HAL
open science

Questions méthodologiques autour de l'étude de deux plateformes interactives multimodales : de la communauté de contenu à l'application intimiste

C. Combe Celik

▶ **To cite this version:**

C. Combe Celik. Questions méthodologiques autour de l'étude de deux plateformes interactives multimodales : de la communauté de contenu à l'application intimiste. Línguas e Instrumentos Linguísticos (Brésil), 2016, 37. hal-01615562

HAL Id: hal-01615562

https://hal.science/hal-01615562v1

Submitted on 12 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

QUESTIONS MÉTHODOLOGIQUES AUTOUR DE L'ÉTUDE DE DEUX PLATEFORMES INTERACTIVES MULTIMODALES : DE LA COMMUNAUTÉ DE CONTENU À L'APPLICATION INTIMISTE

Christelle Combe
Aix Marseille Université, CNRS

Resumo: *Este capítulo trata das dificuldades metodológicas que se pode confrontar um pesquisador quando ele estuda uma plataforma interativa multimodal. Dois exemplos de plataformas completamente diferentes foram analisados: uma comunidade de conteúdo (Youtube) e um aplicativo intimista (Couple). Foi posta a questão do posicionamento do pesquisador dentro destes ambientes digitais ecológicos de observação distanciada e como ele interage no interior deles. A constituição do corpus a partir de dados digitais massivos e instáveis em vista de uma análise micro qualitativa também foi abordada.*

Résumé: *Ce chapitre traite des difficultés méthodologiques auxquelles peut se trouver confronté un chercheur lorsqu'il étudie une plateforme interactive multimodale. Deux exemples de plateformes complètement différentes sont envisagées: une communauté de contenu (YouTube) et une application intimiste (Couple). La question du positionnement du chercheur au sein de ces environnements numériques écologiques de l'observation distanciée au vécu de l'intérieur est posée. La constitution du corpus à partir de données numériques massives et labiles en vue d'une analyse micro qualitative est également abordée.*

Abstract: *This chapter discusses the methodological difficulties a researcher may encounter when studying an interactive multimodal platform. Two distinct platforms are considered: an open content community (YouTube) and an intimate application (Couple). Two questions are explored. On the one hand, a researcher's position as an*

observer in an ecological digital environment, and on the other, as a participant directly taking part in the interaction, while attempting to distance herself as a researcher. Another aspect taken into account is the selection of the corpus, for micro qualitative analysis purposes, from massive sets of digital data which is constantly enriched by contributors.

Depuis la fin du vingtième siècle, dès l'apparition du minitel en France puis l'avènement d'Internet, le langage d'Internet et la communication médiée par ordinateur sont étudiés par les chercheurs en linguistique (ANIS, 1998; CRYSTAL, 2001; HERRING, 2004[1]). La généralisation progressive à l'échelle mondiale et sociale d'Internet, la multiplication des écrans au sein de nos vies quotidiennes contribuent à susciter des recherches qui se diversifient. Les paradigmes évoluent, se spécifient : il est désormais question d'interactions multimodales par écran (DEVELOTTE, 2011) et d'analyse du discours numérique (PAVEAU, 2015), émergence d'un nouveau champ qui prend en compte la spécificité des discours natifs en ligne et qui élabore peu à peu ses outils méthodologiques. En effet, les avancées technologiques contribuent à rendre ce champ mouvant et déstabilisant pour le chercheur qui doit sans cesse adapter sa démarche méthodologique. Dans cet article, j'ai donc choisi de m'intéresser aux questions méthodologiques portant sur l'étude de deux plateformes interactives multimodales tout à fait différentes : la communauté de contenu et l'application intimiste. Après avoir rappelé brièvement le cadre théorique de mes recherches, je présenterai les deux environnements numériques, objets de recherche, avant d'évoquer les principales questions méthodologiques auxquelles je me suis trouvée confrontée: le positionnement du chercheur, la question de la collecte des données et de la constitution du corpus, les questions éthiques et juridiques seront également évoquées.

1. Le cadre théorique

Comme Crystal (2011) le rappelle, les linguistes ont pour vocation d'étudier le langage où qu'il se trouve et Internet à ce titre est un terrain privilégié. Ainsi, depuis plusieurs années, la communication médiée par des outils technologiques est un champ de recherche qui:

évolue rapidement porté par les avancées technologiques. Je rappellerai ici les principales évolutions de ce courant.

1.1. De la communication médiée par ordinateur au discours numérique et interactions multimodales

En France, c'est Jacques Anis, précurseur du domaine, qui s'est intéressé dès les années 90, tout d'abord à l'écrit sur Minitel, CD-Rom, mais aussi sur messagerie électronique et Internet. Aux Etats-unis, c'est Herring (2004) qui a développé, quant à elle, une méthodologie adaptée à la « *Computer Mediated Communication* », tandis que Crystal s'intéressait au « *Netspeak* » (2001). Dans ces travaux précurseurs et fondateurs du domaine, c'est principalement vers le discours écrit que se sont orientées les recherches. Mais depuis plusieurs années Develotte s'intéresse aux interactions multimodales tandis que Paveau étudie les discours natifs en ligne.

La première, s'appuyant sur la linguistique interactive (Develotte, 2012 : 512) étudie les conversations en ligne (Develotte, Kern et Lamy, 2011) et ce qu'elle désigne désormais du nom d'interactions multimodales par écran[2]. Quant à Paveau, elle élabore un cadre original, postdualiste, consacré aux « discours natifs en ligne » qu'elle définit comme :

> Tout discours produit en ligne, c'est-à-dire dans un écosystème numérique connecté, tout particulièrement sur le web, mais également dans l'ensemble des internets (2015 : 3).

Elle s'intéresse ainsi à la nature composite des productions technologagières qu'elle s'applique à définir dans un dictionnaire d'analyse du discours numérique[3]. C'est donc dans ce champ en pleine émergence que j'ancre mes travaux de recherche en m'intéressant ici plus spécifiquement aux plateformes interactives multimodales.

1.2. Les plateformes interactives multimodales

Ce sont, en effet, aujourd'hui ces plateformes qui sont en plein développement. Elles permettent à l'utilisateur de communiquer dans des environnements multimodaux par des canaux multiples et posent, comme le souligne Herring, un certain nombre de questions théoriques et pratiques sur la manière de communiquer :

> To what extent does the choice of text, audio, video, and/or images affect the nature of users' communication? […] What impressions do messages in each mode make on their recipients? Does communication in one mode influence communication in other modes? (CRYSTAL, 2015, p.2).

Amenées à se développer, à se multiplier, des recherches dans cette direction sont nécessaires, souligne cette auteure (HERRING, 2015) y compris de manière empirique et dans le cadre de l'étude d'une plateforme particulière. C'est ce que j'ai modestement essayé de faire et que je relate dans ce chapitre. Que le lecteur ne s'attende donc pas à trouver un mode d'emploi, mais plutôt le récit d'une manière de faire avec toutes les interrogations qui l'émaillent. Les différentes expériences pourront donner lieu à quelques modestes conseils.

2. Les deux objets: communauté de partage et application intimiste

Les deux objets sur lesquels je vais m'appuyer pour expliquer les questions méthodologiques auxquelles je me suis confrontée sont deux plateformes interactives multimodales tout à fait distinctes. La première est la bien connue plateforme *YouTube* et la seconde est une application intimiste moins divulguée, *Couple*.

2.1. *Youtube*: une communauté de partage

YouTube se présente comme un site web d'hébergement de vidéos. Crée en 2005, il a été racheté depuis par *Google*. Sur cette plateforme, les internautes peuvent non seulement visionner des vidéos mais également, après avoir ouvert gratuitement un compte, déposer, partager, évaluer et commenter des vidéos de toutes sortes. Depuis l'avènement du web social et dans la large catégorie des réseaux socionumériques, *YouTube* est défini comme une communauté de partage (KAPLAN; HAENLEIN, 2010). Avec le développement technologique des objets connectés (ordinateur, tablette, téléphone) et la facilité désormais à produire du contenu vidéo, un genre particulier de vidéo s'est développé : le vlogue (COMBE, 2014). Dans un vlogue, un internaute s'adresse via sa chaîne régulièrement aux autres dans des vidéos qu'il poste sur une plateforme multimodale interactive comme *YouTube*. Les internautes peuvent "aimer" ou "ne pas aimer"

ses vidéos en cliquant sur un pouce levé ou baissé, commenter ses vidéos sous formes de messages écrits, voire s'abonner à sa chaîne. Il existe différents sous-genres de vlogue (commentaires de jeu vidéo, saynètes humoristiques, conseils esthétiques, journaux adolescents, etc.), dans cet article, je m'appuierai sur le sous-genre du vlogue "pratique d'une langue en ligne" (COMBE, 2014).

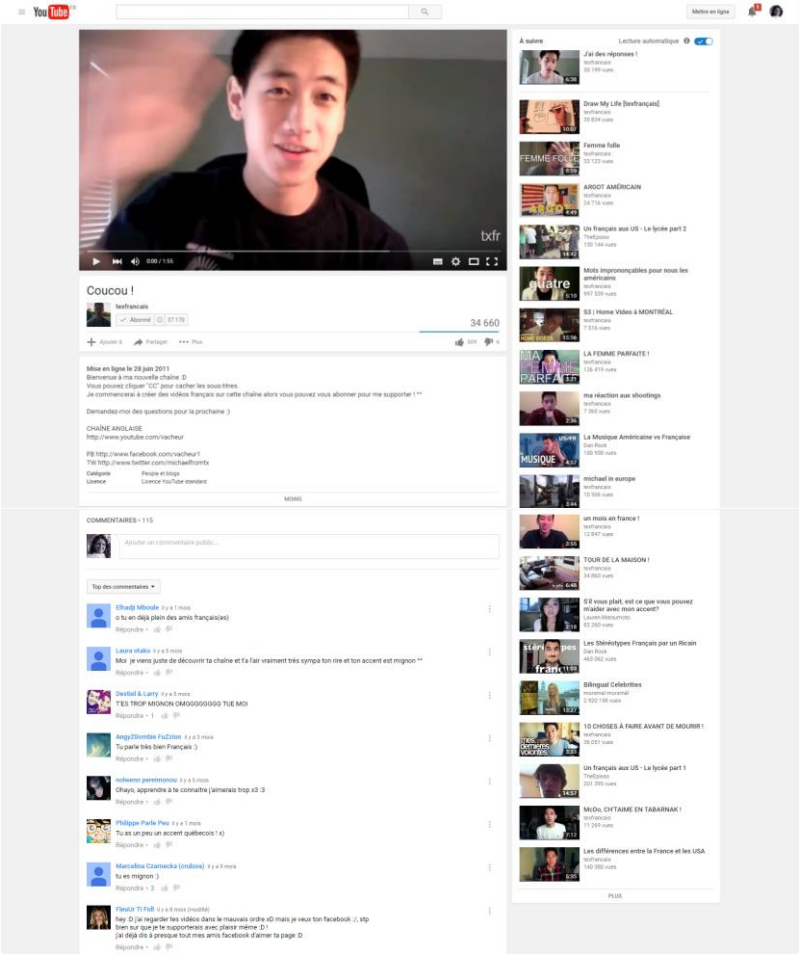

Figure 1: Un vlogue "Pratique d'une langue en ligne" extrait de la chaîne texfrancais du Youtubeur Michael.

2.2. *Couple*: une application intimiste

Couple est une application que je désigne comme "intimiste" puisqu'elle est destinée à deux interlocuteurs uniques. Cette application dont le sous-titre est "*The App for Two*" existe depuis 2012 et a été téléchargée plus d'un million de fois[4]. Dans un espace exclusivement destiné à deux personnes qui se sont identifiées et unis au sein de l'application ("*paired*"), elle offre la possibilité de s'exprimer par texte, vidéo, message audio, photos ou de réaliser un appel téléphonique. Elle possède un calendrier, un espace de stockage de photos et vidéos, une liste de choses à faire et des téléchargements et achats de stickers. Elle permet d'envoyer un "*thinking of you*", de dessiner seul ou ensemble, de se géolocaliser, de voir la musique que le partenaire est en train d'écouter et même de vivre un baiser virtuel. L'application est l'espace premier et le plus abouti pour communiquer, mais on peut également accéder à son compte et échanger *via* un site web[5]. Les développeurs communiquent aussi avec leurs utilisateurs au sein d'un blog[6], d'une page *Facebook*[7] et d'un compte *Twitter*[8].

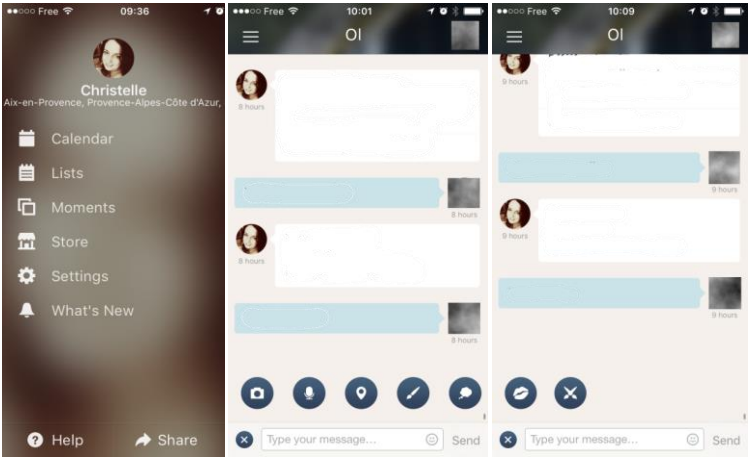

Figure 2: L'application Couple et ses fonctionnalités (copies d'écran réalisées à partir d'un Iphone)

3. Le positionnement du chercheur

Comme le souligne Paveau (2015, p.7), “pour rendre compte des fonctionnements langagiers des univers discursifs numériques natifs, il faut que le chercheur ait lui-même des pratiques du discours web natif”. Il est donc important que le chercheur ait une expérience personnelle de ces environnements, qu’il en soit familier, afin notamment d’en proposer une description savante valide. Le chercheur sera donc à priori impliqué dans son objet. C’est ce que j’ai fait pour ces deux environnements particuliers.

3.1. De l’observation distanciée

Dans le cas de la plateforme *YouTube*, le chercheur pourra accéder aux données publiques sans nécessairement être inscrit sur cette plateforme toutefois s’il veut avoir accès aux autres espaces numériques du vlogueur, il devra être présent sur les différents réseaux (page *Facebook* par exemple). En revanche, il pourra conserver une position tout à fait neutre et extérieure vis-à-vis de celui qui a ouvert une chaîne publique. S’il veut contacter le vlogueur pour un entretien ou un questionnaire, il pourra le faire via *YouTube* ou les autres réseaux sociaux mais il n’est pas certain que le vlogueur répondra. J’ai ainsi cherché à contacter Michael à plusieurs reprises via sa page *Facebook* et son compte *Twitter* sans succès. Le YouTubeur apparaît de fait davantage comme un personnage public inaccessible, comme peut en attester en quelque sorte sa page (et non son compte) *Facebook*, délibérément inscrite dans la catégorie “personnage fictif”.

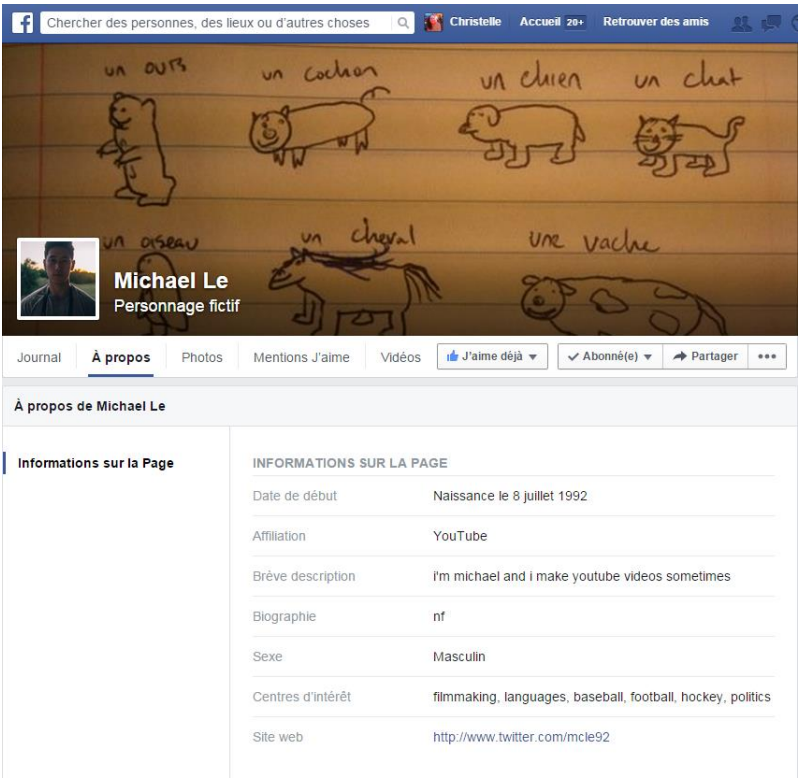

Figure 3 : Page Facebook du Youtubeur Michael

Sur le plan juridique, les données appartiennent à *YouTube* (et par extension à *Google*), toutefois, le chercheur peut considérer que la chaîne de Michael étant publique, elle est librement exploitable dans un contexte de recherche en sciences humaines et dans le respect des différents participants. Il peut ainsi prendre la plateforme *YouTube* pour terrain et étudier les discours publiquement émis *in situ*, comme le ferait un anthropologue ou un sociolinguiste. Sur le plan éthique, le chercheur peut anonymiser les données en totalité ou en partie, toutefois l'anonymisation ne sera pas forcément de règle, comme le souligne Côté:

> En dehors des situations où la sensibilité éthique du chercheur le conduit à privilégier une approche de publication anonyme,

> les messages publics sur la Toile et pour lesquels les auteurs n'ont pas d'attentes en matière de vie privée doivent être cités en respectant la Loi sur le droit d'auteur et les méthodes normalisées de citations pour les documents électroniques (CÔTÉ, 2012, p.8).

Un vlogue étant constitué de données vidéos incluses dans une entité commerciale, *YouTube*, qui est une compagnie à responsabilité limitée propriété de *Google*, la question de l'anonymisation des images se pose. Ainsi, comme le rappelle également Côté:

> Quant aux impressions d'écran incluses dans une publication, des enjeux légaux sont à prévoir pour les logos commerciaux et les images qui s'y retrouvent. Soit le chercheur obtiendra le consentement des compagnies ou individus auxquels ces logos et images appartiennent, soit il brouillera les images ou les personnes qui y figurent (McKee; Porter, 2009a, *apud* CÔTÉ, 2012, p.8).

Dans le cas de données vidéos, l'anonymisation peut non seulement s'avérer délicate à réaliser mais surtout paralysante si le chercheur souhaite étudier la mimogestualité (regards, mimiques, gestes, etc.). Avec le développement des éditions numériques, on peut imaginer que les extraits de corpus issus d'Internet seront directement incrustés dans la revue en ligne et donc s'interroger à nouveau sur les aspects juridiques liés à ces pratiques par partage:

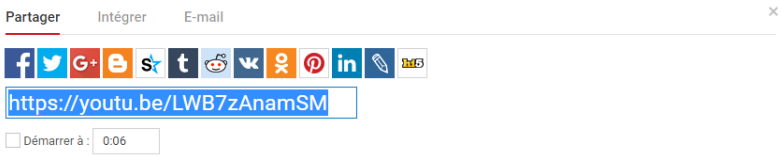

Figure 4 : Exemple d'une zone de partage du vlogue de Michael

Ou intégration:

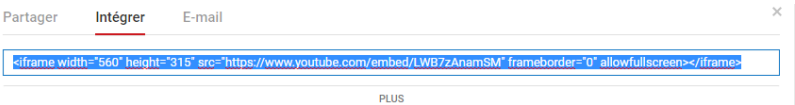

Figure 5 : Exemple d'une zone d'intégration du vlogue de Michael

Il n'en va évidemment pas de même lorsque les individus ont des attentes en matière de vie privée, et Côté rappelle que le chercheur doit alors protéger la vie privée et l'anonymat des internautes en prenant toutes les précautions:

> Bien qu'un participant puisse consentir à la publication de ses messages ou de ses photos et vidéos, d'autres membres de la communauté pourraient s'y opposer: le chercheur doit alors "maintenir l'anonymat de tous les membres du groupe" (ÉPTC 2: 60), car la publication de l'un pourrait engendrer l'identification de l'autre et affecter toute la communauté virtuelle (McKEE; PORTER, 2009a). L'utilisation des moteurs de recherche pourrait permettre de retrouver n'importe quelle bribe d'information, briser l'anonymat, voire contribuer à identifier d'autres bases de données anonymes. Même un pseudonyme peut être associé à nombre de renseignements identificatoires archivés sur le Web (JOHNS; CHEN; HALL, 2004). (CÔTÉ, 2012, p.8).

Ainsi, dans le cas d'un vlogue public, tant que la recherche ne relève pas d'une situation délicate, mais d'une analyse sémiodiscursive à visée didactique par exemple, le chercheur pourra ne pas anonymiser les vidéos du vlogueur, mais en revanche anonymiser les commentaires si nécessaire. *YouTube* met à disposition une aide consacrée à la gestion des droits et des droits d'auteur[9]. Dans le cas des copies d'écran *YouTube* de ce chapitre, elles ont été réalisées dans mon propre environnement *YouTube*, la visée de mon propos étant descriptive, sans usage commercial et sans conséquence pour les personnes citées, j'ai fait le choix de ne rien anonymiser pour conserver l'écologie de l'environnement.

3.2. Au vécu de l'intérieur

Dans le cas de l'application intimiste, la question est plus délicate. En effet, je pense qu'il est important d'expérimenter soi-même l'application et les échanges que l'on aura avec un partenaire au sein de cet espace pourront constituer des données premières. Cette immersion dans cet environnement, cette expérience vécue de l'intérieur permettra au chercheur de faire émerger des interrogations et des hypothèses qu'il lui faudra ensuite vérifier.

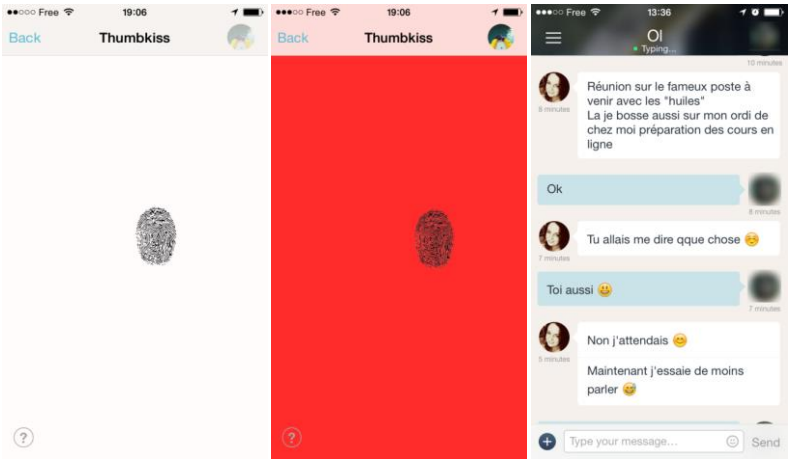

Figure 6 : Expérimentation du baiser virtuel et de la "parole à distance en action" (*Typing*).

Toutefois, compte tenu de l'intimité de ces espaces et des individus qui s'y retrouvent (deux individus en situation amoureuse à priori), il semble difficile d'obtenir ces échanges. La page "privacy[10]" de l'application explique clairement que personne n'a accès au contenu déposé sur la plateforme, excepté la personne avec qui l'on forme un couple virtuel et les employées de la start-up pour le support technique. Il se peut alors que le seul corpus numérique auquel le chercheur ait affaire soit ses propres échanges, mais il semble encore plus difficile de les exploiter. Quelle position adopter par rapport à son partenaire dans l'interaction? Quelle objectivité par rapport à son propre discours? Tant que le chercheur vit l'expérience en tant qu'individu, il a un rapport neutre à ce qui pourra être ensuite des

données, mais dès que l'individu et le chercheur se rencontrent, le discours risque évidemment d'être, même inconsciemment, manipulé. C'est alors que le chercheur devra se constituer un corpus à partir de données autres qui viendront enrichir ses premières données.

4. Quelles données pour quels corpus?

Les données numériques présentent généralement deux particularités: elles sont nombreuses et en apparence facilement accessibles. Mais ce qui peut sembler au départ un atout s'avère rapidement un problème que le chercheur doit solutionner, et n'est pas sans poser un certain nombre de questions que nous allons explorer toujours à partir de nos deux exemples.

4.1. Collecter des données numériques

Comme le souligne Paveau :

> La nature technodiscursive de ce type d'énoncé, définie par une complète intégration de la technique à la matière langagière, impose de modifier nos regards et de prendre pour objets d'analyse, non plus les éléments langagiers dans une perspective logocentrée, mais l'ensemble de l'environnement technodiscursif dans la perspective écologique nécessaire à une analyse du discours numérique. C'est en cela que les corpus de discours natifs du web sont "sensibles" (PAVEAU, 2015, p.2).

Le chercheur qui étudie ce genre de données écologiques se trouve donc confronté à un profond dilemme : recueillir ou ne pas recueillir ces données? En effet, recueillir ces données, c'est les extraire de leur environnement et donc les dénaturer. Ne pas les recueillir, c'est risquer de les perdre, c'est aussi se priver de la possibilité d'utiliser certains outils pour les étudier. Il semble donc indispensable d'effectuer les observations *in situ*, c'est-à-dire en ligne, afin de conserver au mieux l'écologie de la production discursive; toutefois, il ne faut pas se priver de les recueillir parfois en partie. Ainsi, en téléchargeant les vidéos d'un vlogue via le logiciel "Download helper" de Mozilla Firefox, on pourra transcrire et annoter grâce au logiciel ELAN[11] certains moments "choisis". En effet, comme le souligne Drissi (2011), ces transcriptions prennent un temps

considérable, c'est pourquoi le checheur peut ne transcrire que les extraits sur lesquels il veut opérer au niveau "micro", une analyse fine concernant par exemple la mise en relation des éléments mimogestuels, prosodiques, verbaux (oraux et écrits) et graphiques (incrustations au sein de la vidéo).

Figure 7: Copie d'écran d'un "moment multimodal".

En copiant-collant dans Word les commentaires associés à la vidéo, on pourra également opérer plus facilement sur un nombre important de commentaires une recherche lexicale, par exemple, en utilisant la fonction "Rechercher" de Word.

Dans le cas de l'application *Couple*, la consultation des données numériques sera plus aisée si le chercheur est impliqué dans l'interaction (avec toutes les questions que cela soulève cf. *infra*). Le recueil des données en revanche ne pourra s'effectuer que par copie d'écran et plutôt à partir de l'écran d'un téléphone, le site web de l'application n'étant pas aussi complet et le copier-coller s'avérant impossible.

4.2. Constituer le corpus numérique de référence

Si la constitution du corpus s'avère toujours une question importante en recherche, elle se révèle cependant particulièrement

complexe dans une étude du discours numérique et des interactions en ligne.

Comme le souligne Develotte:

> Les corpus multimodaux en ligne ont en commun d'être foisonnants et de comporter beaucoup plus de données que le chercheur sera capable d'en analyser. Dès lors, s'impose la nécessité de sélectionner des moments particuliers, des "espaces-temps" propres à pouvoir mettre au jour les éléments que l'on cherche à éclairer. (DEVELOTTE, 2012, p.518)

Si l'on prend l'exemple du vlogueur Michael, on se trouve face à des données numériques multiples: trois chaînes *Youtube*, une page *Facebook*, un compte *Twitter*, un compte *Instagram*, un compte *Google* + et un compte *Snapchat*.

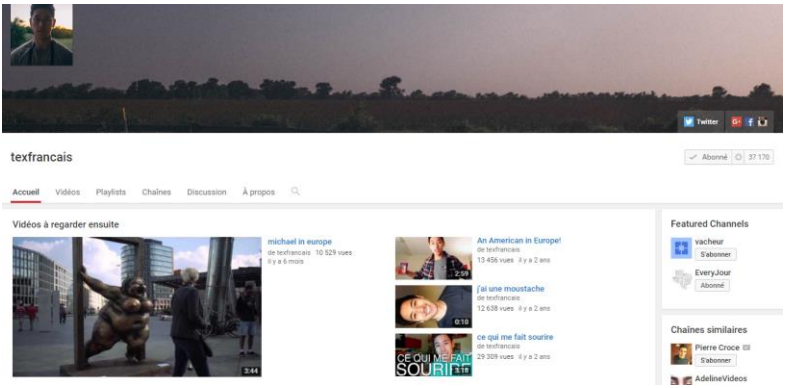

Figure 8 : Copie d'écran de la chaîne "texfrancais" de Michael avec des liens vers ses autres réseaux sociaux.

Ces données s'avèrent non seulement massives mais également labiles: en effet, les comptes, pages, chaînes sont en permanence alimentées par le vlogueur ainsi que par les internautes qui interviennent sur ces espaces publics. Ainsi, dans le cas du vlogue, Michael peut à tout moment ajouter ou supprimer des vidéos à sa chaine. On note aussi, par exemple, que le premier commentaire à la première vidéo de sa chaîne "texfrancais" a été posté en 2011 et le

dernier “il y a un mois” par rapport à la date de rédaction de cet article (en avril 2016). Si le chercheur veut exploiter les données chiffrées d’un vlogue, comme par exemple le nombre d’abonnés, de vues, de “j’aime” et “je n’aime pas”, de commentaires, il n’aura de cesse de les actualiser.

Dans un premier temps, il semble donc indispensable d’effectuer en milieu écologique une première lecture flottante (VAN DER MAREN, 2003) du corpus existant. Le chercheur pourra ainsi voir émerger des questions de recherche et ensuite déterminer dans un second temps le corpus de référence (BOMMIER-PINCEMIN, 1999) qui sera d’abord composé de données numériques. Un ensemble formant un tout comme une chaîne *YouTube* pourra être par exemple un critère de choix.

4.3. Croiser les données et définir le corpus d’étude

Reste ensuite au chercheur à définir son corpus d’étude, c’est-à-dire celui qu’il va effectivement étudier plus en profondeur, de manière auqlitative notamment.

Si l’on prend l’exemple de Michael, on peut considérer que le corpus numérique existant est l’ensemble des données numériques publiques de Michael, donc l’ensemble de ses chaînes *YouTube* et de ses pages et comptes (*Facebook*, *Instagram*, *Twitter*, etc.) accessibles par tout un chacun dès lors que l’on a soi-même un compte sur le réseau social concerné ou même sans dans le cas de *YouTube* par exemple. Le corpus numérique de référence pourra par exemple n’être composé que d’une seule de ses chaînes *YouTube* (“texfrancais” par exemple). A partir de là, le chercheur peut effectuer une sélection de certains vlogues en fonction de sa question de recherche. La thématique de la vidéo initiale pourra être par exemple un critère de choix. Ainsi, la chaîne ”texfrancais”de Michael traite de thématiques très différentes comme sa vie amoureuse, les attentats de Charlie ou encore ses difficultés d’apprentissage du français. Dans le cas d’une question centrée autour de la didactique des langues, le chercheur choisira les vlogues traitant de son apprentissage du français. Par ailleurs, certains vlogues comportent un nombre considérable de commentaires. Afin de procéder à une analyse qualitative, le chercheur devra à nouveau procéder à une sélection. La méthode

proposée par Herring, en 2004, me semble toujours pertinente. Selon cette auteure, le chercheur peut choisir selon les critères suivants:

> Random (e.g., each message selected or not by a coin toss), By theme (e.g., all messages in a particular thread), By time (e.g., all messages in a particular day/week/month), By phenomenon (e.g., only instances of joking; conflict negotiation), By individual or group (all messages posted by an individual or members of a demographic group, e.g., women, students), Convenience (whatever data are available to hand) (HERRING, 2004, p.11).

Dans le cas de la plateforme *YouTube*, le chercheur pourra aussi exploiter le tri effectué automatiquement par *YouTube* et apparaissant comme "Top des commentaires", c'est-à-dire "ceux qui ont suscité des discussions auprès des spectateurs et ceux qui ont enregistré des clics sur "J'aime" au sein de la communauté[12]". Enfin, il me semble important de croiser les données et si le chercheur ne peut réaliser un entretien avec le vlogueur, il peut en revanche, enrichir et éclairer son analyse à partir de discours du vlogueur recueillis sur les autres réseaux sociaux de ce dernier, comme sa page *Facebook* par exemple.

Dans le cas de l'application intimiste, les premières données personnelles (à partir de sa propre expérience donc) que l'on recueille peuvent s'avérer très abondantes, voire continues (si le chercheur continue d'échanger au sein de cette application). Elles sont cependant insatisfaisantes car elles ne sont que le reflet de la perception du chercheur. Cette expérience de l'application vécue de l'intérieur permet cependant d'élaborer un questionnaire ainsi qu'un guide d'entretien dans la perspective de croiser une expérience avec d'autres expériences. Le web et les réseaux sociaux peuvent alors s'avérer à nouveau source de données. Ainsi, l'on peut chercher à atteindre des utilisateurs en diffusant ces questionnaires sur la page *Facebook* et le compte *Twitter*, supports de l'application *Couple*. Cependant, le fait de passer par des réseaux sociaux ne donne que l'apparence d'une diffusion massive (la page *Facebook* est aimée par 119 519 personnes et le compte *Twitter* suivi par 7152 abonnés) et la possibilité d'entrer en contact avec les utilisateurs. Il faudra donc poster d'autres messages à d'autres endroits, j'ai personnellement cherché à envoyer

manuellement le questionnaire via leur compte *Twitter* aux différents utilisateurs de *Couple*, mais *Twitter* m'a rapidement bloquée soupçonnant un robot et dans l'ensemble, ces démarches se sont avérées relativement peu fructueuses (8 réponses à un questionnaire en anglais et 3 à un questionnaire en français). J'ai donc parcouru le web et trouvé un site dédié à l'amour à distance[13] qui offrait des forums pour les amoureux à distance et notamment un intitulé "Conseils pratiques" dans lequel l'auteur du site avait posé la question suivante: "Quels outils utilisez-vous le plus pour communiquer ensemble?", discussion qui a recueilli 23 messages. Il avait également organisé un sondage autour des deux moyens les plus utilisés en couple. Ce site m'a donc fourni quelques données intéressantes : les réponses au sondage ainsi que le discours des acteurs de ce mode de communication directement recueilli sur le forum. Il m'a aussi permis d'entrer en contact avec des personnes susceptibles d'utiliser ce genre d'application. Via le forum du site "Amour à distance", j'ai intégré cette petite communauté et obtenu deux entretiens avec un couple utilisateur de l'application. Aux discours natifs de l'internet peut donc venir s'ajouter d'autres discours, formant un tout de données hétéroclites (un sondage, des échanges sur des forums, des données multimodales personnelles issues de notre expérimentation, des réponses à des questionnaires et des entretiens) mais offrant dans le cadre d'une application intimiste où les données numériques s'avèrent trop confidentielles pour être partagées une perspective plus objectivée de la recherche.

Conclusion

Dans cet article, j'ai cherché à relater quelques questionnements méthodologiques auxquels se trouve confronter un chercheur en sciences du langage qui a fait le choix d'étudier le discours numérique et les interactions en ligne au sein de deux plateformes interactives multimodales tout à fait différentes, voire même opposées sur le plan du concept: à savoir la communauté de partage *YouTube* et l'application intimiste *Couple*. Comme j'ai tenté de le montrer la question du positionnement du chercheur reste centrale et ce plus les données sont intimes, le chercheur, tel un cinéaste, varie ses focalisations, interne ou externe, pour mieux rendre compte d'une réalité qui doit être nécessairement vécue mais aussi observée de

manière distanciée. La question des données et de leur masse est également un point sur lequel s'interroger, et la constitution du corpus est un acte particulièrement important notamment dans une perspective d'analyse micro et qualitative. Croiser les données numériques avec d'autres données recueillies auprès des acteurs peut permettre de multiplier les points de vue et ainsi d'objectiver la recherche.

Références bibliographiques

ANIS, J. *Texte et ordinateur: L'écriture réinventée ?*. Paris; Bruxelles: De Boeck Université, 1998.

BOMMIER-PINCEMIN, B. *Diffusion ciblée automatique d'informations: conception et mise en œuvre d'une linguistique textuelle pour la caractérisation des destinataires et des documents*. Thèse de Doctorat en Linguistique, Université Paris IV Sorbonne, 6 avril 1999, chapitre VII : "Caractérisation d'un texte dans un corpus : du quantitatif vers le qualitatif", §A "Définir un corpus", 1999, p.415 427. Disponible sur Internet (consulté en mars 2010): <http://www.revue-texto.net/Inedits/Pincemin/Pincemin_these.html>.

COMBE, C. "Vlogues sur YouTube: un nouveau genre d'interactions multimodales". In: I. COLON DE CARJAVAL et M. OLLAGNIER-BELDAME (eds), Actes du colloque Interactions Multimodales Par ECrans 2014, Lyon 2 au 4 juillet, 2014.

CÔTÉ J. "Les enjeux éthiques de l'utilisation d'internet en recherche : principales questions et pistes de solutions", *Éthique publique* [En ligne], vol. 14, n° 2 | 2012, mis en ligne le 22 juillet 2013, consulté le 14 novembre 2013. URL : <http://ethiquepublique.revues.org/997>.

CRYSTAL, D. *Language and the Internet: Second Edition*. Cambridge : University Press, 2001.

_____. *Internet Linguistics*. New-York : Routledge, 2011.

DEVELOTTE, C.; KERN, R.; LAMY, M.-N. *Décrire la conversation en ligne*. Lyon, ENS Editions, 2011.

DEVELOTTE, C. "L'analyse des corpus multimodaux en ligne: état des lieux et perspectives". In: *Actes en ligne du Congrès Mondial de Linguistique Française* (CMLF), 2012, téléchargeable à l'adresse : http://www.shsconferences.org/index.php?option=com_article&access=doi&doi=10.1051/shsconf/20120100213&Itemid=129

DRISSI, S. “Apprendre à enseigner par visioconférence: Étude d’interactions pédagogiques entre futurs enseignants et apprenants de FLE”. Thèse de doctorat en sciences du langage, ENS de Lyon, 2011.
HERRING, S.-C. “Computer-Mediated Discourse Analysis: An Approach to Researching Online Behavior”. In: S.-A. Barab; R. Kling; J.-H Gray (éds.). *Designing for Virtual Communities in the Service of Learning*. New Yor : Cambridge University Press, 2004, p.338-376.
HERRING, S.-C. “*New frontiers in interactive multimodal communication*”. In: A. GEORGAPOULOU, A.; SPILLOTI, T. (Eds.). *The Routledge handbook of language and digital communication*. London: Routledge, 2015, p.398-402. <http://info.ils.indiana.edu/~herring/hldc.pdf>.
KAPLAN, A. M.; HAENLEIN, M. “Users of the world, unite ! The challenges and opportunities of Social Media”. *Business Horizons*, 53. Elsevier, 2010, p.59-68.
PAVEAU, M.-A. “L’intégrité des corpus natifs en ligne. Une écologie postdualiste pour la théorie du discours”. In: *Cahiers de praxématique*, Publications de l’Université Paul Valéry, 2015, Corpus sensibles, p.65-90. <hal-01185710>.
VAN DER MAREN, J.-M. *La recherche appliquée en pédagogie: des modèles pour l’enseignement*. 2.ed. Bruxelles: De Boeck, 2003.

Palavras-chave: plataforma interativa modal, metodologia, corpus.
Mots-clés: plateforme interactive multimodale, méthodologie, corpus.
Keywords: multimodal interactive platform, methodology, corpus.

Notes

[1]Les travaux de Susan Herring dans le domaine sont très importants et je ne cite là qu’un article qui nous semble fondateur sur le plan méthodologique parmi tant d’autres écrits, cf. son site Internet http://info.ils.indiana.edu/~herring/pubs.html.
[2]http://impec.sciencesconf.org/resource/page/id/23.
[3]https://technodiscours.hypotheses.org/category/dictionnaire-dadn.
[4]Source : http://blog.couple.me/about/.
[5]https://couple.me/.
[6]http://blog.couple.me/.
[7]https://www.facebook.com/Couple/?fref=ts.
[8]https://twitter.com/CoupleApp.
[9]https://support.google.com/youtube/topic/2676339?hl=fr&ref_topic=6151248.

[10]https://couple.me/privacy.
[11]https://tla.mpi.nl/tools/tla-tools/elan/.
[12]Source: Aide YouTube.
[13]http://amouradistance.fr/.