

Rapport sur le projet « Pour une archive des langues parlées en interaction.
Statuts juridiques, formats et standards, représentativité » financé par le *Programme Société de
l'Information / Archivage et patrimoine documentaire*, mars 2005.

Constitution de corpus de parole-en-interaction et respect de la vie privée des enquêtés : une démarche réflexive

Lorenza Mondada

Laboratoire ICAR (UMR5191, CNRS, Univ. Lyon2, ENS LSH)

Les questions juridiques se posent de plus en plus en linguistique à propos de la constitution et de l'exploitation de grands corpus de données orales. Dans quelles conditions peut-on enregistrer la parole d'informateurs ou de participants ? dans quelle conditions peut-on exploiter ces données ? quels traitements particuliers faut-il leur faire subir avant de diffuser les analyses et les données elles-mêmes ? quelles contraintes observer lors de cette diffusion ? Telles sont quelques-unes des questions que se posent aussi bien aux groupes de recherche expérimentés qu'aux doctorants qui débutent le recueil de données pour leur thèse.

Ce texte a pour double objectif de clarifier et de problématiser les questions *éthiques et juridiques de respect et protection de la vie privée* des enquêtés dans le cadre de pratiques de recherche visant la *constitution de corpus de parole-en-interaction*¹. Cette problématisation est issue d'expériences de terrain en linguistique interactionnelle : en effet, la constitution des corpus y joue un rôle fondamental, qui

¹ Ce texte a été rédigé suite à un travail de 2 ans dans le cadre du projet « *Pour une archive des langues parlées en interaction. Statuts juridiques, formats et standards, représentativité* » proposé par le laboratoire ICAR (UMR5191 du CNRS, Univ-Lyon2 et ENS LSH) en réponse à l'appel à projet du « Programme Société de l'Information / Archivage et patrimoine documentaire. Apports des sciences de l'information et de la cognition » (2003-2005).

En premier lieu, ces questions ont été élaborées et débattues au sein du groupe des chercheurs d'ICAR – Lukas Balthasar, Michel Bert, Sylvie Bruxelles, Lorenza Mondada, Christian Plantin, Véronique Traverso – en collaboration avec Isabelle de Lamberterie (CNRS), spécialiste des questions juridiques se posant dans les projets et les pratiques de la recherche scientifique.

En un deuxième temps, ce travail a été poursuivi dans un groupe national de réflexion sur les questions légales et les corpus oraux soutenu par le DGLF-LF et dont le résultat du travail sera publié sous la forme d'un *Guide de bonnes pratiques*. Participent à ce groupe Olivier Baude, Claire Blanche-Benveniste, Marie-France Calas, Paul Cappeau, Pascal Cordereix, Isabelle de Lamberterie, Michel Jacobson, Christiane Marchello-Nizia, Fabrice Mollo, Lorenza Mondada, Jean Sybil. La première version provisoire du guide est disponible sur le site : <http://www.culture.gouv.fr/culture/dglf/cotelecharger.htm>

nourrit les problématiques éthiques et juridiques grâce à des exigences fortes en matière de données « naturelles » et de degré de granularité des transcriptions. Ces exigences rendent à la fois sensibles les chercheurs aux questions éthiques et juridiques et en même temps permettent une problématisation de ces questions à travers un regard analytique et réflexif (voir pour différentes positions à ce sujet Pollner, 1993 ; Watson, 1987 ; Woolgar, 1988) qui en montre la complexité et les injonctions parfois paradoxales.

C'est pourquoi ce texte commence par introduire aux enjeux de la linguistique interactionnelle (1.) montrant qu'ils produisent un regard analytique particulier envers la définition des modes de participation à l'enquête et la notion de « participant » elle-même (2.). Sur ces bases, nous traiterons de deux questions juridiques et éthiques fondamentales, celle du consentement éclairé (3.) qui est au fondement de l'établissement des relations avec les enquêtés et celle des opérations d'anonymisation des données (4.) qui est au fondement de l'instauration d'une certaine image des enquêtés dans les transcriptions.

1. Introduction : linguistique interactionnelle, constitution de corpus et pratiques de terrain

La linguistique interactionnelle est une discipline relativement jeune, qui apparaît comme telle à la fin des années 90 (le label de *linguistique interactionnelle* a commencé à circuler dans les textes de Selting & Couper-Kuhlen, 2000, Mondada, 2001), mais qui a émergé à partir de la diffusion de l'analyse conversationnelle (Sacks, Schegloff, Jefferson, 1974 ; pour une présentation voir ten Have, 1998 ; Gülich & Mondada, 2002) en linguistique dès les années 70.

L'analyse conversationnelle est un courant à son tour issu, dès les années 60, de l'ethnométhodologie (Garfinkel, 1967, pour une présentation voir Heritage 1984, 1992) ; il développe une « mentalité analytique » qui a pour objectif la description des procédés régissant les activités sociales en contexte. Tout en étant visées dans leur diversité, ces pratiques sociales ont été abordées par Harvey Sacks et Emanuel Schegloff dès les débuts de leur travail par des *enregistrements audio* de conversations en face à face ou au téléphone, voire d'autres interactions (comme des appels d'urgence à un centre pour la prévention du suicide, ou des sessions thérapeutiques avec des adolescents). Cette attention pour les données enregistrées en audio ne dérive pas à l'origine d'un intérêt particulier pour la langue parlée, mais relève du fait que les enregistrements permettent une réécoute infinie de ce qui s'est passé, et donc un retour constant sur les événements à étudier dans leur intégralité. C'est ainsi qu'une description fine des procédés par lesquels les participants organisent leur parole – entendue comme une activité sociale parmi d'autres – a vu le jour et a intéressé de plus en plus, outre que les sociologues, les linguistes (Sacks, 1984, 1992 – voir Schegloff 1992 pour des éléments de contextualisation historique de la démarche et Ochs, Schegloff, Thompson, 1996 ; Selting & Couper-Kuhlen, 2002 pour les premiers recueils de linguistique interactionnelle).

Dans ce cadre, plusieurs problématiques analytiques constituent un apport important à la sensibilisation aux questions juridiques et éthiques – même si celles-ci n'ont pas fait

jusqu'à présent l'objet de traitement explicite en analyse conversationnelle. Nous en releverons trois : en analyse conversationnelle (AC), l'importance attribuée aux données naturelles et l'approche qui en découle (1.1), dans la démarche de Harvey Sacks, cofondateur de l'AC, l'importance de l'analyse des dispositifs de catégorisation (1.2), en ethnométhodologie, courant dont est issue l'AC, un regard qui traite l'enquête comme une activité sociale ordinaire que l'on peut soumettre à l'analyse (1.3).

1.1. Analyse conversationnelle et données « naturelles »

L'analyse conversationnelle est un des courants qui a contribué le plus, dans le contexte international des SHS des dernières décennies, à défendre et illustrer la nécessité de travailler sur des corpus de données enregistrées authentiques. Cette nécessité va de pair avec une série d'exigences analytiques qui définit une attention cohérente pour les détails de la parole des participants ; celle-ci nous semble être un socle fondamental pour aborder les questions juridiques et éthiques relatives au traitement de cette parole.

- l'exigence de travailler sur des « données naturelles » signifie que les chercheurs en AC choisissent d'analyse des activités enregistrées dans leur contexte social d'accomplissement, qui auraient eu lieu même en absence du chercheur, qui se répètent éventuellement de manière ordinaire au fil du temps et qui surtout n'ont pas été orchestrées en tant que telles par l'enquêteur. Cette exigence se distancie d'autres manières de constituer des corpus de données, par exemple par entretien ou autres formes d'élicitation, par protocole expérimental en laboratoire, ou par imposition de tâches à effectuer dans un contexte qui tout en étant plus « écologique » que le laboratoire reste « expérimental » en étant contrôlé et imposé par le chercheur.
- Cette exigence de travailler sur des données naturelles va de pair avec la reconnaissance du caractère *situé* de l'organisation des conduites langagières et sociales, c'est-à-dire de leur ancrage dans un contexte social qui exerce des effets réflexifs sur leur configuration. Autrement dit, les modes d'organisation de la parole sont sensibles au contexte et changent avec celui-ci (ils sont *indexicaux*).
- En outre, l'AC s'intéresse aux données naturelles parce qu'elle considère, comme l'ethnométhodologie, que les *détails* des pratiques – que les enregistrements rendent observables et donc analysables – sont significatifs pour l'organisation de l'action non seulement pour l'analyste mais aussi – et avant tout – pour les participants eux-mêmes. Autrement dit, les participants s'orientent vers ces détails pour comprendre ce qui est en train de se passer et pour s'y ajuster. Il en découle que la question de la visibilité, observabilité, reconnaissabilité et interprétation des détails est au cœur de la démarche de l'AC, et qu'elle est un problème pratique sans cesse abordé et résolu par les participants. Les détails de l'action sont ainsi traités de manière endogène (ou *emic*) dans la perspective des participants – et non de manière exogène (ou *etic*) dans la perspective unique externe des analystes.

Il découle de ces exigences celle, fondamentale, de préserver les détails de l'interaction en contexte ; celle-ci rend attentif à la fois à l'importance et à la pertinence du détail et aux effets interprétatifs que ces détails peuvent générer. Lorsqu'on traite des problèmes

éthiques et juridiques du respect de la vie privée des personnes concernées par l'enquête, ces questions sont fondamentales pour évaluer les détails intervenant dans la production du sens de leurs actions et de leur identité.

1.2. Les procédés de catégorisation

On considère généralement que Harvey Sacks a développé en pionnier deux types d'approches qui vont marquer l'analyse conversationnelle et les courants qui en sont issus : l'analyse séquentielle et l'analyse catégorielle. Cette dernière s'appuie sur un constat simple : pour catégoriser les personnes, alors même qu'il existe un nombre élevé de possibilités (pour identifier une personne on peut recourir à des catégories ethniques : « c'est un français », professionnelles : « c'est un prof », religieuses : « c'est un catho », politiques : « c'est un homme de gauche », ou autres : « c'est un fin gastronome ») une seule de ces possibilités fait sens et est généralement utilisée à la fois (selon la règle d'économie de Sacks) (ainsi, en réponse à la question de savoir qui a tué un jeune dans une banlieue on répondra par une seule catégorie : « c'est un policier », « c'est un dealer », etc., qui produira une compréhension spécifique de l'événement). Sacks montre ainsi à la fois les enjeux du choix de la catégorie pertinente (qui produit des inférences éclairant l'action) et les procédés par lesquels les participants rendent pertinents, en situation et en relation avec l'activité, une catégorie plutôt qu'une autre (Sacks, 1972a, 1972b, 1992).

Cette réflexion est importante pour la discussion des enjeux éthiques et juridiques de l'enquête dans la mesure où elle montre que le choix des catégories pour désigner les participants n'est pas anodin et produit des effets massifs en termes d'interprétation - voire de stigmatisation - des actions effectuées par les participants ainsi catégorisés. Ces interprétations ont un statut différent quand il s'agit de catégories émergeant du terrain lui-même, permettant de comprendre ce qui s'y passe, ou de catégories venant du chercheur, pouvant ne pas correspondre avec les catégories des participants et imposer, dans ce cas, des contraintes artéfactuelles sur la description de l'action.

En outre, la manière dont les participants font référence à des personnes - co-présentes ou absentes - dans l'interaction obéit à des pertinences et à des contraintes à la fois locales et systématiques (voir Sacks & Schegloff, 1979, Schegloff 1996) : elles interviennent dans la façon dont les participants eux-mêmes traitent la référence aux personnes, en choisissant les formes pertinentes par rapport à leur position dans la séquence et par rapport au récipiendaire à qui ils s'adressent. Ces orientations des participants peuvent avoir une importance cruciale au moment où il s'agit d'anonymiser les références aux personnes et tout ce qui les manifeste comme étant connues ou non partagées, aproblématiques ou délicates.

1.3. L'ethnométhodologie et l'analyse des pratiques scientifiques

L'analyse conversationnelle a prolongé le programme d'études de l'ethnométhodologie, qui comme elle se focalise sur l'étude des pratiques sociales. Dans ce contexte, ce ne sont pas uniquement les pratiques langagières à avoir fait l'objet d'analyses, mais aussi - et de plus en plus ces dernières années - les pratiques professionnelles et expertes. De

nouveaux courants ont ainsi vu le jour, comme les *studies of work* (Garfinkel, 1986) ou les *workplace studies* (Luff, Hindmarsh, Heath, 2000), qui se sont penchés sur les activités au travail, impliquant non seulement la parole mais souvent la coordination de l'action dans des collectifs où le rôle des technologies et des artefacts est fondamental (Button, 1993, Heath & Luff, 2000).

Dans ce contexte, une approche des pratiques scientifiques – et en particulier des pratiques de terrain et d'analyse en linguistique – est envisageable : Lynch (1985) l'a fait pour les laboratoires, nous l'avons fait pour des groupes de travail en SHS (Mondada, 2005) ou en linguistique (Mondada, 1998, à paraître a, à paraître b). L'objectif est de comprendre comment les objets de savoir émergent spécifiquement de manières de faire, de modes d'enquêtes, de choix technologiques, en étant imbriqués dans des pratiques situées.

Ce regard *réflexif* permet de se pencher sur les problèmes juridiques et éthiques d'une manière qui n'est pas simplement normative et prescriptive, mais d'une manière qui comprend, décrit et problématise les pratiques de recherche, vues comme des objets à soumettre à l'analyse et non simplement comme des ressources – tacites, échappant à la thématization - pour produire du savoir scientifique (Mondada, à paraître a).

Sur la base de ces acquis, les questions juridiques et éthiques se posent d'une façon qui est étroitement imbriquée dans les enjeux scientifiques – et qui ne constitue d'aucune façon une simple couche supplémentaire de problèmes étrangers à l'analyse.

2. Modes de participation à l'enquête et définition des participants

Afin de mieux situer ensuite les problèmes éthiques et juridiques qui se posent au chercheur, nous allons ici nous pencher sur les acteurs intervenant dans l'enquête sur le terrain pendant laquelle sont constitués des corpus. Cela nous permettra de réfléchir au réseau d'acteurs concernés avant de nous focaliser sur les questions de relation aux enquêtés et de constitution de l'image de ces participants.

Notre approche privilégie la démarche qui consiste à enregistrer soi-même son corpus – et non à utiliser des corpus déjà constitués ou à les constituer de manière indépendante du *fieldwork*. Ce choix n'a rien d'anodin : de nombreux chercheurs en linguistique travaillent sur des données qu'ils n'ont pas recueillies eux-mêmes – qui présupposent donc que les questions abordées ici ont été résolues en amont – ou qu'ils collectent sans nécessairement effectuer une enquête de terrain ou contacter des participants (c'est le cas notamment des travaux portant sur les émissions de radio ou de télévision enregistrées directement du poste – et sans procéder à des enquêtes qui concerneraient la production de ces émissions, comme le fait en revanche Broth, 2003). En outre, tout en privilégiant les enquêtes de terrain qui enregistrent des activités non provoquées par le chercheur (« données naturelles » utilisées en AC), l'analyse qui suit tient compte

aussi des autres techniques d'enquête recueillant des données orales (comme les entretiens par exemple).

2.1. Informateurs, témoins, locuteurs, participants ? problèmes de catégorisation

Un des enjeux éthiques et juridiques qui se posent durant l'enquête et son prolongement dans les analyses qui en découlent est la question de savoir comment est traité l'enquêté au fil du projet de recherche. En préalable à cette question (sur laquelle nous reviendrons plus bas, notamment au chap. 4), nous nous arrêterons sur le problème de savoir comment catégoriser les participants à l'enquête.

Les personnes concernées par l'enquête sont catégorisées dans la littérature par des termes divers et variés qui ont des implications importantes mais qui sont rarement thématiques. On trouve ainsi « informateur », « sujet », « témoin », « interviewé », « locuteur », « enquêté », etc. Chacune de ces catégories présuppose des pratiques spécifiques de recherche :

- de nombreuses catégories semblent fonctionner seules, focalisées sur le pôle de l'enquêté – en mettant entre parenthèse le rôle du chercheur. Par exemple :
 - « informateur » renvoie à une conception de l'enquête comme ayant pour objectif de recueillir des « informations » et donc à identifier les meilleures sources pour cela. Cette conception informationnelle de la communication a été fortement critiquée en linguistique et en sociologie mais semble avoir été beaucoup moins problématisée en considérant le terrain et l'enquête comme une situation d'interaction où s'établissent de nombreuses relations sociales – qui ne réduisent jamais à la transmission d'informations. A noter qu'en général on parle de l'informateur mais pas de l'« informé ».
 - « locuteur » est une catégorie qui, parmi les activités des personnes enquêtées, en privilégie la parole. Alors que l'on parle dans de nombreuses activités, il est relativement rare que « locuteur » soit une catégorie pertinente pour les participants, qui peuvent considérer le fait de parler comme allant de soi et s'orienter vers d'autres catégories pertinentes (p.ex. « expert dans un domaine particulier »). « Locuteur » semble donc être d'abord une catégorie de linguiste ; elle peut être pertinente pour les participants dans un sens un peu différent, par exemple dans des activités où ils s'orientent, voire thématisent, des catégories comme « beau parleur », « bavard » ou « orateur ».
 - Dans certains cas, on peut considérer la personne comme ayant un statut proche de celui d'« auteur », comme dans les récits de vie, les recueils de chants, les explications et témoignages d'un savoir particulier.
- d'autres catégories fonctionnent plutôt par couples :
 - tel est le cas de « enquêté »/ « enquêteur », d'« interviewé » / « intervieweur » qui rétablissent une sorte de réciprocité entre les deux pôles – et en tout cas n'escamotent pas celui du chercheur (voir les travaux en AC portant sur cette situation : Houtkoop, 2000 ; Maynard & alii, 2002).

- Certaines catégories manifestent une plus grande symétrie (« collaborateur », ou « partenaire ») que d'autres, impliquant une approche participative de l'enquête.

2.2. Catégories d'appartenance et activités d'enquête

Sacks nous rappelle que le choix des catégories pertinentes par les participants les associe souvent à des activités typiques de ces catégories (ce sont des *category-bound activities*) : ainsi l'activité d'entretien a lieu entre un intervieweur et un interviewé ou entre un questionneur et un questionné. Ces catégories personnelles incorporent les catégories actionnelles – rendant ainsi manifestes certaines activités comme typiques de certaines méthodologies d'enquête (l'entretien comme étant basé essentiellement sur une série de questions – plutôt que sur d'autres activités communicationnelles). Dans ce sens, un lien étroit existe entre pratiques d'enquête (ou méthodologies d'enquête) et catégories des participants.

On parlera ainsi plutôt d'« informateurs » lorsque l'enquête est orientée vers la production d'un savoir, recueilli par le moyen de questions, alors que l'on parlera de « participants » lorsque l'enquête est orientée vers l'enregistrement de pratiques non provoquées par le chercheur.

L'implication et le statut du chercheur lui-même sont variables et variablement reconnus (il est absent lorsqu'on parle de « témoin » ou de « locuteur » mais il est présent lorsqu'on parle d'« interviewé »), dans des descriptions de l'enquête qui le gomme ou au contraire le considère comme un participant à part entière. On peut reprendre utilement les catégories de Goffman (1981) pour penser l'enquête comme une activité admettant différents modes de participation du chercheur :

- l'enquêteur est « *principal* » dans des activités comme l'entretien, où il agit ouvertement en son nom de chercheur (mais où il peut aussi être simple « *animator* » d'une enquête qui le salarie et qui dépend d'autres) ;
- il est un *participant* comme un autre s'il fait partie du groupe qui est enregistré (voire s'il se fait passer pour un acteur social autrement catégorisé que « enquêteur ») ;
- il peut être un « *représentant* » ou « *porte-parole* » (« *figure* » dans la terminologie de Goffman) de la communauté lorsqu'il fait une recherche engagée et militante ;
- il tend à adopter une posture de « *hearer* » attentif dans des activités qu'il contribue à définir comme monogales (recueil de contes, de récits de vie, de témoignages) ;
- mais il est un « *overhearer* » lorsqu'il enregistre sans être présent. Les participants le traitent ainsi – y compris quand il parle au micro pour s'adresser à l'enquêteur lointain et pourtant présent par le biais de cet objet technique.

Le statut des participants peut lui aussi être varié

- ils participent bien sûr à l'activité enregistrée,
- ils peuvent être des participants qui *collaborent* à l'enregistrement (qui changent les cassettes ou enclenchent le dispositif, par exemple) voire qui dans certains cas le prennent en charge (ce qu'on appelle parfois l'auto-enregistrement) ;

- ils peuvent être des participants qui collaborent à l'enquête tout en le dissimulant aux autres et en s'y engageant comme eux : ce sont ceux qu'on appelle parfois des « *compères* », ceux qui portent le micro sur eux et qui déclenchent une interaction avec des tiers ;
- ils peuvent être des participants *focaux* de l'enquête (p.ex. en tant qu'interviewé) ou *périphériques* (p.ex. en tant que membre co-présent du même foyer que l'interviewé, participant de manière marginale à l'entretien – comme p.ex. l'enfant, la grand-mère, l'époux de l'interviewée).

2.3. Catégories liées à l'activité enregistrée

Lorsqu'on se focalise sur l'activité concernée elle-même (et non en tant qu'enregistrée dans le contexte d'une enquête), les catégories sont celles qui sont rendues pertinentes par les participants eux-mêmes dans le cadre de cette activité. L'identification de ces catégories – que ce soit dans les prises de notes ethnographiques, les transcriptions, les descriptions de la situation – nécessite un travail préalable d'analyse (cf. infra 4 pour quelques conséquences).

Dans ce sens, les « participants » à une activité sociale enregistrées seront catégorisés de manière endogène comme « amis d'enfance » ou « collègues de travail » - mais non comme « enquêtés ». Ces catégories sont aussi à distinguer des celles de la sociolinguistique classique, qui classe les locuteurs selon leur tranche d'âge, leur sexe, leur revenu, etc.

3. L'établissement de la relation avec les participants et le *consentement éclairé*

Les enjeux éthiques et juridiques sont omniprésents dans toutes les étapes qui constituent la démarche de terrain finalisée à la constitution de corpus de données orales interactives et multimodales. Nous insisterons ici surtout sur les phases *préparatoires* de l'enquête, préalables à l'enregistrement des données, où il s'agit notamment établir une relation avec les personnes concernées (3.1) : ces modes d'approche sont étroitement liés non seulement aux méthodologies d'enquête mais aussi aux possibilités et limitations techniques du dispositif d'enregistrement choisi (3.2), dont dépendent les contraintes spécifiques pour les autorisations à enregistrer (3.3.). Une fois terminée l'enquête et analysées les données, il s'agit d'organiser le retour sur le terrain pour différentes formes de « rendu » des résultats et des expériences – retour qu'il vaut mieux prévoir au plus vite sur le terrain et qui configure le type d'engagement pris envers les personnes concernées (3.4).

3.1. Modes d'approche des personnes concernées par l'enquête

Les enquêtes dont la finalité est le recueil de données enregistrées dépendent nécessairement de la qualité de la relation avec les personnes ressources – qu'on appelle des « informateurs » ou des « partenaires » (cf. supra). La mobilisation de ces personnes varie selon la méthode d'enquête choisie : nous insisterons ci-dessous sur la temporalité des différentes approches (3.1.1) des personnes directement concernées ou de leur hiérarchie (3.1.2) et sur la question de savoir comment organiser le retour, le contre-don, éventuellement la rémunération de ces personnes (3.1.3 – cf. 3.4).

3.1.1. Temporalité des modes d'approche et typologie des relations avec les informateurs

On peut considérer que la façon dont les personnes sont approchées sur le terrain - la façon dont une relation personnelle et sociale est établie - est un acte ayant immédiatement des implications éthiques et juridiques. L'établissement de la relation avec les informateurs a d'une part des effets sur la qualité de leur collaboration et donc, en définitive, sur la qualité des données ainsi constituées ; d'autre part, elle a des effets sur les relations de confiance, d'acceptation, voire d'intérêt ou de curiosité scientifique que les informateurs nourriront envers les enquêteurs.

On peut esquisser une typologie des relations établies avec les informateurs en l'articulant au moment où ils sont approchés dans le processus de l'enquête :

- a) quand l'enquête procède en laboratoire par convocation nominale des informateurs, les modalités de leur engagement sont généralement explicitées *préalablement*, au moment où les personnes acceptent de collaborer à l'enregistrement, effectué dans des lieux et à des moments convenus à l'avance. Les personnes sont alors soit sélectionnées et contactées par le chercheur (ou par une institution travaillant pour lui), soit elles répondent à un appel à volontaires. L'appel ou l'annonce de recrutement est le premier acte de communication qui manifeste (ou suscite des attentes quant à) la forme du contact voire du contrat qui s'établit avec le chercheur.
- b) quand l'enquête procède sous la forme d'un *fieldwork* impliquant une permanence plus ou moins prolongée sur le terrain de l'enquêteur et des formes d'*observation participante* – classiquement discutées au sein des méthodes ethnographiques empruntées par les linguistes comme par d'autres chercheurs en SHS (voir pour des discussions touchant à l'analyse conversationnelle Moerman, 1988 ; Duranti, 1997 ; Depperman, 2000) – la relation aux informateurs s'établit dans la *durée* de cette permanence et est souvent associée à la construction de relations personnelles impliquant entre autres une confiance réciproque.
- c) quand l'enquête procède par des entretiens, des « micro-trottoirs », des enregistrements d'activités réalisés de manière *aléatoire* dans des espaces publics, sans viser des témoins particuliers mais des passants choisis simplement à cause de leur présence sur les lieux au moment de l'enregistrement, une rencontre préalable avec les informateurs est par définition impossible. C'est

donc *juste avant, pendant ou juste après* la réalisation de l'enregistrement qu'ont lieu l'explication des finalités et la demande d'autorisation.

- d) dans certains cas, il est possible d'envisager un contact *postérieur* à l'enregistrement : tel est le cas d'enregistrements réalisés à l'insu d'une partie des participants dont l'entrée sur la scène enregistrée n'était pas prévisible (c'est le cas des conversations téléphoniques par exemple, où une partie collabore à l'enquête et l'autre n'est pas toujours au courant de l'enregistrement et est recontactée ensuite pour donner son accord).

La forme du contact, de l'engagement, de la crédibilité, de la confiance varie énormément selon que l'établissement de la relation d'enquêteur à enquêté se fasse au préalable, durant le terrain de manière durable, au même moment que l'enregistrement voire après.

3.1.2. Les personnes contactées

Dans la présentation que nous venons de faire, nous avons considéré pour des raisons de simplicité que le contact s'établissait avec la/les personne(s) directement concernées par l'enregistrement ; or souvent elles sont à considérer comme faisant partie d'un groupe ou d'une institution – ce qui implique des prises de contacts multiples. Il s'agit ainsi de distinguer

- le cas où l'informateur agit en son propre nom, de manière individuelle ;
- le cas où l'informateur est contacté dans le cadre de ses activités professionnelles ou institutionnelles et intervient donc en tant qu'appartenant à une organisation. La hiérarchie des personnes visées par l'enquête est aussi contactée au préalable : tel peut être le cas de la direction d'une entreprise, ou du chef d'une tribu, ou des parents d'élèves. Il convient de remarquer que la relation entre la personne et sa hiérarchie ne va souvent pas de soi et invite à différencier ce qui sera promis, expliqué, montré, etc. aux personnes et à leur hiérarchie.

3.1.3. Rémunération

Lors de l'approche des personnes concernées par l'enquête, des promesses peuvent être faites, de véritables contrats peuvent être proposés, des contre-parties, rémunérations, remboursements peuvent être proposés. Ces engagements peuvent être à la fois éthiques et juridiques, sociaux, matériels voire financiers. De toute façon, la question se pose d'une forme de « dédommagement » des informateurs – qui est très différente si on la catégorise comme « contre-don », « rémunération », « dédommagement », « service rendu »...

Plusieurs cas de figure sont envisageables :

- *durant* voire *avant* l'enquête :
 - rémunérations financières promises dès l'établissement du premier contrat,
 - contre-dons en nature,
 - contre-dons symboliques,
 - prestations pour la communauté concernée,
- *après* l'enquête :

- reconnaissance de l’informateur sous des formes allant du remerciement ou de la citation à la mention comme co-auteur ou comme collaborateur voire comme partenaire de la recherche ;
- restitution des résultats ;
- restitution des données / corpus sous forme d’archives ;
- diffusion de savoir-faire ;
- retours bénéfiques attendus pour la communauté au sens large et sur le long terme (sur le modèle des bénéfices attendus d’une recherche médicale)

Pour une discussion de ces formes de « rendu » nous renvoyons (infra, 3.4) à la discussion du « retour » sur le terrain. La question reste celle de savoir ce qu’on peut/doit promettre aux informateurs lors de l’établissement de la relation, en tenant compte que :

- cette relation se modifie dans le temps (notamment si l’enquête de terrain implique une durée) ;
- cette relation peut plus ou moins reconnaître l’ « informateur » comme un « partenaire » du projet de recherche (et non seulement comme un « objet »), dans des projets participatifs où le « natif » apporte plus que ses propres performances (par exemple en collaborant aux transcriptions, aux traductions, aux gloses des données).
- la rémunération financière peut être moins problématique pour des informateurs recrutés (parfois par des organismes spécialisés) dans le cadre d’un contrat formel ; elle peut être plus problématique sur le terrain où elle implique une mise en concurrence non seulement entre les informateurs possibles, mais aussi entre les chercheurs pouvant y avoir accès (tel est le problème par exemple pour des linguistes d’universités moins dotées de moyens face à des chercheurs venant d’universités davantage dotés – et pouvant de ce fait être privilégiés par les informateurs ou générer chez eux des demandes difficilement satisfaisables).
- la rémunération financière n’est qu’un cas parmi d’autres de retour, qui pour les enquêtes de terrain se fait toujours de manière plus ou moins implicite, au fil de la vie quotidienne et de la négociation des relations mutuelles.

3.2. Choix du dispositif d’enregistrement : modalités et contextes

Le choix du dispositif d’enregistrement des corpus a des effets sur la manière dont les personnes concernées vont être traitées, dont leur consentement va être obtenu, dont l’acceptation ou l’acceptabilité de l’enregistrement va se négocier.

Nous allons ici discuter quelques dimensions qui peuvent se révéler pertinentes, allant du choix des contextes dans lesquels effectuer l’enregistrement (3.2.1) aux modalités de l’enregistrement (3.2.2).

3.2.1. Contexte de l’enregistrement

Par définition, il n’est pas possible de *tout* enregistrer et les chercheurs sont obligatoirement obligés de faire des choix. Ceux-ci dépendent de l’objet de recherche visé, des contraintes techniques (par exemple, difficulté d’enregistrer en vidéo la nuit ou en audio dans des lieux très bruyants) et aussi du respect des personnes enregistrées.

Interviennent notamment :

- le choix du *moment* à enregistrer : il s'agit de trouver un équilibre entre les moments intéressants pour l'enquêteur et le respect de la vie privée de l'enquêté ;
- le choix des *activités* à enregistrer : celles-ci peuvent être davantage publiques et sociables ou bien intimes et privées ;
- le choix du *lieu* où enregistrer : là aussi il y a une tension entre des lieux publics détachés de la vie privée ordinaire des personnes et des lieux intimes ;
 - o le laboratoire est un lieu totalement détaché de l'espace de vie des informateurs – et c'est d'ailleurs ce qui fait que les chercheurs voulant travailler sur les pratiques sociales situées l'évitent.
 - o le domicile des personnes est leur lieu de vie, lui-même articulé en lieux plus « publics » ou plus « intimes » (un repas pris à la salle à manger, à la cuisine ou au lit n'a pas la même teneur, ainsi qu'un entretien effectué au salon ou autour de la table de la cuisine).
 - o les espaces de travail sont eux aussi, quoique de manière différente, structurés par des questions de confidentialité qu'il s'agit de respecter ; leur non-respect peut risquer d'impliquer pour les données recueillies un devoir de confidentialité qui signifie l'impossibilité de leur exploitation.

L'équilibre à trouver se situe donc entre contextualité et naturalité des données enregistrées et voyeurisme – le choix des moments à enregistrer pouvant avoir des conséquences importantes sur la suite de l'enquête (sur les autorisations à exploiter les données et sur le droit de rétractation *post hoc* des sujets).

3.2.2. Modalités d'enregistrement

Les modalités d'enregistrement interviennent souvent dans le choix des contextes à enregistrer (cf. supra), des activités visées ainsi que dans les modalités d'acceptation ou de résistance des personnes concernées. Différentes dimensions techniques peuvent intervenir sur l'acceptabilité de l'enregistrement par les personnes enregistrées :

- le fait que l'enregistrement soit réalisé en audio ou en vidéo : pour certaines activités, les personnes concernées peuvent préférer l'audio à la vidéo – jugée plus invasive –, quitte à passer de l'audio à la vidéo dans un deuxième temps, une fois constatés les modalités et les effets de l'enregistrement sur l'activité.
- le fait que l'enregistrement soit réalisé par l'enquêteur présent, par des techniciens ou par un dispositif pré-installé et fonctionnant en absence du chercheur a des effets sur son acceptation : même si la caméra ou le micro sont souvent traités comme des « prothèses » ou des prolongements du chercheur (p.ex. quand les participants s'adressent directement à eux), l'absence du chercheur derrière la caméra peut être préférée par certains participants.
- le fait que l'enregistrement soit réalisé par le chercheur ou par les participants eux-mêmes : d'une part, la délégation de l'enregistrement aux participants peut être vue comme une forme de contrôle de leur part sur ce qui est enregistré ; d'autre part, cette délégation peut être refusée comme une forme trop poussée de collaboration détournant le participant de son activité.
- le fait que l'enregistrement soit réalisé par un dispositif voyant ou discret, voire caché : il existe de nombreux débats sur le fait de recourir à un micro caché et sur les conséquences de ce choix sur les relations possibles avec les participants;

par ailleurs, même lorsque les participants sont au courant de l'enregistrement, le fait de recourir à un dispositif voyant peut aussi bien être perçu comme un gage de transparence que comme une gêne. Souvent la miniaturisation des dispositifs permet de les installer d'une manière qui, sans du tout les dissimuler, en fait rapidement des éléments intégrés dans le décor.

- le fait que l'enregistrement dépende de moyens techniques nécessitant une intervention à brève échéance (relative par exemple à la durée de la batterie ou à la durée de la cassette) implique des perturbations de l'activité par le chercheur (ou par les participants qui effectueraient le remplacement de la cassette) qu'évitent d'autres dispositifs dotés d'une plus grande autonomie (enregistreur par exemple directement sur des disques durs). Cela peut avoir des conséquences sur la naturalité des conduites enregistrées aussi bien que sur la gêne ou le dérangement occasionné dans l'activité elle-même (selon les activités – comme opérer un patient, effectuer une consultation en thérapie, discuter d'un contrat délicat, être engagé dans un processus de création – ces interruptions peuvent être considérées comme à éviter absolument).
- le fait que l'enregistrement offre ou non des angles morts aux participants qui voudraient lui échapper un instant : par exemple, alors que le cadre et le champ délimités par une seule caméra permettent d'inférer des zones qu'elle ne couvre pas – de même que la puissance imaginée d'un micro –, le fait de recourir à plusieurs caméras sur la même scène peut donner l'impression d'un dispositif de surveillance auquel on ne peut se soustraire.
- le fait de pouvoir arrêter ou imposer des coupures à l'enregistrement peut intervenir comme une matérialisation de la possibilité de rétractation ; le fait que l'effacement ou la coupure de l'enregistrement puissent être effectués par les participants quand ils le désirent ou bien doit être effectué plus tard, ou par des tiers, peut donner l'impression d'une plus ou moins grande latitude à intervenir sur lui et suppose des relations de confiance différentes. Cette question – comme bien d'autres – est là aussi liée aux contraintes techniques sur l'enregistrement et à la sophistication du dispositif. On pourra en tenir compte dans le choix de supports permettant ou non un effacement immédiat des données ou bien permettant ou non un visionnement sur place de ce qui a été enregistré.

Ces considérations montrent bien l'imbrication des questions techniques et des questions juridiques – le respect à la fois personnel, éthique et juridique des participants étant matérialisé dans les choix techniques mis en œuvre.

3.3. Information aux enquêtés et demande d'autorisation (consentement éclairé)

La définition du « consentement éclairé » et sa traduction dans des formes de relation sociale (le contact avec les informateurs) et des formes matérialisées (les documents échangés et signés) est sensible au contexte et aux objets de l'enquête, ainsi qu'aux conditions socio-culturelles du groupe dans lequel elle se déroule. Nous esquissons ici quelques pistes de réflexion, en partant de la définition même du « consentement éclairé » (3.3.1), en reprenant la question de la temporalité à laquelle ces questions se posent (3.3.2), ainsi que la question des personnes que l'on informe et à qui on demande l'autorisation (3.3.3), des formes que prend cette information (3.3.4) et des objets à

propos desquels on choisit d'informer (3.3.5) ainsi que sur les formes du consentement (3.3.6).

3.3.1. Définition du « consentement éclairé »

On parle souvent de formulaires d'autorisation à soumettre aux informateurs ; il est cependant important de faire dépendre cette autorisation de l'information préalable donnée aux personnes concernées : sans *information*, la *demande d'autorisation* n'a pas d'objet ni de sens. C'est pourquoi on parle de *consentement éclairé* (*informed consent*), dans le sens où l'acceptation de l'enregistrement est étroitement dépendant de la compréhension des finalités pour lesquelles il est effectué.

3.3.2. Moment de l'information et de la demande

La demande d'autorisation dépend du mode d'approche des personnes enregistrées. Elle peut différer selon la temporalité à laquelle elle a lieu :

- information/demande préparée à l'avance durant une permanence sur le terrain et dépendant de la relation d'interconnaissance et de confiance avec l'enquêteur,
- information/demande faite juste avant l'enregistrement,
- information/demande faite juste après l'enregistrement,
- information/demande orale effectuée avant et demande écrite effectuée après l'enregistrement (avec possibilité de rétractation).

L'information est plus abondante lorsqu'elle bénéficie de la présence prolongée de l'enquêteur sur le terrain ; elle est plus limitée lorsque la demande d'autorisation se fait rapidement avant ou après l'enregistrement, sans autre forme de contract entre les enquêteurs et les enquêtés.

La temporalité à laquelle se situe l'information et la demande d'autorisation peut être choisie en relation avec ses effets envisagés sur la structuration de l'activité enregistrée : souvent le moment de l'information et de la demande d'autorisation est choisi de manière à ne pas perturber l'activité du point de vue des participants (p.ex. une demande d'autorisation à un client au moment de la vente peut provoquer un risque de perturbation de la vente pour le vendeur et donc être refusée par lui à l'enquêteur qui désirerait documenter cette activité), ou du point de vue des enquêteurs (p.ex. une demande d'autorisation en ouverture de conversation modifie l'organisation du déroulement séquentiel de cette ouverture).

Si l'information et la demande interviennent *après* l'enregistrement, l'information peut apparaître comme un « dévoilement », une « révélation » qui *a posteriori* qualifie l'enregistrement de « dissimulation » : cela peut faire intervenir des recatégorisations des participants et des activités (celui qui s'était présenté comme un touriste perdu dans la ville demandant son chemin devient un enquêteur travaillant sur les descriptions spatiales dans les demandes d'itinéraire) (cf. Mondada, à paraître a).

3.3.3. Qui énonce l'information/la demande et qui y répond

Même si le chercheur est celui qui informe et demande habituellement l'autorisation d'enregistrer, différents cas de figure sont envisageables :

- le cas le plus classique est celui de l'enquêteur se chargeant de l'information et de la demande d'autorisation,
- souvent toutefois le chercheur envoie sur le terrain des étudiants ou des collaborateurs qui sont autant de porte-paroles du projet.

- dans certains cas, il est envisageable que les participants deviennent eux-mêmes les porte-paroles du projet : cela est le cas lorsque le chercheur demande à un participant d'informer d'autres participants (p.ex. l'hôte qui invite chez lui des amis à un repas qui sera enregistré ; le commerçant qui demande à ses clients d'accepter de se laisser enregistrer ; l'enseignant qui demande l'autorisation à ses élèves ou étudiants, etc.). Cette délégation fait partie des collaborations sur le terrain entre enquêtés et enquêteurs ; elles peuvent toutefois être la source de malentendus et de difficultés.

De même, l'autorisation peut concerner les signataires eux-mêmes ou des personnes qui dépendent d'eux (subalternes, enfants, étudiants, etc.). Dans ce dernier cas, il est important de tenir compte du fait que *autorisation* ne se confond pas toujours avec *acceptation*.

3.3.4. Qu'est-ce qu' « informer » ?

Au cœur du consentement éclairé il y a l'exigence d'informer les participants enregistrés. Toutefois, dès que l'on interroge cette exigence, les questions surgissent. Qu'est-ce qu' « informer » ? et informer « à propos de quoi » ? à quelles conditions peut-on dire que cette information produit le statut « éclairé » de son destinataire ?

La notion même d'« information » peut laisser penser à un simple transfert de messages et de contenus ; elle tend à gommer les processus, les contextes et les contingences qui caractérisent cette activité communicationnelle par laquelle un enquêteur explique l'objet de son enquête à ses partenaires sur le terrain. Dès que l'on réfléchit en termes de type d'activité, l'« information » aux enquêtés pose une série de problèmes à résoudre :

- *l'adéquation au destinataire* : l'explication du projet de recherche, pour être comprise et partagée, demande à être ajustée aux compétences, au niveau de langue et de compréhension du destinataire. Cet ajustement concerne aussi le contexte et les modalités de l'enquête, prenant en compte l'adéquation entre ce que les partenaires voient faire sur le terrain et les explications qu'on en donne.
- l'explicitation des *finalités de l'enquête* doit se faire sans *nuire* à celle-ci : cela pose la question de l'équilibre à trouver entre la transparence de l'enquête et les transformations éventuelles des conduites des participants qu'elle pourrait provoquer.
- l'explication du projet de recherche peut se faire à des *niveaux de généralité différents* (p.ex. de « c'est une enquête sur les façons de parler des gens » à « c'est une enquête sur la fréquence et les contextes de la liaison non obligatoire en français »).

L'information aux enquêtés comprend non seulement des explications du projet scientifique mais aussi des informations précises concernant par exemple :

- les responsables de l'enquête et leur affiliation institutionnelle, ainsi que les financeurs ;
- une adresse de contact,
- les personnes qui auront accès aux données et qui travailleront sur elles,
- la façon dont les enquêtés ont été choisis et la population dont ils font partie,
- la façon dont les données seront anonymisées,
- le fait que les données seront transcrites selon des conventions particulières (sous forme d'un exemple d'extrait),

- la façon dont les données seront archivées une fois l'enquête terminée (conservation ou destruction à la fin de l'enquête, conservation auprès de quel garant, modalités de réutilisation éventuelle, transmission à d'autres chercheurs),
- les modalités d'accès aux informations relatives au projet et concernant tout particulièrement les données/analyses faisant référence à la personne (possibilité d'accès aux fichiers et informations concernant tout particulièrement la personne),
- les droits de la personne, notamment le droit de rétractation,
- les risques éventuels ainsi que les retombées positives, morales ou matérielles, de l'étude,

Les modalités d'information peuvent elles aussi varier selon la culture des destinataires, en particulier

- elle peut se faire de manière orale
 - o individuellement dans des conversations casuelles
 - o collectivement dans des réunions d'information
- elle peut se faire de manière écrite (par une brochure, un dépliant,) ou par email

Dans le contexte d'une culture écrite il est recommandé de laisser un texte ; de même l'indication d'un site internet où suivre l'évolution du projet (éventuellement avec des modes d'accès particuliers) peut être utile.

De manière générale, l'information comme la demande d'autorisation émane d'un choix délicat sur ce que le chercheur décide de dire ou de ne pas dire - choix qui a nourri de nombreux débats dans la littérature (cf. Fine, 1993 ; Hodkinson & Sparkes, 1993 ; Mitchell, 1991 ; Welland & Pugsley, 2002 ; Kulick & Wilson, 1995).

-> Voir en annexe un exemple de brochure d'information

3.3.5. L'objet de la demande d'autorisation

Une fois informé l'enquêté, ayant produit le caractère « éclairé » de celui-ci, il s'agit de lui demander l'autorisation de collecter le corpus. La question qui se pose est de savoir comment circonscrire l'objet de cette autorisation.

L'autorisation concerne en effet les dimensions suivantes, qui peuvent interagir et se superposer les unes aux autres :

- les *actions* effectuées par les chercheurs dans le cadre du projet :
 - o l'enregistrement,
 - o la préparation du corpus (transcription, traduction, annotation, etc.),
 - o l'analyse dans le cadre des objectifs annoncés,
 - o les usages des données de manière intégrale/non,
 - o la diffusion des résultats de l'analyse,
 - o la conservation/destruction des données une fois terminée l'enquête ;
- les *formats* et les conditions de l'enregistrement :
 - o audio/vidéo,
 - o avec plusieurs caméras/micros,
 - o à des moments connus ou non des enquêtés, bien circonscrits ou couvrant de longues durées,

- tout choix technique intervenant dans la façon dont la personne figurera dans les données peut être explicité voire négocié ;
- les *conditions de diffusion* des données et des résultats :
 - sous forme intégrale ou partielle (courts extraits dont la longueur maximale peut être prévue),
 - sous forme uniquement textuelle (transcriptions) ou audio-visuelle (dans des documents *power point* par exemple) ;
- les *contextes* de diffusion des données et des résultats :
 - des contextes de recherche (workshops, colloques, congrès),
 - des contextes d'enseignement universitaire,
 - des contextes de formation et de vulgarisation plus larges,
 - des contextes liés au terrain (par exemple il faut demander explicitement l'autorisation de réutiliser les données dans le contexte d'une formation dans la même institution où elles ont été recueillies – où elles peuvent se révéler très sensibles).

L'explicitation de ces contextes se superpose avec celle des activités dans lesquelles les données seront utilisées ; l'enjeu dans les deux cas est celui des personnes qui auront accès aux données dans le cadre de ces activités.

- Il est envisageables de laisser la possibilité pour l'enquêté d'ajouter de contraintes qui lui seraient personnelles ; toutefois cette éventualité pose le double problème de sa légalité ainsi que celui de son interprétabilité.
- Un des problèmes majeurs qui se posent dans la demande d'autorisation – comme d'ailleurs pour l'information – concerne l'évolution toujours possible des finalités de l'enquête, qui peuvent ne pas être totalement fixées à son début et surtout se transformer au fil du travail sur le terrain et sur les corpus. Pour cela il est important de formuler les finalités de manière suffisamment générale pour intégrer leurs évolutions éventuelles ayant émergé du travail de recherche. Par contre tout changement de finalité devra faire l'objet d'une nouvelle demande.

3.3.6. Les formes de l'autorisation

La demande d'autorisation peut prendre différentes formes, dépendant elles aussi du contexte socio-culturel dans lequel se déroule l'enquête : ainsi par exemple l'exigence de demander la signature de l'enquêté n'a de sens que dans les cultures de la *literacy* où cette procédure a un sens, n'effraie pas et n'est pas liée à d'autres pratiques avec lesquelles elle pourrait être confondue (comme la signature de chèques).

On peut donc différencier les formes de la demande selon le support sur lequel elles sont consignées :

- demande écrite et signée
- demande orale
 - il est possible et utile de prévoir que l'autorisation orale soit elle-même enregistrée, sous forme audio ou vidéo. Cela permet d'en assurer la traçabilité.
 - il est possible de prévoir une première autorisation orale et de ne faire signer les formulaires que lorsque les personnes concernées auront visionné les données et octroyé leur consentement en pleine connaissance de cause.

Dans le cas de la demande écrite, celle-ci peut se présenter sous différentes formes – dans un texte préformé (formulaire) :

- un texte compact qui synthétise les différents aspects de la demande d'autorisation et qui demande un accord (ou un refus) global,
- un texte présentant des cases à cocher et donc des choix : cette forme a l'avantage sur la première de matérialiser des choix véritables pour l'enquêté et donc de lui laisser la possibilité de refus partiels (p.ex. il peut accepter l'enregistrement audio mais refuser l'enregistrement vidéo) voire d'ajouts de contraintes (p.ex. il peut demander l'anonymisation de la vidéo et non seulement de l'audio). La question qui se pose alors est celle de la formulation des alternatives, de manière à ce qu'elles ne soient pas redondantes et qu'elles ne soient ni trop compliquées ni trop longues à traiter pour l'enquêté.

Un problème peut se poser lors des demandes collectives, lorsque des groupes sont concernés (par exemple dans le cadre d'enregistrements de réunions) : si de trop nombreuses alternatives sont laissées au choix des participants, il est possible que les réponses mènent à des résultats contradictoires où n'émerge aucun dénominateur commun ; dans ce sens les demandes à des groupes présentent des problèmes et des contraintes qui ne sont pas les mêmes que pour des individus.

-> Voir en annexe des exemples de formulaires d'autorisation

3.4. Prévoir l'après de l'enquête : retours, *debriefings*

On insiste souvent sur la préparation du terrain, mais il est également important de préparer le départ et le retour sur le terrain. Cela présente une importance à la fois scientifique et éthico-juridique : le retour sur le terrain peut se révéler nécessaire à tout moment pour une vérification, un complément d'enquête, une reprise de contact avec les informateurs. Si le départ du terrain s'est mal passé, le retour sera impossible. Par ailleurs, la présence sur le terrain produit non seulement des relations de confiance, mais aussi des attentes qui engagent dans la durée : quitter le terrain en disparaissant tout simplement, après avoir pratiqué une immersion qui souvent noue des relations étroites avec les participants et leur demande de l'aide et des prestations, peut produire de grosses déceptions. Une fois « pris » du savoir, des réponses, des corpus sur le terrain, il s'agit donc de savoir comment « rendre » quelque chose aux personnes sans lesquelles l'enquête aurait été impossible (cf. aussi les questions de rémunération traitées supra, 3.3.1.3).

Il convient en outre de signaler que les « feedbacks », les « debriefings », les retours d'expérience peuvent se faire déjà pendant le terrain, sous la forme de compte-rendus de résultats partiels par exemple. La distinction entre le « pendant » et l'« après » du terrain peut ainsi être relativisée.

Plusieurs types de pratiques sont envisageables pour assurer un « retour » auprès des populations enquêtées. Nous en énumérons quelques unes, allant de la présentation de résultats la plus proche du contexte académique à la formulation de savoirs et savoir-faire la plus proche du terrain. C'est sans doute dans l'évaluation de la distance entre le « retour » et l'académie ou le terrain que se situent les choix de « politique du terrain » :

- *présentation des résultats* à la fin du projet :

- la formulation des résultats peut être plus ou moins vulgarisée, plus ou moins proche des préoccupations des enquêtés.
- la présentation des résultats peut comporter notamment des exemples de *transcriptions* et d'analyse de transcriptions : les participants réagissent de manières très différentes (parfois surprise, parfois choquée) à la représentation de leur voix.
- démarche d' *empowerment* : elle consiste à ne pas simplement penser le « retour » en termes d' « information » mais aussi en termes d'apport en savoirs et savoir-faire à la communauté des enquêtés :
 - on peut ainsi songer non seulement à présenter des analyses mais à permettre aux participants de continuer de collecter des données et d'analyser leurs propres données pour leurs propres fins,
 - on peut formuler les retombées de l'analyse dans les termes de l'agenda, des thèmes, des préoccupations des acteurs,
 - on peut mettre au service de la communauté les savoirs produits par l'enquête en les matérialisant dans d'autres formes que les écrits universitaires traditionnels (p.ex. sous forme d'expositions, ou d'autres produits culturels dérivés),
 - on peut offrir une formation basée sur les résultats/les méthodes de l'enquête ; de manière plus générale, on peut songer à transmettre des outils d'analyse, transférer des compétences qui pourraient être utiles sur le terrain.
- la question du « retour » *des données* elles-mêmes sous forme de *corpus* ou *d'archives* peut se révéler délicate : elle peut s'imposer dans certains cas (ainsi pour les langues en danger on pourra constituer un patrimoine légué à la communauté) mais aussi devoir être évitée pour protéger les informateurs (ainsi dans le cas d'enquêtes dans des entreprises ou des institutions, les données collectées pourraient intéresser certains niveaux de la hiérarchie mais nuire à des subalternes). Le retour des archives, s'il est pertinent, pose donc souvent des questions
 - d'accès limité des personnes pouvant consulter ces archives, en tenant compte des risques et des avantages que produit la mise à disposition sur le terrain,
 - de modes et de technologies d'accès aux archives : si les archives sont formatées pour que la population concernée puisse y avoir accès, les technologies doivent être adaptées aux usages et aux possibilités de ces populations (il ne sert à rien de produire un DVD si personne n'a de lecteur, ou de monter un site internet si personne n'a d'accès à l'informatique). Se pose ici la question de la gestion de l'asymétrie entre l'académie et le terrain.
- la garantie *d'accès aux publications* pose des questions analogues à celle de l'accès aux données, quoique de manière souvent moins problématique.

4. La construction de l'image des enquêtés : procédés d'anonymisation et modes de représentation dans la transcription

La possibilité ou la garantie (que nous relativiserons plus bas) de rendre les données recueillies anonymes est importante pour la protection de la vie privée des personnes concernées par l'enquête et pour la légalité des corpus recueillis par les chercheurs. L'anonymisation des données n'est toutefois ni un processus simple ni une garantie aproblématique, faisant surgir de nombreux problèmes à la fois techniques, scientifiques et sociologiques.

L'anonymisation des données est une garantie importante en matière de légalité des données et de leur usage ; dans certains cas (et si elle garantit véritablement la non-identification des personnes concernées) elle peut permettre d'utiliser des données même en absence de demande d'autorisation préalable. Il convient toutefois d'être prudent sur ce point – en considérant toutes les limitations et les difficultés auxquelles on se heurte dans l'anonymisation (cf. infra).

4.1. Définition

Bien qu'on parle souvent d'anonymisation, la question légale qui se pose est celle de la *non-identifiabilité des personnes* : l'enjeu est que, sur la base des données recueillies et de leurs modes de représentation (transcription par exemple) on ne puisse pas identifier les personnes concernées.

Les aspects touchés par ces considérations sont

- tout ce qui permet d'identifier *directement* une personne
 - o par référence au locuteur ou à un tiers et à sa sphère privée,
 - o sur la base des manifestations du locuteur, comme sa voix ou son apparence physique ;
- tout ce qui peut lui porter préjudice,
- tout ce qui peut *indirectement* permettre, par recoupement d'informations, de remonter au locuteur concerné.

Les opérations qui suppriment ces références ou ces manifestations sont appelées des procédés d'« anonymisation » des données.

4.2. Données concernées par l'anonymisation

L'anonymisation ne concerne pas uniquement les enregistrements ou les transcriptions, mais un ensemble de données qui sont contenues dans les corpus et qui se différencient selon divers supports et formats – dont dépendront les techniques d'anonymisation.

- les données primaires vidéo,
- les données primaires audio,
- les données primaires textuelles : documents officiels ou non recueillis sur le terrain,

- les données secondaires : transcription, notes de terrain, métadonnées, analyses, descriptions ethnographiques,
- les données secondaires visuelles: copies d'écran (*screen shots*), voire représentations de la voix (oscillogrammes, spectrogrammes...).

On remarquera que certaines données personnelles échappent à l'anonymisation : tel est le cas des hommes et des femmes publics, dans des interventions à caractère public (par exemple des hommes politiques dans un débat à la télévision), où ils interviennent en connaissance de cause en ce qui concerne la diffusion de leur image.

4.3. Temporalités auxquelles peut intervenir l'anonymisation

On peut distinguer différentes temporalités auxquelles peut intervenir l'anonymisation. Selon les finalités de l'étude et les contextes de l'enquête, on peut considérer que l'anonymisation doit se faire le plus *tôt* possible ou bien le plus *tardivement* possible. La première solution augmente les garanties de confidentialité pour la personne, la seconde maximise les possibilités d'analyse pour le chercheur. Les temporalités peuvent varier selon les types de données aussi :

- on évite l'anonymisation sur les données premières originales de référence car elle pourrait endommager les données elles-mêmes ; par contre les données non anonymisées doivent être conservées dans un lieu sûr.
- les données peuvent/doivent/ne doivent pas (selon les politiques adoptées) être anonymisées lors de leur dépôt pour conservation. Le rôle de garant des institutions assurant la conservation est ici crucial.
- on peut travailler (dans un groupe de recherche bien délimité et qui garantit la non circulation des données en dehors de lui) sur des données non anonymisées et garantir en revanche une anonymisation de tout extrait diffusé, figurant par exemple dans un écrit ou une présentation orale.
- on effectue toujours l'anonymisation sur les copies destinées à circuler entre chercheurs extérieurs au projet et parfois entre chercheurs internes au projet (c'est le cas notamment pour de grands consortiums de recherche ou des projets articulant des réseaux d'équipes importants).

4.4. Modes d'anonymisation

Les modes d'anonymisation touchent à la fois les supports et les formats des données et mettent ainsi en jeu des possibilités et des contraintes technologiques ; ils concernent aussi des formes et des manifestations symboliques de l'identité des personnes et mettent ainsi en jeu des questions d'analyse.

4.4.1. Formes pouvant être concernées par l'anonymisation

Comme nous allons le voir, il est difficile – voire impossible – de constituer une liste finie des formes concernées par l'anonymisation. On peut toutefois souligner les formes principales :

- formes nominatives (nom, prénom, surnom ou petit nom, sigle d'entreprise...),
- données personnelles (adresse, numéros de téléphone, numéro de passeport, numéro de compte, âge, lieu de naissance...),
- profession, statut, titres,
- activités sociales,
- parenté, réseaux,
- référence à des lieux (toponymes, institutions, services...),
- référence à des caractéristiques de la personne (physiques, culturelles, médicales...) uniques ou rares dans son milieu identifié,
- caractéristiques physiques : voix, visage, caractéristiques corporelles, ...
- etc.

L' « etc. » clôturant cette liste souligne le fait que tout élément, selon les contextes d'enregistrement et de réception de cet enregistrement, peut devenir un porteur d'informations sur l'identité des personnes. L'identification des formes concernées par l'anonymisation suppose donc une compétence sociologique et culturelle qui puisse imaginer les usages, les connaissances et les associations qui pourraient permettre l'identification d'une personne sur la base d'une forme donnée.

4.4.2. Formes de remplacement

Une fois identifiées les formes pouvant porter à l'identification des personnes, il s'agit de les transformer pour effectuer les opérations d'anonymisation.

On fera remarquer que la forme la plus radicale d'anonymisation est la *suppression* pure et simple des données – bien que l'on cherche souvent d'autres moyens d'assurer l'anonymisation qui puissent mieux les préserver. On notera cependant que la suppression peut être partielle (on peut envisager de détruire des extraits qui seraient porteurs de trop d'éléments problématiques et confidentiels, portant à leur inutilisabilité).

La forme généralement adoptée d'anonymisation procède par *substitution* d'éléments confidentiels par des formes de remplacement. Celles-ci varient selon les supports techniques concernés : nous distinguerons ici entre le texte, l'audio et la vidéo.

a) *Texte*

Les textes concernés sont d'abord la transcription et toutes ses mentions dans des articles, exempliers, cours, conférences... D'autres textes devant être anonymisés sont les données primaires textuelles (documents recueillis sur le terrain). Celles-ci peuvent se présenter d'ailleurs sous une forme textuelle ou sous la forme d'image (tel est le cas d'une lettre, d'un document administratif, d'un manuscrit qui est conservé sous forme photocopiée ou scannée). Le principe de la substitution consiste à rendre visible la portion de texte qui a été remplacée, et ainsi à donner des informations générales sur elle (concernant au moins sa durée).

- a. *remplacement par un « blanc »* : c'est la solution la moins informative et surtout la moins visible.
- b. *remplacement par un hyperonyme*, tel que NN ou NVILLE ou NHOPITAL pour nom, nom de ville, nom d'hôpital etc. Cette solution peut être très informative (on précise le type de référence de la forme anonymisée). Elle est utile dans les cas où la substitution par pseudonyme (cf. infra ici-même)

est impossible, difficile ou non vraisemblable. Cette solution implique le développement de conventions spécifiques pour la notation de ces hyperonymes (soit (N de ville) ou NVILLE ou (ville A)...), qui ne sont pas de même nature que le texte qu'ils remplacent (c'est pourquoi l'emploi des majuscules ou de parenthèses est parfois proposé, quand il n'entre pas en contradiction avec d'autres emplois de majuscules prévus dans les conventions de transcription).

c. *remplacement par un pseudonyme* : c'est la solution la plus souvent utilisée, du moins pour les noms de personnes, qui permet une bonne intégration de la forme de remplacement dans le fil du discours, n'attire pas l'attention sur elle, est vraisemblable et garde un certain nombre d'indications contenues dans la forme initiale. Cela n'est toutefois possible que si le choix des pseudonymes est réfléchi et répond aux problèmes suivants :

- i. le pseudonyme est choisi dans le même champ paradigmatique que la forme qu'il remplace. Par exemple « Ahmed » sera substitué par « Moustapha » plutôt que par « Albert ». Cet exemple montre que le pseudonyme tentera de conserver des traits d'ethnicité. Dans certains cas, notamment si l'interaction enregistrée le rend pertinent, on veillera à conserver : les connotations possibles du nom, p.ex. s'il est à la base de plaisanteries ou de jeux de mots (ainsi par exemple dans un corpus un témoin fait un jeu de mots avec Klapisch, nom d'un commerçant spécialisé dans la préparation et la vente de saumons, déformé en Klapfisch – dans ce cas l'anonymisation devrait conserver les deux syllabes finales ; toutefois le maintien même du jeu de mots serait susceptible de mettre à mal l'effort d'anonymisation lui-même); le nombre de syllabes (on ne remplacera pas « Marc » par « Maximilien » mais plutôt par « Yves ») et certaines caractéristiques phonétiques et prosodiques, si elles sont exploitées dans l'interaction.
- ii. le pseudonyme est choisi de manière à éviter de pouvoir reconstituer le nom initial (dans ce sens, le choix d'un pseudonyme commençant par les mêmes lettres que l'original est à éviter, même s'il présente des avantages pour sa mémorisation).
- iii. le pseudonyme est choisi de manière à éviter de ridiculiser la personne (dans ce sens, sont à éviter les pseudonymes qui renverraient à des caractéristiques de la personne – p. ex. « Monsieur Gros »).
- iv. les noms des rues, les numéros de téléphone, etc. peuvent être remplacés de la même manière que les noms de personne.

On remarquera qu'il est plus facile de choisir un pseudonyme pour les personnes que pour les noms de villes (on peut imaginer un nom de petite ville ou de quartier ou encore de rue mais beaucoup moins un nom de grande ville ou de capitale) ; il est parfois envisageable mais pas toujours possible de penser à des pseudonymes pour des noms de services institutionnels (cela n'a souvent pas de sens de remplacer « département de chirurgie » par « département de dermatologie » dans le cas d'un

hôpital). Dans le cas où le choix d'un pseudonyme est difficile ou invraisemblable on recourra à la solution b.

En outre, on peut se poser la question de savoir jusqu'où on doit anonymiser les références à des lieux. L'anonymisation de la référence géographique n'est pas uniquement un problème d'échelle : c'est un problème de pertinence par rapport à l'activité ou au contexte (P.ex. la mention de la ville de Bielefeld peut être anodine dans un récit de quelqu'un qui va y trouver sa tante, mais devient révélatrice dans la comparaison entre deux cliniques spécialisées dans le traitement de l'épilepsie – dès que l'on sait qu'il existe à Bielefeld un des meilleurs centres de traitement de cette maladie d'Allemagne)².

b) *Audio*

- a. *remplacement par du silence*. Cette solution a comme désavantage le fait que le remplacement peut être confondu avec une pause.
- b. *remplacement par un beep ou un autre bruit* qui ne se confond avec aucun signal pouvant intervenir dans l'enregistrement.
- c. *remplacement par le signal original filtré et déformé*. Cette technique est surtout utilisée dans les médias pour rendre la voix non identifiable. Quand elle est pratiquée par des non spécialistes, elle peut poser des problèmes quant à son irréversibilité (possibilité de rétablir le signal original).

c) *Image*

L'image concernée est surtout celle dynamique des enregistrements vidéo. Mais on peut penser aussi aux images fixes, par exemple à des photographies sur des documents et à des captures d'écran dans les transcriptions. De même, on peut songer à l'anonymisation d'une représentation visuelle du flux sonore (dans un spectrogramme par exemple) lorsqu'elle pourrait rendre reconnaissable la prononciation d'un nom ou d'un numéro.

- a. pour ces données, la *suppression* est envisageable sous forme de coupures lors du montage. Dans ce cas, il est conseillé de marquer la durée du segment coupé sur la bande et de ne pas donner l'impression d'une continuité.

² Autre exemple (transcription simplifiée) :

A : et vous à la République vous trouvez que c'est plus beau là où vous êtes ?

B : ah ben moi je suis très bien\ je suis dans une rue
très tranquille la rue Pierre Levée

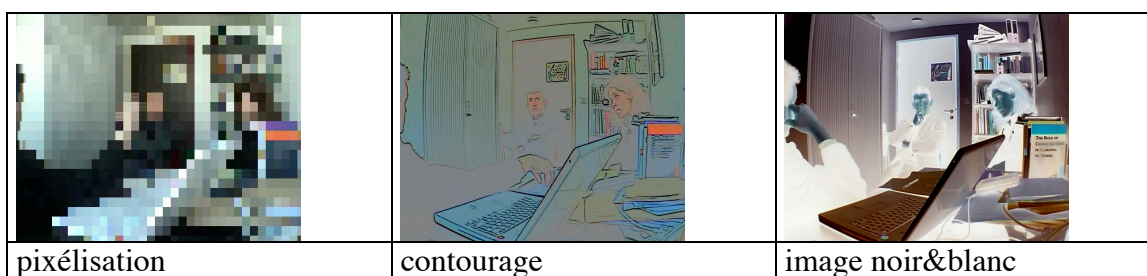
A : alors c'est que vous aimez les choses anciennes/ c'est dans cette rue-là
qu'on a trouvé un menhir

B : ah tiens

A : c'est pour ça qu'elle s'appelle la rue de la Pierre Levée

On peut se demander comment anonymiser en conservant les derniers tours. On pourrait imaginer un autre nom de rue mais ce n'est pas suffisant – car il faudrait conserver la référence plausible au menhir, sous peine de rendre le passage incompréhensible (la suppression ou correction du « menhir » impliquant une « correction » des données étant inenvisageable)

- b. *remplacement par un brouillage du signal* : soit par floutage, par pixélisation ou par contourage de l'image ou par application d'autres types de filtres. Ce traitement peut concerner toute l'image ou un détail uniquement. Dans ce dernier cas, elle est technique plus complexe à réaliser quand ce détail est en mouvement.
- c. *placement d'un bandeau noir sur les yeux de la personne*



4.4.3. Un exemple

Nous proposons ci-dessous un exemple des différentes possibilités d'anonymisation d'un extrait qui montre les enjeux que cela pose à chaque fois : il s'agit d'un appel téléphonique à un *call centre* pour une demande de dépannage.

version 1 (r16/0.44-1.08)

```

1 OPE j'peux avoir votre nom monsieur s'il vous plaît
2 APP je m'appelle monsieur xxllar: xxhar
3 OPE euh excusez-moi j'vous ai j'vous ai pas entendu
4 APP bellar
5 (0.2)
6 OPE .houi\
7 APP mihar
8 <(0.8) ((OPE tape au clavier))>
9 APP bellar . bé eu deux^el a er
10 OPE o[ui\
11 APP [mihar èm i ache a er
12 <(2.9) ((OPE tape au clavier))>
13 APP et mon numéro de téléphone est le: . zéro trois/
14 cinquante-huit/ cinquante-cinq/ quarante-huit trente-quatre\
15 OPE quarante-huit trente-quatre hein
16 APP tout à fait oui
17 OPE okay ben écoutez j'contacte un technicien/
    
```

Un tel extrait – typique pour la situation qu'il documente où la demande d'adresses et de numéros de téléphone est routinière - présente un grand nombre de données devant être anonymisées. D'autres solutions peuvent être envisageables, comme la suivante :

version 2 (r16/0.44-1.08)

```

1 OPE j'peux avoir votre nom monsieur s'il vous plaît
2 APP je m'appelle monsieur NOM PRENOM
3 OPE euh excusez-moi j'vous ai j'vous ai pas entendu
4 APP NOM
5 (0.2)
6 OPE .houi\
7 APP PRENOM
8 <(0.8) ((OPE tape au clavier))>
9 APP NOM . ((épelle))
10 OPE o[ui\
11 APP [PRENOM ((épelle))
    
```

12 <(2.9) ((OPE tape au clavier))>
 13 APP et mon numéro de téléphone est le: . NO TELEPHONE

Nous allons brièvement contraster ces deux versions du point de vue de l'analyse qu'elles permettent ou non de faire :

- dans la version 1 (V1) ligne 2 nous avons la première occurrence du nom. Celui-ci est prononcé alors que la qualité de la communication est perturbée. Ce bruit occasionne une réparation, associée à une explication de la part de l'opérateur li 3. Dans la V1, la transcription rend compte du fait que les deux noms sont partiellement audibles : elle rend ainsi disponible les circonstances qui conduisent à la réparation. Dans la V2, ces circonstances ne sont pas présentes (bien qu'on aurait pu noter : « <NOM PRENOM ((difficilement audible))>).
- les deux noms sont répétés li 4 et li 7. Cette fois, ils le sont d'une manière plus distincte et audible ; ils sont surtout présentés par l'appelant dans un format différent du premier : il les sépare et attend (li 5) que l'opérateur en ait ratifié la compréhension (6) avant de continuer.
- cela toutefois ne semble pas suffire à l'appelant qui redit son nom, en l'épelant cette fois. Le format adopté est celui du nom entier, suivi éventuellement d'une pause, plus les lettres successivement (li 9, li 11). Il les épelle sans hésitations, d'une traite, d'une voix uniforme. Cette dernière information n'est pas disponible sur la V2. Elle est pourtant importante si on s'intéresse à la façon dont l'appelant montre à ce stade de l'interaction qu'il s'oriente explicitement vers la facilitation de la tâche de l'opérateur, qui est de taper le nom dans son ordinateur.
- La même considération peut être faite pour les numéros de téléphone. Ceux-ci peuvent être anonymisés de différentes façons :

NO TEL

NO TEL A 8 CHIFFRES

03 58 55 48 34

L'avantage de la solution retenue dans la V1 est qu'elle permet de rendre compte du format particulier utilisé par l'appelant : c'est lui qui prend l'initiative d'enchaîner du nom au numéro de téléphone, qu'il propose avant que l'opérateur ne le lui demande ; il marque un arrêt juste avant de commencer le numéro – s'orientant donc vers le fait que l'opérateur est en train de sélectionner le bon champ où l'insérer et se prépare à le noter – puis formule son numéro par couples de chiffres, les deux derniers couples étant prononcés sans aucune démarcation prosodique. Ce sont d'ailleurs ce deux derniers couples qui sont repris de la même façon par l'opérateur.

Une activité comme donner son numéro de téléphone peut sembler extrêmement simple et banale : elle ne l'est pas si on prend en compte le fait qu'elle est un accomplissement pratique coordonnée et orienté vers l'autre, permettant des choix de formats ajustés à la situation et à l'activité (noter, taper le numéro) telle que les participants la définissent ensemble. Cet accomplissement peut être fondamental dans le cas des appels d'urgence, où un malentendu sur le numéro de téléphone peut occasionner une perte de temps pendant l'appel même voire l'impossibilité de rejoindre l'appelant plus tard – avec des conséquences éventuellement tragiques (Whalen, Zimmerman & Whalen 1988 analysent un appel problématique suite auquel la personne nécessitant de l'aide est décédée avant l'arrivée des secours).

Ce que préserve le type d'anonymisation présenté ici – par remplacement par pseudonymes et pseudonuméros – est la forme de ce qui est anonymisé, permettant une analyse qui s'appuie sur les détails formels pour décrire l'activité en cours. Cela reste possible même sur ce type de séquence, qui contient de nombreuses informations confidentielles.

4.5. Les limites de l'anonymisation

Même si l'anonymisation est une opération fondamentale pour assurer la circulation légale des données, il convient d'être prudent par rapport aux promesses et garanties faites aux enquêtés et affirmées face au public concernant l'anonymisation des données. Les limitations sont essentiellement de deux ordres très différents, le premier concernant les contextes qui augmentent ou diminuent la reconnaissabilité des personnes (4.5.1.), le second concernant les contraintes que l'anonymisation fait peser sur les objets même de la recherche (4.5.2.).

4.5.1. Limitations issues des contextes de production et de circulation des données

L'anonymisation est relativisée par différents facteurs intervenant soit lors de la production des données – et selon les spécificités de ce qui se passe durant l'enregistrement – soit lors de la réception de ces données :

- l'anonymisation opère d'abord sur une série de formes censées contenir les indications principales permettant l'identification de la personne (cf. 4.4.1 supra) ; néanmoins n'importe quelle référence ou forme peut, selon les contextes, conduire à l'identification de la personne et souvent d'une manière qui passe au premier abord inaperçue de l'enquêteur. Ainsi, par exemple, la mention d'un détail rare dans l'interaction (une pathologie rare de la personne, un attribut extraordinaire, une caractéristique unique et connue dans l'environnement de la personne...) peut se révéler significative pour certains.
- le caractère reconnaissable de ces détails dépend de manière cruciale du contexte de réception et plus spécifiquement du public qui consultera ou prendra connaissance des corpus. Ainsi les membres d'un département d'anesthésie reconnaîtront facilement un de leur collègue sur la base d'expressions typiques, d'expertises spécifiques ou de façons propres de parler ou d'agir ; en revanche les mêmes détails passeront inaperçus chez les professionnels d'un autre hôpital ou a fortiori chez des étudiants de linguistique d'une université. Mais, là encore, la reconnaissabilité ne dépend pas simplement de l'éloignement géographique ou social du contexte dans lequel ont été enregistrées les données : les personnes sont mobiles dans l'espace et dans les milieux sociaux et il n'est pas impossible que le fils d'un patient puisse reconnaître son père dans un cours universitaire portant sur des consultations thérapeutiques. La valeur identifiante d'un détail dépend donc du contexte de réception des données.
- selon les cas, la référence à une institution ou à un organisme peut rendre nécessaire ou non l'anonymisation : par exemple la référence à une grande enseigne doit être anonymisée s'il s'agit du lieu de travail d'un employé, n'a pas besoin d'être anonymisé si elle intervient comme élément du paysage dans une

indication d'itinéraire, et doit à nouveau être anonymisée si elle est citée dans des propos diffamatoires.

- D'autres aspects sont liés au *recoupement* d'informations venant de plusieurs sources (cela peut concerner par exemple la relation entre données anonymisées et métadonnées).

C'est pourquoi le degré d'anonymisation est toujours lié au contexte d'exploitation des données.

Il dépend aussi de la longueur des extraits cités (des extraits courts sont moins reconnaissables que les extraits longs).

Les dimensions concernées par l'anonymisation ne sont pas décidables a priori, ce qui rend vain l'établissement d'une *check-list*. Elles sont liées à des processus de pertinentisation qui émergent de l'interaction elle-même. Ceci montre bien que l'anonymisation repose sur une analyse fine des matériaux. En outre, cela rend prudents quant aux promesses d'anonymisation faites aux partenaires sur le terrain : l'anonymisation ne peut jamais être assurée complètement et il est important de le signaler.

4.5.2. Limitations issues du travail d'analyse

Les limitations de l'anonymisation peuvent venir d'un autre type de considérations, davantage liées aux pratiques d'analyse des chercheurs.

Le problème fondamental est posé par la contradiction éventuelle entre anonymisation et disponibilité des détails pour l'analyse (sur le principe de disponibilité voir Mondada, 2003). En effet, les enregistrements et les transcriptions visent à produire la disponibilité des détails observables pour qu'ils puissent être exploités par l'analyse ; l'anonymisation au contraire a pour rôle de rendre indisponibles certains de ces détails en les effaçant ou en les transformant.

Cela peut être le cas par exemple de l'anonymisation par beepage d'un nom qui est prononcé en chevauchement avec un autre tour de parole et qui rend impossible l'analyse de ce chevauchement.

Cela peut être le cas de l'anonymisation de numéros de téléphone lors d'appels d'urgence qui rend indisponible la manière dont l'appelant donne son numéro de téléphone dans une situation de stress et d'émotion pouvant affecter de manière cruciale cette information.

Cela peut être le cas de l'anonymisation des visages sur une bande vidéo qui rend impossible une analyse des regards. De manière analogue, le filtrage de la voix (tel que pratiqué par les médias) n'est pas envisageable pour la plupart des études linguistiques qui se basent sur les qualités intrinsèques du signal sonore.

C'est pourquoi les chercheurs revendiquent souvent la nécessité de travailler – en garantissant la sécurité et l'inaccessibilité des données – sur des données non anonymisées, de les conserver sous cette forme et de faire intervenir l'anonymisation le plus tardivement possible et d'une manière qui tienne compte des pertinences pour l'analyse.

4.6. Transcription

La transcription est une pratique qui, loin de se limiter à être un exercice techniques de reproduction, intègre de nombreux enjeux théoriques et interprétatifs (cf. déjà Ochs, 1979). Dans le passage de l'oral à l'écrit graphico-visuel, de nombreuses opérations de catégorisation sont effectuées, soit quant aux formes linguistiques, segmentées visuellement en unités (Blanche-Benveniste & Jeanjean, 1987 ; Mondada, 2000), soit quant aux identités des locuteurs eux-mêmes (Mondada, 2002). Du point de vue de la protection de l'image et de l'identité des personnes enquêtées et enregistrées, il convient d'apprécier ces effets pour éviter la surinterprétation, la stéréotypisation et la stigmatisation des locuteurs et de leurs façons de parler. Nous nous limiterons ici à considérer ces enjeux de la transcription.

4.6.1. Introduction de la transcription : description ethnographique

La transcription est souvent précédée d'une brève description ethnographique qui esquisse le contexte dans lequel elle a été recueillie ainsi que le type d'activité et l'identité des participants. Cette description, qui intègre des éléments issus des métadonnées du corpus, peut avoir plusieurs effets sur la lecture (ou sur la réception d'un exposé oral) :

- elle peut contenir des informations permettant l'identification des personnes et qui entrent en contradiction avec les principes de l'anonymisation ;
- elle peut contenir des indications qui forcent la lecture ou l'interprétation des données. En rendant telle catégorie d'appartenance ou telle autre dimension pertinente elles peuvent construire une image particulière de l'activité et des locuteurs.
- en particulier, elle peut contenir des allusions, permettre des inférences qui renforcent certains stéréotypes (voire qui les utilisent pour provoquer des effets comiques pour conquérir le public – cela n'étant pas rare dans les exposés oraux).

Ces remarques ne concernent pas uniquement la description des données mais aussi les noms des corpus, qui peuvent parfois intégrer des éléments confidentiels. Dans ce sens, même si cela a souvent une fonction mémorielle, il convient d'éviter d'intégrer le nom des acteurs concernés dans le nom du corpus.

4.6.2. L'identification des locuteurs dans la transcription

La transcription intègre les résultats de l'anonymisation (cf. supra), que ce soit dans les identifiants en marge des lignes de la transcription ou que ce soit dans le corps de son texte.

La transcription attribue les paroles transcrites à des locuteurs dont l'identité est indiquée à gauche au début de la première ligne de chaque prise de parole. Ces identifiants reprennent généralement les pseudonymes attribués aux locuteurs, mais pas toujours. D'autres possibilités sont en effet envisageables, qui ont néanmoins des effets variables sur l'interprétation du texte qui les suit. Tout choix effectué en la matière pose le problème de la manière dont est traité le locuteur. Par exemple :

- A, B, C, ... : solution qui est la moins connotée mais qui en adoptant l'ordre alphabétique ordonne les locuteurs en premier, deuxième, troisième....
- E1, E2, E3... (pour des élèves) : choix qui homogénéise les personnes au sein d'une même classe, désignée par une catégorie unique. La même chose vaut pour L1, L2, L3 où L renvoie au Locuteur : si le linguiste peut considérer que tous les locuteurs sont égaux et que les acteurs sociaux l'intéressent avant tout en tant qu'êtres parlants, du point de vue de l'activité en cours, ceux-ci participent d'abord sous d'autres catégories, que ce soit enquêteur/enquêté, père/fils, médecin/patient, etc.
- H, F (pour homme et femme) : là encore, le choix privilégie la catégorie du genre sur toute autre catégorie, en postulant ainsi la pertinence généralisée de cette catégorie pour la compréhension des activités en cours.

Ces remarques invitent à se demander quels effets interprétatifs produisent les choix des identifiants. Il convient dans ce sens de se demander quels sont les identifiants pertinents pour les participants – surtout dans des démarches analytiques qui se préoccupent de la perspective des participants (comme l'analyse conversationnelle). C'est pourquoi les solutions alternatives peuvent être les suivantes :

- EVA, MAR, ROB, AND... : indication des 3 premières lettres des pseudonymes, que ce soit des prénoms ou des noms propres – selon la tonalité de la conversation.
- APP/OPE pour appelant/opérateur ou DOC/PAT pour docteur/patient, ou encore INTE/IEUR pour interviewé/intervieweur lorsque l'activité institutionnelle est régie par des paires catégorielles de ce type. Sur ces questions, on peut renvoyer aux réflexions de H. Sacks sur les catégorisations des personnes et sur la pertinence des catégories selon l'activité et le contexte en cours (une personne qui est médecin dans un contexte peut très bien être père de famille dans un autre ; la manière de l'identifier dépend donc de l'activité en cours) (Sacks, 1972a, 1992).

4.6.3. Enjeux liés aux choix effectués dans le corps du texte

Lorsqu'on transcrit, on prend sans cesse des décisions quant à la manière de représenter les locuteurs et leurs manières de parler. De cette manière, l'analyse – et parfois le jugement – se glisse immédiatement dans la pratique de la transcription. Nous soulignerons quelques enjeux des choix effectués dans la transcription elle-même.

4.6.3.1. Quelle orthographe ?

De nombreuses discussions ont eu lieu en France sur l'emploi de l'orthographe standard, de l'orthographe adaptée et de l'API dans les transcriptions.

Ces débats ont montré que le choix de l'orthographe adaptée peut conduire, lorsqu'il n'est pas maîtrisé, non seulement à des incohérences mais surtout à des effets caricaturaux et stéréotypisants (Blanche-Benveniste & Jeanjean, 1987 ; Jefferson, 1983).

Par ailleurs, l'API est un choix qui ne résout pas tous les problèmes et qui affecte de manière très importante la lisibilité des transcriptions.

Le choix de l'orthographe standard a comme effet de « normaliser » les locuteurs et leur parler, d'en permettre l'« édition » par le linguiste qui ainsi « nettoie » voire « corrige »

leurs dires pour mieux en permettre l'analyse (notamment automatique) ou pour en sauvegarder l'image : la normalisation du locuteur obéit ainsi à deux enjeux majeurs, celui de la standardisation des données pour une standardisation des analyses (cf. infra 3.6) et celui de l'intervention du linguiste pour non seulement en protéger mais aussi sauvegarder ou rétablir l'image. Ainsi Pierre Bourdieu défend, dans les entretiens publiés dans la *Misère du monde* (1993), le choix de standardiser voire de corriger les dires de ses interlocuteurs pour ne pas les ridiculiser et ne pas leur manquer de respect. A cette posture, en répond une autre (défendue notamment par Lahire, 1996) qui considère que les façons de parler appartiennent à l'identité sociale du locuteur et ne sauraient être effacées – leur effacement conduisant à la production d'une sorte de fiction ou d'artefact sociologique (cf. Mondada, 2002). Cette controverse a le mérite de souligner les effets de ces choix orthographiques ; elle montre aussi qu'ils peuvent avoir des effets très différents en fonction du lectorat qui aura accès aux transcriptions.

4.6.3.2. La représentation du parler exolingue

Le choix de transcrire en API certains passages ou certains locuteurs plutôt que d'autres permet certes une plus grande précision dans la représentation des détails de leur parler mais risque aussi de provoquer des effets d'asymétrie non maîtrisés.

Ainsi le recours à l'API et à l'orthographe adaptée peut produire des effets de stigmatisation et d'asymétrie à l'encontre de locuteurs « non-natifs » – lorsque ces derniers sont représentés de manière différente par rapport aux locuteurs « natifs » (ceux-ci par des notations standard, les « non-natifs » par des orthographe spéciales qui en mettent en relief non seulement la différence mais aussi l'« anormalité », l'« anormativité »).

Dans le même sens, la notation explicite, par convention, de la variété de langue du locuteur (différenciation grâce à des polices spécifiques de plusieurs langues dans une conversation bilingue ou de l'interlangue de l'apprenant dans une conversation exolingue) opère une précatégorisation de cette variété, qui de fait se révèle souvent être un élément négocié par les participants et changeant au fil de la conversation (où par moment certaines formes sont thématiques comme « étrangères » ou « étranges » et où à d'autres moments elles ne sont pas du tout prises en considération).

Ces questions se posent aussi pour la traduction de la transcription :

- le fait de traduire les dires de certains locuteurs plutôt que d'autres peut les traiter d'opaques voire de non corrects ou d'incompréhensibles ;
- la façon dont on traduit, plus ou moins littéralement, peut amener à produire une version appauvrie de la parole du locuteur, et à en effacer ou au contraire à en souligner la différence.
- différents formats existent pour la traduction (fournie en note, à la suite de l'original, ligne par ligne ; de manière à proposer un équivalent à la forme originale, de manière à respecter un lien quasi littéral à l'original, de manière à en fournir une glose grammaticale) qui produisent chacune une image différente de la culture et de la langue de l'autre (Traverso, 2002).

4.6.3.3. Enjeux du multimodal et du détail de la transcription

Le fait de ne noter que les activités verbales et d'ignorer d'autres ressources – comme c'est actuellement le cas dans la plupart des transcriptions – peut produire une image aberrante de certains comportements de locuteurs. Cela peut être le cas notamment de

locuteurs aphasiques ou d'enfants s'exprimant par d'autres moyens que les moyens linguistiques standards : ne pas tenir compte de la totalité des ressources mobilisées par ces locuteurs signifie en donner une image réduite, qui pathologise ou anormalise leur comportement (Mondada, 2002).

De même, différents degrés de granularité de la transcription (cf. Jefferson, 1985) peuvent nuire à la représentation de conduites non-standards (p.ex. la vocalisation prononcée par un patient aphasique peut être significative et demander une transcription adéquate ; mais elle peut aussi être réduite à un simple « bruit » sans aucun sens dans une transcription superficielle).

Le caractère approfondi ou détaillé de la transcription ne répond donc pas uniquement à des exigences scientifiques ; elle répond aussi à des exigences éthiques et juridiques, qui permettent de nuancer et de complexifier l'image que l'on donne des locuteurs – en s'éloignant d'autant plus du risque de le caricaturer et de le stigmatiser dans des comportements stéréotypés.

Conclusion

Ce texte a voulu passer en revue les principaux problèmes qui se posent à différents moments du processus de recherche et qui concernent le respect de la vie privée des enquêtés. Deux problèmes ont ainsi été discutés : a) les conditions et les pratiques au fil desquelles s'établit le « consentement éclairé », comportant l'information aux enquêtés et la demande d'autorisation à enregistrer leurs paroles et leurs activités ; b) les modes possibles d'anonymisation des données, empêchant l'identification des personnes enregistrées. Ces problèmes ont été discutés par rapport à une démarche qui se fonde, en linguistique, sur l'enregistrement de données interactionnelles « naturelles », i.e. non provoquées par le chercheur. Cette démarche pose, par ses exigences théoriques et analytiques, des problèmes spécifiques de confidentialité des données et offre des façons d'y réfléchir elles aussi originales - toutes deux pouvant inspirer les sciences du langage dans leur ensemble. L'approche retenue ici refuse de donner des recettes résolvant a priori les problèmes mais choisit plutôt de problématiser des aspects souvent considérés comme allant de soi, en considérant que les difficultés soulevées par les questions éthiques et juridiques sur le terrain et dans la recherche sont significatives des processus constitutifs de la démarche du chercheur et ont donc à être analysés en tant que tels.

Bibliographie

- Blanche-Benveniste, C., & Jeanjean, C. (1987). *Le français parlé. Edition et transcription*. Paris: INALF.
- Bourdieu, P., & alii, e. (1993). *La misère du monde*. Paris: Seuil.
- Broth, M. (2003). Analyse de l'interaction à la télévision. *Moderna Sprak*, 2(2), 193-202.
- Button, G. (Ed.). (1993). *Technology in Working Order: Studies of Work, Interaction and Technology*. London: Routledge.
- Deppermann, A. (2000). Ethnographische Gesprächsanalyse: Zur Nutzen und Notwendigkeit von Ethnographie für die Konversationsanalyse. *Gesprächsforschung* (www.gespraechsforschung-ozs.de), 1, 96-124.

- Duranti, A. (1997). *Linguistic anthropology*. Cambridge: Cambridge University Press.
- Fine, G.A. (1993). Ten lies of ethnography : moral dilemmas of field research. *Journal of Contemporary Ethnography*, 22-3, 267-294.
- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.
- Garfinkel, H. (Ed.). (1986). *Ethnomethodological Studies of Work*. New York: Routledge.
- Goffman, E. (1981). Footing. In E. Goffman (Ed.), *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Gülich, E., & Mondada, L. (2001). Analyse conversationnelle. In G. Holtus, M. Metzeltin & C. Schmitt (Eds.), *Lexikon der romanistischen Linguistik* (Vol. Band I,2, pp. 196-250). Tübingen: Niemeyer.
- Have, P. t. (1998). *Doing Conversation Analysis. A Practical Guide*. London: Sage.
- Heath, C., & Luff, P. (2000). *Technology in Action*. Cambridge: Cambridge University Press.
- Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.
- Hodkinson, P., Sparkes, A.C. (1993). To tell or not to tell ? Reflecting on ethical dilemmas in stakeholder research, *Evaluation and Research in Education*, 7/3, 117-132.
- Houtkoop-Steenstra, H. (2000). *Interaction and the Standardized Survey Interview*. Cambridge: Cambridge University Press.
- Jefferson, G. (1983). Issues in the transcription of naturally occurring talk: caricature versus capturing pronunciational particulars. *Tilburg Papers in Language and Literature, Tilburg University, Tilburg, Netherlands*, 34.
- Jefferson, G. (1985). An Exercise in the Transcription and Analysis of Laughter. In T. A. v. Dijk (Ed.), *Handbook of Discourse Analysis Volume 3* (pp. 25-34). New York: Academic Press.
- Lahire, B. (1996). Variations autour des effets de légitimité dans les enquêtes sociologiques. *Critiques sociales*, 8-9, 93-101.
- Luff, P., Hindmarsh, J., & Heath, C. (Eds.). (2000). *Workplace Studies. Recovering Work Practice and Informing System Design*. Cambridge: Cambridge University Press.
- Maynard, D. W., Houtkoop-Steenstra, H., Schaeffer, N. C., & Zouwen, J. v. d. (Eds.). (2002). *Standardization and Tacit Knowledge. Interaction and Practice in the Survey Interview*. New York: John Wiley.
- Mitchell, R.G.Jr. (1991). Secrecy and disclosure in fieldwork. In : W.B. Shaffir, R.A. Stebbins (eds.), *Experiencing Fieldwork : An inside View of Qualitative Research*. London : Sage.
- Moerman, M. (1988). *Talking Culture: Ethnography and Conversation Analysis*. Philadelphia: University of Pennsylvania Press.
- Mondada, L. (1998). Technologies et interactions sur le terrain du linguiste. *Le travail du chercheur sur le terrain. Questionner les pratiques, les méthodes, les techniques de l'enquête. Actes du Colloque de Lausanne 13-14.12.1998. Cahiers de l'ILSL*, 10, 39-68.
- Mondada, L. (2000). Les effets théoriques des pratiques de transcription. *Linx*, 42, 131-150.
- Mondada, L. (2001). Pour une linguistique interactionnelle. *Marges Linguistiques* (<http://www.marges-linguistiques.com>), 1(mai).
- Mondada, L. (2002). Pratiques de transcription et effets de catégorisation. *Transcrire l'interaction, No spécial (B. Bonu, éd.). Cahiers de Praxématique*, 39, 45-75.
- Mondada, L. (2005). *L'émergence des savoirs. Comment interagissent les chercheurs*. Lausanne: Presses Polytechniques Fédérales.
- Mondada, L. (à paraître a). La demande d'autorisation comme moment structurant pour l'enregistrement et l'analyse des pratiques bilingues. *Tranel (Université de Neuchâtel)*.
- Mondada, L. (à paraître b). La pertinence del dettaglio : registrazione e trascrizione di dati video per la linguistica interazionale. In Y. Bürki, E. de Stefani, (eds.) *Transcriptio*. Bern : Lang,
- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 43-72). New York: Academic Press.
- Ochs, E., Schegloff, E. A., & Thompson, S. A. (Eds.). (1996). *Grammar and Interaction*. Cambridge: Cambridge University Press.
- Pollner, M. (1993). The Reflexivity of Constructionism and the Construction of Reflexivity. In J. A. Holstein & G. Miller (Eds.), *Reconsidering Social Constructionism: Debates in Social Problems Theory* (pp. 199-212). New York: Aldine de Gruyter.
- Sacks, H. (1972a). An initial investigation of the usability of conversational materials for doing sociology. In D. Sudnow (Ed.), *Studies in Social Interaction* (pp. 31-74). New York: Free Press.

- Sacks, H. (1972b). On the Analyzability of Stories by Children. In J. J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics: The Ethnography of Communication* (pp. 325-345). New York: Holt, Rinehart and Winston.
- Sacks, H. (1984). Notes on methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 21-27). Cambridge: Cambridge University Press. (Edited by Gail Jefferson from various lectures).
- Sacks, H. (1992). *Lectures on Conversation [1964-72] (2 Vols.)*. Oxford: Basil Blackwell.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696-735.
- Sacks, H., Schegloff, E.A. (1979). Two preferences in the organization of reference to persons in conversation and their interaction. In: G. Psathas (ed.). *Everyday Language: Studies in Ethnomethodology*. (pp. 15-21). New York: Irvington.
- Schegloff, E. A. (1992). Introduction. In G. Jefferson (Ed.), *Harvey Sacks, Lectures on Conversation (Vol.1: Fall 1964-Spring 1968)* (Vol. 1, pp. ix-lxii). Oxford: Blackwell.
- Schegloff, E.A. (1996). Some practices of referring to persons in talk-in-interaction: a partial sketch of a systematics. In B. Fox (ed.). *Studies in Anaphora*. (pp. 437-485). Amsterdam: Benjamins.
- Selting, M., & Couper-Kuhlen, E. (2000). Argumente für die Entwicklung einer 'interaktionalen Linguistik'. *Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion*, 1, 76-95.
- Selting, M., & Couper-Kuhlen, E. (Eds.). (2002). *Studies in Interactional Linguistics*. Amsterdam: Benjamins.
- Traverso, V. (2002). Transcription et traduction des interactions en langue étrangère. *Cahiers de Praxématique*, 39, 77-99.
- Watson, G. (1987). Make me reflexive, but not yet: Strategies for managing essential reflexivity in ethnographic discourse. *Journal of Anthropological Research*, 43, 29-41.
- Welland, T., Pugsley, L. (eds.) (2002). *Ethical Dilemmas in Qualitative Research*. Hants : Ashgate,
- Whalen, J., Zimmerman, D. H., & Whalen, M. R. (1988). When words fail: A single case analysis. *Social Problems*, 35(4), 335-362.
- Woolgar, S. (Ed.). (1988). *Knowledge and Reflexivity: New frontiers in the Sociology of Knowledge*. London: Sage.

ANNEXE 1 – exemples d’information aux enquêtés

1) Exemple de brochure d’information réalisée dans le cadre des enquêtes du laboratoire ICAR

[logos des différentes institutions]

[Titre]

Un projet de l’Université NN

direction : NN

chercheurs : NN, NN

stagiaire : NN

PRESENTATION DU PROJET



La question de la coordination des équipes dans les réunions professionnelles est un thème de plus en plus étudié pour comprendre les conditions de travail assurant la qualité et l’efficacité des soins. Cette enquête a pour objectif de décrire les processus et les pratiques qui rendent cette coordination possible, en insistant sur le rôle du langage et de la communication verbale. On s’intéressera notamment à l’organisation des interactions durant les réunions d’équipe durant lesquelles les cas cliniques sont discutés et des décisions thérapeutiques sont collectivement prises.



Ce projet s’inscrit dans une collaboration entre la Clinique NN et le laboratoire NN.

L’enquête sera menée sur le terrain par NN, chercheur au laboratoire NN et par NN étudiante en Master de sciences du langage à l’Univ. NN et stagiaire dans ce projet. Le projet est supervisé par NN, professeur à NN et chercheur au laboratoire NN.



Cette étude met en oeuvre une *méthodologie* innovante:

- des *enregistrements* audio seront effectués durant les réunions hebdomadaires où le travail de la semaine, ainsi que le traitement des patients, sont discutés.
- ces enregistrements seront ensuite *numérisés et transcrits*;
- ces enregistrements et les transcriptions seront *anonymisés* : cela signifie que les noms des patients comme des soignants seront remplacés par un beep sur la bande et par un pseudonyme dans la transcription. D'autres informations permettant l'identification des personnes comme des adresses ou des numéros de téléphone seront aussi beepées et remplacées dans le texte.
- pour faire ensuite l'objet d'*analyses* qualitatives portant sur l'organisation de l'action collective des équipes.



Afin de donner une idée de la manière dont se présentent les transcriptions qui seront effectuées des données enregistrées, voici un extrait - tiré d'une enquête précédente dans un autre hôpital :

(extrait de l'enregistrement p2/12014/9'35)

1 (4.4)
2 DrD voilà\
3 (2.8)
4 DrD donc madame coli[n/
5 DrL [madame colin\ donc c'est cette dame-là
6 qui est avec madame robinson/ là qui a été un peu (0.3) euh
7 que le ((NN institution)) nous a hospitalisé sur ma demande puisque
8 on avait un peu des places libres/ donc j'appelé euh ((NN
institution))
9 pour euh faire vivre la fédération/ et donc xx nous l'a envoyée/
10 une dame qui fait une seconde dépression/ pas d'antécédents
11 psychiatriques/ elle a fait euhm un épisode dépressif il y a/ .hh
12 deux ou trois ans/ euh qui s'est bien résolu c'est ça non/ amandine
13 AMA mhm
14 DrL qui s'est résolu bien sous effexor/ ça s'est bien passé

Remarques :

- les noms des personnes citées ainsi que des personnes qui parlent ont été anonymisés
- la parole des participants est transcrite selon des conventions particulières

- la parole des participants est transcrite de la manière la plus détaillée possible, ainsi que les indications de pause (lignes 1, 3) et de chevauchement de la parole (lignes 4-5).



Au terme de l'enquête, un *feed-back* sera organisé pour les équipes : une présentation des analyses effectuées sur les transcriptions et des résultats du projet sera assurée par NN et NN.



Contacts :

NN + affiliation + adresse, no de tél., email

Une telle recherche est impossible sans la collaboration active du personnel de la clinique et des personnes qui l'entourent, sur qui nous comptons et que nous remercions dès maintenant pour leur participation.

ANNEXE 2 – exemples de demandes d'autorisation

1) Exemple de formulaire type de demande d'autorisation mis au point au laboratoire ICAR

[papier avec entête officiel]

autorisation pour l'enregistrement audio/video et l'exploitation des données enregistrées

Présentation de l'enquête

[Peut se présenter sous forme de brochure séparée laissée aux enquêtés]

[Préciser l'institution d'où émane la recherche, la personne qui dirige / qui est responsable du projet, les chercheurs concernés sur le terrain.

Préciser le thème général du projet, le type de corpus qui est recueilli de manière générale, le type d'enregistrement qui est recueilli auprès de ces informateurs en particulier, son traitement et utilisation prévus.

Souligner les apports du projet, valoriser la collaboration de l'informateur, expliciter les bénéfices éventuels qu'il peut en tirer et les risques éventuels qu'il peut courir.]

Ces recherches ne sont possibles que grâce au consentement des personnes qui acceptent d'être enregistrées. Nous vous demandons par conséquent votre autorisation à procéder aux enregistrements.

Autorisation (biffer les paragraphes qui ne conviennent pas)

Je soussigné(e) _____

- autorise par la présente NN et NN à enregistrer en audio/vidéo le [préciser le type d'événement enregistré].

- autorise l'utilisation de ces données, sous leur forme enregistrée aussi bien que sous leur forme transcrite et anonymisée (cf. infra) :

- a) à des fins de recherche scientifique (mémoires ou thèses, articles scientifiques, exposés à des congrès, séminaires).
- b) à des fins d'enseignement universitaire (cours et séminaires données à des étudiants avancés, à partir du niveau maîtrise, en sciences du langage et en sciences sociales).
- c) pour une diffusion large dans la communauté des chercheurs, sous la forme d'éventuels échanges et prêts de corpus à des chercheurs, moyennant la signature d'une convention de recherche.
- d) pour une diffusion sur un site Internet dédié à la recherche.

- prends acte que pour toutes ces utilisations scientifiques les données ainsi enregistrées seront *anonymisées* : ceci signifie

- a) que les transcriptions de ces données utiliseront des pseudonymes et remplaceront toute information pouvant porter à l'identification des participants ;
- b) que les bandes audio qui seront présentées à des conférences ou des cours (généralement sous forme de très courts extraits ne dépassant pas la minute) seront « beepées » lors de la mention d'un nom, d'une adresse ou d'un numéro de téléphone identifiables (qui seront donc remplacés par un « bruit » qui les effacera) ;
- c) en revanche, pour des raisons techniques, le projet ne peut pas s'engager à

anonymiser les images *vidéo* mais s'engage à ne pas diffuser d'extraits compromettant les personnes filmées.

- souhaite que la contrainte supplémentaire suivante soit respectée :.....

Lieu et date: _____

Signature : _____

[Prévoir un double ou un autre document qui sera laissé à la personne, comportant une adresse de contact et éventuellement une adresse internet où consulter les résultats publiés du projet]

2) Exemple de demande d'autorisation pratiquée par la Talkbank

(source : <http://talkbank.org/share/consent.html>)

Part I: Explanation of Informed Consent

We are asking you to be a subject in a research study. If you agree to participate in this study, we would like to ask you to sign this informed consent document. Informed consent is a written agreement that you sign indicating your willingness to participate in this research. This document will tell you about the purpose, risks, and benefits of this research study. You should consent only after you have received all the necessary information and have had enough time to decide whether you wish to participate. Your signature on this form is voluntary and does not waive any of your legal rights or make any institutions or persons involved in this research any less responsible for your well being.

Part II: Explanation of Current Study

Why have I been asked to be a part of this study?

(Sample Text: We are asking your permission to videotape you, because we are hoping that you will take part in a communicative interaction that we hope to study.)

Who is responsible for the data collected in this study?

(Name and Address of Project Director.)

Why is this study being done?

(The purpose of this study is to Statement of Project Goals.)

What is involved in this study?

(Sample Text: We will videotape your participation in the (Description of the Type of Interaction Being Filmed). We will try to avoid interfering with the activity during filming. However, you will always be aware when videotaping is taking place. In some circumstances, you maybe contacted after the activity and asked to help us make sense of the activity captured on the tape.)

Who will look at the resulting data?

These data are only available for research purposes. All uses for commercial or other non-research purposes are prohibited. Researchers using these data must subscribe to a Code of Ethics that requires them to respect the views of all participants. Segments of the tape with accompanying transcriptions may be presented in the context of scholarly publications, academic symposia, university classes, and professional training activities. Data will be made available to researchers through CD-ROM or web distribution by password.

What level of access am I permitting?

(The researcher should choose one of these three levels to present here and below):

1. I am allowing full access to these data to researchers who subscribe to the TalkBank Code of Ethics.
2. I am allowing access to these data to researchers who subscribe to the TalkBank Code of Ethics and who have received a special password allowing them to use these data.
3. I am allowing access to these data to researchers who subscribe to the TalkBank Code of Ethics and who are sent the data on CD-ROM.

How long will I be in this study?

Your consent extends only for the duration of the following activity: _____ . The video acquired (Specify activity and dates) will be preserved indefinitely unless you request otherwise (see risks). Even after the interaction has been filmed, you have the right to revoke your agreement to participate in this study and to remove your data from inclusion.

What are the risks of this study?

Because of the nature of the data being gathered (i.e., video and audio recordings), it may not be possible to conceal your identity as a participant. There is a potential risk that people known or unknown to you will formulate negative opinions of you or your behavior on the basis of their viewing of these data. We have three safeguards against this:

1. Researchers using these data must subscribe to a Code of Ethics that requires them to respect your views, as well as those of all other participants. Any written or verbal public criticism of individuals or groups must first be reviewed by a committee of scholars. If you wish, you can will be provided with a copy of this scholarly Code of Ethics.
2. You can request that videotaping be stopped at any point during the activity, thereby preventing a video record from ever being produced.
3. You may review the recording. If you wish, you may file a request with the Principal Investigator to have particular segments containing your likeness deleted or erased.
4. As discussed in the section on confidentiality below, all additional data, apart from the videotape, will be confidential.

Are there benefits to taking part in this study?

There are no direct benefits to you personally for participating in this study. The primary benefits from this work are for the advancement of scientific understanding of communication processes. The availability of these data may lead to improvements in Field of Study of which you or others may be a direct beneficiary. If you are interested in these scientific outcomes, arrangements can be made with the Principal

Investigator to acquire digitized files of the interaction, along with transcripts and analyses, as they become available.

Will I receive any payment or other monetary benefits?

You will receive no payment for being videotaped. This data will not be used for commercial purposes. Therefore, you should not expect any royalties or payments in the future.

What other options are there?

You can choose not to participate in this study.

What about confidentiality?

The nature of the data being gathered precludes concealing your identity as a participant. Any researcher who happens to know you personally would be able to identify you from the videotape or the audiotape. If this occurs, our Code of Ethics requires that this researchers refrain from making further use of these data. In addition,, we will keep any additional information that we have about you confidential through the use of pseudonyms. These data will not be made available for any purposes outside of research activities.

What are the costs?

There are no costs to you for participating in this study.

What are my rights as a participant?

Taking part in this study is voluntary. You may choose to not take part or subsequently cease participation at any time. Leaving this study will not result in any penalty or loss of benefits to which you are entitled.

Part III: Informed Consent Approval

Project Title: Project Title

Conducted By: Project Director

I have read the above explanations and I agree to participate in the research study conducted by (Project Director). I understand that the proposed research has been reviewed by the University's Institutional Review Board and that, to the best of their ability, they have determined that, given my informed consent, the filming and analysis of these conversational interactions involve no procedures or requirements which may be found morally or ethically objectionable. They have also determined that the benefits from this research outweigh any potential risks. If, however, at any time I wish to terminate my participation in this study I have the rights to do so without penalty, even after the data have been collected. I also have the right to review the audio and video data and can request that specific segments be omitted, if they contain material that might embarrass me.

If I have any question about this study, I realize that I am encouraged to ask them now or at anytime during the study by contacting:

(Project Director's name, building, office number phone, fax, and email. Include possible alternate contacts too.)

I realize that I may report any objections to this study, either orally or in writing, to:

(Director of the University Office of Human Subjects Review.)

I agree to allow the following level of research access to these data from the conversational interaction in which I have participated.

(Insert materials from the list of access levels above.)

I understand that, in signing this consent form, I give Project Director, and his associates, permission to present this work in written and oral form, without further permission from me.

Signed, Dated

3) Exemple de demande d'autorisation de Susan Ervin-Tripp, Univ. de Californie, Berkeley

Researcher name _____

LETTER OF CONSENT

PHOTOGRAPHIC, AUDIO, AND/OR VIDEO RECORDS RELEASE CONSENT FORM

As part of this project we have made a photographic, audio, and/or video recording of you while you participated in the research. We would like you to indicate below what uses of these records you are willing to consent to. This is completely up to you. We will only use the records in ways that you agree to. In any use of these records, names will not be identified.

1. The records can be studied by the research team for use in the research project.

Photo _____ Audio _____ Video _____

[Please use initials]

2. The records can be shown to subjects in other experiments.

Photo _____ Audio _____ Video _____

[Please use initials]

3. The records can be used for scientific publications.

Photo _____ Audio _____ Video _____

[Please use initials]

4. The written transcript can be kept in an archive for other researchers.

Photo _____ Audio _____ Video _____

[Please use initials]

5. The records can be used by other researchers.

Photo _____ Audio _____ Video _____

[Please use initials]

6. The records can be shown at meetings of scientists interested in the study of _____

Photo _____ Audio _____ Video _____

[Please use initials]

7. The records can be shown in classrooms to students.

Photo _____ Audio _____ Video _____

[Please use initials]

8. The records can be shown in public presentations to nonscientific groups.

Photo _____ Audio _____ Video _____

[Please use initials]

9. The records can be used on television and radio.

Photo _____ Audio _____ Video _____

[Please use initials]

I have read the above description and give my consent for the use of the records as indicated above.

Date _____

Signature _____

Signature of Guardian, if Applicable _____

Native language(s) _____

Where native language learned (city or region) _____

Languages used on the tape _____

Where language(s) used on tape were learned _____

Age at which each language used on tape was learned _____

Education _____ Occupation _____

Name _____ Age _____ Sex _____