Comment constituer et traiter des *corpus* issus des médias sociaux à l'aide des CAQDAS? NVivo : un instrument pour l'analyse des discours numériques

How to build and analyse social media corpus with the help of CAQDAS? NVivo: an instrument for digital discourse analysis

Giovannipaolo Ferrari^{1,a}

¹Università di Salerno & Université Paul-Valéry Montpellier III

Résumé. Avec l'apparition des médias sociaux, le chercheur en Sciences Humaines et Sociales a vu s'ouvrir des possibilités de création de *corpus* à partir des données disponibles en ligne. Le risque dans ces modalités de création de *corpus* est de créer de grandes bases de données difficilement gérables avec les outils traditionnels et surtout dans le cadre d'une recherche qualitative. Pour cette raison, il semble important de mobiliser des outils numériques pour l'aide à la recherche. Ces logiciels sont appelés *CAQDAS*. Cette contribution montre comment les utiliser pour une enquête de terrain en ligne ou numérique afin d'organiser et d'analyser un corpus. Ce *corpus* est constitué en vue d'une analyse des discours produits dans le contexte professionnel de la radio utilisant les nouveaux médias.

Abstract. The rise of social media gave the chance for scholars in Humanities and Social Sciences to create corpus from data available online. The danger, in this way, is collecting big data very difficult to manage with traditional tools and, overall, in the frame of qualitative research. For this reason, it seems important to mobilize digital tools to assist the research in a reflexive perspective. These kind of software are called *CAQDAS*. This paper shows how analyse these software in an online or digital fieldwork to organise and analyse a corpus. This latter is built to develop a discourse analysis in a radio professional environment using social media.

1 Introduction

Avec l'apparition des médias sociaux, le chercheur en Sciences Humaines et Sociales (SHS) a vu s'ouvrir des possibilités de création de corpus à partir de données en ligne (Lejeune, 2010 [1]). En effet, les médias sociaux, aujourd'hui, représentent un terrain de recherche très vaste où les chercheurs des différentes disciplines des SHS peuvent observer des situations et des interactions, participer directement *in situ*, recueillir des données et des matériaux de recherche et donc, créer des corpus complexes pour les analyser avec différentes techniques et approches méthodologiques. Le danger dans ces modalités de création de corpus en ligne est de créer de grandes bases de données difficilement gérables avec les outils traditionnels du chercheur en SHS. Pour cette raison, il semble important de faire appel à des outils numériques pour l'aide à la recherche dans une perspective réflexive (Schön, 1987 [2], Schön, 1994 [3]). L'enjeu principal de cet article, donc, sera d'articuler une réflexion sur les choix méthodologiques d'une recherche portant sur les médias sociaux au sein d'une production radiophonique et l'utilisation d'un logiciel destiné à accompagner deux phases importantes de la recherche : l'organisation des matériaux empiriques et l'analyse. Cette contribution montre, à cet égard, comment recueillir, organiser et analyser avec un logiciel CAQDAS (Nvivo10), un corpus de matériaux de recherche multimodal sur les médias sociaux en vue d'une « analyse du discours numérique »

^a Auteur de correspondance : gifferrari@unisa.it ou giovannipaolo.ferrari@univ-montp3.fr

(Paveau, 2013a [4]). Par *corpus*, il est entendu, non seulement des données mais dans le contexte de la recherche qualitative, il semble plus pertinent d'utiliser l'expression « corpus de matériaux de recherche multimodaux » (Pacificio & Coppola, 2010 : 32 [5]) ou « base empirique multimodale de la recherche » (Giuliano & La Rocca, 2008 : 11 [6]).

Nous présenterons dans un premier temps brièvement quelques problématiques du débat sur les logiciels d'aide à l'analyse qualitative et leurs enjeux pour les Sciences du Langage. Nous prendrons l'exemple d'un logiciel CAQDAS dont nous présenterons les fonctionnalités principales. Dans un deuxième temps, nous ferons état d'une recherche en cours de réalisation sur les médias sociaux et la radio en exposant son contexte, ses objectifs et l'observation participante qui eu lieu à la Radio 1 RAI en mai 2015 en Italie. Enfin, nous rendrons compte des enjeux pour l'analyse de l'utilisation du logiciel.

2 Le débat sur les CAQDAS

Dans la recherche qualitative surtout et dans l'enquête de terrain, le chercheur peut notamment utiliser des outils numériques qui l'aident à recueillir et organiser son *corpus*. Ces logiciels sont appelés *CAQDAS*^b. Les différentes approches et les différents objectifs de la recherche qualitative ont amené à la création, depuis la diffusion des ordinateurs personnels dans les années 80, un nombre important de logiciels qui aident le chercheur à organiser et analyser des corpus de données (textuelle, audio et vidéo)^c avec des approches et des techniques propres de la recherche qualitative. Nous reportons ci-dessous un schéma récapitulatif (cf. [figure 1](#)) [7:73] des principales approches dans l'analyse des données qualitatives nées grâce au débat développé à l'intérieur des *SHS* en particulier.

Figure 1. Situer l'ADQAO (Voynnet-Fourboul, 2012 : 73 [7]).

Ce que l'on observe ici, c'est une distinction entre les recherches dont l'enjeu est la description des caractéristiques du langage^d, les recherches qui ont comme objectif de proposer des constructions théoriques, entre autres, à partir d'une unité empirique et enfin, les recherches qui se veulent principalement descriptives et interprétatives. A partir de ces différents objectifs, les logiciels CAQDAS ont été conçus pour accompagner le chercheur dans sa démarche scientifique, que sa recherche soit

^b *Computer-Assisted Qualitative Data Analysis Software*. Parfois, ils sont appelés *QDAS* (*Qualitative Data Analysis Software*) ou *QDA* (*Qualitative Data Analysis*). La version française de l'acronyme est *ADQAO* (*Analyse de Données Qualitative Assistée par Ordinateur*). Dans cet article, nous utiliserons plutôt l'acronyme anglais plus connu et utilisé dans la littérature.

^c Les données textuelles, audio et vidéo peuvent être analysées également selon une approche quantitative. Dans ce cas également, des logiciels ont été développés afin d'assister la recherche.

^d « Cette taxonomie présentée dans la [Figure 1](#) sépare deux courants : l'analyse de contenu et la construction théorique. Précisons que le sens attribué à l'analyse de contenu diffère entre la recherche française ou anglo-saxonne. L'analyse de contenu dans la littérature française est un ensemble de différentes méthodes, mais n'est pas assimilable au terme anglo-saxon « content analysis ». En effet, selon l'ouvrage de référence en la matière, la « content analysis » est l'équivalent en France de l'analyse lexicologique. L'analyse lexicologique ou lexicométrique constitue une méthodologie pertinente pour détecter des structures à la fois apparentes et sous-jacentes d'un texte » (Voynnet-Fourboul, 2002 : 73 [8]).

centrée sur la description des caractéristiques du langage, qu'elle ait pour objectif de proposer une conceptualisation à partir de situations étudiées et de données complexes, ou encore qu'elle associe plusieurs approches (analyse du langage et conceptualisation dans un second temps). Les [figures 2 et 3](#) illustrent les différents objectifs, qu'on pourrait également qualifier « d'exigences » de recherche (et du chercheur) ainsi que les logiciels créés à cet égard pour atteindre ces objectifs et pour satisfaire les exigences exposées ci-dessus.

Figure 2. Identification des outils informatisés pour l'analyse des données qualitatives (Voynnet-Fourboul, 2002 [8]).

Au-delà des différents objectifs de recherche mais également en lien avec ceux-ci, on peut distinguer parmi les *CAQDAS* des « fonctions de base » (fonctions d'archivage, de recherche, de codification et d'annotation) et des « fonctions avancées » (soutien à l'élaboration théorique à travers l'identification et la visualisation des liens entre les concepts) communes à de nombreux programmes comme le montre également la [figure 3](#).

Figure 3. Identification des outils informatisés pour l'analyse des données qualitatives (Voynnet-Fourboul, 2002 [8]).

Enfin, l'innovation technique proposée par les dernières^e versions des logiciels *CAQDAS* (Paveau, 2013b [9] ; Paveau, 2012a [10] ; Diana & Montesperelli, 2005 [11]) permet de capturer des bases de données multimodales provenant des médias sociaux, c'est-à-dire, des principales plateformes d'échanges textuels et multimédias (*Facebook*, *Twitter* etc.), les principales plateformes en ligne professionnelles (*LinkedIn*, *Viadeo* etc.) et sur les forums de discussion.

^e Les objectifs initiaux de l'utilisation de *CAQDAS* (cf. note de bas de page n. b) sont :

- structurer des données, rendre compte du détail, du vivant de façon organisée ;
- trouver des régularités, des logiques d'articulation entre concepts ou idées pour produire une synthèse ;
- isoler, bâtir, définir des concepts *in-vivo* ;
- traiter à la fois les cas et les variables.

3 Présentation du logiciel

Il existe une multitude de *CAQDAS* qui privilégient différentes approches de la recherche qualitative. Cependant, l'objet de cet article n'est pas de faire une présentation exhaustive de tous les outils numériques et de la littérature concernant ces derniers (Weitzman & Miles, 1995 [12] ; Dicks & al., 2005 [13] ; Lejeune, 2007 [14] ; Lewins & Silver, 2009 [15] ; Bournois & al., 2002 [16] ; Gilbert & al., 2014 [17]) mais de montrer les enjeux et limites de l'usage de l'un d'entre eux, l'un des plus complets, a priori, pour la constitution d'un *corpus* issu d'une recherche *in situ* sur les médias sociaux. Ce *corpus* est constitué en vue d'une analyse des « matières technolangagières » (Paveau, 2014 : 8 [18]) dans un environnement^f de « technologie discursive^g » (Paveau, 2013b [9] ; Paveau 2013a [4]), autrement dit l'analyse des « technodiscours^h » (Paveau, 2012a : 106 [10]) produits par les nouvelles figures professionnelles qui utilisent les médias sociaux. C'est dans cette perspective que le logiciel *Nvivo10* a été choisi afin de réaliser un travail quotidien sur notre projet de recherche de l'organisation des matériaux de recherche à la transcription en passant par l'annotation et la catégorisation des discours et documents (cf. [figure 4](#)). Ici, la figure 4 montre toutes les fonctionnalités de l'interface principale de Nvivo qui permettent, dans ce cas, l'importation d'une vidéo réalisée avec une caméra 360° (4 vues), la réalisation d'une transcription avec la synchronisation vidéo, l'importation de fichiers audio et des différents formats (word, pdf, excel, NCapture etc.).

Figure 4. Présentation de l'interface et des fonctionnalités principales de *Nvivo10* pour l'organisation d'un corpus de matériaux de recherche multimodaux.

^f Il est fait référence, ici, à la notion d'environnement dans le sens d'un « ensemble des données humaines et non humaines au sein desquelles les discours sont élaborés. La notion d'environnement est ... une alternative critique à celle de contexte courante en analyse du discours, plutôt centrée sur les paramètres sociaux, historiques et politiques » (Paveau, 2013a : 141 [4]).

^g « La technologie discursive est l'ensemble de processus de mise en discours de la langue dans un environnement technologique. C'est un dispositif au sein duquel la production langagière et discursive est intrinsèquement liée à des outils technologiques en ligne ou hors ligne (ordinateurs, téléphones, tablettes, logiciels, applications, sites, blogs, réseaux, plateformes) » (Paveau, 2013b : 141 [9]), ces outils laissant des traces dans les matières langagières.

^h Désigne la mise en acte en contexte numérique de la technolangue qui désigne le système de formes disponibles : communes, apprises, évolutives et régies par des règles [10].

Il nous semble, en effet, que les fonctionnalités de *Nvivo10* (Lejeune, 2014a [19] ; Lejeune, 2014b [20]) permettent d'accompagner une analyse du discours. En effet, toutes les fonctionnalités des requêtes (*Queries*) permettent d'interroger le logiciel sur le contenu du texte. Par exemple, il est non seulement possible de faire une requête sur la fréquence de mots (*wordfrequency*) mais également de faire tout l'étiquetage (réflexif) (Lejeune, 2014b [20]) du texte par des nœuds (*Nodes*) et donc de sélectionner les différentes séquences textuelles et de les distinguer selon différentes catégories, sous-catégories et/ou dimensions, sous-dimensions. Il est possible également d'avoir un espace d'*editing* (édition) où sont traités des documents audio et vidéo, où l'on retrouve tous les outils nécessaires pour mettre en place des transcriptions pour une analyse du discours.

4 Présentation de la recherche

La recherche que nous présentons ici a comme objectif de rendre compte des nuances relatives aux activités de la figure professionnelle du gestionnaire des médias sociaux dans le contexte radiophonique public (*Radio France, Radio RAI*) et commercial (*Virgin Radio France, Virgin Radio Italia*) [21]. Ce dernier est en général intégré dans l'équipe chargée de la réalisation d'une émission de radio. Corrélativement au déroulement de l'émission, ce professionnel « saisit » les moments forts de cette dernière afin d'en rendre compte sur les médias sociaux.

La recherche que nous présentons ici a eu lieu dans le cadre du Projet Web Radio Interactions, financé grâce à une action du Programme Vinci de l'Université Franco-Italienne (UIF). Dans ce contexte, deux équipes de recherche, l'une du département de *Dipartimento di Scienze Politiche, Sociali e della Comunicazione* de l'Université de Salerne en Italie et l'autre du laboratoire Praxiling UMR 5267 CNRS Praxiling de l'Université Paul-Valéry Montpellier en France, ont conduit pendant trois ans une recherche dont l'enjeu principal portait sur l'analyse des différentes formes de langages et des différentes formes de communication radiophonique à la lumière des nouvelles technologies et des médias sociaux. Le projet a été intégré aux activités de l'Observatoire des Pratiques Médiatiques Emergentes (OPME) qui s'intéresse à l'observation et à l'étude des nouvelles pratiques professionnelles émergentes et des nouvelles professions liées aux nouveaux médias. Suite à une première recherche à France Inter, les deux équipes ont choisi deux émissions de radios italiennes qui pourraient être les correspondantes de la Matinale de France Inter : "Radio Anch'io" sur Radio 1 RAI et e "Tutta la città ne parla" de Radio 3 RAI. L'objectif principal du projet, dans sa dimension méthodologique et théorique, est lié au développement d'une recherche ensemble et en l'observation des évolutions des nouvelles pratiques professionnelles en relation avec les mutations technologiques : déclinaisons multi-médiatiques, rétroactions des usagers en particulier sur les "médias sociaux" associés aux transmissions radiophoniques, nouvelles pratiques de recueil, de traitement et de formatage de l'information et formation des nouveaux professionnels des nouveaux médias, et dans ce cas, analyse des productions du discours radiophonique à travers les médias sociaux.

Cette recherche a ainsi comme objectif d'observer une situation a priori « complexe » dans le sens où les formats de production radiophonique sont ici multiples et hétérogènes. Ils impliquent en effet :

- le flux sonore de l'émission ;
- les productions sur les médias sociaux qui sont relatives à l'intégration du gestionnaire des médias sociaux au sein de l'équipe au cours de l'émission.

Dans cette optique, la première phase de l'enquête a été caractérisée par une stratégie ethnographique qui a principalement eu lieu en mai 2015 dans les studios de la radio italienne Rai (Radio 1 et Radio 3). Les chercheurs engagés dans le projet ont mobilisé des techniques classiques telles que l'observation participante, semi-participante ou non participante ; ils ont mené des entretiens avec des techniques « non-standards » (entretien non structuré avec ou sans trace, entretien qui suit les trajectoires biographiques, entretien herméneutique (Diana & Montesperelli, 2005 [11]). Enfin, ils ont réalisé des enregistrements audio et/ou vidéos des entretiens et des observations ; les notes ethnographiques ou les comptes rendus (Lejeune, 2014c : 13 [22]) écrits sur le terrain durant l'observation, c'est-à-dire *in situ*. Par la suite, ils ont inséré et organisé ces données dans le logiciel *Nvivo10*, qui a permis par ailleurs, l'importation automatisée des productions du gestionnaire des médias sociaux.

Enfin, l'ensemble de ces matériaux ont traités, transcrits et connectés entre eux à l'aide du logiciel, ce qui a constitué les « matériaux de recherche élaborés », comme le montre la [figure 5](#).

Figure 5. Organisation des matériaux de la recherche sur la radio grâce à Nvivo10.

5 Des enjeux d'une visualisation en « environnement natif » pour une analyse du discours numérique au sein de la production radiophonique

Revenons sur l'innovation de la dernière version de *NVivo*ⁱ qui a nous a semblé assez importante pour notre recherche. Elle réside dans les fonctionnalités qui permettent d'exporter directement de *Google Chrome*^j par le biais d'une application qui s'appelle *NCapture* et qui est toujours visible via un bouton activable sur la barre d'outils de *Google Chrome*. Grâce à cette petite application, il est possible d'exporter une copie d'écran d'un profil complet sur un réseau social en format *PDF* avec l'ensemble des matériaux technologiques (Tweets, logos, images retweets, informations) postés de l'activation des comptes à aujourd'hui. Mais la fonctionnalité la plus intéressante, comme le montrent les [figures 6](#) et [7](#), consiste à pouvoir exporter des « dataset » (base des données) complets de *Facebook* et *Twitter*.

ⁱ Courant octobre 2015, il est prévu la sortie d'une nouvelle version de *NVivo* : des versions *demo* et *beta* ne sont pas disponibles mais de la campagne publicitaire de la compagnie *QSR International*, il semble que cette nouvelle version présentera des nouvelles fonctionnalités encore plus avancées pour l'extraction et l'analyse de matériaux de recherche multimodaux.

^j Nous citons ici le nom exact du navigateur. En effet, il n'est pas possible d'installer cette application sur d'autres navigateurs. L'opérabilité est donc limitée à *Google Chrome*.

Figure 6. Capture d'écran d'une base de données avec NCapture d'un profil Twitter d'une radio publique italienne.

Figure 7. Capture d'écran d'une base de données avec NCapture d'un profil Facebook d'une radio publique italienne.

La [figure 8](#) présente un exemple de la manière dont un profil *Twitter* apparaît après exportation dans une *dataset* via *NCapture* : outre une quantification des données importées disponibles dans la matrice, il y a une colonne qui, plus que les autres, nous apparaît importante pour développer une analyse des « matières technolangagières » (Paveau, 2012a : 106 [10]) dans un contexte de discours en ligne comme mentionné ci-dessus. Cette colonne reporte le texte complet des *Tweets* et des *Retweets*, dans le cas ici de *Twitter*, et les textes complets des *posts* et des commentaires sur les *posts* dans le cas de *Facebook*. Une telle fonctionnalité permet, d'organiser et d'accélérer la phase de recueil et de constitution du « corpus de matériaux de recherche multimodaux » (Pacífico & Coppola, 2010 : 50 [5]) ou « base empirique multimodale de la recherche » (Giulano & La Rocca, 2088 : 11 [6]).

Figure 8. Importation et visualisation de la capture avec NCapture d'un profil Twitter d'une radio publique italienne par PDF.

Cette possibilité a un intérêt, il nous semble, dans le cadre de l'observation des pratiques professionnelles sur les médias sociaux et du gestionnaire des médias sociaux et dans le cadre d'une analyse des technodiscours produits par les nouvelles figures professionnelles qui utilisent les médias sociaux. Dans une perspective d'analyse des matières technolangagières, le *CAQDAS* comme *NVivo10* permet une évolution au niveau méthodologique parce qu'il offre la possibilité de reproduire un « environnement natif » (Paveau, 2014 : 8 [18] ; Roy, 2008 [23] non seulement par des « captures d'écran ou avec des outils d'archivage de page web rendant compte de la délinéarisation et de l'augmentation et pour les réseaux sociaux numériques de la contextualisation technorelationnelle » comme le suggère Paveau (2014 : 8-12 [18]) ; il permet ainsi d'exporter n'importe quel contenu en ligne en format *PDF*, comme souligné ci-dessus, permettant de « capturer » les matières technolangagières du début de leur production en ligne. De plus, il est possible de codifier et d'annoter directement sur le *PDF*, ce qui rend possible une perception majeure de la part du chercheur des matérialités technolangagières. Le chercheur ne peut néanmoins se retrouver au moment de l'organisation, et plus tard de l'analyse, face à des matériaux différents à traiter avec des modalités différentes mais il en est conscient et réussit à gérer de manière appropriée un *corpus* de matériaux de recherche multimodaux : il réussit à les travailler, les modeler à travers une approche multimodale et interdisciplinaire. Il réussit également à connecter les textes avec les images, les vidéos avec les notes, les traces audio avec les références bibliographiques et les données décrivant une communauté sur *Facebook* avec la richesse des « formes technodiscursives de discours rapportés » (Paveau, 2011 [24] ; Paveau, 2012b [25]).

La capture suivante montre ainsi deux modalités de traiter des données issues des médias sociaux : la première dans leur environnement natif et la seconde déjà organisée pour une analyse plus fine du positionnement d'un *Tweet* dans l'environnement « *Twitter* » grâce à une pré-structuration déjà réalisée par *Nvivo10*.

Figure 9. Importation et visualisation de la capture avec NCapture de la base des données d'un profil Twitter d'une radio publique italienne.

Dans l'analyse, la corrélation de deux vues permet de mettre en exergue les matérialités technolangagières dans leur environnement natif. Elle permet par ailleurs de corréler les transcription de l'émission radiophonique, la transcription des interactions et actions en cours au sein du studio et de la régie (enregistrés lors de l'observation par les chercheurs) et enfin les productions réalisés sur les médias sociaux. Dans cette perspective, il est possible de rendre compte, dans l'analyse, de l'imbrication des différents modes de production d'une part au niveau de la production radiophonique et des discours médiatiques mais également au niveau des discours numériques et de voir la corrélation entre eux.

6 Conclusion

En conclusion, dans cet article, nous avons voulu rendre compte des modalités de recueil et d'analyse d'un matériau empirique recueilli sur des médias sociaux et *in situ* lors d'une recherche sur les radios. L'enjeu principal de cette recherche reste principalement l'analyse du discours numérique, en corrélation avec la production radiophonique. Faute d'espace ici, nous n'avons pas pu rentrer dans le détail de l'analyse de la circulation, des reprises et des reformulations des différentes informations et éléments discursifs de l'émission radiophonique sur les médias sociaux et de montrer dans quelles modalités, il y a une augmentation de la production radiophonique. Il n'en reste pas moins que l'utilisation d'un logiciel CAQDAS permet la mise en corrélation de différentes sources et modalités de production dans la perspective de leur « environnement natif ». Enfin, même s'il existe une multitude d'instruments en ligne qui peuvent être utilisés dans cet objectif, à titre gratuit à travers des démonstrations mensuelles ou payantes, nous avons décrit principalement l'utilisation du logiciel CAQDAS *Nvivo10* que nous avons utilisé pour cette recherche. Il nous semble, ici, que la vraie innovation de *Nvivo10* réside dans son interopérabilité avec de nombreuses applications et des logiciels pour la recherche (par exemple, avec les logiciels de *reference management*) et l'intégration au sein de son interface d'outils de transcription et de codification textuelle. Ces derniers sont capables de dialoguer directement avec les médias sociaux et d'exporter des matériaux de recherche multimodaux directement dans l'interface du logiciel les rendant immédiatement disponibles pour le travail d'analyse du chercheur.

7 Bibliographie

1. C. Lejeune, *Montrer, calculer, explorer, analyser. Ce que l'informatique fait (faire) à l'analyse qualitative*. Recherches qualitatives, **9**. (2010)
2. D. A. Schön, *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco. (1987).
3. D. A. Schön. *Le praticien réflexif. À la recherche du savoir caché dans l'agir professionnel*. Montréal: Éditions Logiques (1994).

4. M. A. Paveau, *Technodiscursivités natives sur Twitter. Une écologie du discours numérique*, Epistémè **9**, 139-176. (2013a)
5. M. Pacifico & L. Coppola. *NVivo: una risorsa metodologica: procedure per l'analisi dei dati qualitativi*. F. Angeli. (2010)
6. L. Giuliano & G. La Rocca. *L'analisi automatica e semi-automatica dei dati testuali. Software e istruzioni per l'uso*. LED Edizioni Universitarie. (2008)
7. C. Voynnet-Fourboul. *Ce que "analyse de données qualitatives" veut dire*. Revue internationale de Psychosociologie, **18(44)**, 71-88. (2012).
8. C. Voynnet-Fourboul. *L'analyse des données qualitatives assistée par ordinateur*. Université Paris II Panthéon-Assas. (2002).
9. M. A. Paveau. *Genre de discours et technologie discursive. Tweet, twittécriture et twittérature*, Pratiques, 157-158, **7-30**. (2013b)
10. M. A. Paveau. *Réalité et discursivité. D'autres dimensions pour la théorie du discours*. Semen, (34), 95-115. (2012a)
11. P. Diana & P. Montesperelli. *Analizzare le interviste ermeneutiche*. Carocci. (2005)
12. E. Weitzman & M. B. Miles. *Computer programs for qualitative data analysis: A software sourcebook*. Sage Publications. (1995)
13. B. Dicks, B., B., Mason, A. Coffey & P. Atkinson. *Qualitative research and hypermedia: Ethnography for the digital age*. Sage. (2005)
14. C. Lejeune. *Petite histoire des ressources logicielles au service de la sociologie qualitative*. Humanités numériques (Tome 1). (2005)
15. A. Lewins & C. Silver. *Choosing a CAQDAS package*. <http://goo.gl/IP11Gc> (2009)
16. F. Bournois, S., Point & C. Voynnet-Fourboul. *L'analyse des données qualitatives assistée par ordinateur: une évaluation*. Revue française de gestion, (137), 71-84. (2002)
17. L. S. Gilbert, K. Jackson & S. di Gregorio. *Tools for Analyzing Qualitative Data: The History and Relevance of Qualitative Data Analysis Software*. In *Handbook of Research on Educational Communications and Technology*. Springer New York, 221-236. (2014)
18. M. A. Paveau. *L'alternative quantitatif/qualitatif à l'épreuve des univers discursifs numériques. Corela. Cognition, représentation, langage HS-15* (2014).
19. C. Lejeune. *Logiciels d'analyse de textes: typologie et approfondissement des outils d'étiquetage réflexif*. Editions de Boeck Supérieur (2014a)
20. C. Lejeune. *Analyse de discours et sciences sociales. Rencontre et malentendu*. Texto, Toulouse, France. (2014b).
21. G. Ferrari. *The profession of Social Media Manager in the radio: A case study of public and commercial radio in Italy and France*. In ISA Interim meeting 2015 "Professions, bonds and boundaries" Università Cattolica del Sacro Cuore - Milano, 19th – 21st March 2015.
22. C. Lejeune. *Manuel d'analyse qualitative. Analyser sans compter ni classer*. Louvain-la-Neuve, De Boeck. 152 p., ISBN : 978-2-8011-1749-1 (2014c)
23. C. S. Roy. *The Linguists: Petit précis d'ethnographie numérique* The Linguists, États-Unis 2008, 70 minutes The Linguists, États-Unis 2008, 70 minutes. Séquences: La revue de cinéma, (255), 20-20. (2008)
24. M. A. Paveau. *Les technologies discursives. Altérisation, cognition, narcissisme* », intervention au Séminaire COSTECH consacré aux « Altérités technologiques », 21 janvier. (2011).
25. M. A. Paveau. *Les amitiés numériques. Formes technolangagières et polysémie relationnelle*, La pensée du discours, <http://penseedudiscours.hypotheses.org/?p=8334> (2012b).