



HAL
open science

**”Transcrire un corpus audio dans la perspective de la
préservation du patrimoine culturel immatériel. Les
obstacles rencontrés, les méthodologies retenues, les
questions restées en suspens”**

Marie-Madeleine Bertucci, Mounia Illourmane

► **To cite this version:**

Marie-Madeleine Bertucci, Mounia Illourmane. ”Transcrire un corpus audio dans la perspective de la préservation du patrimoine culturel immatériel. Les obstacles rencontrés, les méthodologies retenues, les questions restées en suspens”. Le corpus audiovisuel : Quelles approches ? Quels usages ?, Institut National des Langues et Civilisations Orientales (INALCO). Equipe PLIDAM (Pluralité des Langues et des Identités : Didactique, Acquisition, Médiations). Axe 6 : SemioMed. INA. Université de Liège. FNRS. Campus Tech. Laboratoire Paragraphe. ANR MémoMlnes, Jun 2019, Paris, France. hal-03704393

HAL Id: hal-03704393

<https://cyu.hal.science/hal-03704393v1>

Submitted on 29 Jun 2022

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.

Copyright

Marie-Madeleine Bertucci et Mounia Illourmane
Université de Cergy-Pontoise
EA 7518 LT2D

Coordonnées

Marie-Madeleine Bertucci
Professeure de Sciences du langage
Université de Cergy-Pontoise, Les Chênes 2, UFR LSH,
33, boulevard du Port, 95 011, Cergy-Pontoise cedex
EA 7518 LT2D
marie-madeleine.bertucci@u-cergy.fr

Mounia Illourmane
Doctorante en Sciences du langage
Université de Cergy-Pontoise, Les Chênes 2, EA 7518 LT2D
33, boulevard du Port, 95 011, Cergy-Pontoise cedex
mounia.illourmane@u-cergy.fr

Communication au colloque des 11 et 12 juin 2019, INALCO, Paris

Le corpus audiovisuel : Quelles approches ? Quels usages ?

Titre de la proposition de communication : *Transcrire un corpus audio dans la perspective de la préservation du patrimoine culturel immatériel. Les obstacles rencontrés, les méthodologies retenues, les questions restées en suspens.*

Mots clés : corpus audio, transcription, reconnaissance vocale, indexation, archivage, conservation du patrimoine immatériel.

Marie-Madeleine Bertucci

La communication a pour objet « l'étude de la mise en mots de la mémoire des « chibanis » du quartier de la Goutte d'or à Paris » à partir d'un corpus d'entretiens biographiques audio réalisés par Jean-Marc Bombeau, dans les locaux de la galerie associative l'Échomusée, située rue Cavé dans le XVIII^e arrondissement de Paris, avec un groupe de *chibanis* ou *anciens* du quartier de 2010 à 2012. Cette étude discursive mémorielle est le sujet de la thèse de Mounia Illourmane. Ces entretiens font émerger la mémoire de ces acteurs sociaux en les amenant à jeter un regard rétrospectif et réflexif sur leur passé, de leur arrivée à Paris dans les années soixante, dans le cadre d'une immigration de travail jusqu'à leur retraite qui prend effet dans la période contemporaine. Ils autorisent l'émergence d'un processus de patrimonialisation de données mémorielles, qui constituent le socle d'un patrimoine immatériel. Les entretiens ont visé à la fois le passé et le présent et ont cherché à mettre en lumière, à travers le travail sur la mémoire, le retour rétrospectif sur le passé, l'historicité du présent et les types de liens sociaux tissés à travers l'expérience de l'exil. Au plan théorique, la réflexion s'inscrit dans la perspective des cultural et des subaltern studies pour mettre en évidence le caractère fragmentaire et discontinu de ces récits mémoriels par le biais d'une étude du lexique privilégié

par les énonciateurs. Le patrimoine culturel immatériel, porté par la convention de l'UNESCO du 17 octobre 2003, et sa conservation supposent deux types de questionnements. D'une part, on l'a vu, il s'agit de concevoir différemment les politiques patrimoniales, en les envisageant du point de vue des détenteurs du patrimoine et non plus seulement du point de vue des experts, d'autre part d'élaborer des modalités de préservation des données autres que celles du patrimoine matériel, ce sur quoi va porter la communication. Cette recherche a donc pour objet de sauvegarder une expression culturelle encore vivante mais néanmoins menacée, la mémoire des Chibanis, et de mettre à la disposition de la communauté scientifique et d'un plus large public le cas échéant un ensemble de données peu accessibles matériellement et s'inscrivant dans le cadre du patrimoine culturel immatériel. Par ailleurs, ce projet s'inscrit dans le contexte de la réflexion menée actuellement en SHS sur la constitution de corpus d'archives numériques et de la pérennisation de leur accessibilité par la numérisation et l'indexation des sources.

La communication se déroulera en deux temps. Dans un premier temps, Marie-Madeleine Bertucci présentera la problématique générale des modalités de conservation du patrimoine mémoriel d'une communauté minorée, les chibanis du quartier de la Goutte d'Or dans une perspective sémiotique et sociolinguistique. Dans un second temps, Mounia Illourmane abordera les questions liées à la transcription du corpus qui se sont posées au fil de l'avancée du travail.

Marie-Madeleine Bertucci

I. Problématique générale des modalités de conservation du patrimoine mémoriel d'une communauté minorée, les chibanis du quartier de la Goutte d'Or dans une perspective sémiotique et sociolinguistique.

I.1 Position du problème. L'émergence du sujet migrant à travers le prisme de la mémoire

Le patrimoine culturel immatériel désormais (PCI), porté par la convention de l'UNESCO du 17 octobre 2003, et sa conservation supposent deux types de questionnements. D'une part, il s'agit d'élaborer des modalités de préservation des données autres que celles du patrimoine matériel, d'autre part, le PCI suppose de concevoir différemment les politiques patrimoniales, en les envisageant du point de vue des détenteurs du patrimoine et non plus seulement du point de vue des experts. Pour ce faire, on soulignera d'emblée la nécessité de faire émerger une mémoire de l'immigration qui permette à la fois i de changer le regard sur les migrants et ii d'éviter l'obstacle de l'enfermement dans des récits mémoriels qui constitueraient une assignation identitaire clivante, au sujet de laquelle on fera l'hypothèse qu'elle est un frein à l'intégration.

Les entretiens, qui constituent le corpus permettent de mettre en mots la mémoire de ces existences modestes et de les rendre visibles et audibles. Ils constituent un ensemble de connaissances, un savoir sur une forme d'immigration, discrète, voire invisible, alors que cette génération de migrants a été le témoin de transformations importantes et notamment de l'émergence d'une néo-traditionalisation des mœurs, voire d'un retour du religieux chez les jeunes générations. L'un des intérêts des entretiens, et non de moindres est de mettre en évidence ces effets de contraste entre ces deux générations. Par ailleurs, les entretiens constituent un témoignage sur le quartier de la Goutte d'Or et montrent en quoi celui-ci est à un lieu à la fois privé et public auquel s'attachent les mémoires de ces migrants et qui fait l'objet de discours. Ils constituent une méthodologie pertinente pour appréhender certaines trajectoires de migrants et pour faire émerger leur degré d'insertion dans la société française contemporaine à travers les types de liens sociaux qu'ils ont noués. La recherche tentera de mettre en évidence la présence d'éléments tangibles récurrents dans les entretiens, qui prennent la valeur de traces ou de supports de mémoire, comme la valeur des lieux, les formes de vécu, le potentiel émotionnel et thématique qui s'y rapportent (Bertucci, 2006). On fera l'hypothèse que l'identité des migrants se dit à travers des lieux et thèmes (topoi) privilégiés : par exemple, le rapport à la langue et aux lieux d'origine et le rapport à la langue française, la mobilité,

l'arrivée en France, l'insertion professionnelle et l'intégration, les liens avec les autres migrants, avec les français, l'hypothèse du retour au pays, la mémoire et l'oubli du passé.... On postulera donc des similitudes dans l'expérience des sujets concernés.

I.2 Les entretiens biographiques

On supposera qu'une impulsion narrative est sous-jacente aux entretiens et qu'elle fait apparaître chez les informateurs une impulsion à se représenter en se racontant (Bertucci, 2006). Les informateurs se situent parfois dans des situations de marginalité, voire de conflit ou de crise et sont confrontés à l'altérité. On peut penser que cette situation de distance peut les amener à mettre en mots et à relier des éléments isolés de leur histoire. Les entretiens biographiques rendent possible l'émergence d'un sujet qui ne peut être visible que par un détour par sa vie personnelle, laquelle en s'articulant au champ du social fait émerger une forme spécifique d'identité. Or, l'entretien biographique, donne à lire la biographie des individus. Il est un moyen d'accéder à la connaissance de l'autre parce qu'il met en place un ordre du sens (Adam, 1995) et lui donne une forme, une forme qui fait sens. En cela, il est particulièrement apte à exprimer l'altérité. Se saisissant comme objet et *quasi personnage*, le narrateur ou *quasi narrateur* (Ricœur, 1983) se pose à la fois en sujet et en objet, donnant de ce fait au lecteur ou à l'auditeur une possibilité d'appréhender des éléments identitaires au nombre desquels figurent les lieux selon notre hypothèse. La finalité de la recherche est de mettre en lumière l'expérience qu'a le sujet et le type d'identité qui en résulte dans une démarche herméneutique (Bertucci, 2006). Le processus réflexif est au centre de la méthodologie dans la mesure où il permet la construction et la représentation de l'autre dans la confrontation avec lui dans un espace social, historicisé et contextualisé. Cette approche paraît d'autant plus pertinente lorsque la recherche porte, comme ici, sur des identités potentiellement clivées entre les langues, les cultures, les appartenances, car la dynamique réflexive de l'entretien biographique amène les informateurs à construire le sens par eux-mêmes. On admettra également que l'organisation des discours mémoriels, malgré la diversité des informateurs, n'est pas forcément aléatoire et qu'il existe un ordre narratif signifiant, résultant d'une expérience de l'espace et du temps singulière, qui est celle de la mobilité, qu'elle soit exil ou migration (Bertucci, 2009 ; 2017). Enfin, on supposera que de par son caractère rétrospectif, les discours des informateurs impliquent un processus mémoriel particulier signifiant par ses souvenirs mais aussi par ses silences et qu'il y a une part de l'histoire du narrateur qui n'est pas remémorée.

I.3 Principes méthodologiques

La démarche prendra appui de manière très souple sur les méthodes mises en œuvre par la socio-histoire (Noiriel, 2006 a), les cultural studies (Thompson, 1963), les subaltern studies (Merle, 2004) avec l'objectif d'aller vers une conception de l'histoire perçue comme « une histoire anthropologique qui s'intéresse surtout aux liens que tissent les individus entre eux dans le cadre de leur vie quotidienne » (Noiriel, 2006 b : 40). Elle s'appuiera également sur une forme d'observation participante et une typologie des médiations sociales.

Cette démarche suppose un engagement particulier de l'observateur, une réflexivité en retour par la pratique d'une observation participante, supposant un décentrement socioculturel et ce que Lévi-Strauss appelle une « distanciation sociale » (1955 : 416). Comprendre l'autre et le connaître implique de se distancier de soi pour appréhender son expérience de l'exil, interprétée comme une expérience de rupture avec la culture d'origine. Les entretiens biographiques peuvent être interprétés comme trace de cette expérience. Une typologie des médiations sociales est également requise (Ferrarotti, 1990 : 61). On postulera donc qu'un individu reflète un ou plusieurs pôles sociaux et inversement que le système social se projette vers l'individu et ce à travers des médiations. Chaque comportement individuel est à la fois la synthèse d'une structure sociale et une individualisation de l'histoire sociale collective (Ibid. : 21). Il est la représentation, l'attestation de l'altérité, et donc il affirme ce que Bakhtine appelle la relation exotopique

(1978), laquelle rend possible de poser en tant que telle l'existence de l'Autre et son aptitude à parler en son nom.

Je passe maintenant la parole à Mounia Illourmane qui va aborder dans une seconde partie les différents problèmes posés par la transcription du corpus au fil de l'avancée de son travail de thèse.

Mounia Illourmane

« *La mise en mots de la mémoire des chibanis du quartier de la Goutte d'Or à Paris* », un sujet qui m'a été proposé par Marie-Madeleine Bertucci, ma directrice de thèse. L'idée est de sauvegarder et de mettre en mots la mémoire de ces vieux immigrés, en partant d'un corpus d'archives audio, de documents authentiques et d'enregistrements vocaux qui n'ont fait l'objet d'aucune manipulation au préalable. En effet, l'identification de certaines pistes de recherches est une étape qui s'avère nécessaire, suite à l'exploitation des données transcrites avec l'objectif de mettre au point des méthodologies de recherches et des approches pour mieux appréhender le traitement et l'analyse du corpus.

1. Les données

La récupération des données auprès de Jean-Marc Bombeau a eu lieu en décembre 2014 dans les locaux de la galerie de « l'Echomusée » dans le XVIII^{ème} arrondissement de Paris. Il s'agit de documents authentiques, séquences vidéos, bandes son et photos, peu accessibles matériellement. Ces données sont brutes et représentent une manne de ressources et de matériaux pour la constitution de notre corpus.

Mise au point d'un outil de classement, d'indexation et d'arborescence

Les données sont sauvegardées sur un disque dur. Avec ma directrice de recherche, on a réalisé un pré-test sur un échantillon afin de trouver une méthodologie d'organisation des données reçues. Par la suite, j'ai conçu un outil de mise en ordre des données en utilisant une méthodologie d'inventaire qui consiste à établir une liste exhaustive des éléments figurants dans le disque dur, en élaborant un inventaire physique qui correspond à un comptage manuel de chaque type de document dont les procédés sont le regroupement, le classement, l'indexation et l'arborescence.

Le disque dur est présenté ici comme le support de stockage des données à exploiter, sauvegardant ces dernières d'une manière hiérarchique, en subdivision, c'est là le principe de l'arborescence, cela s'apparente à des dossiers qui peuvent en contenir à leur tour d'autres sous dossiers subdivisés. Ce schéma de stockage et d'arborescence peut être recrée autant de fois que possible. Afin de déterminer avec exactitude la nature et la quantité des matériaux détenus, l'arborescence permet donc de voir la hiérarchisation du contenu, il s'agit de schémas qui permettent de mieux appréhender l'organisation des données. L'inventaire est donc composé de 132 pages format A4 avec 191 tableaux, présentés sous la forme suivante (colonnes de gauche à droite) :

- Document, type du document, remarque ;
- Bande son, durée, remarque ;
- Séquence vidéo, durée, remarque ;
- Photo, remarque.

À l'issue de cette démarche on a sélectionné les données à exploiter dans notre étude, ce qui a orienté notre choix sur les bandes son.

2. Génération d'un corpus

La recherche à partir des corpus oraux est très répandue en sociolinguistique. C'est une démarche empirique traitant des données authentiques. L'étape suivante consiste à transcrire et à baliser le corpus.

Les bandes son sont transcrites en orthographe standard. La partie transcrite comporte 213 pages et repose sur des enregistrements d'archive audio, composé de 18 bandes son. La durée totale des enregistrements est de 6 heures et 58.32 minutes. La durée varie d'une bande son à une autre, la plus courte est de 03.58 minutes et la plus longue est de 1 heure et 2.04 minutes. Dans le corpus on distingue au total 33 « quasi personnages » (Ricoeur 1983), dont 04 intervieweurs et 29 chibanis. Toutefois, une anonymisation totale du corpus s'avère nécessaire pour la protection des données personnelles des chibanis. De ce fait, nous avons choisi de ne pas constituer d'archive sonore. L'exploitation des bandes son se limite seulement à la version transcrite de celles-ci, ce qui sous-entend une non utilisation de la version sonore.

Prise en main des logiciels

Il s'agit d'une transcription mot à mot, j'ai mis par écrit exactement tout ce qui est dit à l'oral incluant des phrases incomplètes, des reformulations et des intonations, on y lira aussi les remarques d'ordre général « tu voulais que je te filme aussi », les mots répétés à l'oral « avait des problèmes, des problèmes avec leur enfant », des arrêts fréquents sont nécessaires pour ajouter les éléments extralinguistiques comme les rires et les hésitation. La transcription manuelle reste donc incontournable pour notre corpus. Ce type de transcription a principalement pour but la réalisation d'une retranscription authentique, tout en respectant les enchaînements et les chevauchements de l'enregistrement oral. Les transcriptions qu'utilisent les linguistes conservent soigneusement toutes les particularités des productions orales : répétitions, hésitations, amorces de mots, retouches. (Baude 2006 : 30)

C'est une technique très utile mais coûteuse en temps. Selon les estimations courantes, un minimum de trente minutes de travail est nécessaire pour transcrire une minute d'enregistrement. (Baude 2006 : 30). Cette opération nécessite l'utilisation d'un lecteur audio et d'un traitement de texte, entre lesquels il est essentiel de faire des va-et-vient successifs de Play/stops, pour les transcrire par saisie sur l'éditeur de texte « Word office ». Néanmoins, j'ai également utilisé les logiciels « Express Scribe » et « E-Speaking ».

Express Scribe¹ est un logiciel professionnel de lecture audio qui permet de lire des fichiers audio au format Wav ou MP3. Ce programme ne permet pas la conversion automatique des fichiers audio en texte, mais il contient d'autres fonctionnalités pour rendre le processus plus facile. Tel que la transcription, un outil conçu pour faciliter la transcription des enregistrements audio au clavier dans un éditeur de texte. La fonction de contrôle permet de contrôler la vitesse de lecture, qui se gère entièrement à partir des

¹ <https://www.nch.com.au/scribe/fr/index.html>

touches d'accès rapides du clavier. On peut sauvegarder, lire ou revenir en arrière en utilisant les touches retour, avance rapide. Le logiciel dispose d'une option de chargement automatique des fichiers audio. Elle est accessible à partir d'un périphérique (lecteur cd ou dvd, clé USB...) ou d'un réseau local.

E-Speaking², est un logiciel shareware³ de contrôle et de reconnaissance vocale, disponible en téléchargement. Il a pour fonction la dictée et l'exécution des tâches sur ordinateur par voix de l'utilisateur; Il assure de meilleurs résultats en s'intégrant à Word Office, il contient également des commandes intégrées comme *Voice Dictation* pour la dictée, elle consiste à reformuler sous forme intelligible et grammaticalement correcte les propos de l'utilisateur.

Pourquoi l'emploi de ces deux outils ?

Les systèmes de transcription doivent être adaptés en fonction des corpus à transcrire. (Baude 2006 : 31). Partant des différentes difficultés techniques rencontrées liées à la transcription du corpus, nous abordons ici certaines solutions pour y remédier. L'emploi de ces deux logiciels a permis la transcription des enregistrements des chibanis en restant le plus fidèlement possible à l'original tout en tenant compte des nombreuses contraintes d'ordre syntaxique, orthographique et phonologiques résultant du caractère oral des données.

Difficultés rencontrées

La transcription des entretiens présente des difficultés d'ordre pratique. La première difficulté est liée au caractère oral des données d'un point de vue syntaxique, orthographique et phonologique. Les logiciels de transcription automatique ne correspondent pas à notre corpus, il a donc fallu tirer profit des logiciels à notre disposition, cependant nous sommes obligées de recourir à un traitement majoritairement manuel, même si ces logiciels facilitent grandement la tâche de transcription des données. Afin de contrôler au mieux la transcription intégrale une telle démarche est nécessaire, le travail manuel et l'assistance des outils informatique est incontournable.

De plus, la reconnaissance vocale n'a pas pour option d'interpréter et de retranscrire d'autres voix que celle de l'utilisateur. Dès lors, la seule utilisation possible est de répéter à voix haute devant un microphone les propos tenus par les chibanis pour que le logiciel les retranscrive en texte. Étant donné que les logiciels de reconnaissance vocale ne prennent pas en compte les différents accents des locuteurs, il est essentiel de revoir et de corriger les textes retranscrits.

3. Analyse et traitement des données recueillies

Il s'agit de réaliser une analyse de contenu des entretiens, cette étude est menée selon une méthodologie d'analyse visant à extraire les éléments pertinents qui constitueront la base du sujet. L'analyse de contenu cherche à savoir ce qui est derrière les paroles sur lesquelles

² <https://www.nch.com.au/scribe/fr/index.html>

³ Payant, disponible avec souscription à un abonnement.

elle se penche. La linguistique est une étude de la langue, l'analyse de contenu est une quête, à travers des messages, de réalité autre. (Bardin 1977 : 48).

Une fois les entretiens transcrits, je procède au décortilage des données brutes, le discours des personnes interviewées contient des informations à exploiter. Tout document parlé, écrit ou sensoriel contient potentiellement une quantité d'informations sur la personne qui en est l'auteur, sur le groupe auquel elle appartient, sur les faits et événements qui y sont relatés, sur les effets recherchés par la présentation de l'information, sur le monde ou sur le secteur du réel dont il est question. (Mucchielli 1984 : 17). En effet, l'analyse de contenu s'appuie sur la catégorisation quantitative et qualitative. C'est une technique de recherche pour la description objective, systématique et quantitative du contenu manifeste de la communication. (Berelson 1952 cité dans Ghiglione & Matalon 1978 : 155). Elle permet de concilier les deux méthodes et de réaliser un tri statistique de l'ensemble des données recueillies.

En effet, c'est un travail de synthèse, afin d'établir un inventaire des mots les plus courants et des thèmes les plus saillants, pour réaliser une analyse lexicale d'une part et une analyse thématique d'autre part. L'idée d'appliquer sur le corpus la méthode d'analyse lexicale est proposée par Marie-Madeleine Bertucci et réalisée par Mounia Illourmane. À partir des résultats de cette première démarche, j'ai réalisé l'étude thématique, ainsi, l'analyse qualitative se fonde sur les résultats quantitatifs. L'analyse de contenu peut être une analyse de "signifiés" (ex : l'analyse thématique) mais elle peut être aussi une analyse des "signifiants" (analyse lexicale) (Bardin 1977 : 39).

Conclusion

Mounia Illourmane

Pour conclure, on notera que certaines questions autour de l'exploitation de l'inventaire et de l'arborescence hiérarchique du disque dur restent en suspens, finalement, comment peut-on exploiter au mieux cette base de données ? Quelle méthodologie d'étude adopter ?

Marie-Madeleine Bertucci

L'intérêt de ce travail est qu'il permet de conserver les traces de voix minorées, celles des chibanis du point de vue du patrimoine culturel immatériel et d'analyser l'intérêt du support audio dans cette perspective. Ce type de support ne va pas sans poser de problèmes comme a essayé de le montrer Mounia Illourmane. La question restée en suspens est celle de l'utilisation de l'arborescence réalisée pour indexer et classer les données de départ dans la perspective de l'analyse du corpus et l'impact éventuel de ce type de description sur les modalités de conceptualisation de la thématique et le déroulement de la recherche.

Bibliographie

- Adam, J.-M. 1995. « Aspects du récit en anthropologie ». In Adam, J.-M., Borel, M.-J., Calame, C., Kilani, M. *Le discours anthropologique, description, narration, savoir*. Lausanne : Payot. pp. 251-283.
- Bakhtine, M. 1978 [éd. fr.]. *Esthétique et théorie du roman*. (Olivier, D. trad.) Paris : Gallimard.
- Bardin, L. 2003. *L'analyse de contenu*. Paris : Presse universitaire de France.
- Baude, O. *Corpus oraux : les « bonnes pratiques » des linguistes*. Laboratoire Ligérien de Linguistique UMR 7270. pp. 1-104.
- Baude, O. (dir) 2006. *Corpus oraux. Guide des bonnes pratiques*. Paris : CNRS Éditions.
- Berelson, B. 1952. « Content Analysis in Communication Research » in Ghiglione, R., Matalon, B. 1978. *Les Enquêtes Sociologiques*. Paris : Armand Colin.

- Bertucci, M.-M. 2006. « La fonction du récit de vie dans la constitution de l'identité plurilingue chez les migrants : quelques hypothèses ». In Molinié, M., Bishop, M.-F. (coord.). *Autobiographie et réflexivité*. Amiens : UCP / CRTF. Encrage. pp. 127-140.
- Bertucci, M.-M. 2006. *Plurilinguisme et altérité. Français, école, politiques linguistiques-éducatives*. Document de synthèse présenté pour l'obtention de l' Habilitation à diriger des recherches. Université François Rabelais de Tours.
- Bertucci, M.-M. 2009- 2017, (J. Assier et E. Chemblette, coll. pour la transcription du corpus) *Mémoires de l'immigration, vers un processus de patrimonialisation ?* Rapport de recherche universitaire pour le Ministère de la Culture et la Cité nationale de l'Histoire de l'immigration. Université de Cergy-Pontoise. 1 vol. (170 p.). Disponible à la BNF. URL <http://catalogue.bnf.fr/ark:/12148/cb42101812c>. Rapport mis en ligne en 2017 sur le site du Ministère de la Culture et de la Communication, Page du Patrimoine Ethnologique. <http://www.culturecommunication.gouv.fr/Thematiques/Patrimoine-ethnologique/Soutien-a-la-recherche/Travaux-de-recherche/Liste-des-travaux-de-recherche-par-mots-cles/Migration>
- Ferrarotti, F. 1990. [1ère éd. 1983]. *Histoire et histoires de vie. La méthode biographique dans les sciences sociales*. Paris : Méridiens Klincksieck.
- Lévi-Strauss, C. 1973. [1ère éd. 1955]. *Tristes Tropiques*. Paris : Plon.
- Merle, I. 2004. « *Les subaltern studies*. Retour sur les principes fondateurs d'un projet historiographique de l'Inde coloniale ». In *Genèses*. 56. pp. 131-147.
- Mucchielli, A. (dir) 1996. *Dictionnaire des méthodes qualitatives en sciences humaines et sociales*, Paris : Armand Colin.
- Mucchielli, R. 1984. *L'analyse de Contenu des Documents et Communications* ». Paris : ESF.
- Noiriel, G. 2006 a. *Le creuset français. Histoire de l'immigration XIX^e-XX^e siècles*. Paris : Seuil.
- Noiriel, G. 2006 b. *Introduction à la sociohistoire*. Paris : La Découverte.
- Paillé, P., Mucchielli, A. 2012. *L'analyse qualitative en sciences humaines et sociales*. Armand Colin.
- Ricœur, P. 1983. *Temps et récit*. T.1. Paris : Seuil.
- Thompson, E.P. 1975. [1ère éd. 1963]. *The making of the english working class*. Londres : Penguin book.

Sitographie pages actives en mai et juin 2019

<http://catalogue.bnf.fr/ark:/12148/cb42101812c>

<https://www.nchsoftware.com/fr/index.html>

<https://e-speaking.com/>

www.unesco.org

<http://www.culturecommunication.gouv.fr/Thematiques/Patrimoine-ethnologique/Soutien-a-la-recherche/Travaux-de-recherche/Liste-des-travaux-de-recherche-par-mots-cles/Migration>