Véronique Ginouvès

# *Citation et éditorialisation des enquêtes orales*

Le fonds sonore de la phonothèque de la Maison méditerranéenne des sciences de l'homme a rejoint l'institution en 1997, après s'être constitué en 1979 au sein du Centre de recherches et d'études sur l'histoire orale et les parlers régionaux (CREHOP)[1]. À la suite de ce déménagement, les collections sonores se sont élargies à l'espace méditerranéen et à l'ensemble des disciplines de sciences humaines et sociales. Les chercheurs des laboratoires constitutifs de l'institution ont alors découvert l'existence de la phonothèque. Certains s'en sont emparés pour déposer avec intérêt et une certaine insouciance leurs entretiens enregistrés, d'autres s'en sont inquiétés, affirmant l'impossibilité du dépôt de leurs sources. Le temps passant, et le dispositif mis en œuvre devenant plus performant ou mieux compris, plusieurs d'entre eux ont pris la décision d'archiver les entretiens issus de leur recherche. Ainsi, la phonothèque de la MMSH est aujourd'hui riche de plus de 8 000 heures déposées par les chercheurs en SHS dont plus de 3 000 accessibles en ligne. Ce texte s'interroge sur l'évolution de la prise en compte de la citation scientifique des sources orales et de son accès, à

1. Crivello, 2014.

partir de l'expérience de la gestion d'un fonds sonore d'une phonothèque de recherche.

## Quarante années d'évolutions

Lorsque la phonothèque intègre la MMSH, les archives des chercheurs en sciences humaines et sociales sont de diverses natures. Elles se constituent non seulement d'entretiens enregistrés mais aussi de carnets et de notes prises sur le terrain, de papiers, de photographies, de films, etc., c'est-à-dire de toutes les sources qu'ils ont recueillies ou créées au cours de leur réflexion et de leur travail sur les sociétés étudiées[2]. Intégrée dans une médiathèque lorsque la MMSH est bâtie, la phonothèque a placé la question du statut de l'oral comme source pour les chercheurs au cœur de ses missions dès sa création[3], avec un triple objectif : tout d'abord, permettre l'administration de la preuve par le retour à la source, les entretiens enregistrés, car, jusque-là, seules les transcriptions – pourtant souvent partielles – réalisées par l'auteur semblaient suffisantes aux yeux des pairs ; ensuite, conserver la source sur laquelle s'appuient les publications et la citer correctement avec l'idée qu'on ne détruit pas une source constituée, fût-elle orale et exploitée ; enfin, faciliter une potentielle comparaison des données en documentant les archives déposées et en constituant une masse critique conséquente pour une science cumulative grâce à l'usage d'autres chercheurs. À ces missions s'en est très vite ajoutée une quatrième, amplifiée en 2003 par la signature de la Convention du Patrimoine culturel immatériel de l'UNESCO, celle de la nécessité de prendre en compte le retour d'expérience auprès des communautés ou des populations étudiées, explicitant clairement le double engagement de la phonothèque vis-à-vis des chercheurs et, tout autant, de ceux qui avaient été enregistrés[4].

2. Molinié et Mouton (eds), 2008.
3. Pelen, 1990.
4. Voir https://ich.unesco.org/fr/convention (consulté en juin 2020).

La traversée de ces deux périodes, le passage d'un laboratoire à une Maison des sciences de l'homme, a été aussi celui du passage de l'analogique au numérique qui, à l'aube du XXIe siècle, modifie radicalement les usages des documents sonores, leur archivage et donc leur citation[5]. Avec l'analogique, si l'accès aux données était chronophage, le support lui-même permettait aux chercheurs de ranger et classer les entretiens et de les annoter directement sur les emballages des boîtiers. Avec le numérique, difficile pour le chercheur de lier de façon intrinsèque l'information signalétique et le fichier sonore. L'ordre des bandes et leur organisation, pensée ou souvent cotée par le chercheur, disparaissent également. Petit à petit, de nouveaux réflexes savants se mettent en place pour organiser le nommage des fichiers et pour faire en sorte que la citation sonore dématérialisée revienne dans les habitudes des chercheurs. Les citations se codifient et le logiciel Zotero – créé en 2006 – a rapidement intégré dans ses typologies les documents sonores[6]. Par exemple, le logiciel a développé les différentes fonctions du *creator* des *audio recording* (responsabilité auteur du document sonore) proposant celles de *performer*, *composer*, *contributor* ou *words by* (artiste-interprète, compositeur, contributeur ou voix de).

## Une dynamique retrouvée

Les citations sonores trouvent aujourd'hui une place plus légitime au cœur des usages scientifiques. Les revisites des grandes enquêtes en sciences sociales qui se sont multipliées ont impliqué de nouvelles exigences de réflexivité sur cette question[7]. Plusieurs thèses soutenues en 2018 illustrent cette dynamique, s'appuyant sur des sources orales déjà constituées et déposées dans des centres d'archives : Bénédicte Girault a réécouté plus de 1 100 heures issues de 284 témoignages de témoins archivés au sein du Service d'histoire de l'éducation ;

5. Crivello, 2014.
6. Descamps, 2001 et 2019.
7. Laferté, Paquali et Renahy (eds), 2018.

Ariane Mak a réécouté 130 témoignages issus pour la plupart des archives du Mass Observation (université de Sussex) ; enfin Maëlle Mériaux a réécouté plus de 200 heures issues de la phonothèque de Dastum, en Bretagne[8]. Chacune cite les sources sonores dans sa bibliographie, aucune ne permet véritablement l'écoute mais propose *a minima* le lien vers une référence dans le centre d'archive et, parfois, une URL dans une base de données. Ariane Mak pose des questions spécifiques qui portent, en sus de la transcription, sur la traduction des entretiens et sur leur citation dans les deux langues. Des préambules précisent les usages et codifient l'interprétation de la lecture de cette oralité. Bénédicte Girault, quant à elle, explique : « Les transcriptions d'entretiens apparaissent en retrait du texte. Elles sont en italique pour distinguer leur caractère oral des autres sources ou citations. Afin de conserver le contexte, les citations transcrites sont parfois longues. Les passages ciblés par l'analyse sont alors signalés en gras pour faciliter la lecture. Sauf indications contraires, tous les passages en gras sont soulignés par l'auteure[9] ». Maëlle Mériaux indique pour sa part : « Les citations des informateurs figurent en italique. Au-delà de trois lignes, elles sont également mises en exergue, dans un paragraphe distinct. Elles sont suivies du nom de l'informateur et de la date de l'échange pour les entretiens enregistrés que nous avons réalisés. Elles sont anonymisées et suivies de la localité géographique et du nom du fonds d'archives orales lorsqu'elles proviennent des enregistrements de témoignages étudiés. Nous avons effectué des choix de transcription, en tentant de rester au plus proche des propos de nos interlocuteurs. Les seules interventions que nous nous sommes permises sur les citations concernent la restauration des négations (« Elles sont pas » devient « Elles [ne] sont pas ») et la suppression des contractions de langage (« Les enregistrements qu'étaient » devient « Les enregistrements [qui étaient] »). Parfois, nous avons également ajouté des mots ou segments de phrases

8. Girault, 2018 ; Mak, 2018 ; Mériaux, 2018.
9. Girault, 2018, p. 10.

entre crochets, lorsque ceux-ci pouvaient expliciter quelque chose d'implicite dans le discours[10] ».

Le mouvement de la citation accompagnée de la source sonore est donc en cours et la facilitation de la manipulation des objets sonores numériques commence à se mettre en place. À l'ère analogique, repérer une séquence et la réécouter au même minutage nécessitait une écoute en temps réel et l'annotation restait compliquée à mettre en œuvre car, d'une copie sonore à l'autre, sur un même support, les durées pouvaient varier de quelques secondes et les éventuels compteurs des machines ne correspondaient pas sur tous les lecteurs. Pour pouvoir citer ces sources et les réutiliser, nombre d'institutions culturelles ou associations qui détenaient des fonds sonores ont mis en place un système de « découpage » des enquêtes enregistrées qui soulevaient des questions méthodologiques mais aussi déontologiques[11]. Le Musée des arts et traditions populaires (MNATP) a été l'une de ces institutions qui ont poussé la méthode assez loin, modifiant la source en créant une forme de focus sur ce qui était considéré comme « œuvres » et gommant le contexte de production[12]. L'annotation d'une séquence précise est sans nul doute ce qui est le plus difficile à matérialiser et ce qui a ralenti l'usage direct de la source sonore : sortie de son contexte, elle est parfois devenue muette.

## Des questions légales et éthiques vives

Le *blogging* scientifique, en facilitant l'éditorialisation de la source orale, a permis la citation contextualisée. À la phonothèque de la MMSH, l'exemple le plus souvent mis en avant est celui des enquêtes ethnologiques réalisées auprès de plusieurs communautés de luthiers par Hélène Claudot-Hawad[13]. Une vie éditoriale des extraits sonores rassemblés

10. Mériaux, 2018, p. 11.
11. Mériaux, 2018.
12. Gétreau et Colardelle, 2003.
13. Claudot-Hawad et Ginouvès, 2017.

pour cette recherche est encore à écrire, qu'il s'agisse des commentaires du carnet, des discussions sur les forums de lutherie ou d'instrumentistes, des sites de luthiers, des articles wikipedia, etc., car les réusages culturels sont nombreux. Il faudrait également affronter la question de la durée de la citation du fichier sonore (ou vidéo) qui n'est pas toujours très facile à caractériser. Le sujet est abordé sous plusieurs angles sur le carnet de recherche *Questions éthiques et juridiques*. Le billet du 17 juillet 2017, « L'exception pédagogique et de recherche », en particulier, revient sur la question de la durée de citation des extraits musicaux et des vidéos protégés[14]. Pour les entretiens de terrain, la durée des extraits ou la possibilité de mettre en ligne tout ou partie de l'entretien sont soumises aux autorisations des témoins et des enquêteurs. Là encore, plusieurs billets sur le site informent sur la question de la rédaction d'un contrat d'utilisation[15].

L'une des raisons qui empêche le dépôt des sources dans un centre d'archive est, d'une part, l'inquiétude que posent les questions légales et éthiques et, d'autre part, l'impossibilité du centre d'archive de permettre l'utilisation de ces sources parce que ces questions légales et éthiques n'ont pas pu être éclaircies[16]. Le Plan national pour la science ouverte va peut-être permettre une évolution[17]. Jusqu'ici aucune injonction n'obligeait les chercheurs de sciences humaines et sociales à déposer leurs sources ; ils vivaient dans la quiétude de savoir que leurs notes de terrain et leurs entretiens ne seraient jamais consultés, ce qui leur donnait une grande liberté. Pourtant – sauf si elles étaient détruites – ces données avaient déjà toutes les chances d'être un jour archivées dans un centre de ressource. Sans que des règles ne soient réellement établies, un chercheur sait qu'il sera confronté au

14. Ainsi, chaque extrait vidéo ne doit pas dépasser 15 % de la durée totale de l'œuvre audiovisuelle et, pour les extraits musicaux protégés, 30 secondes sachant que la totalité des extraits ne doit pas dépasser 15 % de la durée totale de l'œuvre : Stérin, 2017a.
15. Fellous-Sigris, 2018 ; Stérin, 2017b ; Ginouvès, 2017.
16. Zeitlyn, 2019.
17. Vidal, 2018.

partage des sources, et ce pas seulement parce que le vent de la patrimonialisation a soufflé. Il s'agit bien de demander aujourd'hui le dépôt de leurs matériaux de terrain pour une science cumulative, une historisation de la discipline et son articulation avec la preuve... La résolution des questions juridiques et éthiques permet de mieux comprendre le contexte de production et la façon dont ces sources pourront être citées et réutilisées.

En 2011, la phonothèque a mis en place un groupe de travail pour apporter des éléments de réponses à ces problématiques et proposer des solutions concrètes aux questions qui se posent en matière de production et d'exploitation des sources numériques[18]. Chercheurs comme professionnels de l'information scientifique et technique et du droit ont ainsi croisé leurs approches en matière de diffusion des données. Un guide de bonnes pratiques a été publié en novembre 2018 aux Presses universitaires de Provence. Il n'ambitionne pas de donner une seule réponse mais il essaie plutôt de réfléchir à cette question par projet, par type de fonds, en fonction du contexte. Le carnet de recherche précédemment cité publie en parallèle des billets très concrets sur les questions que se posent les chercheurs, depuis la collecte jusqu'à la diffusion des informations[19].

En conclusion, il n'aura jamais été aussi simple avec le numérique de diffuser et de mobiliser des contenus de toute sorte en même temps, mais citer et publier ses sources demeure compliqué si une réflexion sur les questions éthiques et juridiques n'a pas été menée en amont. En ce sens, la phonothèque de la MMSH est devenue un lieu de référence déontologique pour les chercheurs qui s'appuient sur des entretiens dans leur démarche et souhaitent que leur méthode soit prise en compte à part entière. Ces questionnements viennent croiser ceux qui se posent de façon cru-

18. Le groupe de travail « Questions d'éthique et Droit » s'est constitué au sein du réseau des Maisons des sciences de l'homme et de la TGIR (Très grande infrastructure de recherche) Huma-Num, en particulier, le consortium des ethnologues.

19. Voir https://ethiquedroit.hypotheses.org/ (consulté en juin 2020).

ciale aujourd'hui sur l'intégrité scientifique, en particulier sur le partage des sources et leur citation de façon correcte, comme on le voit dans le dossier 2018 de *Nature* « How to grow a healthy lab », ou comment construire un laboratoire productif, rigoureux et en bonne santé.