quand la vidéo devient silencieuse : analyse sémio-historique du sous-titrage dans les productions audiovisuelles des médias en ligne français (2014-2020)

When the video becomes silent. Semio-historical analysis of subtitling in audiovisual productions of French online media (2014-2020)

Guylaine Guéraud-Pinet

Univ. Grenoble Alpes, GRESEC, 38000 Grenoble, France

Résumé. Depuis le milieu des années 2010, la vidéo informative en ligne (Brut., Konbini, Culture', Loopsider, etc.) se développe en France. L'une de ses particularités réside dans la mobilisation systématique du sous-titrage ou du légendage. Production audio-visuelle, la vidéo devrait solliciter, par définition, aussi bien la vue que l'ouïe. Cependant, ce recours au texte questionne la place des formes sonores dans la vidéo en ligne. Cet article étudie alors la construction « audio-scripto-visuelle » de ces contenus. À partir d'une analyse sémio-historique jointe à une analyse de contenu de vidéos en ligne, les résultats montrent que le recours au texte (sous-titrage et légendage) dans ces vidéos provient d'un héritage médiatique ancien (cinéma, radio, télévision). Ce recours au texte montre aussi que le son semble de plus en plus dispensable à la compréhension des contenus. Le format de vidéo en ligne étudié se multiplie et semble se standardiser. Par conséquent, son industrialisation est questionnée au regard des enjeux sonores.

Abstract. Since the middle of 2010, online informative video (Brut., Konbini, Culture', Loopsider, etc.) has been developing in France. One of its particularities lies in the systematic mobilisation of subtitling or captioning. By definition, as an audio-visual production, video should appeal to both sight and hearing. However, this recourse to text questions the place of sound forms in online video. This article studies the "audio-scripto-visual" construction of these contents. It based on a semio-historical analysis combined with an analysis of online video content. The results show that the use of text (subtitling and captioning) in these videos comes from an ancient media heritage (cinema, radio, television). This recourse to text also shows that sound seems to be increasingly dispensable to the comprehension of content. The online video format studied is multiplying and seems to be becoming standardised. Consequently, its industrialisation is questioned with regard to the sound stakes.

Fondé en 2016, en France, le média en ligne Brut. propose une diffusion de l'information sous la forme de vidéos au format carré ou rectangulaire agrémentées systématiquement d'un sous-titrage. Diffusé sur les réseaux sociaux numériques (Facebook, Instagram ou Twitter) et sur des plateformes d'hébergement de vidéos (YouTube et Dailymotion), ce type de productions audiovisuelles semble s'être adapté aux pratiques des utilisateurs d'écrans d'ordinateur, et plus encore, d'écrans de smartphone. Selon une enquête menée par le site d'information Digiday.com en 2017, 85 % des utilisateurs de Facebook y regarderaient des vidéos sans le son¹. Par conséquent, le son, pourtant lié à l'image animée depuis les prémisses du cinéma (Chion, 1995), serait de plus en plus dispensable à la compréhension des contenus. Le format audiovisuel proposé par Brut., où le texte supplanterait les formes sonores, s'est depuis étendu, progressivement, à d'autres médias en ligne aussi bien privés (Kobini, Melty, etc.) que publics (Culture').

1 La vidéo informative en ligne : une approche communicationnelle

1.1 La vidéo informative en ligne, une production audio-visuelle sans « l'audio » ?

Cet article interroge la place des matériaux sonores, iconographiques et textuels nécessaires à la compréhension des vidéos informatives spécifiquement conçues par les médias en ligne. Ces productions audiovisuelles, qui par définition devraient solliciter aussi bien la vue que l'ouïe, semblent se défaire du préfixe « audio ». Jean-Yves Chevalier, en 2018, remarquait une tendance à la multiplication de l'usage des sens face aux écrans : le toucher serait de plus en plus impliqué avec l'utilisation des tablettes et des smartphones, le recours à la voix des utilisateurs se développerait avec la commande vocale, etc. (Chevalier, 2018). Cependant, dans la vidéo informative en ligne, le son s'effacerait, dans une mesure à qualifier, au profit de la vue. Le texte, et en particulier le sous-titrage, accompagnerait ou remplacerait les voix usuellement entendues dans la vidéo. Les légendes, les images mouvantes et fixes ou encore les icônes accentueraient cet appel à la vue en se substituant aux formes sonores ou en les illustrant.

Les questionnements qui guideront cette recherche découlent de ces constats. Nous tâcherons tout d'abord de comprendre comment se sont construits les agencements visuels et sonores du contenu informationnel de la vidéo en ligne. La place du sous-titrage, du légendage, mais aussi des voix des intervenants et des voix off seront décrites puis mises en regard d'agencements « audio-scripto-visuels » d'ores et déjà remarqués dans les médias audiovisuels dits traditionnels. En effet, ces pratiques d'insertions textuelles ou iconographiques ne sont pas nées de la production de vidéos informatives en ligne. Héritées aussi bien du cinéma, de la télévision de flux² ou encore du documentaire, elles tendent cependant à s'autonomiser dans l'objet qui nous intéresse, voire à témoigner d'« une manière de faire » spécifique à la vidéo informative en ligne. Par conséquent, ces éléments nous permettront d'appréhender la manière dont le son semble être de moins en moins nécessaire à la compréhension des contenus informationnels diffusés sous la forme de vidéos en ligne. Enfin, nous nous interrogerons sur la standardisation de ces formes audiovisuelles. La circulation du format d'un média en ligne, à un autre, mais aussi sa vente ou sa récupération par divers acteurs entraînent un questionnement sur l'industrialisation des contenus informationnels et l'incidence que cela pourrait avoir sur les formes sonores et leur dépréciation.

1.2 Une approche communicationnelle de l'agencement « audio-scripto-visuel » des vidéos informatives en ligne

Peu de travaux ont été engagés en sciences de l'information et de la communication (SIC) sur le sujet du sous-titrage dans les vidéos en ligne. Très récent, il suscite pourtant l'intérêt d'étudiants, comme ce fut le cas cette année, avec le mémoire de fin de master professionnel de Valentin Ribas (2020) sur « La place du texte dans les vidéos d'information sur les réseaux sociaux ». Afin de proposer une lecture communicationnelle des agencements « audio-scripto-visuels » de ces vidéos et en accord avec nos questionnements, nous proposons une approche croisée entre travaux sur l'intermédialité (Altman, 1999 ; Gaudreault et Marion, 2000 ; Müller, 2006) et sémiotique des écrits d'écrans (Jeanneret et Souchier, 2005).

Selon Rick Altman, l'approche intermédiaire, ou l'intermédialité, désigne « une étape historique, un état transitoire au cours duquel une forme en voie de devenir un média à part entière se trouve encore partagé entre plusieurs médias existants, à un point tel que sa propre identité reste en suspens » (Altman, 1999 : 38). Nous comprenons alors que l'intermédialité permet d'envisager la construction d'un média avant qu'il ne soit proclamé comme tel, par l'étude des relations et associations dont il émane. La notion que nous mobiliserons particulièrement est enracinée dans cette approche intermédiaire. Il s'agit des « séries culturelles » qui sont définies par André Gaudreault comme des « sous-systèmes ou unités de signification composant le plus grand ensemble » (Gaudreault, 2008 : 114). Comme l'explique l'auteur, si le « grand ensemble » est le cinéma, les « unités de significations » ou encore les « séries culturelles » seront

¹ <https://digiday.com/media/silent-world-facebook-video/>

² Le programme de flux (magazine, journal télévisé, télé-réalité, etc.) se distingue du programme de stock (long métrage, série télévisée, documentaire, etc.) parce qu'il n'est pas considéré comme une œuvre audiovisuelle, mais plutôt comme un programme éphémère lié à l'actualité (Jost, 2005). Dans le cas qui nous intéresse, nous pensons particulièrement aux reportages diffusés dans des journaux télévisés ou encore dans des magazines.

« le café-concert, le théâtre d'ombres, les sketches magiques, la féerie, les arts du cirque, etc. » (*Ibid.*). Par conséquent, cette notion offre la possibilité de comprendre comment les formes médiatiques de plusieurs médias d'ores et déjà établis en composent un autre. En d'autres mots et rapproché de notre objet d'étude, travailler à partir des « séries culturelles » permet de comprendre comment les circulations intermédiaires des formes comme le sous-titrage ou encore la voix off au cinéma et à la télévision ont mené à cet agencement hybride audio-visuel que l'on retrouve sur les écrans numériques, et ce, dans un temps long. De plus, cela nous permet de questionner l'autonomisation des agencements les plus récents, ceux remarquables dans les vidéos informatives en ligne.

L'approche sémiotique des écrits d'écran se fonde, dans cet article, sur les travaux d'Emmanuel Souchier et d'Yves Jeanneret et s'articule particulièrement autour de la notion « d'énonciation éditoriale » (Jeanneret et Souchier, 2005). Cette notion, développée dès la fin des années 1990 dans un contexte d'utilisation croissante de logiciels modelant le texte permet de mettre en évidence : « l'ensemble de ce qui contribue à la production matérielle des formes qui donnent au texte sa consistance, son 'image de texte' » (*Ibid.*). Les auteurs ajoutent, pour la définir, qu'il s'agit :

« d'un processus social déterminé, qui demeure largement invisible du public, mais qui peut néanmoins être appréhendé à travers la marque qu'impriment les pratiques de métiers constitutives de l'élaboration, de la constitution ou de la circulation des textes » (*Ibid.*).

En nous appuyant sur la définition ci-dessus, nous comprenons qu'elle permet d'interroger conjointement le mode d'existence des textes, leurs circulations, mais aussi les acteurs qui concourent à la production de ces textes. Par conséquent, la notion « d'énonciation éditoriale » nous permettra d'analyser les formes textuelles tout en mettant en lumière les mouvements de standardisation de ces formes produites industriellement.

Les deux approches convoquées nous aideront à mettre en lumière l'évolution des formes « audio-scripto-visuelles » numériques qui nous intéressent dans un temps long, tout en considérant ces dernières comme des productions industrielles standardisées.

Enfin, comme nous l'avons déjà souligné, la vidéo informative en ligne n'a pas été étudiée en détail en SIC. Par conséquent, des travaux d'autres disciplines seront convoqués et particulièrement ceux traitant du sous-titrage. Nous nous référerons donc régulièrement à la littérature issue de la traductologie et particulièrement à la traductologie audiovisuelle (Petitjean, 2001 ; Gambier, 2004 ; Baldo De Brebisson, 2016) pour comprendre certaines spécificités et certains enjeux du sous-titrage.

1.3 Une analyse de contenu « intermédiaire » de vidéos produites par des médias en ligne français

1.3.1 Méthodes engagées

L'approche méthodologique découle directement de ce cadre théorique. Dans un premier temps, une analyse intermédiaire du sous-titrage dans l'audiovisuel a été menée. Elle consiste à construire des « séries culturelles » visant à comprendre « d'où » peut provenir ce recours au texte. Notons, en effet, que l'intermédialité peut être envisagée comme la méthode pour « étudier des relations médiatiques » et pas uniquement en tant que théorie ou objet de recherche (Müller, 2000). Dans un second temps, nous avons mis en place une analyse de contenu (Bardin, 1977) afin de rendre compte objectivement et systématiquement des constructions audiovisuelles des vidéos d'information en ligne. Ces dernières ont été diffusées à différentes périodes et proviennent de médias différents. Par conséquent, bien que l'analyse de contenu soit davantage rapprochée des méthodes quantitatives, dans une visée comparative et longitudinale, elle semble opportune (Coulomb-Gully, 2002). Deux grandes étapes composent cette analyse. La première consiste à construire des indicateurs traduisant nos questionnements sur les héritages médiatiques du texte dans la vidéo en ligne. Ce faisant, nous avons recherché différentes manières d'utiliser le sous-titrage, l'icône ou la voix off dans le temps à partir de l'analyse de la littérature académique existante sur le sujet. Nous avons aussi examiné des contenus audiovisuels dits traditionnels (film de cinéma, reportage TV, documentaires). Voici quelques exemples d'indicateurs mobilisés lors de l'analyse de contenu : sous-titrage intra ou inter – linguistique³ ; doublage⁴ ; voix off⁵ ; légendes⁶ ; icônes⁷, etc. Dans les contenus proposés par les vidéos en ligne, nous avons alors observé si ces éléments étaient présents ou absents et s'ils étaient utilisés de la même manière que dans l'un des médias traditionnels cités précédemment. Des critères descriptifs des vidéos, comme la date de mise en ligne, ont aussi été recueillis. Ils

³ Le sous-titrage intralinguistique correspond au sous-titrage d'un élément oral dans sa langue d'expression. Il est souvent rapproché du sous-titrage effectué pour les sourds et malentendants. *A contrario*, le sous-titrage interlinguistique correspond au sous-titrage d'un élément oral dans une autre langue. On parle souvent ici d'adaptation.

⁴ Le doublage consiste à remplacer une voix dans sa langue originale par une voix dans la langue parlée du pays diffuseur.

⁵ La voix off correspond à une voix qu'on a ajoutée lors du montage de la vidéo et qui ne représente pas une personne visible à l'écran.

⁶ Les légendes sont tous les éléments textuels permettant de décrire un élément de l'image. Il peut s'agir d'un titre, de la présentation d'un personnage ou encore de la source d'un document.

⁷ Les icônes correspondent, dans ce travail, à de petites images ajoutées lors du montage de la vidéo et symbolisant un élément audible dans la production.

permettent en effet d'obtenir des premières informations sur une éventuelle standardisation des contenus. La vidéo en ligne semble avoir connu un véritable essor au milieu des années 2010. La vidéo informative produite par des médias en ligne, elle, s'est multipliée à la fin des années 2010. Nous y reviendrons. Montrer que l'utilisation croissante des formes textuelles a accompagné ce mouvement serait un résultat à prendre en considération afin d'aborder la standardisation de la vidéo en ligne et ses enjeux pour le son.

La seconde étape de notre analyse consiste à collecter les données. Pour cela, le logiciel d'analyse qualitative NVivo a été utilisé. Développé sur le modèle de la Grounded Theory (Glaser et Strauss, 1967), cet outil permet un « étiquetage réflexif » (Lejeune, 2010) des contenus. Les indicateurs préconstruits peuvent être annotés de manière fine sur les vidéos en pourcentage de temps occupé. De plus, si des indicateurs « oubliés » lors de la mise en place de l'analyse de contenu surgissent, ils peuvent facilement être intégrés au recueil des données. Une fois encodées, ces données ont été extraites. Une analyse statistique des pourcentages de couverture de chaque indicateur a été réalisée à partir d'un tableur, et dans notre cas, à partir du logiciel Excel.

1.3.2 Terrain et corpus

Un corpus de cinquante-six vidéos, soit environ 3 heures et 20 minutes de visionnage, a été soumis à l'analyse. Ces productions audiovisuelles proviennent de cinq médias en ligne français à destination d'individus de moins de trente-cinq ans : Melty, Konbini, Brut., Culture' et Loopsider. Ces derniers sont présentés dans le tableau 1 ci-dessous.

Tableau 1. Présentation des médias en ligne français sélectionnés pour l'analyse

| Nom du média | Date de création | Cible principale | Modèle économique | Types de contenus proposés en ligne | Thématiques principales des contenus |
|--------------|------------------|------------------|---|-------------------------------------|--------------------------------------|
| Melty | 2008 | 15-35 ans. | Hybride (<i>production de contenus sponsorisés ; marketing de contenu ; utilisation de la data pour la production de contenus</i>). | Articles et vidéos. | Infodivertissement. |
| Konbini | 2008 | 18-30 ans. | Hybride (<i>production de contenus sponsorisés, marketing de contenu</i>). | Articles et vidéos. | Infodivertissement. |
| Brut. | Nov. 2016 | Moins de 35 ans. | Hybride (<i>partenariat avec France TV ; production de contenus sponsorisés ; agence de conseil en communication</i>). | Vidéos. | Information généraliste. |
| Culture' | Nov. 2018 | Moins de 35 ans. | Média du service public. Production sur les archives de 6 partenaires de l'audiovisuel français ⁸ . | Vidéos. | Information culturelle. |
| Loopsider | Janvier 2018 | Moins de 35 ans. | Hybride (<i>production de contenus sponsorisés, utilisation de la data pour la production de contenus</i>). | Vidéos. | Information généraliste. |

Le premier média à avoir été retenu est Brut.. Cité dès la première ligne de cet article, il est le premier acteur du domaine en France à produire une vidéo uniquement diffusée sur le web et à mobiliser le sous-titrage depuis ses débuts. Comme présenté dans le tableau, Brut. est partenaire de France Télévisions, et ce, depuis 2017. L'un des éléments clés de ce partenariat correspond à la possibilité, pour le service public audiovisuel français, de se servir de son format pour diffuser de la vidéo. C'est d'ailleurs de cette manière que Culture' a commencé à formater ses différentes productions, ce qui justifie sa place dans notre corpus. Inscrivant notre recherche dans une démarche comparative qui vise à montrer l'évolution du recours au texte dans la vidéo en ligne, nous avons sélectionné deux médias plus anciens : Melty et Konbini tous deux créés en 2008. Leur analyse permet de mettre en évidence le début de leur production vidéo, mais aussi leur utilisation des formes textuelles et iconographiques dans leurs contenus. Le dernier média sélectionné est Loopsider. Bien que plus récent que Brut., il lui ressemble beaucoup : mêmes cibles concernées, même modèle économique, thématiques abordées dans les vidéos similaires, etc. Pour chaque média et pour chaque année, une vidéo par trimestre a été sélectionnée de manière aléatoire. L'enquête a débuté en avril 2020. Nous sommes donc remontés dans le temps à partir de cette date jusqu'à ne plus trouver de contenus. Par conséquent, comme on peut le voir sur le tableau 2 ci-dessous, les cinquante-six vidéos qui composent le corpus se répartissent sur une période comprise entre juillet 2014 et février 2020.

⁸ Les six partenaires de l'audiovisuel français sont France Télévisions, Radio France, France Médias Monde, Arte, l'Ina et TV5 Monde.

Tableau 2. Répartition du corpus des vidéos informatives diffusées sur les réseaux sociaux numériques par dates de diffusion et par acteurs

| | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | Total |
|---|------|------|------|------|------|------|------|-------|
| Melty | | | | | 1 | 4 | 1 | 6 |
| Konbini | 1 | 0 | 2 | 2 | 4 | 4 | 1 | 14 |
| Brut. | | | 1 | 4 | 4 | 4 | 1 | 14 |
| Culture' (et France Culture* ⁹) | | | | *4 | 4 | 4 | 1 | 13 |
| Loopsider | | | | | 4 | 4 | 1 | 9 |

Les résultats de cette recherche seront présentés dans trois parties distinctes. La première proposera de revenir brièvement sur les héritages intermédiaires du sous-titrage dans la vidéo d'information en ligne. À partir de la présentation du rôle du texte et de la voix off dans l'audiovisuel, au sens large, nous décrirons l'agencement « audio-scripto-visuel » de notre objet. La deuxième partie de l'article s'intéressera plus particulièrement au légendage et au recours aux icônes, aux images, pour remplacer le son. L'évolution de ces « inserts » sera présentée et illustrée. Enfin, la dernière partie de l'article reviendra sur l'industrialisation et la standardisation d'un format « à la Brut. » mis en évidence dans nos analyses, mais aussi dans les discours d'acteurs. Les enjeux circonvenant le son, ou plutôt le remplacement du son par le texte et l'image, seront interrogés.

2 Le sous-titrage dans la vidéo en ligne : d'une pratique héritée à son autonomisation

Dans les cinquante-six vidéos de notre corpus, nous avons remarqué que toutes étaient agrémentées d'un sous-titrage sauf six, comme on peut le voir sur la figure 1, ci-dessous. Les six vidéos non concernées par le sous-titrage ont toutes été diffusées de 2014 à 2018 et cinq d'entre elles sont des productions de Konbini. Hormis ce cas, ce premier résultat montre une généralisation croissante du sous-titrage dans la vidéo informative en ligne.

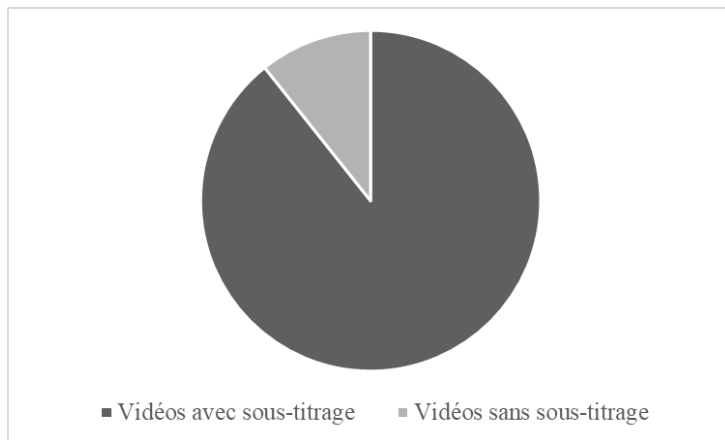


Figure 1. Répartition des vidéos selon la présence de sous-titrage.

L'un des questionnements principaux de cette recherche est de comprendre comment le sous-titrage a investi la vidéo en ligne et sur quels héritages cette pratique, aujourd'hui très répandue, s'est construite. À partir de la mise en évidence des « séries culturelles » liées à l'utilisation du texte pour accompagner ou remplacer des formes sonores, nous avons relevé deux héritages distincts qui se chevaucheraient. Le premier est caractérisé par une utilisation du texte pour expliquer l'image animée ou rendre lisible un message formulé oralement. Cela correspond principalement à l'intertitre et aux sous-titrages inter et intralinguistiques. Ils seront décrits dans le premier point de cette partie. Le second héritage est caractérisé par l'utilisation de la voix pour raconter cette image animée, pour l'expliquer. La voix et la voix off seront alors évoquées dans le deuxième point de cette partie. La définition de ces éléments et de leurs fonctionnements nous permettra alors de décrire les statuts du texte et des voix des vidéos informatives en ligne, et ce, dans un troisième point.

⁹ La collecte des vidéos a été réalisée sur YouTube. En remontant dans le temps pour les vidéos de Culture', nous avons rencontré les vidéos produites par France Culture. Produites sous le même format, nous avons décidé de les inclure au corpus. Par conséquent, les quatre vidéos recensées en 2017 ont été diffusées sous le nom de France Culture.

2.1 De l'intertitre au sous-titrage : quand le texte explique l'image

Jean-François Cornu, en 2014, publie un ouvrage fondateur sur l'histoire du doublage et du sous-titrage. On apprend de ce travail que l'écrit est un élément incontournable du cinéma dès 1900 (Cornu, 2014 : 223). L'écrit, lisible parfois sur des décors ou des accessoires, permettra de narrer l'histoire à une époque où la parole ne peut pas encore être enregistrée. Par conséquent, Jean-François Cornu nous apprend que de 1900 à 1930, il va falloir trouver un moyen pour commenter les images qui ne sont pas encore synchronisées au son. Des intertitres seront alors montrés en direct, ou parfois incorporés directement au film. Ceux-ci détiennent une fonction d'explication de l'intrigue ou présentent une partie des dialogues (*Ibid.* : 226). Une fois le cinéma devenu parlant, l'intertitre tend à disparaître. Cependant, cela ne sonne pas le glas de l'écrit sur le grand écran. L'exportation croissante des films suscitera de nombreuses interrogations sur la traduction des langues. Plusieurs étapes auront lieu. On parlera de versions multiples des films, de doublages aussi, qu'on retrouve d'ailleurs encore aujourd'hui. Mais, l'écrit, matérialisé par le sous-titrage, perdurera avec comme grande fonction de traduire les dialogues et se généralisera dans les années 1930 (*Ibid.*). Cette fonction de traduction sera par la suite rejointe par une fonction d'explication des formes sonore pour les sourds et malentendants. La traductologie et particulièrement des chercheurs spécialistes de la traduction audiovisuelle (Gambier, 2004 ; Dumas, 2014) définissent ces deux types de sous-titrage de la sorte :

- le sous-titrage interlinguistique correspond à un passage d'une langue parlée à une autre langue écrite. La fonction du sous-titre est donc la traduction ;
- le sous-titrage intralinguistique correspond à la retranscription de propos oraux à l'écrit, le tout dans une même langue. Initialement, cette fonction d'explication est utilisée pour rendre compréhensible l'information diffusée aux personnes sourdes et malentendantes.

Ces deux types de sous-titrage se retrouvent aussi dans d'autres médias comme la télévision. Une interview dans un journal télévisé ou un magazine peut être traduite par un sous-titre. Un accès au sous-titre pour aider à la compréhension aux sourds et malentendant est obligatoire en France depuis 2005 pour toutes les chaînes. Comme Valentin Ribas le remarque bien dans son mémoire (2020 : 46), les allocutions du président de la République, Emmanuel Macron, dans le cadre de la crise sanitaire liée à la Covid-19, sont aussi toutes sous-titrées, en direct par un vélotypiste, de manière intralinguistique.

De cette brève présentation de la place de l'écrit au contact de l'image animée, nous retenons que trois grands objectifs, trois grandes fonctions, sont rattachés aux intertitres puis aux sous-titres depuis les années 1900 :

- expliquer l'image, l'action en cours ;
- traduire un propos dans une langue étrangère ;
- aider la compréhension de publics spécifiques.

En nous focalisant sur le sous-titrage, nous comprenons qu'il n'existe qu'au regard d'une voix, d'une forme sonore à sous-titrer au cinéma ou à la télévision. C'est d'ailleurs ce que remarque Louise Dumas (2014 : 133) : « En effet, le sous-titre n'existe par définition qu'en présence de l'original. L'un et l'autre coexistent dans un même mouvement d'appréhension de l'œuvre ». Cependant, notre analyse de vidéos informatives en ligne montre que le sous-titre n'est pas toujours lié à une voix, n'a pas toujours « un original ». C'est pourquoi nous proposons un détour par la deuxième « série culturelle » remarquée pour comprendre les origines du sous-titrage dans la vidéo en ligne, celle de la place de la voix off dans les productions audiovisuelles.

2.2 Voix et voix off : quand la parole raconte l'image

L'écrit n'est pas le seul élément, dans l'histoire de l'audiovisuel, à pouvoir raconter l'image. Au cinéma, au temps du muet, des bonimenteurs offrent leur voix pour « suppléer les carences expressives de l'image » (Albera, 2004). Bien qu'ils usent de leur voix pour attirer un public ou légitimer une institution naissante, les bonimenteurs participent à la narration orale de ces « vues animées » (Lacasse, 2000). Selon François Albera, le bonimenteur peut être entrevu comme un parent de la voix off : « Sous un certain angle, on peut donc spontanément lui trouver des rapports d'homologie avec l'instance de la voix off qui aurait, en quelque sorte, intériorisé sa place en le faisant entrer dans l'espace filmique » (Albera, 2004 : 45). Cette voix off, toujours en tant que concept narratif, serait, elle, un emprunt du cinéma effectué à la radio (Chion, 1982). On la retrouvera sur le grand écran dès la fin des années 1930 dans *Le Roman d'un tricheur* de Sacha Guitry (1936) ou encore dans *Citizen Kane* d'Orson Welles (1941). Elle sera aussi entendue dans les salles de cinéma pour conter les actualités françaises. Puis, dès les débuts de la télévision, on l'utilisera dans les sujets de journaux télévisés ou de magazines. Depuis les années 1990, on la retrouve aussi régulièrement dans des jeux télévisés ou dans des programmes de la réalité¹⁰. Les productions où l'on entend une voix off pourraient aussi être complétées par le documentaire ou encore les jeux vidéo. Le point commun entre ces différents types de production reste que la voix off est celle qui fait voir (Alibert, 2008 : 80) sans apparaître à l'image. C'est celle qui donne l'information que souhaite diffuser l'émetteur sans la montrer de manière précise.

¹⁰ Télé-réalité, jeux réalité ou encore télécoaching.

2.3 Statuts du texte, de la voix et de la voix off écrite dans la vidéo informative en ligne

Les différents éléments soulevés dans la mise en lumière des héritages du texte et de la voix en tant qu'agents médiateurs de l'audiovisuel se retrouvent enfin dans la vidéo en ligne. Le sous-titrage, introduit dès le début de l'article, est un élément distinctif de ce type de vidéo. La voix off, elle, est assez rare. Cependant, nous allons montrer que c'est sur ses racines qu'un grand nombre de sous-titres semble s'être construit.

La figure 2, ci-dessous, présente les pourcentages de couvertures des différentes fonctions du sous-titrage relevées dans les vidéos en ligne de notre corpus.

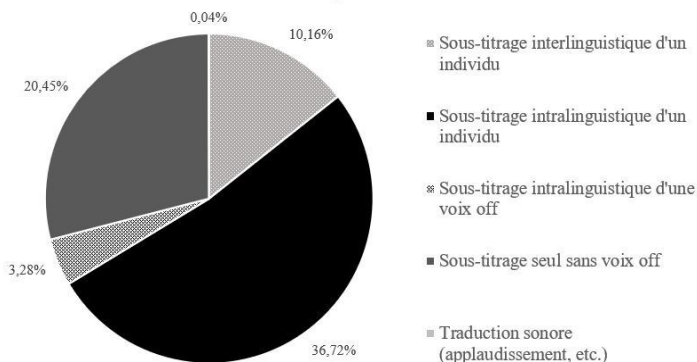


Figure 2. Pourcentages de couverture des différents types de sous-titrage remarqués dans les vidéos en ligne.

Le sous-titrage est utilisé pour doubler à l'écrit une même langue orale à 36.72 % (sous-titrage intralinguistique), ou encore pour traduire une langue à l'écrit à 10.16 % (sous-titrage interlinguistique). Plus rarement, à 3.28 %, il est utilisé en tant que sous-titrage d'une voix off (sous-titrage intralinguistique de la voix off). De manière anecdotique dans notre corpus, soit à 0.04 %, le sous-titrage représente une traduction sonore de ce qui se passe dans la vidéo, comme des applaudissements par exemple. Dans ces quatre premières catégories, le texte accompagne toujours un son, surtout la voix. Il traduit le contenu vocal des vidéos et aide à la compréhension des publics n'utilisant pas le son. Par conséquent, ces fonctions du sous-titrage sont bien héritées des pratiques mises en lumière dans le premier point de cette partie sur la place de l'écrit à l'image.

Cependant, reste une catégorie à décrire. Statistiquement, il s'agit même de la deuxième plus importante de notre corpus. Sur la figure 2, l'aplatissement qui représente 20.45 % de couverture des vidéos analysées correspond ainsi à l'utilisation du sous-titrage seul, sans voix off. Ce sous-titrage n'accompagne pas de voix, il la remplace. En reprenant les propos tenus dans le point précédent, nous pourrions dire que c'est le texte qui fait voir, qui fait comprendre le message diffusé. On assisterait alors à une certaine verbalisation du récit, mais par l'écrit. Selon nous, cette manière d'utiliser le sous-titrage seul est l'un des résultats majeurs de l'enquête. Il montre que certains médias en ligne, principalement Brut., Culture' et Loopsider, ont mobilisé des formes préexistantes, les ont modelées, pour les adapter aux pratiques des utilisateurs. Le format de Brut. repose beaucoup sur ce système. Une musique de librairie compose le fond sonore de la vidéo. Lorsqu'il n'y a pas d'interview, le sous-titrage seul narre l'information au rythme d'images provenant souvent de banques d'images en ligne, comme Getty Images. Culture', de son côté, est partenaire de Brut.. Il utilise son format pour mettre en valeur les archives audiovisuelles du service public français. Ce système de sous-titrage seul accompagne alors des gravures, des peintures ou encore des extraits de livres traitant d'un sujet culturel particulier. Enfin, pour Loopsider, le plus jeune des médias étudiés, on entrevoit un calque de la « manière de faire » de Brut.. Ce recours au sous-titrage seul en témoigne. Peu d'individus sont interviewés, contrairement aux vidéos de Konbini et Melty, et le contenu ainsi que l'information doivent être diffusés et compris par le texte.

Ce recours au « sous-titrage seul » trahit une volonté de s'économiser de la voix off. En nous intéressant brièvement à la pratique professionnelle du sous-titrage, nous pouvons en effet entrevoir certains avantages de cette « manière de faire ». L'activité de sous-titrage est une activité sous contrainte et très codifiée lorsqu'il s'agit de transcrire la voix au cinéma ou à la télévision :

« Le texte est découpé en unités qui deviendront autant de sous-titres (de 1 à 2 lignes, de 35-36 caractères en général, 31 pour Arte). [...] Ces futurs sous-titres tentent de respecter d'autres unités : les interventions d'un personnage, les unités de sens quand c'est possible, les pauses respiratoires et les contraintes propres aux films, comme les changements de plan par exemple. [...] À partir de ce découpage, on calcule le nombre de caractères autorisés pour un sous-titre [...] pour un caractère, mais il faut tenir également compte de la longueur totale du sous-titre, de la complexité du texte. Il faut aussi penser qu'un sous-titre de deux lignes ne peut rester à l'écran plus de 3.5 à 4 secondes ». (Petitjean, 2001 : 51).

Carine Petitjean, dans ces phrases, montre qu'il est impossible dans la pratique de sous-titrage de « tout dire » ou de « tout traduire ». De fait, le spectateur doit « puiser en plusieurs endroits les informations qui lui sont habituellement données en un même lieu et [...] se saisir du moindre indice » (*Ibid.*) pour comprendre l'intégralité du contenu. Cependant, pour le « sous-titrage seul » de la vidéo en ligne, sans voix à traduire ou sans voix off à doubler, ces contraintes sont moindres. Comme le texte défile, le rythme à son importance, de la même manière que le nombre de mots à l'écran. Néanmoins, l'objet n'est pas le contenu vocal de la vidéo, ses personnages ou ses images. Le contenu principal est bien le texte, le sous-titrage. L'image animée devient illustration du texte et l'utilisateur peut atteindre le sens de la vidéo en ligne par la lecture. La catégorie « sous-titrage seul » montre donc bien une autonomisation des « manières de faire » des médias en ligne.

3 La multiplication des inserts visuels dans la vidéo en ligne : le légendage et l'iconographie pour remplacer le son par « l'image »

L'utilisation généralisée du sous-titrage dans la vidéo en ligne est accompagnée d'une mobilisation croissante d'éléments iconographiques et textuels pour pallier la non-écoute du son de la part des utilisateurs. En ce qui concerne l'iconographie, prenons l'exemple du coassement d'une grenouille représenté par l'association d'une image de l'animal et d'une représentation graphique d'un signal sonore. Une voix initialement diffusée à la radio et réutilisée dans la vidéo pourra, elle, être accompagnée d'icônes représentant une onde radiophonique pour montrer l'origine de la source sonore visible à l'écran. Concernant les éléments textuels, l'indication des sources, les titres, la présentation des individus à l'écran, etc. sont des formes ajoutées à la production qui illustrent une pratique de « légendage ».

Cette mobilisation de formes textuelles ou iconographiques pour illustrer ou ajouter de l'information à l'image se rapproche de ce que Francesco Casetti et Roger Odin appelaient des « inserts » en 1990, lorsqu'ils analysaient l'évolution des formes télévisuelles. En d'autres termes, de nombreux éléments externes à la captation audiovisuelle sont ajoutés à l'écran. Pour la néo-télévision, « le royaume de l'insert » (Casetti et Odin, 1990), il pouvait s'agir d'inserts spatiaux soit des incrustations visant à légendiser des éléments du programme (numéro de téléphone à appeler, bandeau de présentation d'un individu, etc.). Il pouvait s'agir aussi d'inserts indépendants de l'émission d'accueil, comme le logo de la chaîne de diffusion ou encore des encarts annonçant les programmes à venir. Dans le cas des vidéos en ligne, on retrouve cette idée d'ajouter, d'insérer, des éléments et des informations qui ne sont ni de l'image filmée et montée ni du son. Dans cette troisième partie, nous interrogerons donc ces mobilisations de formes visuelles qui tendent à remplacer ou compléter les formes sonores en présentant tout d'abord l'omniprésence des inserts textuels dans la vidéo en ligne, puis en donnant des exemples d'inserts iconographiques.

3.1 L'omniprésence de « l'insert » textuel dans la vidéo en ligne

L'analyse de contenu menée pour cet article a révélé dix manières différentes de mobiliser le texte dans la vidéo que nous pouvons regrouper en deux catégories. Notons que ces mobilisations du texte diffèrent dans cette partie du sous-titrage d'ores et déjà étudié. La première catégorie comprend les éléments de « légendage » d'un document. On retrouve des formes textuelles présentant un individu ou une équipe de rédaction ou informant du titre de la vidéo. On retrouve aussi des inserts invitant à la redirection vers les réseaux socionumériques du média, etc. Cette pratique, développée initialement dans le documentaire ou les reportages télévisés, est nommée par les professionnels de la post production « la synthé » (Baldo De Brébisson, 2016). La deuxième catégorie d'insert textuel que nous proposons correspond au remplacement de formes traditionnellement orales. L'exemple de la question posée par un journaliste et sa substitution par une phrase apparaissant à l'écran illustre bien cette seconde catégorie. C'est d'ailleurs ce que l'on peut voir sur la figure 3 ci-dessous.

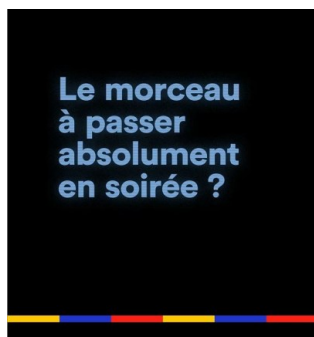


Figure 3. Capture d'écran d'une question remplacée par du texte dans une vidéo Konbini consacrée à Vege dream le 13/08/2019.

Bien que la période analysée soit très courte en raison de l'émergence récente de l'objet que nous étudions, nous remarquons une tendance à la multiplication du recours aux formes écrites, autres que le sous-titrage. Le tableau 3, ci-dessous, en témoigne. Il présente le nombre d'inserts textuels différents relevés dans l'ensemble des vidéos analysées au fil des ans et selon les médias de diffusion.

Tableau 3. Nombre d'inserts textuels différents (sur 10 relevés) utilisés par chaque média tous les ans

| | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|---|----------|------|----------|----------|-----------|-----------|
| Brut. (2016) | | | 6 | 3 | 3 | 6 |
| Konbini (2008) | 1 | | 5 | 2 | 5 | 8 |
| Melty (2008) | | | | | 0 | 1 |
| Culture' (2018) (et France Culture) | | | | 7 | 8 | 8 |
| Loopsider (2018) | | | | | 7 | 7 |
| Nombre d'inserts différents relevés par année | 1 | | 7 | 8 | 10 | 10 |

En 2018 et 2019, les dix types d'inserts textuels différents sont visibles dans notre corpus, ce qui correspond à une légère augmentation au regard des autres années étudiées. Pour aborder la multiplication de ces formes, l'exemple de Konbini est assez parlant. En 2019, les vidéos du média mobilisent beaucoup le texte en remplacement de questions posées, en compléments de l'image, etc. Cela devient même une sorte de « marque de fabrique » dans ses séries de vidéos comme le *Fast and Curious*, le *Track ID* ou encore le *Fast Life*. Les questions posées sont toutes visibles à l'écran, mais non audibles. Ce recours au texte en évolution se recoupe d'ailleurs avec un résultat évoqué dans la partie précédente : avant 2017, Konbini n'utilisait même pas de sous-titrage dans ces vidéos. Les cas des médias Culture' et Loopsider, les plus récents du corpus, sont eux aussi intéressants. Les deux utilisent très rapidement, dans leurs récentes histoires éditoriales, les inserts textuels. Ceux-ci font principalement partie du légendage.

3.2 Trois types d'inserts iconographiques

Pour accompagner les inserts textuels, des icônes vont aussi être mobilisées dans la vidéo en ligne. De la même manière, elles vont jouer un rôle de médiation entre la vidéo et l'internaute, et ce, en nous renseignant sur les sources sonores mobilisées au sein de la vidéo (la langue parlée, le bruit d'un réveil, une radio, etc.), mais aussi en décrivant une forme sonore particulièrement (le cri d'un animal, d'une machine, etc.).

3.2.1 Iconographies sonores mobilisées dans les vidéos en ligne analysées

Dans notre corpus, ces inserts sont peu nombreux. Nous en avons relevé uniquement dans six vidéos sur cinquante-six. Ils servent à décrire des ondes radio, à décrire la langue parlée, à signaler un chant, ou encore, simplement, à signaler à l'individu qu'il doit mettre le son pour comprendre l'information. En sortant du corpus, nous avons relevé une manière intéressante d'utiliser des inserts iconographiques. Il s'agit de l'exemple présenté sur la figure 4 ci-dessous. Cette illustration provient d'une vidéo Brut. focalisée sur le chant de deux oiseaux : les ménures superbes et les ménures d'Albert. La particularité de leur chant est qu'il peut ressembler à des bruits de leur environnement ou encore à des cris d'animaux d'autres espèces. Dans l'extrait illustré par la figure 4, le chant de l'oiseau ressemble au cri d'un singe. Pour agrémenter l'image, deux iconographies sonores ont été insérées : une enceinte en haut à gauche de l'image qui sert à inciter l'utilisateur à enclencher le son et trois images de singe, la gueule ouverte « imitant » le chant de l'oiseau.



Figure 4. Capture d'écran représentant le remplacement des formes sonores par des icônes dans une vidéo Brut. diffusée le 08/01/2020.

Même sans pouvoir écouter cette portion de la vidéo, on peut imaginer le chant de cet oiseau grâce à la signification de l'icône.

3.2.2 Matérialiser la provenance des sources radiophoniques

Un autre exemple cité précédemment est la matérialisation des ondes radio. À trois reprises, dans notre corpus, nous avons remarqué que pour veiller à ce que l'utilisateur ne confonde pas le sous-titre remplaçant une voix off et le texte traduisant une archive sonore ajoutée, le média mobilisait l'image d'une onde radio, comme on peut le voir sur la figure 5 ci-dessous.



Figure 5. Captures d'écran représentant la matérialisation des sources radiophoniques sur des vidéos Brut. et Culture' respectivement diffusée les 16/11/2016 et 04/04/2018.

Ainsi, si la mise en forme du texte sous-titré est identique à celle du sous-titre de l'archive, c'est l'onde qui va permettre la différenciation. L'onde renseignera aussi l'utilisateur sur la provenance de l'archive.

3.2.3 Matérialiser les origines géographiques des discours

Nous proposons un dernier exemple sur les inserts iconographiques. Il est relatif à la matérialisation des origines géographiques des discours. La figure 6, ci-dessous, est composée de quatre captures d'écran provenant d'une vidéo Brut. sur des manifestations mondiales contre la violence faite aux femmes. Au même moment, toutes les femmes présentes à l'écran chantaient les mêmes paroles : « Le violeur, c'est toi », mais dans des pays et des langues différents. Les deux langues chantées dans la vidéo sont l'espagnol et le français. Cependant, les scènes se passent dans quatre pays différents. Le sous-titrage en français est le même pour les quatre pays représentés, la France, la Belgique, le Chili et l'Espagne. Pour les différencier et renseigner l'internaute de la localisation géographique de chaque manifestation, le drapeau du pays en question est affiché dans l'angle gauche de la vidéo.

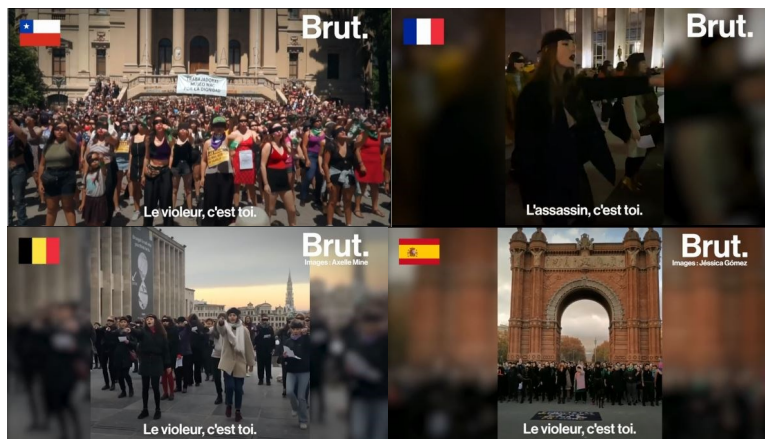


Figure 6. Captures d'écran réalisées sur une vidéo Brut. diffusée le 03/12/2019 représentant le recours aux drapeaux pour matérialiser les origines géographiques des discours.

Cette partie sur les inserts textuels et iconographiques nous informe sur leur présence multipliée dans la vidéo en ligne depuis 2014. Formes visuelles, elles sollicitent la vue de l'utilisateur en accompagnement du sous-titrage décrit dans la partie 2. Ces inserts résultent d'une part d'une facilitation technique engagée par l'informatisation de la production audiovisuelle, procès établi depuis les années 1990. Comme pour le sous-titrage spécial (Baldo De Brébisson, 2016), cette mise en forme du texte et de l'image pouvant apparaître à n'importe quel moment dans la vidéo participe de la recherche de captation des publics. Par conséquent, et d'autre part, ces inserts offrent à la vidéo « la plus grande dynamique visuelle et la plus grande force attractive » (Casetti et Odin, 1990 : 19), comme le remarquaient Casetti et Odin pour la télévision, il y a une trentaine d'années. Cette recherche de captation visuelle prend tout son sens dans le flux des réseaux socionumériques où le nombre de posts vidéos croît depuis quelques années.

4 Le format « à la Brut. » : standardisation, industrialisation et contamination

Les résultats présentés jusqu'ici ont montré une tendance à la mobilisation croissante du texte dans la vidéo en ligne. Le sous-titrage s'est généralisé à tous les médias étudiés de 2014 à 2020. Le texte et l'icône ont eux aussi été intégrés de manière croissante aux différents contenus durant cette période. La forme de la vidéo informative en ligne a donc eu tendance à se standardiser. Bien qu'il ne s'agisse pas du seul facteur explicatif, nous le verrons, nous soumettons l'hypothèse que le format de Brut. a, en quelque sorte, contaminé les productions de ses concurrents directs (Konbini, Melty, Loopsider, etc.), des productions publicitaires, mais aussi celles d'autres médias, comme la télévision. Ce format, souvent nommé par la presse ou par des entreprises de post production « format à la Brut. », semble au prisme d'un processus d'industrialisation menant à une dépréciation des formes sonores. Cette dernière partie tâchera donc d'esquisser les enjeux socio-économiques circonvénant le format de vidéo analysé dans cet article.

4.1 Généralisation d'un format au regard d'une offre de vidéo en ligne grandissante

Depuis 2019, les cinq médias analysés dans cette étude diffusent des vidéos quotidiennement. Avant, nous avons remarqué que ce n'était pas le cas pour Melty et Konbini par exemple, pourtant acteurs du web depuis 2008. Si l'on prend l'exemple de Konbini, nous remarquons qu'avant 2018, les contenus vidéos diffusés sont rares et qu'ils ne répondent à aucune logique éditoriale. Le succès rapide de Brut. peut être une piste d'explication de cette généralisation de la vidéo, et surtout, de son format. En effet, après son lancement, Culture' et Loopsider sont entrés sur le marché en 2018 et ont commencé à mobiliser des codes éditoriaux très proches et particulièrement centrés sur l'utilisation du texte. Un autre argument d'ores et déjà soumis en introduction de cet article et qui expliquerait cette généralisation de la vidéo en ligne pourrait être sa consommation par les internautes. Cet élément va de pair avec le contexte socio-économique et industriel de la diffusion de la vidéo en ligne à partir des années 2000.

En 2005 naît la plateforme vidéo YouTube. Racheté par Google dès 2006, ce site web d'hébergement vidéo enregistre un nombre de vues croissant sur ses contenus depuis sa création. En 2013, l'action de Google passe à 1000 dollars, une première historique¹¹. La société avance que YouTube ainsi que sa consommation de plus en plus régulière sont les plus grands facteurs de réussite de l'entreprise à ce moment-là. Cette période témoigne alors d'un intérêt grandissant pour le format vidéo sur le web, comme en attestent les stratégies de diffusion développées par les réseaux sociaux numériques. En 2014, Facebook intègre à son fonctionnement l'*AutoPlay*. Il s'agit de la mise en lecture automatique et silencieuse des vidéos lorsque l'utilisateur passe sa souris dessus. Initialement conçu pour accueillir des vidéos publicitaires (*trailer* de film, campagne vidéo de grandes marques, etc.), l'outil a très vite permis la lecture de vidéos produites par les utilisateurs, puis les médias.¹² En 2015, l'*AutoPlay* est aussi lancé par Twitter. Ces éléments, davantage que de montrer que la vidéo est devenue un type de production incontournable du web et des réseaux sociaux numériques, montrent bien qu'une attention est portée à la manière dont elles sont consommées par les utilisateurs. La suppression, bien que momentanée du son, incitera, par la suite, les acteurs de la production vidéo au sens large (annonceurs, journalistes, voire amateurs), à développer un système de sous-titrage et à mobiliser le texte, comme nous l'avons vu tout au long de cet article.

4.2 Le format « à la Brut » un format journalistique et publicitaire

La généralisation de la vidéo, mais aussi d'un format spécifique, semble donc résulter de pratiques et d'enjeux socio-économiques perçus aussi bien en analysant la production (médias en ligne), la diffusion (plateformes et réseaux sociaux numériques) et la consommation des vidéos (internautes). Pour en revenir au « format à la Brut. », et donc à celui qui mobilise le plus le texte, il semble avoir dépassé le cadre journalistique pour migrer, en partie, dans un espace publicitaire.

Tous les acteurs privés que nous avons étudiés proposent du *Brand Content*. Il s'agit de produire des contenus similaires à leur production, mais réalisés pour des marques qui les sponsorisent. En d'autres termes, ces médias vendent leur format, leur savoir-faire, à des marques. Ces dernières vont sponsoriser ces contenus tout en s'affichant dans un format « en vogue » au moment de la diffusion. Ce type de pratique ressemble à ce que Valérie Patrin-Leclère appelle la publicitarisation du média soit « l'adaptation de la forme, des contenus, ainsi que d'un ensemble de pratiques professionnelles médiatiques à la nécessité d'accueillir la publicité » (Patrin-Leclère 2010). Cette pratique va permettre de « réduire la rupture sémiotique entre le contenu éditorial et le contenu publicitaire » (Berthelot-Guiet, *et al.*, 2013). En d'autres termes, la publicité et le format du média, en s'unissant, tendraient à être reconnaissable par un individu comme étant propriété éditoriale du média. C'est ce que l'on peut constater dans les deux exemples figurant sur la figure 7.

¹¹ Rauline, Nicolas, le 21/10/2013, « Google s'envole en Bourse, porté par les performances de YouTube et du mobile », *Les Échos*.

¹² Gattaz, Pierre, le 08/12/2014, « Facebook fait de l'ombre à YouTube sur le marché de la vidéo », *Les Échos*.



Figure 7. Captures d'écran réalisées sur Facebook de deux partenariats rémunérés réalisés entre Brut. et Paypal en juin 2020 et Loopsider et Arkéa en septembre 2020.

Brut., Loopsider et les autres, par leurs succès respectifs, sont donc sollicités pour diffuser des contenus publicitaires. Cette production accroît la circulation du format vidéo analysé et donc de ses particularismes sonores et textuels. Ainsi, elle contamine, en quelque sorte, le support publicitaire.

Dépassant la publicité, ce format « à la Brut. » se retrouve aussi depuis quelques années sur les chaînes de télévision du service public français et particulièrement lors de journaux télévisés. M. Chion, dès les années 1990, décrivait la télévision comme étant une radio illustrée. Par conséquent, le téléspectateur, même sans regarder l'écran, pouvait comprendre un programme :

« Dire que la télévision – sans intention péjorative, d'ailleurs – qu'elle est une radio illustrée [...], c'est rappeler que le son, principalement le son de la parole y est toujours premier, qu'il n'est jamais hors-champ et qu'il est toujours là, à sa place, n'ayant pas besoin de l'image pour se localiser. » (Chion, 2013 : 135).

Avec l'investissement de courts reportages formatés comme une vidéo Brut., c'est-à-dire sans paroles et avec une information diffusée à partir d'un sous-titrage, nous voyons bien que la place du son n'est plus première. Sans regarder l'écran, le téléspectateur ne peut plus consommer l'information proposée.

4.3 Le format « à la Brut. », un argument de vente dans la création audiovisuelle

Un dernier point concernant la contamination de ce format « à la Brut » peut être décrit pour comprendre toute l'importance qu'il a obtenu ces dernières années dans la production audiovisuelle. Nous avons remarqué que des sociétés de post production comme EasyMovie fournissent des services de création de vidéos « à la Brut. » L'expression « à la Brut. » est convoquée par les acteurs de cette même post production. De la même manière, Animoto ou Wave.video proposent des services similaires, mais directement par le prisme d'une plateforme web. Ainsi comme remarqué par Yves Jeanneret et Emmanuel Souchier (2005), cités au début de cet article, l'industrialisation de la pratique de vidéo en ligne, et donc d'un accompagnement systématique d'un sous-titrage, témoignerait de l'élaboration de « matrices de textes de nature à créer du profit, non par la dissémination des produits culturels, mais par la dissémination des moyens de production culturelle » (Jeanneret et Souchier, 2005 : 14). En d'autres mots, le fait que le format vidéo circule, soit vendu ou acheté, n'est pas le seul argument à retenir en ce qui concerne la création de profit. Le fait que ses moyens de production soient disséminés et multipliés engage alors de véritables questionnements sur la place de la valeur de ces produits et sur leur industrialisation.

5 Conclusion

L'analyse des contenus menée montre que la vidéo informative en ligne semble de plus en plus être une production visuelle, voire textuelle. Les formes sonores, qu'elles soient vocales ou musicales, tendent à être complétées

systématiquement par du texte ou des icônes ajoutées à la production audiovisuelle. Complétées, voire remplacées, comme c'est le cas avec la voix des journalistes ou encore la mise en évidence de cette forme de sous-titrage sans voix. Fondée sur des héritages intermédiatiques venant aussi bien du cinéma, de la radio ou encore de la télévision, cette tendance au sous-titrage semble circuler de manière accélérée depuis quelques années et se généraliser sur le web. La prise en considération de la non-écoute du son par les utilisateurs lors de la consommation de vidéo d'information mériterait alors d'être approfondie pour rendre complète l'analyse intermédiaire et comprendre les enjeux liés à la prise en considération des publics dans la conception même des vidéos.

Les stratégies économiques des acteurs du web, et particulièrement de Google et Facebook, nous informent aussi sur leur implication dans l'évolution des formes audiovisuelles numériques, de la même manière que les pratiques des utilisateurs. L'ensemble de ces éléments montre qu'un format, proche de celui du média Brut, semble se distinguer et surtout s'industrialiser. Par conséquent, pour revenir sur les propos de Yves Jeanneret et d'Emmanuel Souchier lorsqu'ils parlent d'énonciation éditoriale, nous ne sommes pas uniquement face à une dissémination rapide des biens culturels eux-mêmes depuis l'investissement de l'informatique dans nos sociétés, mais bien face à une dissémination des moyens de production culturelle. Ce qui, dans le cas de cette étude, a une implication directe sur le son et son (dés)intérêt dans la création audiovisuelle.

Notons tout de même que cette dépréciation sonore concerne particulièrement la vidéo informative en ligne. Sur le web, les contenus audiovisuels restent très sonores. En 2018, YouTube, où l'on trouve de nombreux vidéo-clips, était utilisé à 98 % dans la pratique de *streaming* musical par exemple. De plus, le son sur les réseaux socionumériques devient un véritable moteur de création alliant justement « image et musique ». L'exemple de TikTok et de tous ces défis de danse, ces récupérations de musiques pour rejouer des scènes de la vie quotidienne, etc. montre bien que le son a encore un grand rôle à jouer sur nos écrans.

Références

- Albera François. (2004). Du bonimenteur à la voix off. « Le silence d'or », de René Clair. *Vertigo*, n°26, p. 45-49.
- Alibert Jean-Louis. (2008). *Le son de l'image*. Grenoble : PUG.
- Altman Rick. (1999). De l'intermédialité au multimédia : cinéma, médias avènement du son. *Cinéma*, vol. 10, n°1, p. 37-53.
- Baldo De Brébisson Sabrina. (2016). Formes, sens et pratiques du sous-titrage spécial. *Signata*, n°7, p. 255-284. [<http://journals.openedition.org/signata/1229>].
- Bardin Laurence. (1977). *L'analyse de contenu*. Paris : PUF.
- Berthelot-Guiet Karine, Marti de Montely Caroline, Patrin-Leclère Valérie. (2013). Entre dépublicitarisation et hyperpublicitarisation, une théorie des métaphores du publicitaire. *Semen*, n°36. [<https://journals.openedition.org/semes/9645>].
- Casetti Francesco, Odin Roger. (1990). De la paléo- à la néo-télévision. Approche sémio-pragmatique. *Communication*, n°51, p.9-26.
- Chevalier Jean-Yves. (2018). Le numérique à contre-sens. *Médium*, n°55, p.107-126.
- Chion Michel. (1982). *La voix au cinéma*. Paris : Cahiers du cinéma.
- Chion Michel. (1995). *La musique au cinéma*. Paris : Fayard.
- Chion Michel. (2013). *L'audio-vision, son et image au cinéma*. Paris : Armand Colin. (3^e édition).
- Cornu Jean-François. (2014). *Le doublage et le sous-titrage. Histoire et esthétique*. Rennes : Presses universitaires de Rennes.
- Coulomb-Gully Marlène. (2002). Propositions pour une méthode d'analyse du discours télévisuel. *Mots. Les langages du politique*, n°70, p. 103-113. [<https://journals.openedition.org/mots/9683>].
- Dumas Louise. (2014). Le sous-titrage : une pratique à la marge de la traduction. *ELIS- Revue des jeunes chercheurs en linguistique de Paris-Sorbonne*, vol. 2, p. 129-144. [halshs-01090467].
- Gambier Yves. (2004). La traduction audiovisuelle : un genre en expansion. *Meta*, vol.49, n°1, p. 1-11. [<https://id.erudit.org/iderudit/009015ar>].
- Gaudreault André. (2008). *Cinéma et Attraction. Pour une nouvelle histoire du cinématographe*. Paris : CNRS Éditions.
- Gaudreault André, Marion Philippe. (2000). Un média naît toujours deux fois. *Sociétés & Représentations*, n°9, p. 21-36.
- Glaser Barney, Strauss Anselm. (1967). *The discovery of Grounded Theory : Strategies for Qualitative Research*. Chicago : Aldine.
- Jeanneret Yves, Souchier Emmanuel. (2005). L'énonciation éditoriale dans les écrits d'écran. *Communication & Langages*, n°145, p. 3-15.
- Jost François. (2005). *Comprendre la télévision*. Paris : Armand Colin.

- Lacasse Germain. (2000). *Le bonimenteur de vues animées. Le cinéma « muet » entre tradition et modernité*. Québec/Paris : Nota Bene / Méridiens.
- Lejeune Christophe. (2010). Montrer, calculer, explorer, analyser. Ce que l'informatique fait (faire) à l'analyse qualitative. *Recherches Qualitatives*, Hors-série 9, p.15-32. [<http://hdl.handle.net/2268/25352>].
- Müller, E. Jürgen. (2000). L'intermédialité, une nouvelle approche interdisciplinaire : perspectives théoriques et pratiques à l'exemple de la vision de la télévision. *Cinémas*, vol. 10, N°2-3, p. 105-134. [<http://id.erudit.org/iderudit/024818ar>].
- Müller, E. Jürgen. (2006). Vers l'intermédialité. Histoires, positions et options d'un axe de pertinence. *MédiaMorphoses*, n°16, p.99-110. [<http://documents.irevues.inist.fr/handle/2042/23499>].
- Patrin-Leclère Valérie. (2010). Les traces d'un fourvoiement journalistique : « Kibboutz spirit », Libération Next, juin 2009. *Communication & Langages*, n° 164, p. 117-126.
- Petitjean Carine. (2001) Le sous-titrage, un mode de déplacements In Irmtraud Behr (dir.) *Télévision et Internet : le parlé et l'écrit*. Paris : Presses Sorbonne Nouvelle. [<http://books.openedition.org/psn/6443>].
- Ribas Valentin (2020). *La place du texte dans les vidéos d'information sur les réseaux sociaux, du sous-titrage au légendage*. Mémoire en sciences de l'information et de la communication. Sorbonne Nouvelle Université.