

Corpus oraux Guide des bonnes pratiques 2006. Version allemande

Olivier Baude, Claire Blanche-Benveniste, Marie-France Calas, Paul Cappeau, Pascal Cordereix, Laurence Goury, Michel Jacobson, Isabelle de Lamberterie, Christiane Marchello-Nizia, Lorenza Mondada

▶ **To cite this version:**

Olivier Baude, Claire Blanche-Benveniste, Marie-France Calas, Paul Cappeau, Pascal Cordereix, et al.. Corpus oraux Guide des bonnes pratiques 2006. Version allemande. Presses universitaires d'Orléans; CNRS Éditions, 106 p., 2010. halshs-01165896

HAL Id: halshs-01165896
https://shs.hal.science/halshs-01165896

Submitted on 20 Jun 2015

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

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

VERSION ALLEMANDE

CORPUS ORAUX

Guide des bonnes pratiques 2006

coordonné par **Olivier BAUDE**

PRESSES UNIVERSITAIRES ORLÉANS

CNRS EDITIONS

CORPUS ORAUX

Guide des bonnes pratiques 2006

CORPUS ORAUX

Guide des bonnes pratiques 2006

coordonné par **Olivier BAUDE**

PRESSES UNIVERSITAIRES ORLÉANS

CNRS EDITIONS

Délégation générale à la langue française et aux langues de France
6, rue des Pyramides75001 PARIS
http://www.dglflf.culture.gouv.fr

Dieses Werk ist das Ergebnis der Überlegungen einer um
Isabelle **de LAMBERTERIE** gesammelten Arbeitsgruppe.
Es wurde von Olivier **BAUDE** koordiniert.

Olivier **BAUDE** (*DGLFLF und CORAL – Universität Orléans*)
Claire **BLANCHE-BENVENISTE** (*EPHE und Universität der Provence*)
Marie-France **CALAS** (*DMF*)
Paul **CAPPEAU** (*Universität Poitiers*)
Pascal **CORDEREIX** (*BnF)*
Laurence **GOURY** (*CNRS – CELIA*)
Michel **JACOBSON** (*CNRS – LACITO*)
Isabelle **de LAMBERTERIE** (*CNRS-CECOJI*)
Christiane **MARCHELLO-NIZIA** (*CNRS-ILF und ENS-LSH-Lyon*)
Lorenza **MONDADA** (*ICAR, CNRS, Universität Lyon2*)

In Zusammenarbeit mit :

Gilles **ADDA** (*für das Kollectiv COPTE LIMSI-CNRS*), Michel **ALESSIO** (*DGLFLF*), Alain **CAROU** (*BnF*), Ibrahim **COULIBALY** (*CDF – Universität Grenoble*), Valérie **GAME** (*BnF*), Fabrice **MOLLO** (*CNRS-CECOJI*), Michel **RAYNAL** (*INA*), Jean **SIBILLE** (*DGLFLF*), Dominique **THERON** (*BnF*), Luc **VERRIER** (*BnF*).

Übersetzung : Gisèle **DESOTEUX**

VORSTELLUNG DER VERFASSER

OLIVIER BAUDE
Dozent in Sprachwissenschaften an der Universität Orléans, Mitglied des *CORAL* (EA-3850), Forschungszentrum von Orléans in Anthropologie und Linguistik. Sekretär des wissenschaftlichen Rates der Überwachungsstelle der sprachlichen Praktiken, *DGLFLF*, oberster Rat für die Überwachung des Französisches und der Sprachen Frankreichs.

CLAIRE BLANCHE-BENVENISTE
Verdiente Professorin, Praktische Hochschule von Paris und Universität der Provence. Forscherin im Gebiet der französischen Linguistik : geschriebene und gesprochene Sprache, Syntax, Morphologie, Aufbau von Korpora der gesprochenen Sprache.

MARIE-FRANCE CALAS
Generalkonservatorin des Erbes. Generalinspektorin der Museen, Direktion der Museen Frankreichs. Spezialist für das Tongebiet als umfangreiches interdisziplinäres Gebiet verstanden, das Geschichte, Verwaltung, Bewahrung und Aufwertung der gesprochenen und musikalischen Aufnahmen, der Töne der Umgebung, die heute ein integrierender Bestandteil des immateriellen Erbes sind, einbezieht.

PASCAL CORDEREIX
Oberkonservator der Bibliotheken. Verwalter der Tondokumente in der audiovisuellen Abteilung der *BnF*, Nationalbibliothek Frankreichs ; außerdem Vizepräsident des französischen Vereins der Besitzer von audiovisuellen und Tondokumenten (*AFAS*). Seine Grundtätigkeit richtet sich nach der Problematik der Archivistik des Tones.

LAURENCE GOURY
Forschungsbeauftragte im *IRD*, Forschungsinstitut für die Entwicklung, Mitglied des *CELIA*, Forschungszentrum der einheimischen Sprachen Amerikas), Terrainlinguistik und Typologie (insbesondere Kreolsprachen).

MICHEL JACOBSON
Informatikingenieur im Labor für « Sprachen und Kulturen mit verbaler Tradition » des nationalen Forschungszentrums. Mitverantwortlicher des Programms « Archivierung ». Spezialist für die Verwaltung der Korpora der gesprochenen Sprache.

ISABELLE DE LAMBERTERIE
Forschungsdirektorin im *CNRS*, verantwortlich für den Stab « Normativität und Informationsgesellschaft » des Forschungszentrums über die juristische internationale Zusammenarbeit (*CECOJI* – UMR 6224), Mitglied des Ethikkomitees des *CNRS*.

CHRISTIANE MARCHELLO-NIZIA
Professorin in Sprachwissenschaften an der *ENS-LSH* (Lyon), Leiterin des Instituts der Französischen Sprache (*CNRS*) : Linguistik, Geschichte, Geschichte des Französisches, Theorien der Sprachentwicklung.

LORENZA MONDADA

Professorin in Sprachwissenschaften an der Universität Lyon 2 und Mitglied des Labors *ICAR* (*UMR CNRS* 5191). Arbeitet im Gebiet der interaktionnellen Linguistik über die Korpora der gesprochenen Sprache in Interaktion so wie über die multimodale Analyse von Videokorpora.

VORWORT VON XAVIER NORTH,
GENERALDELEGIERTER FÜR DIE FRANZÖSISCHE SPRACHE UND DIE SPRACHEN FRANKREICHS

In der Wissenschaftsgeschichte oder in den kulturellen Politiken kommt es selten vor, dass gesamte rohe Angaben und ungenaue Stoffe sich in ein Wissensobjekt verwandeln. Die Veröffentlichung dieses Handbuches zählt zu solchen Momenten, denn es bietet jedem Forscher die Hilfsmittel, die « guten Praktiken » an, die ihm ermöglichen werden, diese Metamorphose durchzuführen : die Verwandlung mündlicher Produktionen in ein Korpus der gesprochenen Sprache, das studiert und aufbewahrt werden und infolgedessen seinen Platz in das kulturelle Erbe der Nation nehmen kann.

Die sprachlichen Produktionen von literarischen Werken oder geschichtlichen Dokumenten in ihrer *geschriebenen*, festen und endgültigen Form waren zweifellos immer im Kern der vom Kultusministerium eingesetzten Politiken, ob es sich um das Buch oder um die Archivalien handele. Erst ganz neulich kam man aber auf den Gedanken, dem lebendigen Aspekt der Sprache in ihr spontanes Sprudeln, in ihre tägliche, gewohnte Formulierung und in die außerordentliche Vielfältigkeit ihrer Mundarten Interesse zu widmen... So zeichnet sich für das erste Mal die Möglichkeit ab, ein echtes Archiv des Wortes auf eine gesicherte Grundlage anzulegen. Der Fortschritt der Technologien sollte dazu beitragen.

Ein Korpus der gesprochenen Sprache ist eigentlich keine einfache Sammlung von Aufnahmen der mündlichen Äußerungen des Menschen, es ist ein « entworfenes » Objekt : die Ausarbeitung der Daten (Digitalisierung, Transkribierung, Indexierung) erlaubt nicht nur ihre Bewahrung sondern gibt ihnen den neuen Status von Forschungs- und Aufwertungsobjekt. Man sollte sich wenigstens auf zusammenhängende und leicht anwendbare Methodevorschriften stützen können.

Das « Handbuch der guten Praktiken » öffnet der Neugierde der Forscher von nun an ein neues und umfangreiches Gebiet. Durch seine *Überwachungsstelle der sprachlichen Praktiken* hat der oberste Rat der Überwachung des Französischen und der Sprachen Frankreichs den Impuls gegeben und er hat dann die Energiequellen und die verschiedenen Ressourcen zusammengebracht und koordiniert, ob sie aus der Welt der Forschung kommen oder aus den durch diese Initiative betroffenen Horizonten des Kultusministeriums.

Die Entwicklung der Korpora der gesprochenen Sprache, ihre Verbreitung und ihre Bewahrung zu sichern, heißt auch, sie zugänglich zu machen, das französische linguistische Erbe in seiner Vielfältigkeit, in seinem Reichtum und in seiner Wahrheit hören zu lassen. Das heißt dazu, sich ein kostbares Instrument der Kenntnis der sprachlichen Praktiken zu geben, die für die Bestimmung der Sprachpolitiken aber auch der Erziehungs- und Sozialpolitiken nützlich sind.

Dieses Unternehmen hat mehrere Monate lang Juristen, Sprachwissenschaftler, Konservatoren und Informatiker zusammengebracht, die die neuen Wege der Kultur und der Forschung mit dem Respekt des Rechtes in Einklang bringen wollten. Das

Ergebnis gemeinsamer Denkbemühungen stellen wir heute vor, in der Hoffnung, dass es seinerseits manche Arbeiten befruchtet.

VORWORT VON BERNARD MEUNIER, PRÄSIDENT DES CNRS

Das Mündliche und das Schriftliche. Diese zwei Wörter haben eine starke Aussagekraft. Wir denken an die Art, wie die Kulturen sich durch die gesprochene Praxis und dann durch das Schaffen von Schriften gebildet haben, die es erlaubten, die Worten von den einen oder den anderen durch Raum und Zeit besser zu übermitteln.

Mein Forscherblick über die respektive Rolle des Mündlichen und des Schriftlichen in die Verbreitung der wissenschaftlichen Kenntnisse lässt mich nicht vergessen, dass die mündliche Darstellung vor Berufskollegen oder vor ein breites Publikum weit über die erstgangige Rolle des Geschriebenen immer wesentlich ist, um zu verbreiten, überzeugen, Ideen teilen zu lassen. Die gesprochene Sprache behält eine Überzeugungskraft, die es erlaubt, die Mehrzahl zu erreichen, sobald sie aufgenommen und durch die aktuellen audiovisuellen Mittel übertragen werden kann.

Das Einsammeln und die Verwendung der Korpora der gesprochenen Sprache sollen im Respekt der « guten Praktiken » gemacht werden, wie für die Korpora der gechriebenen Sprache gemacht wird. Wir wissen alle, wie ein Satz, der aus seinem Kontext entnommen und hemmungslos verbreitet wird, für seinen Autor, eine Gruppe Leute oder eine Gemeinschaft gefährlich werden kann.

Die Verfasser dieser merkwürdigen Arbeit haben alle juristische Aspekte des Einsammelns und der Verwendung der Korpora der geschriebenen Sprache angesprochen. Ich wünsche, dass dieses Werk bei den Verfassern und Benutzern der Korpora der gesprochenen Sprache, die wir alle eines Tages sind, die beste Verbreitung genießt.

VORWORT VON JEAN-NOËL JEANNENEY, PRÄSIDENT DER NATIONALBIBLIOTHEK FRANKREICHS

Die Nationalbiliothek Frankreichs freut sich darauf, bei der Ausarbeitung dieses *Handbuches* mitgemacht zu haben. Sie steht zwar in altem und engem Kontakt mit den gesprochenen Sprachen, ihrer Erhaltung und ihrer Verbreitung. Ihre audiovisuelle Abteilung ist der Erbe der *Archives de la Parole* (Archiv des Wortes) von Ferdinand Brunot, die schon 1911 geschaffen wurden. Seit diesem Datum hat sich unsere Einrichtung ständig darum gekümmert, die besten Erhalts- und Bewahrungsbedingungen mündlicher Ausdrücke jeder Art zu sichern, wie ihre Verbreitung in die weiteste Öffentlichkeit.

Die Digitaltechnologien verstärken heute dieses historische und wissenschaftliche Band. Was die Bewahrung betrifft, wurde ein hochfliegender Plan der Digitalisierung unserer Sammlungen engagiert, von dem die Ton- und audiovisuellen Dokumente besonders profitieren. Außerdem wird der Verbreitung dieser Reichtümer zwischen unseren Wänden und aus der Ferne der phantastische Aufschwung unserer digitalisierten online-Bibliothek « Gallica » zustatte kommen, der jedem Internauten erlaubt, Zugang zu diesen Grundquellen des Wissens zu erlangen, egal wo er ist und was der Gegenstand seiner Forschung oder seiner Neugierde ist.

Dieses *Handbuch* ist das Ergebnis einer vertrauensvollen Zusammenarbeit und bezeugt die Komplementarität der Wissen der Sprachwissenschaftler, Juristen, Konservatoren und Informatiker, Ton- und Bildtechniker : ich freue mich darauf, dass die Nationalbibliothek Frankreichs zu diesem innovativen und fruchtbaren Unternehmen beigetragen hat.

1 Vorstellung

1.1 Die Ziele
1.2 Die Bedingungen der Ausarbeitung
1.3 Die juristischen Aspekte
1.4 Die anderen Aspekte
1.5 Die Methode
1.6 Der französische juristische Rahmen
1.7 Ein « Handbuch der guten Praktiken » ?
1.8 Einige häufige Fragen

2 Der Kontext

2.1 Linguistik und Korpora der gesprochenen Sprache
2.2 Politische Rahmen der Verbreitung der Forschung
2.3 Juristische Rahmen

3 Das Verfahren

3.1 Das Verfahren erläutern
3.2 Elemente der betroffenen Situation
3.3 Terrainpraktiken
3.4 Anonymisierung
3.5 Transkription

4 Sind die Korpora der gesprochenen Sprache Erbobjekte ?

4.1 Erinnerung an die Situation
4.2 Die privaten Initiativen
4.3 Der Zugang zu den Sammlungen

5 Anhang

Abkürzungen

1 VORSTELLUNG

1.1 DIE ZIELE

Viele Grundlagen- oder angewandte Forschungen beruhen zur Zeit auf der Auswertung von „Korpora der gesprochenen Sprache“ (geordneten Sammelwerken der Aufnahmen von mündlichen und multimodalen sprachlichen Produktionen). Dieses *Handbuch der guten Praktiken* entsteht aus der Erkenntnis von Sprachwissenschaftlern, die darauf bedacht sind, den Fortbestand der Quellen und einen verschiedenartigen Zugang zu den mündlichen von ihnen produzierten Produktionen zu sichern ; es schneidet zuerst die „Korpora der gesprochenen Sprache“ an, die von Sprachwissenschaftlern und für sie geschaffen und verwendet wurden. Die durch die Erschaffung und die dokumentarische Auswertung dieser Korpora hervorgerufenen Fragen trifft man aber in vielen Fächern : die Völkerkunde, die Anthropologie, die Soziologie, die Psychologie, die Demographie, die mündlich überlieferte Geschichte gebrauchen vor allem die verbale Befragung, die Aussage, das Interview, die Lebensgeschichte. Dieses *Handbuch* beruft sich auf das Verfahren der Sprachwissenschaftler, es stimmt aber mit den Beschäftigungen anderer Forscher überein, die Korpora der gesprochenen Sprache (z. B. in Sprachsynthese und -entzifferung) gebrauchen, auch wenn ihre spezifischen Bedürfnisse im vorliegenden Dokument nicht systematisch angeschnitten werden.

Dieses *Handbuch* hat sich als erstes Ziel gesetzt, die für den Korporaaufbau von mündlichen oder multimodalen Daten erforderlichen *Informationen* zu liefern und *Vorschläge* anzubieten, die sich nicht nur auf die juristischen Punkte beziehen sondern auch auf die materiellen Aspekte, was ebenso das Einsammeln betrifft wie die Strukturierung und die Formgebung der Daten, die Auswertung, Mitteilung und Verwaltung dieser Daten.

Das zweite Ziel dieses *Handbuch* besteht darin, den Forschern, die Korpora der gesprochenen Sprache anlegen oder erweitern, dabei zu helfen, bestimmte „verzögerte Schwierigkeiten“ *vorauszusehen*, die die Auswertung und dann die zukünftige Entwicklung ihres Korpus schwer zu belasten drohen. Bestimmte Anfangsentscheidungen, bestimmte Mängel können sich in späteren Etappen des Prozesses als wichtig erweisen, wenn es zu spät ist, um irgendwas zu ändern.

Das dritte Ziel heißt, das Auftauchen *gemeinsamer Praktiken* zu begünstigen, um sowohl beim Aufbau wie bei der Verwendung der Daten den aktuellen Vorschriften der Bewahrung und der Interaktionen zwischen den Korpora, der Bewertung und der Ethik zu entsprechen.

1.2 DIE BEDINGUNGEN DER BEARBEITUNG

In Frankreich z. B. hat sich der wissenschaftliche Rat der Überwachungsstelle der sprachlichen Praktiken (*DGLF*) gewünscht, die Aktionen für die Bewahrung, den Aufbau und die Aufwertung der mündlichen und multimodalen Korpora stark zu fördern, und dies für folgende Gründe :

– die Bewahrung eines reichhaltigen Erbes der Sprachpraktiken zu ermöglichen ;
– zum Aufbau umfangreicher Referenzkorpora beizutragen, die für die Forschung, den Unterricht, die Sprachindustrien aber auch für das Erbe angelegt werden ;
– zur Entwicklung der Informatikmittel beizutragen, die die Bearbeitung, die Erweiterung und die Aufwertung der Korpora ermöglichen ;
– die Bereitstellung dieser Korpora begünstigen.

1.3 DIE JURISTISCHEN ASPEKTE

Die juristischen Aspekte, die mit dem Aufbau und der Verwendung der Korpora der gesprochenen Sprache gebundenen sind, erwiesen sich sehr schnell als ein häufiges und wesentliches Hindernis.

Diese juristischen Aspekte betreffen hauptsächlich die Fragen der moralischen und erblichen Rechte und des Eigentums der Daten, die bei jeder der vier wichtigen Etappen der Korpusarbeit zu begegnen sind :

– das Einsammeln der Daten und die Aufnahme (Recht auf das Bild, auf die Stimme, Befragungssituation, Erlaubnisse...) ;
– die Verwendung und die elektronische Auswertung der Daten (Archivierung, Datenbank für die Forschung, den Unterricht, die Projektplanung...) ;
– die Verbreitung und das In-Umlauf-Bringen der Daten (Rechte, Rechte auf Zitat, Online-Verbreitung...) ;
– die Verwaltung der Daten.

In Anbetracht der vielen betroffenen Bereiche hat die DGLF die Schaffung eines Komitees mit Fachleuten von verschiedenen Fächern hervorgerufen. Dieses Komitee hat eine Arbeitsgruppe errichtet. Sie soll den Forschungsmannschaften dabei helfen, die Praktiken der Korporaeinsammeln und –Auswertung zu normalisieren, vom Standpunkt der Gesetzgebung und unter Berücksichtigung der gesamten mit der Forschung gebundenen Zwänge. Das hierbei vorgestellte Handbuch ist das Ergebnis der fünfzehnmonatigen Arbeit dieser Gruppe.

Diese Arbeitsgruppe sollte natürlich Rechtswissenschaftler einschließen, die Spezialisten für das Forschungsrecht sind, aber nicht nur : der Bedarf an Kompetenzen bezüglich des Aufbaus der Korpora, der Verwendung und der Verwaltung haben dazu geführt, den Rechtswissenschaftlern Linguisten zuzuteilen, die „Korpuslinguistik“ ausüben und mit mündlichen Angaben arbeiten, Vertreter der hohen Institutionen für die Verwaltung des Erbes (IDS, DWDS) und Informatiker, die in Korpusverwaltung spezialisiert sind.

Um ihre Aufgabe zu erfüllen hat sich diese Arbeitsgruppe unter anderem folgende Ziele gesetzt :

– den Bestand der aktuellen Praktiken aufzunehmen und die methodologischen und theoretischen Zwänge zuerst zu bestimmen, die mit der Forschung gebunden sind ;
– eine Synthese über die aktuelle Gesetzgebung zu verbreiten ;

– Empfehlungen zusammenzustellen ;
– wenn nötig, im Falle einer Lücke oder einer Unklarheit, Vorschläge für die Ausarbeitung von juristischen (unter anderem europäischen) Normen und Regelungen zu formulieren.

Deswegen sollte man zuerst :

– den Bestand der betroffenen juristischen Gebiete aufnehmen ;
– die Risiken bestimmen und quantitativ erfassen ;
– die bestehenden Antworten erkennen ;
– und dann diese Antworten in Form von Empfehlungen der guten (juristischen und ethischen) Praktiken aufbauen.

Infolgedessen hat die Gruppe entschlossen, mit mehreren Zeugeteams, die mündliche oder audio-visuelle Daten sammeln oder gesammelt haben, eng zu arbeiten.

Gezielt wurde dadurch das Erlangen einer „Typologie der Situationen“ und das Aufzählen aller schon gebrauchten Praktiken und Lösungen, in Deutschland oder anderswo.

1.4 DIE ANDEREN ASPEKTE

Der Arbeitsgruppe wurde unterwegs klar, daß es keine befriedigende Antwort auf die vorhandenen Schwierigkeiten wäre, lediglich eine Liste Empfehlungen oder Lösungen juristischer Natur vorzuschlagen.

Zwar hat es sich gezeigt, daß die Schwierigkeit oder die Lösung ja oft mit der Art der Einsammeln- oder Verwendungpraktik gebunden waren ; daß einige Lösungen durch technische Wege zu erreichen waren, die die Angaben selbst beeinflussten (Anonymisierung oder Verschwommung) ; daß es überhaupt nicht gleichgültig war, dieses oder jenes juristische Problem in diesem oder im anderen Moment zu lösen. Kurz gesagt, Lösungen zu juristischen Fragen vorzuschlagen, hieß soviel wie den Prozess selbst des Einsammelns oder der Formgebung, der Weiterleitung oder der Verwendung dieser Art Daten anzuschneiden.

Schließlich stellte sich über den Respekt hinaus, den man den Rechten der aufgenommenen Personen schuldet, die Frage der „Urheberschaft“ dieser Art von Daten : welches sind die Rechte der Sammler dieser Daten ? Wer ist juristisch verantwortlich dafür, wer darf sie weiterleiten ? Die juristischen Aspekte, die mit dem wissenschaftlichen Urheberschaft oder der strafrechtlichen Verantwortung gebunden sind, waren auch, wie man sieht, von der Einsammeln- und Verwendungspraktik der Daten nicht zu trennen.

War es infolgedessen nicht besser, die Sachkenntnis des geplanten „Handbuches“ zu erweitern und nicht nur die juristischen Praktiken zu erwähnen, sondern auch die gesamten Praktiken, die diese Art Korpus einsetzt ? Diese Entscheidung wurde getroffen, denn so war es möglich, alle Aspekte gebunden zu halten, so wie sie in der Wirklichkeit sind.

1.5 DIE METHODE

Die von der Arbeitsgruppe gewählte Methode zeichnet sich durch folgende Züge aus :

- die Überzeugung, dass man nicht glauben lassen darf, dass es für jede Situation vorgefertigte Antworten gibt ;
- den Wille, die Forscher nicht zu „zügeln“ (z. B. wenn man bestimmte Praktiken verbietet) ;
- die Einhaltung der Methodologie des Forschers und der mit der Beobachtung gebundenen Zwänge (die Forscher wünschen, Situationen aufzunehmen, ohne dass die vor allem technischen und juristischen Zwänge sie ändern ;
- die Notwendigkeit des Zusammenbringens der Kenntnisse, die bei den verschiedenen Etappen erforderlich sind (Sprachwissenschaftler, Juristen, Verwalter) für die Ausarbeitung und die Verfassung dieses Handbuches ;
- die Kundgebung eines auf dem Respekt des Gesetzes und der Ethik beruhenden Verfahrens ;
- die Notwendigkeit, durch dieses *Handbuch* ein Gutachteninstrument der Risiken (um sie ausfindig zu machen aber auch zu schätzen) zu beschaffen.

1.6 DER FRANZÖSISCHE JURISTISCHE RAHMEN

Etliche Fragen und Lösungen betreffen die *Einwilligung* der Befragten aber auch die Verantwortung der *Copyrightbesitzenden* Instanzen. Da ist sicher ein Hauptpunkt. Er ist aber nicht der einzige auf dem Spiel und die Antworten auf diese Frage haben sich komplex erwiesen.

Die Art und Weise, die Einwilligung und die Erlaubnis heute aufzunehmen, ist verschiedenartig. Es gibt keine anerkannten Normen und viele Schwierigkeiten.

Zuerst soll die Einwilligung *aufgeklärt* sein (Rahmen, Zwecke, „Risiken“ für den Befragten).

Die Aufnahme der Einwilligung a priori kann manchmal die Befragung stören (Paradox des Beobachters), indem sie eine Situation formalisiert, während „natürliche“, dem umgänglichen Gespräch nahe Daten gewünscht sind.

So hat sich zum Beispiel die Praktik interessant und wirksam erwiesen, die darin besteht, den Befragten, (zu der Aufnahme der Einwilligung) ein Dokument zu lassen, das den Rahmen, die Zwecke, die Risiken, den Zugang und die Informationen erläutert, die es erlauben, das zugrunde liegende Bezugssystem der Veröffentlichungen und der Ergebnisse im nachhinein wiederzufinden.

Die Schwierigkeit besteht auch in einem *Widerspruch* zwischen dem Zwang, die Zwecke der Befragung zu erklären, um die Einwilligung aufzuklären, und der Unmöglichkeit, alle Zwecke und *die zukünftigen Gebrauchsmöglichkeiten der Daten* im voraus vorherzusehen, *in Anbetracht der aktuellen Bemühung*, eine *maximale* Interaktionsmöglichkeit zu erlangen.

Dazu ist auch daran zu denken, dass einige verbale Kulturen (und nicht nur am Ende der Welt) keine Möglichkeit geben, eine schriftliche Spur der Einwilligung vorzuschlagen und zu bewahren.

Alle andere Fragen juristischer Natur bieten dieselbe Komplexität an : Anonymität, Verschlüsselung, Verschwommung, Bestimmung der Haftungen, Depotzwang, Mitteilungen, usw., alle Praktiken, die zwangsläufig mit dem Aufbau und dem Bestehen eines Korpus der gesprochenen Sprache gebunden sind. Keiner dieser Aspekte beruht auf einer einzigen, klar definierten und überall anerkannten Praktik.

Jede dieser Etappen wird allmählich mit technischen Wahlen, mit sozialen oder wissenschaftlichen Praktiken eng gebunden, denn das alles ist schwer voneinander zu trennen.

Daher die Entscheidung der Arbeitsgruppe, ein Handbuch anzubieten, das nicht nur ein "juristischer Leitfaden" ist, sondern auch eine praktische, zuverlässige Hilfe, die alle Aspekte des Prozesses in Betracht zieht.

1.7 EIN „HANDBUCH DER GUTEN PRAKTIKEN" ?

Dieses Handbuch berücksichtigt den juristischen Rahmen in Frankreich (und in einigen Ländern in Europa) und stützt sich auf die Befragungen der Forscher, die zu seinem Aufbau beigetragen haben. Diese haben versucht, die Grundlagen der geltenden juristischen Regeln und das, was mit ihrem Respekt und ihrer Umsetzung aufs Spiel gesetzt wird, zu verstehen. Eine *dynamische Vision der juristischen Regulation* bildet also durch das von den Forschern gefolgte Verfahren die Grundlage für dieses Handbuch. Die Autoren dieses Handbuchs sind auch selbst in die betroffenen Forschungsgebiete verwickelt und haben sich darum bemüht, Praktiken und Sitten vorzuschlagen, die das bestehende Recht einhalten. Infolgedessen soll das Verfahren des Forschers darin bestehen, diese Rechte und die Zwänge, die sich davon ableiten lassen, zu kennen. Dann wird es darum gehen, die Konsequenzen dieser Zwänge in die Phase des Einsammelns der Daten so wie in die der Aufwertung zu ziehen.

Damit ein solches Verfahren auf eine strikte und glaubwürdige Art und Weise vorgestellt werden kann, soll es zuerst in seinem wissenschaftlichen, politischen, juristischen oder institutionnellen Kontext dargestellt werden. Die vorgeschlagenen Gebräuche und Praktiken werden im Laufe des Handbuchs durch diesen Kontext „aufgeklärt" werden, so dass man besser verstehen kann, was durch den Respekt oder durch die Respektlosigkeit dieser Gebräuche und Praktiken aufs Spiel gesetzt wird.

1.8 EINIGE HÄUFIGE FRAGEN

Das erste Ziel dieses Handbuchs heißt, Informationen und Elemente der Antworten auf die Fragen zu geben, die sich all die Forscher oder Verantwortlichen für den Aufbau, die Auswertung, die Verwaltung und die Verbreitung dieser Korpora stellen.

Um dieses Ziel zu erreichen, wurde das Handbuch mit vielen Verweisen, die ebenso viele mögliche Lesenswege bilden, konzipiert. Die folgenden Fragen bilden die

traditionnellen Befragungen am Anfang eines Forschungsprojekts und schlagen also ein erstes Wegbeispiel vor.

FRAGEN

1. Welche Erlaubnisse soll ich von den aufgenommenen Sprechern unterschreiben lassen, um dann dieses Korpus benutzen zu können und folgendes zu können:

a. es in einer akademischen Arbeit zitieren ;

b. es in einem Artikel zitieren, der in einer wissenschaftlichen Zeitschrift veröffentlicht ist ;

c. es in einem Werk mit kommerziellen Verbreitung zitieren ;

d. es auf eine Website zur Verfügung stellen ;

e. es auf einer CD verbreiten ;

Unterziehen sich diese verschiedenen Auswertungsmittel denselben Regeln ?

Antwort: Die Fragen a, b, und c unterstehen dem Zitatrecht. Die Elemente der Antworten auf die Fragen d und e sind unter anderem in den Kapiteln 2.1.5, 2.3 und 3.5 vorgestellt

2. *Ich habe Leute aufgenommen, die ich gut kenne*.

a. Unter welchen Bedingungen darf ich die Aufnahme benutzen ? (benutzen im Sinne der Frage 1)

b. Dürfen sie ihre Erlaubnis zurücknehmen ?

Antwort: Das gesamte Kapitel 3.4 ist eine Betrachtung der Einsammelnbedingungen der Daten, die zu den zahlreichen Schwierigkeiten sensibilisieren möchte, die sich im Laufe des Einsammelns ergeben können. Die Tatsache, dass man die betroffenen Leute gut kennt, mindert die juristischen Zwänge nicht (ihnen gegenüber), im Gegenteil (sie stellt Vertrauensfragen, die zu ziemlich komplexen Situationen führen können).

3. *Wenn ich Kinder aufnehme*,

a. Wer kann seine Erlaubnis geben ?

b. Kann das Kind seine Erlaubnis zurücknehmen, wenn es volljährig wird ?

c. Sind besondere Erlaubnisse erforderlich, wenn die Aufnahme im schulischen Rahmen stattfindet ?

Antwort: Dieser Fall stimmt dem allgemeineren Fall der Leute überein, für die eine zusätzliche Erlaubnis von den Betreuern (Eltern und Schule in diesem Fall) erforderlich ist (vgl Kapitel 3.3.2 Gruppe der Teilnehmer).

4. *Im Rahmen einer Arbeit innerhalb eines Labors*,

a. Wer wird als Verfasser des Korpus angesehen ?

b. Welches Recht/ welche Rechte gibt diese Arbeit dem Forscher ?

Antwort: Vgl Kapitel 2.3 (Urheberrecht).

5. *Wer ist als „verantwortlich" für die Verbreitung und die Bearbeitung eines Korpus angesehen ?*

Antwort: Vgl Kapitel 2.3.

6. *Ist es genügend, wenn ich die Eigennamen der Leute verdecke, damit ich ein Transkript frei gebrauchen kann ?*

Antwort: Die Anonymisierung heißt nicht nur, dass die Eigennamen verwischt sind. Vgl Kapitel 3.5 Anonymisierung.

7. Unter welchen Bedingungen darf ich mein Korpus als elektronische Dateien archivieren?

Antwort: Die juristischen Aspekte müssen berücksichtigt werden (Schutz des Privatlebens, Gesetz über Informatik und Freiheit, Genehmigungsantrag.

8. *Darf ich die Aufnahmen (in den Medien oder privat) frei ausnutzen, wenn die aufgenommenen Leute gestorben sind ?*

Antwort: Die Urheberrechte bestehen (in Frankreich) 70 Jahre nach ihrem Tod! Was den Schutz in bezug auf das Privatleben betrifft, er kann erst nach dem Tod der betroffenen Person vorgebracht werden, außer wenn sie zu ihrer Lebzeiten die Verbreitung verboten hat. Die Mitglieder der Familie des Gestorbenen können dazu ihr persönliches Recht auf den Schutz des Privatlebens vorbringen. Vgl Kapitel 2.3.1.

9. In einem Schrank entdecke ich Aufnahmen. Ich möchte sie verwerten können. Ich habe keine Spur mehr von dem, wer aufgenommen hat oder aufgenommen wurde.

a. Darf ich sie gebrauchen ?

b. Welche Sicherheitsvorkehrungen (welche Garantien) muss ich treffen ?

Antwort: Man könnte nicht zu vorsichtig sein und Recherchen sind nötig, damit die Tonaufzeichnungen einschließlich wegen wissenschaftlicher Gründlichkeit identifiziert werden. Vgl Kapitel 2.3 und Kapitel 3.5.

10. Ich nehme eine Radio- oder Fernsehsendung auf.

a. Darf ich das Transkript frei verwenden ?

b. Darf ich die Tonfassung verwenden ?

c. Gibt es einen Unterschied zwischen den Sendungen der öffentlichen Sender und den der Privatsender, was die Erlaubnisse betrifft ?

d. Gibt es einen Unterschied zwischen der Aufnahme bekannter Persönlichkeiten und der „anonymer" (Leute, die etwas berichten, sich frei ausdrücken, Hörer, die Fragen stellen, usw.) ?

e. Sind die Auswertungsrechte verschieden, wenn ich eine Kassette, eine DVD oder eine CD der Sendung kaufe oder wenn ich persönlich die ausgestrahlte Sendung aufnehme ?

Antwort: Die Radiosendungen stehen unter Schutz, ob sie öffentlich oder privat sind. Vgl 3.3.1 über die Wiederholung der Medienaufnahmen und besonders den Begriff der Aktualitätsdokumente.

11. Ich möchte ein Korpus mit authentischen Daten aufbauen. Welche Sicherheitsvorkehrungen soll ich treffen ?

Antwort: Vgl Kapitel 3, wo eine Betrachtung vorgeschlagen wird, die die Methodologie der praktischen Forschung mit den im Laufe des Verfahrens getroffenen ethisch-juristischen Problemen miteinander verbindet.

Und noch viele andere Fragen...

2 DER KONTEXT

DER WISSENSCHAFTLICHE, POLITISCHE, JURISTISCHE UND INSTITUTIONNELLE KONTEXT

Wer vom Kontext spricht, spricht von „Perspektivierung“ . Es ist der Zweck dieses Kapitels, das darstellt, was die wissenschaftliche Arbeit des Sprachwissenschaftlers über das Mündliche ist. Die Perspektivierung sollte auch *politisch* und *juristisch* sein. Der institutionnelle Kontext wird immer wichtiger, da es nötig ist, auf Dauer die „Rückverfolgbarkeit“ und die Fortsetzung der Forschungen zu sichern. Indem der Forscher oder die Institution die Verewigung der Daten, die einem Forscher zu arbeiten erlaubt haben, so wie sie der erreichten Ergebnisse garantiert, nimmt er/sie an der Entwicklung der Kenntnisse in einer nahen oder fernen Zukunft teil.

2.1 LINGUISTIK UND KORPORA DER GESPROCHENEN SPRACHE

Die Studien über die Korpora der gesprochenen Sprachen haben seit ungefähr zwanzig Jahren den Sprachwissenschaften ein völlig neues Gesicht gegeben. Um sich davon zu überzeugen, braucht man nur, in den jüngsten Bibliographien in Frankreich und außerhalb Frankreich nachzuschlagen (zum Beispiel in der *Revue Française de Linguistique Appliquée* oder in den *Recherches sur le Français Parlé*). Diese Studien haben es erlaubt, neue Hypothesen über das normale und pathologische Funktionieren der Sprache zu formulieren und sind eine wesentliche Komponente des Dialogs zwischen Sprachwissenschaftlern und Informatikern geworden. In Frankreich widmete sich das Interesse für die gesprochenen Sprachen bis vor kurzem hauptsächlich den Bereichen, wo es sich „von sich selbst“ bemerkbar machte : zuerst in den Studien über die typisch stimmhaften Aspekte der Sprache (Phonetik, Phonologie und Prosodie), in der Sprechweise der jungen Kinder, oder alles, was unter den „Sprachen ohne verbale Bräuche“ bezeichnet wurde, d. h. in den regionalen Sprachen und lokalen Mundarten in Frankreich und alles, was als „exotische Sprachen“ benannt wurde, außerhalb Frankreich. Dazu kamen in den Fünfziger und Sechziger Jahren einige einzelne Versuche, um für den Unterricht des Französischen als Fremdsprache Musterbeispiele des gesprochenen Französischen zusammenzubringen.

Die Darstellungen der französischen Sprache beruhten immer, insbesondere in den Grammatikbüchern, auf Daten der literarischen oder nicht literarischen gesprochenen Sprache, auf den „Grapholekten“, wie Ong (1988) sagte, oder auf intuitiv erworbenen Daten. Diese Fernhaltung der Daten der gesprochenen Sprache hat zwei bedeutende Konsequenzen mit sich gebracht, einerseits das sehr negative Image, das die Franzosen von ihrer Sprache haben, und andererseits einen sehr wichtigen Einfluss auf die geläufigsten Sprachtheorien. Die neuen von den Korpora der gesprochenen Sprache aufgedeckten Daten haben ja zweifellos das Image der Sprache im breitem Publikum noch nicht ändern lassen, sie haben aber schon die Theorien unter den Fachleuten ändern lassen.

Neue Gebiete, die schon in den Siebziger Jahren in Grossbritannien (Sinclair und Coulthard für die Schule von Birmingham) angeschnitten wurden, sind in Frankreich

erschienen, wie der Aufschwung der Interaktionsmuster und die konversationelle Analyse (Gründungsbericht von Sacks, Schlegloff, Jefferson in den Vereinigten Staaten in 1974, Artikel von Bange und von Quéré in Frankreich, in 1983 und 1984).

Die vor der Informatikära gesammelten Daten der gesprochenen Sprache können mit dem, was heute „Korpus der gesprochenen Sprache" benannt wird, nicht verglichen werden. Jede von den ehemaligen Sammelungen, die den Recherchen nach zerstreut wurden, folgte ihren eigenen Wahl-, Aufnahme-, Transkriptions- und Bewahrungsregeln, so dass es heute schwer ist, dazu Zugang zu haben und sie zusammenzubringen (die Aufnahmen des *Français Fondamental* wurden gelöscht, die des *Corpus d'Orléans* sollen heute transkribiert werden). Keine konnte einen sehr großen Ausmaß erreichen (es handelte sich meistens nur um einige aufgenommene Stunden) und in diesen Daten konnten Informationen nur manuell gesucht werden. Ab den Achtziger und Neunziger Jahren hat die Entwicklung der Informatik es erlaubt, moderne Korpora der gesprochenen Sprache in der gesamten Welt zu schaffen, zuerst in den angelsächsischen Ländern. Ein neues Fach wurde geboren, das der Korpuslinguistik (G. Kennedy hat sie 1998 für das Englisch beschrieben und Habert und seine Mitarbeiter 1997 für das Französisch), die die Hochschullehrer und die Sprachindustrien betrifft und die jetzt als *Language Resources* zu den nationalen Erben gehören. Frankreich war für die Anfertigung der Korpora der gesprochenen Sprache voraus (insbesondere für FRANTEXT, Quelle des *Schatz[es] der französischen Sprache*) und ist in der Fertigung der Korpora der gesprochenen Sprache zurückgeblieben.

Es gibt viele Korpora der gesprochenen Sprache, die für verschiedene Zwecke in verschiedenen Fächern vorgesehen sind. Es handelt sich immer um Aufnahmen von Tondaten, eventuell mit (Video- oder Fernseh-) Daten, die fast immer von Transkripten und elektronischen Ausarbeitungen begleitet sind. Ohne zu behaupten, hier alles vorzutragen, werden wir vier Aspekte vorstellen : die Daten- und Sprechertypen, die Größe der Korpora, die Transkripte und einen schnellen Überblick der Auswertungen und der Ergebnisse.

2.1.1 *DATEN- UND SPRECHERTYPEN*

Einige Daten werden „ersucht". Man lässt zum Beispiel in Phonetiklabore Sprecher kommen, die als „Versuchskaninchen" Aussprache- und Betonungstypen in sehr guten Aufnahmebedingungen liefern. Man lässt sie Wörter und Wörterlisten, Ziffern und Ziffernlisten aussprechen oder man lässt sie Texte oder Texteauszüge lesen. Diese Dokumente dienen zu unterschiedlichen Auswertungen, entweder um die Aussprachen als solche niederzuschreiben und zu studieren, wie J. Durand, B. Laks und Ch. Lyche machen, um die Aussprache des zeitgenössischen Französisches zu studieren (Projekt PFC), oder um ein Sprachbenehmen zu testen (wie man in den Krankenhausabteilungen macht, die Aphasiephänomene beobachten), oder um Analysen aufzustellen, die zu der elektronisch erzeugten Sprache, dem automatischen Lesen geschriebener Texte (*Text-to-Speech data*) oder den Mensch-Maschinen Dialogen (es ist der Zweck von *SpeechDat Exchange*, das 500 bis 5000 Telefonaufnahmen für 28 Sprachen speichert) dienen. In all diesen Situationen

wissen meistens die Sprecher, dass sie aufgenommen werden und sie haben eine genaue oder vage Ahnung des Zweckes ihrer Leistungen.

Andere Daten werden mit verschiedenen Spontaneitätsstufen „von andauernder Rede" gesagt (der Begriff wurde in einer Auflage der *Revue Française de Linguistique Appliquée* speziell studiert). Einige werden in Situationen gesammelt, die nicht vom Forscher verursacht worden sind, und die sowieso ohne ihn passiert wären. Andere, die mehr oder weniger „ersucht" werden, werden von dem Forscher inszeniert und organisiert. Das Ideale des totalen Spontanen wäre, die Sprecher aufzunehmen, ohne dass sie dessen bewusst sind (versteckte Mikros, Raubaufnahmen), und dass es ihnen nachher gesagt wird oder nicht, da das Ziel, ihre Sprache „völlig frei" mit der Mindestkontrolle festzuhalten ist. Die juristischen Bestimmungen begrenzen diese Möglichkeit. Die Anwesenheit des Forschers und der Geräte bremst sowieso diese Freiheit (es ist die Frage des „Paradox[es] des Beobachters", die von W. Labov popularisiert wurde). In der Praxis können verschiedene Zwangstufen identifiziert werden, je nachdem, ob es sich um privates oder um öffentliches Wort handelt, vor bekannten oder fremden Leuten, mit verschiedenen Einverständnisformen oder ohne sie, ob es sich um ein Gespräch in Form eines Interviews handelt oder das durch ein Medium wie das Telefon, den Anrufbeantworter, das Radio, das Fernsehen oder andere technischen Einrichtungen übertragen wird. Eine richtige ethnographische Herangehensweise (wiederholte Aufnahmen) erlaubt es, das Problem der Empfindlichkeit auf das Mikrophon zu lösen. Es verlangt aber, dass mehr Zeit in der Phase des Dateneinsammelns diesem Punkt gewidmet wird.

Selten bestehen die modernen Korpora aus „ungesiebten" Worten. Die Sprecher und die Aufnahmesituationen werden im allgemeinen den am Anfang gegebenen Zwecken entsprechend gewählt. Die Forscher schlagen vor, Gespräche zwischen Erwachsenen, berufliche Verhandlungen, (vorbereitete oder nicht vorbereitete) Unterredungen, Wortergreifen in öffentlichen Einrichtungen, Wahlreden, Erklärungen zwischen öffentlichen Diensten und Gebrauchern, öffentliche Vorlesungen, Predigte, politische Reden, (Fach- oder populärwissenschaftliche) Vorträge, historische Aussagen, Nachrichtberichte, (von Individuen, Gruppen, Gruppenvertretern, Wortführern verfassten) Lebensgeschichten, Dialoge zwischen Müttern und jungen Kindern, Kinder, die im schulischen Kontext oder außer der Schule aufgenommen werden (in ihren Spielen oder in ihren Erzählungen, als Antwort auf Tests oder außerhalb, in schulischen oder nicht schulischen Situationen, in freien oder gezwungenen Spielen, mit Parodie und Rollenspielen), Kranken in den Krankenhäusern, usw. Zum Beispiel wird eine Datenbank CLAPI zur Zeit in Lyon (Frankreich) angelegt, um die diversifiziertesten Korpora der „Sprache in Interaktion" wie möglich zu sammeln, in Situationen, die nicht von den Forschern verursacht sind : Tischgesprächen, Abstimmungen zwischen Notaren, Anrufen an Zentren für Sozialhilfe und an Therapiezentren, usw. Diese Datenbank umfasst 300 Stunden Audio- und Viedeoaufnahmen, Transkripte und „Metadaten", die die Merkmale der Sprecher beschreiben.

Manche Fächer versuchen, die Zusammenhänge zwischen die Produktionen der gesprochenen Sprache und andere Phänomene zu studieren. Die Zusammenhänge zwischen Sprache und sozioökonomische Komponenten waren die Grundlagen für

die Forschungen der Soziolinguistik. In den Vereinigten Staaten hatte W. Labov bekannte Studien über die Schwarzen in den Ostmetropolen verfasst, indem er eine Forschung über sie in den Wohnungen, auf den Straßen oder in den Kaufhäusern (mit oft schlechten Aufnahmebedingungen) führte. Die Studien über die Entwicklung der Sprache werden je nach dem Alter der Kinder, den beobachteten Beschäftigungen, den gegebenen Anweisungen und den Familiendaten geführt. Die Berücksichtigung der „Genera" (wie D. Biber sie für das Englisch konzipiert) führt dazu, die Orte der Aussagen, die betroffenen Themen, die Sprecher- und Austauschtypen (Monologe, Dialoge, Gespräche zu mehreren) in Korrelation zu bringen. Um diese Korrelationen studieren zu können, sind der Inhalt und die Größe der Korpora meistens im voraus bestimmt : so und so viele Situationen- und Sprechertypen (wie die Arbeitsgruppe Sankoff-Cedergren es in den siebziger Jahren gemacht hatte, um die soziale Schwankung in Montreal zu studieren). In anderen Fällen schneiden die Forscher aus existierenden Korpora repräsentative ihrem Studium angepasste Unterkorpora aus (das hat D. Biber vorgeschlagen, um repräsentative Querschnitte im großen *British National Corpus* hervorzubringen). In diesem Fall handelt es sich um „geschlossene" und „quergeschnittene" Korpora.

Die Sprachwissenschaftler haben ihrerseits oft „offene" Korpora eingesammelt, die sie nach dem Fortschreiten ihrer Arbeit ändern, ohne im voraus ein vorherbestimmtes Forschungsobjekt zu begrenzen, denn sie sind sicher, dass sie neue Phänomene entdecken werden, die am Anfang nicht vorherzusehen waren : Verteilung der formalen und unformalen Sprache, Zusammenhänge zwischen Grammatik und Wortschatz, Verknüpfungen zwischen den Komplexitätsgraden der Syntax und den Typen der Sprechsituationen, Gebrauch der mündlichen Morphologie, Funktion der Kontexte in den Aufbau des Sinnes der Aussagen, Funktion der Prosodie in die Strukturierung der Texte, usw.

Die technische Qualität der Aufnahmen hängt natürlich von den gebrauchten technischen Geräten ab, aber auch von den gewählten Situationen- und Sprechertypen (laute Plätze, zu viele Sprecher, Sprecher mit einem Sprachfehler). Diese verschiedenen Situationen haben auch einen Einfluss auf die Einwilligung der Sprecher : es ist leichter, die Einwilligung für die Aufnahme der öffentlichen Sprache zu bekommen als für die der privaten Sprache, die Aussagen eines selbstsicheren Sprechers aufnehmen zu dürfen als die eines, der unruhig ist und für das, was man „die linguistische Unsicherheit" nennen konnte, empfänglich.

Es ist jedoch schwer, die Aufnahmen durch die Studie der Sprache zu rechtfertigen. Wenn man diesen Zweck erwähnt, haben die französischen Sprecher unvermeidlich den Eindruck, dass sie schlecht sprechen und dass die Studie sie lächerlich erscheinen lassen wird. Wenige nehmen es locker mit dieser Frage. Fast alle Forscher haben Strategien entwickelt, um das Problem schräg anzusprechen : sie sagen, dass sie sich für den Inhalt, die Aussagen, die Erklärungen, das besondere Wissen der Sprecher (das ein Sprachwissen sein kann, im Fall der Forschungen über Regionalismen) interessieren.

In den Studien über die Sprache in Interaktion ist es ein bisschen anders : die Forscher können sagen, dass sie sich genau für die Art und Weise interessieren, wie die Teilnehmer interaktiv agieren, für ihre Koordination, für die merkwürdig

treffenden Anpassungen, die sie durch die Sprache, die Gebärden, die Mimik, die Blicke und die gesamten Körperhaltungen (multimodale Mittel, die schwer zu kontrollieren sind, auch von Sprechern, die auf sich achten) in Anspruch nehmen.

2.1.2 *GRÖßE*

Die sinnvolle Größe der Korpora und der bildenden Einheiten ist je nach der vorgesehenen Studie unterschiedlich. Die Phonetik-, Phonologie- und Prosodiestudien können mit ziemlich kurzen Tonaaufnahmen gute Ergebnisse haben. Wenn man aber die Beziehungen zwischen der Sprache und anderen Phänomenen oder den Wortschatz studieren will, braucht man größere, zahlreichere Einheiten in diversifizierteren Tätigkeitsbereichen. Die Größe der Korpora der gesprochenen Sprache und der bildenden Elemente kann man mit zwei Maßeinheiten messen. Zeiteinheiten werden gebraucht, wenn man sich in erster Linie für die Tonaufnahme interessiert, und dabei von dem Transkript ansieht. Als sehre kleine Korpuselemente geordnet werden zum Beispiel diejenigen, die zwischen vierzehn und dreißig Sekunden dauern (vierzehn Sekunden ist die Durchschnittslänge einer Meldung am Radio). Noch kürzere Untereinheiten werden aber berücksichtigt, wenn die Sprechüberschneidungen zwischen Sprechern oder die Pausen (bis zu dem Zehntel einer Sekunde) beobachtet werden. Die kurzen Einheiten werden zum Beispiel durch die Telefongesellschaften gebraucht, die zur Zeit europäische Auskunftsdienste per Telefon in allen Sprachen Europas aufbauen (Eurospeech 2003). Werden als kurze Elemente geordnet diejenigen, die zehn Minuten dauern, und als sehr große Elemente diejenigen, die sechzig oder neunzig Minuten dauern. Beim Zusammenrechnen dieser Elemente wird man zum Beispiel sagen, dass man über 100 oder 500 Aufnahmestunden verfügt.

Diese Maßeinheiten sind aber für die großen Korpuskomponente wenig zuverlässig, denn die Aufnahmedichte hängt von der Redeweise der Sprecher ab. Im Französischen schätzt man, dass die Sprecher, die langsam sprechen, 110 Wörter pro Minute aussprechen, und diejenigen, die schnell sprechen, 350 pro Minute (in einigen Aphasiefällen und unter dem Einfluss der Neuroleptika sinkt die Redeweise unter 100 Wörter pro Minute, was sehr penibel zu hören ist. Über 350 Wörter pro Minute lassen das Zuhören und die Transkription sehr schwer werden). Die Dichte schwankt also zwischen eins und drei, was beachtlich ist. Nach den zwei extremen Redeweisen, die eben erwähnt wurden, kann eine Aufnahmestunde 6600 oder 21000 Wörtern entsprechen. Man täte also gut daran, die großen Korpora nach der Zahl der graphischen Wörter der Transkription abzuschätzen. Die heute in der Welt gesammelten großen Korpora der gesprochenen Sprache haben eine Größenordnung von etwa zehn Millionen aufgeschriebenen Wörtern für das amerikanische oder das britische Englisch. Unglücklicherweise haben die heutigen Korpora des gesprochenen Französisches eine Größenordnung von einer Million Wörtern. Mit einer so beschränkten Größe ist es kaum möglich, lexikalische Forschungen zu machen oder eine zuverlässige Statistik über die Sprachgebräuche aufzustellen.

2.1.3 *TRANSKRIPTIONEN*

Die heute geltenden Transkriptionen der gesprochenen Sprache unterscheiden sich so sehr voneinander, dass es schwer ist, sie unter einer einzigen Bezeichnung zu versammeln. In einigen Fällen wären die Bezeichnungen von *Übertragung* oder *Adaptation* passender, wenn man nur den Inhalt der Aufnahmen berücksichtigt und ihre Form frei ändert. Das machen oft die Journalisten, wenn sie die Worte der befragten Personen wiedergeben, zusammenfassen und ihnen meistens eine normativere Wendung geben. Die Historiker und Soziologen handeln manchmal auch so, wenn sie sich vor allem für den informativen Inhalt interessieren : sie treffen eine Auswahl zwischen den Daten, schneiden die Passagen ab, die sie uninteressant finden, und streichen die von ihnen als störend betrachteten Merkmale der gesprochenen Produktion, die Wiederholungen, Zögern oder Überarbeitungen.

Einige Tätigkeitssektoren wie die Transkripte der parlamentarischen Debatten haben auch diese Arbeiten kodifiziert, indem sie mehrere Adaptationsgrade einführten.

Wenn es sich darum handelt, sich für die Sprache selbst zu interessieren, hängt die Wahl des Transkriptionsprozesses von den Zwecken der Studie ab (europäische und internationale Projekte haben sich Anweisungen für die Korpusausgabe gegeben) und wie E. Ochs es schon 1976 erwähnte, ist die Transkription immer mit einer Theorie gebunden. Einige Studien benötigen phonetische oder phonologische Transkriptionen. Der Standard Unicode, der unter der Norm ISO-10646 synchronisiert wurde, umfasst in seiner Version 4.0 mehr als 96000 Typen, unter denen die des *Internationalen Phonetischen Alphabets*. Es ist eine Notwendigkeit für alle Arbeiten, die die Aussprache betreffen, aber auch für alle Fälle, wo es schwer ist, stabile Morpheme zu finden, die in einer Standard-Rechtschreibung geschrieben werden könnten : Sprache der sehr jungen Kinder (internationales Modell CHILDES), Sprache der Ausländer beim Spracherwerb, Notierung einiger Regionalismusformen, Notierung einiger Aphasieformen wie die Jargons (Abou-Haidar 2002). Diese Transkripte, die nur für kurze Korpusauszüge gemacht werden können, werden oft von aneinanderliegenden Übersetzungen begleitet. Die Darstellung der Prosodie fordert spezifische Muster, die in den jüngsten Techniken hoch entwickelt wurden (Martin 1987). Die Video-Aufnahmen fordern spezielle Notierungen, die man mehr oder weniger vorantreiben kann (Van der Straten 1998, Mondada 2006).

Was die Korpora der gesprochenen Sprache betrifft, so sind sie in einer Standardrechtschreibung abgeschrieben, so dass sie leicht zu lesen sind. Diese Entscheidung zieht mehrere Möglichkeiten mit sich : Standardrechtschreibung mit oder ohne Anpassungen, mit oder ohne Interpunktion, mit oder ohne Hinweis auf Pausen, Verlängerung, Rhythmus, Betonung, Zögern, Husten, Lachen, Gestik, usw. Lange Debatten über diese Punkte fanden statt, um die optimale Transkriptionsbedingungen herauszuarbeiten, die den Forschungszwecken angepasst sind. Zum Beispiel : die Sprachwissenschaftler, die sich für die syntaktischen Einheiten der gesprochenen Sprache interessieren, misstrauen meistens der Interpunktion, weil sie Abgrenzungen durchsetzt, die der geschriebenen Sprache eigen sind, und sich oft als falsch erweist, wenn man sie setzt, bevor man die Texte tief genug analysiert hat. Die nicht interpunktierten Texte verstimmen aber die

Informatiker, deren automatisierte Analysen Interpunktionsmerkmale erfordern. Verhandlungen werden manchmal zwischen den Sprachwissenschaftlern und den Informatikern geführt (ICOR im ICAR-Labor), um Transkriptionskonventionen festzustellen, die diese Probleme und die internationalen Standards (GAT, TEI, Du Bois, Jefferson) berücksichtigen.

Die von den Sprachwissenschaftlern gebrauchten Transkripte bewahren sorgfältig alle Merkmale der verbalen Produktionen : Wiederholungen, Zögern, Wörterbeginn, Überarbeitungen. Sie erfordern, dass der Transkripteur dafür sorgt, seine eigenen Interpretationen auf das Trankript nicht zu projizieren (zum Beispiel bei der Rekonstruktion eines Textstückes nach den erwarteten Stereotypen). Diese Sorge um die kleinste Einzelheit erfordert ein Training und eine spezifische Ausbildung der Transkripteure. Die lange und mühsame Arbeit stellt viele Fallen (Leech 1991). Den üblichen Schätzungen zufolge ist ein Minimum an dreißig Minuten Arbeit erforderlich, um eine Aufnahmeminute aufzuschreiben (die Schöpfer des holländischen Korpus schätzen, dass jedes graphische Wort ein Euro kostet !). Auf Grund ihrer Treue missfallen die Transkripte der gesprochenen Sprache den Laien : darin sehen sie eine Unmenge „Deutschfehler", Wiederholungen, Auflösungen der Information. Einem laienhaften Informationsgeber ein Transkript seiner Sprache zu zeigen, verursacht oft eine Ablehnung. Es ist kein gutes Mittel, um seine Erlaubnis zu bekommen, das Ergebnis der Forschungen zu transkribieren und veröffentlichen.

Informatik hat die Transkriptionsarbeit verändert, einerseits durch die gebrachten Hilfen, andererseits durch die neu eingeführten Zwänge. Die Transkriptionshilfen (Anvil, Clan, Ite, Praat, Transcriber...) erleichtern die Handhabungen und erlauben es, die studierten Aufnahmestücke wiederzuhören. Die Technik der *synchronisierten Korpora* erlaubt es, auf einem Bildschirm Stücke eines geschriebenen Textes zu lesen, indem man die selben Stücke in ihrem Tonablauf hört (*Speech Communication* 33, Sonderaufgabe über die Anmerkungen und Analyseinstrumente der Korpora). Die neuen Erfordernisse betreffen die informatisierten Anmerkungen : morpho-syntaktisches Etikettieren aller Textelemente, Baumdiagramme, Metadaten (die die Aufnahmebedingungen, die Situationen und die Sprecher betreffen). Verschiedene Einstufungen und Codierungen ermöglichen, die nötigen Lemmatisierungen und Konkordanzer herzustellen, um Gesuche über das Gesamtkorpus einzureichen. Eine Polemik hat um 2000 begonnen, was den nötigen Komplexitätsgrad der Anmerkungen betrifft (Sinclair, Teubert). Die Standardisierung wird jetzt auf der europäischen Ebene gemacht (SpeechDat Exchange Format).

2.1.4 *AUTOMATISCHE BEARBEITUNG DER SPRACHE*

Im Gegensatz zu manchen anderen Gebieten der Sprachforschung benötigt die automatische Transkription der Sprache, die auf einer ununterbrochenen akustischen Flut durchgeführt wird, eine Modelisierung der gesamten Phänomene des Tonsignals. Man muss also über die Wörter, mit denen eine phonologische Darstellung im Aussprachwörterbuch gebunden ist, extra-lexikalische Phänomene modelisieren : Atmungen, Zögern, Wörterfragmente, usw.

Je nach dem Typ der bearbeiteten Aufzeichnungen erreichen die Systeme der automatischen Entzifferung sehr unterschiedliche Irrtumsquoten[1]. Wenn aber die Modelle (im großen und ganzen die Kenntnisse) des Systems mit den zutranskribierenden Korpora nicht übereinstimmen, können diese Irrtumquoten rasch zunehmen. Um die besten Ergebnisse zu erreichen, müssen die Transkriptionssysteme nach den zutranskribierenden Korpora angepasst werden.

Die aktuellen Forschungen zeigen, dass die automatische Transkription ein wertvolles Instrument wird, um der Transkription und der Anmerkung der Korpora beizutragen. (Barras *et al.* 2004) erweist zum Beispiel die Nützlichkeit der automatischen Transkription für die teilweise manuelle Generation feiner akustischer Transkriptionen (das heißt, die nicht nur aus allen orthographischen Wörtern, sondern auch aus den „Verschiedenheiten„ und anderen außerlexikalischen Ereignissen bestehen). Die laufenden Forschungen zeigen auch, dass die automatische Transkription der Sprache ein genaues Instrument werden kann, um Korpora zu erforschen, analysieren, linguistische Phänomene quantitativ zu erfassen. Allgemeinerweise kann man denken, dass die Ansichten der Sprachwissenschaftler und die der Informatiker in die Zukunft immer weniger gegenübergestellt werden können. In dieser Hinsicht soll sich das Auftauchen der Linguistik der Korpora der gesprochenen Sprache als Forschungsgebiet auf die Ausbildung der sprachwissenschaftlichen Informatiker und der Linguisten-Informatiker stützen.

2.1.5 AUSWERTUNGEN UND ERGEBNISSE

Die aktuellen grossen Korpora der gesprochenen Sprache sind teuer. Einige Korpora werden besonders im Engineering in Zusammenarbeit mit den Industriellen verwertet : Dialog Mensch-Maschine, Sprachentzifferung und –synthese, Telefonate, usw. (Einrichtungen wie ELRA/ELDA haben sich in die Verbreitung der Korpora und der in diesem Gebiet verfügbaren Mittel spezialisiert).

Die großen Korpora dienen an erster Stelle als allgemeine Dokumentation über die Landessprache. Die großen *Nachschlagekorpora*, deren Sprecher in Berücksichtigung der Gegenden und der sozio-ökonomischen und kulturellen Daten ausgewählt worden sind, ermöglichen eine langfristige Führung der linguistischen Politiken. Das Nachschlagekorpus des gesprochenen Portugiesischen umfasst Aufnahmen aus Portugal, Afrika, Brasilien und Asien und erlaubt es zum Beispiel, die Unterschiede nach der Geographie der Welt zu schätzen und auf diese Forschung einige Gebräuche der schulischen Praktiken und sogar der Regierungsbeschlüsse zu stützen. Das *British National Corpus* war die Basis der Herstellung einer wichtigen Grammatik, der *Longman Grammar of Spoken and Written English*, die auf sehr neuen Grundlagen ausgedacht wurde. Eine große Veröffentlichungstätigkeit hat sich bei der Benutzung dieser Materialien in englischer Sprache entwickelt. So hat der Verleger Collins die englischen Korpora für die Veröffentlichung mancher didaktischer Werke gebraucht, die zu dem Unterricht des Englischen als Muttersprache und als Fremdsprache dienen. Eine Dokumentation über die

[1] Die Arbeiten des LIMSI (Barras 2004) zeigen Ergebnisse vor, die mit optimisierten Systemen für eine gegebene Arbeit von 10 bis 30% Wortirrtümern gehen.

gesprochene Sprache ist manchmal der Ausgangspunkt für neue Tätigkeiten : Korpora der gesprochenen Sprache waren die Basis der Verbreitung kaum (oder gar nicht) geschriebener Sprachen, wie es für die Maori-Sprache gemacht wurde, die als Muster für die Entwicklung von Radio- und Fernsehsendungen (Kennedy 1998 : 72) diente.

Der Vergleich mit den gesprochenen Sprachen der selben linguistischen Gruppe erlaubt es, *in vivo* die Ähnlichkeiten und Unterschiede innerhalb einem großen linguistischen Bereich zu schätzen.

Eine wichtige Auswertung bieten die mehrsprachigen (auch parallel oder gereiht genannten) Korpora, die den Übersetzern, dem Unterricht der Sprachen und der kontrastiven Studie nützen. Es bestehen welche für die gesprochene Sprache :

- Englisch/Französisch an der Universität von Lancaster, Oslo, in Mannheim, an der Universität von Gand in Belgien (Contragam, bank.ugent.be/contragam/newslet.html), an der Universität von Montreal,
- Französisch/Englisch/Niederländisch, an der Universität von Courtrai,
- Französisch/Englisch/Spanisch an der Universität von Pennsylvania.

Eine jüngste Studie, die auf Aufnahmen und Transkripte vier romanischer Sprachen (Italienisch, Französisch, Portugiesisch, Spanisch) gegründet ist, ermöglicht es, die Prosodie (Intonationen, Betonungen, Rhythmen) zu vergleichen, indem verschiedene Situationen und Medien (C-ORAL-ROM, Cresti & Moneglia) berücksichtigt werden.

Auf dieser Weise haben die großen Korpora der gesprochenen Sprache manche linguistische Probleme erneuert. Durch die von diesen Sammlungen gelieferten Daten haben sich neue Disziplinen wie die konversationnelle Analyse und die Analyse der Interaktionen, der Verhandlungen und der Höfflichkeitskodes begründet. Die Forschungen auf dem Gebiet der Pragmatik stützen sich massiv auf diese Daten. Einige Kenntnisse wurden eindeutig geändert, wie zum Beispiel die Studien über die Produktion und die Wahrnehmung der gesprochenen Sprache und infolgedessen über die Schwäche der linguistischen Intuition (Blanche-Benveniste 1997). Man konnte den Grad der geordneten und systematischen Organisation in die Interaktionen zeigen. Man hat ihn gebraucht, um bestimmte Grundeinheiten wie *den Satz* in Frage zu stellen und andere einzuführen, wie die Einheiten der jetzt von mehreren Sprachwissenschaftlermannschaften (Blanche-Benveniste *et al.*, 1999, Scarano 2003, Nolke 2002) gebrauchten *Makrosyntax*. Die Studie der Intonation wurde sehr ernsthaft in die Abgrenzung der Einheiten der Makrosyntax (Cresti & Moneglia 2005, Couper-Kuhlen & Selting, 1996) übernommen. Es wurde gezeigt, dass mehrere ineinandergreifende Organisationsstufen in die Interaktionen zustande kommen (Turn-Constructional Units oder „Einheiten des Reiheaufbaus“, Selting 1995, 1998, 2000, Auer *et al.* 1999, Ochs, Thompson & Schleghoff, 1996). In verschiedenen Sprachen hat man die Rolle der Merkmale der mündlichen Produktionen zeigen können, das heißt der diskursiven Partikeln, Wiederholungen, Zögern oder „Flickungen“, für die zur Zeit die Neurowissenschaften sich interessieren. Die Perspektiven auf die Geschichte der Sprachen wurden sogar dadurch geändert, sofern man jetzt den Einfluss der verschiedenen

Sprechsituationen auf den eingenommenen Grammatiktyp studieren kann (Biber 1987). Man kann zum Beispiel für das Französische zeigen, dass die Erklärungsberichte und die Argumentationen Praktiken einer Syntax mit hohem Einbettungsgrad zeigen, während es wenige in den Konversationen gibt oder während die Unfallberichte komplexe chronologische Organisationen enthalten. Man weiß, dass die als „erhaben" geltenden Themen (Sprache über die Moral, die Religion, den Tod) typische Merkmale der „Feiersprache" auslöst, zum Beispiel auf Französisch viele Liaisons, einen massiven Gebrauch des ersten Wortes der Verneinung und manchmal auch unerwartete Gebräuche des passé simple. Die großen Korpora ermöglichen die Beobachtung bestimmter laufender Grammatikalisierungsprozesse. Sie zeigen die zahlenmässige Bedeutung der Wortlaute in Klammern, der Fokussierungen und der Thematisierungen. Sie zwingen zu bedenken, dass die festen Redewendungen einen sehr wichtigen Platz im Vergleich zu dem freien Aufbau der Wortlaute einnehmen, so dass der Zusammenhang zwischen Grammatik und Wortschatz jetzt klarer auftaucht als früher, da viele grammatische Wendungen nur für einen kleinen Teil des Wortschatzes von den Sprechern gebraucht werden. Daraus ist zu schließen, dass man kein „Wort" beim Sprechen aussucht, sondern ein vorgebautes Ganzes (Sinclair, 1991).

Das stellt ja die linguistischen Theorien wieder in Frage, die danach strebten, die Syntax als eine unabhängige Komponente der Sprache zu betrachten.

Diese großen Korpora, wenn sie existieren, tun entscheidende Dienste : sie dienen als Datenbank für alle Vergleiche, die die Sprache betreffen : um die Kindersprache in verschiedene Spracherwerbstufen zu schätzen, um die Diagnosen in den Sprachpathologien zu unterstützen, um den Ausführungsgrad im Spracherwerb der Mutter- und Fremdsprache zu schätzen, um die Wirkungen der Gruppen- und Berufsprachen (Gadet) abzuwägen, um die Koordinationsarten in einer Mannschaft oder in einer Gruppe zu studieren, um die spezifischen Besonderheiten der Tätigkeitstypen und der damit angepassten Beiträge in verschiedenen institutionnellen Rahmen zu verstehen oder um die Wirkung der regionalen Einflüsse zu kennen. Zum Beispiel : bevor man beurteilt, dass eine Wendung typisch für die Sprache der Kinder in derartigem Alter oder von derartiger Herkunft ist, muss man zu einer Vergleichsdatenbasis greifen, um zu wissen, ob die Wendung spezifisch ist oder nicht.

KORPORA DER SPRACHEN MIT MÜNDLICHER TRADITION

Die Probleme, die bei dem Aufbau, der Auswertung, der Verbreitung und der Verwaltung von Korpora der gesprochenen Sprache in den Gesellchaften „mit mündlicher Tradition", in den „ethnischen" oder „exotischen" Gesellschaften getroffen wurden, stimmen teilweise mit denen überein, die bei der Aufstellung der großen Korpora der Sprachen der westlichen Länder getroffen wurden. Da muss man die Vorkehrungen, die man (wie im Handbuch empfohlen) für den Respekt der Leute treffen muss, dem Kontext der Terrainarbeit anpassen.

In einer Gesellschaft mit mündlicher Tradition kann die Erlaubnis nach Information (auf dem im Handbuch geschriebenen Vorbild der „aufgeklärten Einwilligung", das hier noch der Situation angepasst war) in einigen Fällen nur einen Wert haben, wenn sie mündlich ist und von der Person erteilt, die das

Recht dazu hat (so wie die Erlaubnis nur gültig sein wird, wenn sie von der Ärztekammer erteilt wird, wenn die Befragung in ärztlicher Umgebung gemacht wird). Außerdem stellt die Information des Sprechers einige Probleme in Gesellschaften, wo die Forschungsaktivität, die Zwecke des Aufbaus des Korpus und seine Verbreitungsnetze (Veröffentlichungen, Internet) nichts Konkretem entsprechen.

Der Forscher soll sich dazu über das übliche Recht des Landes informieren, wo er arbeiten wird. Das französische Recht zum Beispiel erkennt das intellektuelle Urheberrecht oder die Persönlichkeitsrechte nicht, wenn es sich um Märchen- oder Mythensammlungen handelt, die als Teil des Erbes angesehen werden und dem Allgemeingut gehören. Dagegen wurden in manchen afrikanischen Ländern Urheberrechteämter gegründet, um diese Art Produktionen und deren Autoren zu schützen. Bestimmte Gemeinschaften erkennen auch das Recht des Staates nicht, wo sie leben. Es ist zum Beispiel der Fall in Französisch-Guayana in einigen Indianergemeinschaften, die nach einem kollektiven und nicht privaten Recht funktionieren (Tiouka 2005, für eine Betrachtung über die Integration des Gewohnheitsrechtes in das französische und europäische Recht). Bestimmte Erlaubnisse werden für diese Gemeinschaften nur gelten, wenn sie das Gewohnheitsrecht einhalten ; obwohl der Forscher sich geschützt fühlt, wenn er das Nationalrecht einhält, können die Gewohnheitsbehörden im Konflikt mit ihm sein und ihm den Zugang zu dem Terrain verwehren.

Die Auswertung des Korpus erfordert meistens die Intervention mehrerer Personen : der Transkripteur, der der Sprecher der Aufnahme sein kann, aber nicht immer, der Übersetzer (id.). Die Rechte dieser Personen auf das Korpus müssen auch nach verschiedenen Elementen festgesetzt werden : nach dem Nationalrecht, wenn es für die Gemeinschaft sinnvoll ist, oder nach dem Gewohnheitsrecht.

In zwei Punkten ist der Aufbau der Korpora der gesprochenen Sprache in bestimmten Gesellschaften unterschiedlich von dem der großen Korpora der Nationalsprachen:

1. *die Größe des Korpus*

In einigen Sprachen kann man schwer planen, Korpora sammeln zu können, die die Größe der großen Korpora der westlichen Sprachen mit mehreren Millionen Wörtern erreichen könnten. Das Problem des repräsentativen Charakters der Korpora stellt sich also anders.

2. *die Rückkehr zur Gemeinschaft*

Die Erfahrung des Terrains (Anthropologie des 20. Jahrhunderts, missionarische Linguistik, usw.) bis vor fünfzig Jahren hat Spuren in den Gemeinschaften hintergelassen, die sich geplündert und ausgebeutet gefühlt haben, ohne je Zugang zu den Ergebnissen der Forschung gehabt zu haben. Diese Gemeinschaften fordern jetzt, dass die Forschungen direkte Nebenwirkungen unter verschiedene Formen haben, und diese Forderungen werden von allen Einrichtungen aufgenommen, die Forschungen über die gefährdeten Sprachen (UNESCO, Projekt DOBES vom Max-Planck Institut, usw.) finanzieren oder veranstalten. Die von den Gemeinschaften kommenden Ansprüche haben mit der Entschädigung der verschiedenen Sprecher nichts zu tun, die in den Aufbau und die Auswertung des Korpus einbezogen sind, und sie verlangen eine Verwicklung des Forschers (durch die Teilnahme an Erziehungsprogramme, die Zurückgabe der Aufnahmen und der eingesammelten Materialien, den Aufbau von Archiven, die den Gemeinschaften zugänglich sind, usw.). Die Notwendigkeit, der Gemeinschaft die gesammelten Materialien zurückzugeben, sollte außerdem die Forscher dazu motivieren, Datenbanken anzulegen und ihre Korpora zu archivieren.

Es scheint jedoch, dass diese Korpora noch oft „persönliche Instrumente" bleiben, die nur als Grundlage für die linguistische Analyse des einzigen Forschers vom Nutzen sind. Es gibt vielfältige Gründe dazu : wie für den Aufbau

irgendwelches Korpus der gesprochenen Sprache sind der technische Aspekt und die für die Formgebung nötige Zeit (Numerisierung, manchmal Synchronisierung, usw.) dem Forscher zuwider, und das um so mehr, dass diese Arbeit durch die wissenschaftlichen Institutionen nicht aufgewertet ist. Außerdem ist auf bestimmten Terrains der Aufbau des Korpus das Ergebnis Vertrauensbeziehungen, die sich in schwierigen Zusammenhänge zwischen den Forscher als Person (und nicht als Vorsteller einer wissenschaftlichen Gemeinschaft) und die Gemeinschaft oder einige Mitglieder der Gemeinschaft entwickeln. Der Entschluss, das Korpus innerhalb der wissenschaftlichen Gemeinschaft oder breiter (Zugänglichkeit durch das Internet) zu verbreiten, befragt die Ethik des Forschers und untersteht dem juristischen Rahmen nicht mehr. In allen Fällen ist es wünschenswert, dass die Verbreitung nach der Zurückgabe des Korpus an die betroffene Gesellschaft stattfindet.

Der Forscher ist also in einer heiklen Lage zwischen dem Willen, eine privilegierte Beziehung mit seinem Terrain zu bewahren und der wachsenden Notwendigkeit, die Ressourcen, die die Basis der Analyse und der Ergebnisse der Forschung sind, zur Verfügung der wissenschaftlichen Gemeinschaft zu stellen.

Mit der Entwicklung der großen aktuellen Korpora der gesprochenen Sprache entwickelt sich die Standardisierung (seit den von EAGLES im 1993 verbreiteten Vorschriften) und die Forschungsgebiete werden immer interessanter. In dieser Perspektive ist es eine wichtige Aufgabe des „immateriellen Erbes", die existierenden Korpora des gesprochenen Französischen oder die jeder anderen Sprache zugänglich zu machen und neue zu schaffen. Die juristischen Probleme des Sprachschutzes, die lange zu Unrecht als nebensächlich betrachtet wurden, sind zur Zeit sehr wichtige Bremsen : viele Forscher lehnen es ab, ihre Korpora in Umlauf zu setzen, weil sie nicht sicher sind, „die richtigen Genehmigungen" zu haben. Viele wagen es nicht, neue herauszubringen, weil der Genehmigungsantrag ihnen wesentlich aber schwer zu stellen scheint. Deswegen ist es jetzt unvermeidlich, kollektiv darüber nachzudenken.

2.2 POLITISCHE RAHMEN DER VERBREITUNG DER FORSCHUNG

DIE VERBREITUNG DER ERGEBNISSE DER FORSCHUNG GEHÖRT ZU DEN AUFGABEN DES FORSCHERS

> *„Die öffentlichen Einrichtungen müssen sich ständig darum bemühen, die nationale Gemeinschaft bestens von den Ergebnissen ihrer Arbeiten profitieren zu lassen...".*
>
> *„Die Politik der Forschung und der industriellen Entwicklung zielt auf die Zunahme der Kenntnisse, auf die Verbreitung der wissenschaftlichen und technischen Information und auf die Beförderung des Französischen als wissenschaftliche Sprache".*[2]

Mit diesen Wörtern definiert der Bericht, der dem Orientierungsgesetzes des 15. Juni 1982 beigefügt ist, die Umrisse der Aufwertung. Zweifellos treffen diese allgemeinen Prinzipien für die Forscher zu, deren Arbeiten zu dem Aufbau von Korpora der gesprochenen Sprache führen. Die Bedingungen der Aufwertung und der Verbreitung werden jedoch auch von den möglichen existierenden Rechten über die eingesammelten Inhalte und über die Ergebnisse ihrer Ausarbeitung von den Forschern abhängen.

DIE DYNAMIK DES AUSTAUSCHES UND DIE DEN RECHTSTRÄGERN ANGEBOTENEN GELEGENHEITEN, UM DEN ZUGANG ZU DER INFORMATIONSGESELLSCHAFT ZU ERLEICHTERN

Zweifellos kann man heute von einer neuen Art sprechen, den Beitrag von jedem in den Informationsaustausch zu sehen. Diese Dynamik des Austausches erzeugt neue Benehmen. Die Zugangsfreiheit, die Kostenlosigkeit und das Recht auf eine Wiederverwendung scheinen selbstverständlich, wenn sie sich in die Gegenseitigkeit einfügen.

Am 22. Oktober 2003 haben die meisten Generaldirektoren der öffentlichen Einrichtungen mit wissenschaftlichem und technologischem Charakter die *Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen* unterschrieben, deren Ziel das Internet zu fördern ist, „als funktionsgerechtes Instrument zugunsten einer Basis für eine globale Kenntnis des menschlichen Denkens".

Indem sie diese Erklärung unterzeichneten, haben sich die politischen Entscheidungsträger für die Wissenschaft, die Forschungsinstitute, die Finanzierungsagenturen, die Bibliotheken, die Archive und die Museen dazu verpflichtet, einige Maßnahmen zu planen. Diese Maßnahmen müssen es erlauben, „die Weiterentwicklung der bestehenden rechtlichen und finanziellen Rahmenbedingungen zu unterstützen, um die Voraussetzungen für eine optimale Nutzung und einen offenen Zugang [des Internets] zu ermöglichen". Der Text erkennt auch, dass die Forderungen des Schutzes und des offenen Zugangs sich

---

[2] Artikel 5 des Gesetzes Nr 82-610 des 15. Juli 1982 für die Orientierung und das Programmieren der Forschung und technologischen Entwicklung Frankreichs, heute Art. L 111-1 des Forschungscodes. *JO* des 16. 07 1982, Seite 2273 und ff.

möglicherweise widersprechen. Schließlich geht aus dieser Erklärung hervor, dass der offene Zugang der aktiven Mitwirkung eines jeden als Erzeuger wissenschaftlichen Wissens oder Verwalter von kulturellem Erbe voraussetzt, indem dieser offene Zugang „im Respekt der Urheber und Rechteinhaber" stattfindet. Der offene Zugang soll also von den Rechteinhabern geregelt und abgeändert werden. Die Autoren (oder Institutionen) können ein „kostenloses, unwiderrufliches und weltweites Zugangsrecht zu diesem Werk" gewähren oder „ihnen die Lizenz geben, diese Veröffentlichungen in jedem beliebigen digitalen Träger und für jeden verantwortbaren Zweck zu kopieren, zu nutzen, zu verbreiten, zu übertragen und offentlich wiederzugeben sowie Ausarbeitungen davon zu erstellen und zu verbreiten, sofern die Urheberschaft korrekt angegeben wird". Kann nur gewährt werden, dass „von diesen Beiträgen eine geringe Anzahl von Kopien zum privaten Gebrauch angefertigt werden". Die Formalisierung dieser Genehmigungen kann unter der Form von Lizenzen wie *creative commons* (Gebrauchsgenehmigungen, die von den Autoren ohne Entschädigung direkt gegeben werden. Die Autoren können dagegen wenn nötig diese Nutzung begrenzen, indem sie sie zur Nutzung zu Erziehungszwecken reservieren).

Wenn sie den Korpora der gesprochenen Sprache angewendet werden, können diese Lizenzen ermöglichen, den zukünftigen Gebrauchern die Verantwortung für den Respekt der Verpflichtungen übernehmen zu lassen, die der Forscher, der das Korpus aufgebaut hat, denjenigen gegenüber, die zu seiner Ausarbeitung beigetragen haben, eingegangen ist.

DIE PROGRAMME DER ERBDIGITALISIERUNG

Der Kontext der Informationsgesellschaft hat etliche öffentliche Initiativen hervorgerufen, um das Weiterbestehen des kulturellen Gedächtnisses zu sichern. 2001 hat eine Gruppe von nationalen Vertretern der Mitgliedsstaaten Europas, die sich für die Digitalisierungsprobleme interessierten, in Lund (Schweden) einen Text erarbeitet, der unter anderem folgendes anpreist : die Bereitstellung von Interoperabilitätsstandards ; die Verbreitung von guten Praktiken, unter denen die Verwaltung der Urheberrechte ; die Organisation von Kompetenzzentren über die Digitalisierung, für die die Fachleute der Information die Verantwortung haben.

Die Frage der Bewahrung der Ergebnisse der Forschung stellt sich heute um so ernsthaft, dass die Ergebnisse und auch die Materialien selbst, die diesen Forschungen dienten, auf digitalen Trägern liegen. Wie kann man die „Rückverfolgbarkeit" der verschiedenen Etappen der Forschungsarbeit sichern ? Was soll man aufbewahren ? Wer wird diese Verwaltung sichern ? Unter welchen Bedingungen ? Diese Fragen müssen heute gestellt und Antworten für jeden Forschungsprozess gefunden werden. Wenn allgemeine Empfehlungen gegeben werden können, so befreit es gar nicht von ihren Verantwortungen diejenigen, die eine Forschung initiieren, deren eines der Ziele oder eine der Etappen der Aufbau eines Korpus der gesprochenen Sprache ist.

2.3 JURISTISCHE RAHMEN

Die Absicht dieses Handbuches, das für die Forscher bestimmt ist, heißt nicht, sich mit allen zu berücksichtigenden juristischen Techniken zu befassen. Es handelt sich darum, den Leser zu sensibilisieren und ihn dazu anzuregen, sich die notwendigen Fragen zu stellen, um seine Verpflichtungen aber auch seine Rechte zu verstehen.

Welches kann der juristische Status sein von jedem der von den Forschern aufgebauten Korpora der gesprochenen Sprache ? Diese Frage kann *a priori* theoretisch erscheinen, aber sie ist nicht zu verschleiern, denn nach den gebrachten Antworten wird man die Auswertungs- und Verbreitungsbedingungen der Korpora bestimmen können. Um diese Frage zu beantworten, muss man zuerst die Ausarbeitungsbedingungen des Korpus und seiner verschiedenen Komponenten kennen. Besteht das Korpus aus Informationen des Allgemeingutes ? Ist es das Produkt einer oder mehrerer intellektueller Schaffungen, die vom Urheberrecht geschützt werden können ? Besteht der Inhalt des Korpus aus persönlichen Daten ? Welches sind dann die Rechte der Sprecher oder der betroffenen Personen ?

Wenn dieser juristische Status festgestellt und die folgenden Rechte einmal gekannt sind, so soll man sich über die Modalitäten der vertraglichen Verwaltung dieser Rechte informieren. Haben sich die Besitzer der Rechte über die Bedingungen der Erschliessung und Wiederverwendung der Korpora geäußert ?

Zum Schluss soll man allen, die in das „Korpusleben" eingreifen werden, Aufmerksamkeit widmen : Verantwortung der Schöpfer, Verantwortung der Beherberger, der Verbreiter, Verwalter...

Um das Verfahren des Forschers zu erleichtern, wird hier ein genereller Überblick über vier große Fragen gegeben, die häufig in den Aufbau und das Leben der Korpora wiederkommen : was ist das Allgemeingut, also das, was man „sich nicht eignen kann" ? Wann geht es um das Urheberrecht in bezug auf die Korpora ? Wie kann man die persönlichen Daten vom Standpunkt der Ausarbeitung der Informationen, die die Korpora der gesprochenen Sprache bilden, schützen ? Welches sind die Verantwortungen der Leute, die die Aufgabe haben, die Korpora auf dem Internet zu verbreiten ?

2.3.1 DAS ALLGEMEINGUT UND DAS URHEBERRECHT

WAS IST DAS ALLGEMEINGUT ?

Wenn das Wort „Allgemeingut" meistens von allen bekannt ist, kann der juristische Sinn des Terminus verschiedenartig verstanden werden und diese verschiedenen Sinne soll man klarer formulieren, um Zweideutigkeiten oder Unverständnisse bei dem Aufbau der Korpora der gesprochenen Sprache zu vermeiden. Im juristischen Sinne ist das Allgemeingut ein mehrfaches Konzept, das sowohl auf einen Ort, eine Regelung oder auf Inhalte verweist.

Das Allgemeingut kann also „der Ort sein, wo die bürgerliche Gesellschaft sich bemüht, die Weise zu beeinflussen, wie die Kollektivgüter verwaltet und verteilt werden". In diesem Sinne ist die UNESCO die Grundlage einer echten Politik der Inhalte und sie entwickelt eine Förderungsstrategie für ein starkes, online und offline zugängliches Allgemeingut. Das Allgemeingut betrifft nicht nur die Ideen der

Zugangsfreiheit und der kostenlosen Verwendung der Daten, sondern auch die Möglichkeit für jeden, sie auszunutzen. Es ist außerdem durch die Monopollosigkeit gekennzeichnet, da die Informationen, die zum Allgemeingut gehören, *de facto* „gemeinsame Dinge" werden.

Dagegen können zwei Typen von Informationen unterschieden werden : diejenigen, die immer zum Allgemeingut gehörten und diejenigen, die es „geworden" sind. Ideen, Sprache, Gesetztexte und alle Elemente, die das gemeinsame Erbe einer gegebenen Gemeinschaft begründen, bilden von Natur aus das „gemeinsame Vermögen" des Allgemeingutes. Dieses gemeinsame Vermögen bleibt aber schwer zu begrenzen. Die Sprachaufnahmen rufen viele Unschlüssigkeiten hervor. Gehören sie zum Allgemeingut oder nicht, über die Rechte hinaus desjenigen, der aufgenommen hat, über den Inhalt einer Sprache, über ihren phonischen Ausdruck ? Die Frage kann auch den Überlieferungen und Sitten gegenüber gestellt werden. Ist außerdem dieses gemeinsame Vermögen universal oder nur einer kleinen Gemeinschaft gemein? Es wird heute als Identitätszeichen gefordert, was neue Befragungen aufwirft.

Nach einer bestimmten Frist kommen schließlich die vom Urheberrecht geschützten Werke, vor allem von dem Urheberrecht oder den Lizenzen, ins Allgemeingut. Das Urheberrecht zum Beispiel schützt die Werke siebzig Jahre nach dem Tod ihres Verfassers. Im französischen Recht können andere Schütztypen nach dieser Frist auf die geistigen Werke bleiben : die Urheberrechte einerseits ; die unverjährbaren Komponenten des moralischen Rechtes andererseits. Infolgedessen können noch einige Elemente des Allgemeingutes den Schutz des moralischen Rechtes bekommen.

Diese Unterscheidungen lassen zwei Situationstypen erscheinen, die anscheinend gegenüberstehen : entweder bestehen die Korpora aus Werken des Allgemeingutes, die (wegen ihrer Natur oder des Ablaufes des Schutzfristes) nicht angeeignet werden können, und infolgedessen gebührenfrei sind, oder die Korpora sind dem Urheberrecht und deswegen den erforderlichen Erlaubnissen unterworfen. In Wirklichkeit existiert, wie wir es gesehen haben, eine Zwischenmöglichkeit : die durch das Urheberrecht geschützten Korpora können ein offenes Zugangsrecht im Rahmen einer Lizenz geniessen, die von den Rechtsinhabern, die die Verwendung und die Auswertung der Ergebnisse sichert, genehmigt wird. Ohne zum Allgemeingut zu gehören, geniessen diese Korpora – durch den Willen ihrer Schöpfer – ein freies Zugangs- und Verwendungsrecht. Wenn die Schöpfer auf die Ausübung ihrer Urheberpersönlichkeitsrechte verzichten können, ist es ihnen trotzdem unmöglich, auf ihr moralisches unverjährbar bleibendes Recht zu verzichten.

DAS URHEBERRECHT UND DIE KORPORA

Welches sind die Bedingungen für den Schutz eines Korpus ? Es gibt drei Bedingungen.

Es muss zuerst eine gewisse *Schöpfungshöhe* erreichen : eine Arbeit, die Informationen kompiliert, wird als solches nicht geschützt.

Um geschützt zu werden, muss auch das Korpus unbedingt *sinnlich wahrnehmbar* sein. Was geschützt wird, ist nicht der Inhalt des Korpus sondern seine Form, sein Aufbau.

Schliesslich muss der Schöpfer des Korpus *einen schöpferischen Beitrag* geleistet haben. Was heißt ein schöpferischer Beitrag ? Die schöpferische Kraft mancher Schöpfungen des Digitalzeitalters wie die Software oder die Datenbanken kann nur nach objektiven Kriterien abgeschätzt werden. Es scheint auch so für die Korpora der gesprochenen Sprache, da diese meistens mit einer Datenbank gleichgesetzt werden können. Dann wird meistens die Tatsache, dass das Korpus kopiert wird oder nicht, und ein Minimum an schöpferischer Aktivität erkennen lässt (und nicht nur die Berücksichtigung der Persönlichkeitsprägung seines Schöpfers), als Beweis dienen, um zu entscheiden, ob es originell ist oder nicht.

„ES GIBT KEINEN PLATZ FÜR URHEBERRECHTE, WENN ES KEINEN URHEBER GIBT"

Der Urheber ist im Prinzip der (oder die) natürliche(n) Person(en), unter derer Name (deren Namen) das Werk veröffentlicht wird. Die wissenschaftliche Arbeit setzt die Anwesenheit mehrerer Schöpfer voraus, von denen viele Anspruch auf die Stellung des Urhebers der Forschungsergebnisse erheben können.

Gewisse Korpora der gesprochenen Sprache können wie andere Forschungsergebnisse das Werk eines einzigen Urhebers bleiben, während andere das Werk mehrerer Schöpfer sein können. Im Fall von mehreren Urhebern unterscheidet das Recht die Zusammenarbeitswerke von den gemeinsamen Werken. Für die ersten genießt jeder Urheber dieselben Vorrechte. Andere Werke – wie die Datenbanken oder die Wörterbücher – können als gemeinsames Werk angesehen werden, wenn sie so geschaffen werden :

> *„aus der Initiative einer natürlichen oder moralischen Person, die sie verlegt, unter ihrem Namen veröffentlicht und verbreitet und wo der persönliche Beitrag der verschiedenen Urheber sich ins Ganzes aufgeht"* [3].

In diesem Fall stehen der persönlichen oder moralischen Person, die die Iniative des Werkes genommen hat, die Urheberrechte zu. Außerdem können der Kontext der Erschaffung oder der Status des Schöpfers Auswirkungen auf die Bestimmung des Urhebers haben. Wurde das Werk im Rahmen eines Auftrags von einem Angestellten oder einem Beamten gemacht ? Welches sind die respektiven Rechte des Urhebers und seines Arbeitsgebers ? Wenn die Frage meistens durch den Arbeitsvertrag beantwortet wird, bleibt sie heikler, wenn der Schöpfer ein Beamte ist. Seit vielen Jahren stehen ja zwei Logikformen gegenüber, einerseits die der Anerkennung des Rechtes des Schöpfers und andererseits die der einzigen Anerkennung des Staatsrechtes über die Schöpfungen der Beamten. Die Übertragung der Weisung über die Urheberrechte in der Informationsgesellschaft hat die Staatsorgane dazu ermutigt, einen Mittelweg vorzuschlagen, der zugleich das Recht der Urheber wie das Recht des „staatlichen" Arbeitsgebers anerkennt, wenn

[3] Artikel L.113-2 des französischen Urheberrechtsgesetzes (UrhG).

die Schaffung des Werkes in die Durchführung eines Auftrags des öffentlichen Dienstes passt. Wenn der Text vom französischen Parlament angenommen wird, könnten (in Frankreich) Urheberrechte für den Beamten zustande kommen.

Andererseits gingen seinem Arbeitsgeber also dem Staat alle Auswertungsrechte für die Bedürfnisse seines Auftrags (Mitteilungs- und Verbreitungsrecht für den Auftrag). Jedoch werden dem Schöpfer als natürliche Person im Falle einer kommerziellen Auswertung seine Rechte wieder zustehen, mit der Auflage, seinem Arbeitsgeber ein Optionsrecht und die Möglichkeit, an der kommerziellen Auswertung beteiligt zu sein, zu gewähren. Dieser Text löst Debatten und Befragungen aus. Wie wird der Umfang des Auftrags der Forscher, die in die Aufstellung des Korpus eingreifen, bestimmt werden ? Wie kann man die Auswertung für den Auftrag von der kommerziellen Auswertung unterscheiden, wenn – wie wir es vorher gesehen haben – der Forscher beauftragt wird, die Ergebnisse seiner Forschung mitzuteilen und sie durch die Veröffentlichung aufzuwerten ?

WELCHE RECHTE GIBT ES FÜR DIE URHEBER VON KORPORA DER GESPROCHENEN SPRACHE, DIE MIT WERKEN GLEICHZUSETZEN SIND?

Es ist angebracht, die Vermögensrechte von den moralischen Vorrechten zu unterscheiden. Dazu soll man daran erinnern, dass das Gesetz den exklusiven Urheberrechten einige Abgrenzungen festlegt.

Die Vermögensrechte werden in dem exklusiven Recht zugunsten des Schöpfers (oder der Rechtsinhaber) oder der Anspruchsberechtigten (Empfängers einer Abtretung, Erben...) zusammengefasst, die Vervielfältigung oder die öffentliche Mitteilung des geschützten Werkes zu erlauben oder zu verbieten. Wenn das Korpus der gesprochenen Sprache ein Werk ist, bedarf jede Vervielfältigung (die Digitalisiserung ist von Rechts wegen eine Vervielfältigung) und jede Bereitstellung für die Öffentlichkeit (auf einer Webseite wie auf jedem anderen Träger) der ausdrücklichen Genehmigung des Schöpfers oder des Rechtsinhabers.

Was die Vorrechte des moralischen Rechtes betrifft, die immer mit der natürlichen Person, die das Werk geschaffen hat, gebunden sind, sind sie zu viert : *das Verbreitungsrecht, das Reu- und Entzugsrecht, das Urheberecht und das Recht auf den Respekt des Werkes*. Jedes dieser Rechte gilt für die Korpora der gesprochenen Sprache. Der Urheber des Korpus kann (nach seinem Verbreitungsrecht) den Moment oder die Bedingungen der Bereitstellung des Korpus für die Öffentlichkeit bestimmen, da die Pflichtabgabe nicht unbedingt als Verbreitung gilt. Ein unveröffentlichtes Korpus kann also ohne die Genehmigung seines Schöpfers nicht für die Öffentlichkeit bereitgestellt werden. Der Schöpfer-Forscher, der weigert, das von ihm geschaffene Korpus zu verbreiten, ist (nach dem Urheberrecht) in seinem Recht, selbst wenn er außerdem administrativ bestraft werden kann, weil er seine Mission im öffentlichen Dienst nicht durchgeführt hat, das heißt die Ergebnisse seiner Forschung nicht mitgeteilt hat. Das Reu- und Entzugsrecht kann auch über ein Korpus der gesprochenen Sprache geltend gemacht werden, die Reue kann nur den geistigen Inhalt und nicht die materiellen Bedingungen seiner Verbreitung betreffen. Wenn

das Urheberrecht leicht zu verstehen ist, kann man sich fragen, was das Recht auf den Respekt des Werkes bedeutet, wenn es auf ein Korpus der gesprochenen Sprache angewendet wird. Dieses Recht entspricht sowohl dem Respekt der Form des Werkes (keine Streichung, Hinzufügung oder Änderung) wie dem Recht auf den Respekt des Sinnes des Werkes (Verfälschung des Zweckes des Korpus).

Wie alle Monopole sind die Urheberrechte begrenzt. Erstens kann man daran erinnern, dass sie zeitlich begrenzt sind und dass das Werk nach Ablauf dieser Frist zum Allgemeingut gehört (cf. *supra*). Diese Grenzen können sich auch durch die Art der Benutzung der Werke rechtfertigen. Da wird man von Ausnahmen vom Urheberrecht sprechen, die von den Zwecken, dem Kontext oder im Sinne des Allgemeinwohls rechtfertigt werden.

Schließlich betreffen das *Recht auf private Kopie* und das Zitatrecht direkt die Korpora der gesprochenen Sprache.

2.3.2 DER RESPEKT DES PRIVATLEBENS

DER RESPEKT DES PRIVATLEBENS IN DER VERFASSUNG, DER AUSWERTUNG, DER VERBREITUNG UND DER VERWALTUNG DER KORPORA

Die Schaffung eines Korpus setzt meistens das Sammeln der Daten voraus. Diese können persönliche Daten sein, so muss dieses Sammeln das Gesetz *Informatik und Freiheit* einhalten : Rechtmäßigkeit und Aufrichtigkeit, Vorinformation, Erhalten der Einwilligung der betroffenen Personen, Respekt der angekündigten Zwecke[4]... Wenn es sich um Forschungszwecke handelt, soll man eine als solche identifizierte Forschung verstehen oder kann man den Ausdruck „Forschungszweck" im breiteren Sinne verstehen ? Das Problem stellt sich, wenn man eine neue Verwendung und neue wissenschaftliche Auswertungen plant, nachdem das Korpus von den Forschern aufgebaut und wissenschaftlich ausgenutzt wurde, die seine Erschaffung in die Wege geleitet haben. Die wissenschaftliche Forschung genießt heute eine Ausnahme vom allgemeinen Prinzip durch die Verwendung des sogenannten *Zweckerweiterung*. Jede neue wissenschaftliche Auswertung wird jedoch im Respekt der Vorformalitäten für jede Ausarbeitung (neues Meldung- oder Erlaubnisverfahren) und der vom Gesetz gestellten Grundsätzen (Information, Einwilligung und/oder andere angemessene Garantie...) gemacht werden.

Selbst wenn die Verbreitung der Korpora und ihre neue Verwertungen unter den erforderlichen Voraussetzungen stattfinden, bleibt jedoch das Problem der Bewahrung der persönlichen Daten.

Wenn die Daten auf nicht rückgängige Art und Weise „anonymisiert" werden, verlassen sie den Boden des Gesetzes und sie können bewahrt werden. Das Bedürfnis der „Rückverfolgbarkeit" in der Forschung macht es indessen oft erforderlich, dass die persönlichen Daten bewahrt werden.

---

[4] Für Frankreich : siehe http://www.cnil.fr/fileadmin/documents/approfondir/textes/CNIL-78-17_definitive-annotee.pdf

Im Prinzip sollen jedoch auf der Grundlage des Vergessensrechtes die persönlichen Daten nicht über die ursprünglich vorgesehene Dauer hinaus bewahrt werden, und wenn der beim Sammeln dieser Daten anfänglich angekündigte Zweck keine Berechtigung mehr hat, müssen diese Daten vernichtet werden. Heißt es, dass man bestimmte Korpora, die persönliche Daten enthalten, nicht bewahren kann, wenn diese nicht anonymisiert werden konnten ? Nein, aber es kann sich nur um Sonderfälle handeln, wo die Aufrechterhaltung der persönlichen Daten sich für wissenschaftliche Gründe rechtfertigt. In diesen Fällen würden die Korpora der gesprochenen Sprache – als öffentliche Archive – von einer Sonderregelung des Vergessensrechtes profitieren, die ihre Bewahrung über die vorgesehene Dauer hinaus erlauben würde mit dem Ziel, eine Bearbeitung mit historischen oder wissenschaftlichen Forschungszwecken zu machen. Dann wird das Archivgesetz die Bedingungen ihres offenen Zugangs bestimmen (mehr oder weniger lange Frist – 60 bis 150 Jahre[5] – je nach der Empfindlichkeit der im Korpus enthaltenen Daten).

WELCHES SIND DIE VERANTWORTUNGEN DER LEUTE, DIE MIT DER VERBREITUNG DER KORPORA AUF DEM INTERNET BEAUFTRAGT SIND?

Die Verbreitung der Korpora der gesprochenen Sprache auf dem Internet kann mit „der Ausgabe eines öffentlichen Kommunikationsdientes online“ gleichgesetzt werden. Es ist also wichtig, die Verpflichtungen und Verantwortungen der Herausgeber eines Kommunikationsdienstes online einzuschätzen.

[5] Siehe Art. 213-2 des Erbrechtes (früherer Artikel 7, Gesetz von 1979).

3 DAS VERFAHREN

AUFBAU, VERWERTUNG, BEWAHRUNG, VERBREITUNG

3.1 DAS VERFAHREN ERLÄUTERN

Die unter anderen wissenschaftlichen Zwecke, die mit dem Aufbau, der Verwertung, der Bewahrung und der Verbreitung der Korpora der gesprochenen Sprache gebunden sind, sind sehr verschiedenartig und ihr Respekt wie ihre Verschiedenartigkeit bedeuten, dass die Verschiedenheit der *Verfahren*, die von den Forschern und von den Verbreitungs- und Bewahrungsverantwortlichen dieser Korpora eingeführt werden, anerkannt wird.

Das *Handbuch der guten Praktiken* ist nicht beauftragt, dieses Verfahren aufzuzwingen, indem es einen Methodologietyp verordnet, es möchte aber alle Informationen liefern, die für die Kennzeichnung der „empfindlichen" juristischen und ethischen Punkte nötig sind. Nur die genaue und ausführliche Identifizierung der Elemente der betroffenen Situation und vor allem der Daten und ihrer Träger, der praktischen Erfahrungen, aber auch der verschiedenen Etappen ihrer Bearbeitung ermöglicht es, zugleich Elemente für juristische, der Situation angepasste Antworten und eine Abschätzung der eventuellen „Risiken" zu geben. Schließlich ist eine reflexive Analyse des mit dem Aufbau und den Bearbeitungen der Korpora der gesprochenen Sprache gebundenen Verfahrens das erste Element der Ausarbeitung einer von einer gesamten Wissenschaftsgemeinschaft anerkannten Ethik.

3.2 ELEMENTE DER BETROFFENEN SITUATION

Die Aufnahmen, die die Grunddaten der linguistischen Befragung darstellen, sind weit davon entfernt, ein einheitliches Objekt zu bilden. Ein Märchen zum Beispiel, das auf einem Tonband während traditioneller Feierlichkeiten auf dem Dorfplatz aufgenommen wurde, ist ein wissenschaftliches und patrimoniales Objekt, das von der digitalisierten Aufnahme eines Textes sehr unterschieden ist, der von einem „bezahlten Informanten" in den Räumen eines Universitätslabors gelesen wird, wie von den Antworten auf einen Fragebogen, die von einem Forscher in der Wohnung der befragten Person auf einer Mini CD aufgenommen werden, oder auch von einem spontanen Gespräch, das von den Forschern nicht ersucht wurde, in einem Café stattfindet und von einem oder mehreren Kameras gefilmt wird.

Man muss also zuerst die Elemente identifizieren, die die im Kontext gesammelten Daten charakterisieren :

- die *Art von Daten*, die das Korpus bilden, und ihre Träger (Aufnahme- aber auch Lagerungsträger für die Verwertung und Bewahrungsträger),
- die *verschiedenen Techniken*, die von den Forschern gebraucht worden sind, um die Daten zu sammeln,
- die Beschreibung der *Teilnehmer* und ihrer Rolle,
- die Kategorisierung der *Orte* des Sammelns.

3.2.1 *KORPUS UND ART DER DATEN*

Wenn der Wille, die Sprache zu „fangen", sehr alt ist, konnten nur vor kurzem die technologischen Fortschritte und die Forschung (vor allem in der Linguistik), an die Aufnahmen wie an echte „Daten" denken. Die internationale Lautschrift ist ein Beispiel eines alphabetischen „Aufnahmesystems", das von den Sprachwissenschaftlern erfunden wurde, um die Codierung der phonetischen und/oder phonologischen Transkription der Sprache zu normen. Die moderne Geschichte der Audio- und Videoaufnahmen verläuft mit den Transformationen der Registrierverfahren wie mit denen der gebrauchten Träger.

DIE AUFNAHMEVERFAHREN

Das digitale Aufnahmerverfahren wurde als erstes für die Tonaufnahme und die Tonbewahrung gebraucht. Es verschlüsselt die Schwankungen, die unter der Form von Signalen gemessen werden, die demselben Variationsgesetz unterworfen sind wie dem, das ihre Fortpflanzung in einem natürlichen Umfeld bestimmt. Seit einigen Jahrzehnten wird eher ein digitales Aufnahmeverfahren begünstigt. Unter dieser Form wird das punktuelle Messen des Luftdrucks ausgeführt (Probenentnahmen). Diese Messungen werden dann in der Form eines digitalen Wertes verschlüsselt, das in einem Maßstab ausgedrückt wird, und sind dann auf den Lagerungsträger unter der Form einer organisierten Folge von binären Einheiten dargestellt.

DIE AUFNAHMETRÄGER

- *Physische Träger*

Die ersten modernen Träger, die die Bewahrung der Sprache erlaubten, waren die physischen Träger. Dieses Wort kommt daraus, dass die vom einem Apparat (Mikrophon) gemessenen Druckschwankungen ins Material des Trägers physisch geschrieben werden. Unter diesen Trägern zählt man die ehemaligen Druckzylinder, die Schallplatten, usw. Diese Träger behalten in ihrem Material (Vinyl, Wachs, usw.) ein digitales Bild der gemessenen Druckschwankungen unter der Form einer gewellten Rille. Diese im letzten Jahrhundert gebrauchten Träger sind praktisch vernachlässigt. Sie stellen heute Zugangs- und Bewahrungsprobleme.

- *Magnetische Träger*

Die magnetischen Träger erschienen später, in der zweiten Hälfte des 20. Jahrhunderts. Verschiedene Lagerungsträger wurden und werden noch heutzutage (Faden, Band, Platte) in verschiedenen Aufmachungen (Spule, Kassette, Patrone, usw.) gebraucht. Das Prinzip beruht hier auf der Remanenz der magnetischen Partikeln, die am Träger entlang verteilt sind (d. h. der Eigenschaft dieser Partikeln, ihre Magnetisierung lange zu bewahren). Diese Magnetisierung der Partikeln wird je nach den Aufnahmeverfahren Informationen unter eine binären Form verschlüsseln können (wie auf den Festplatten-Trägern, DAT Kassetten, Disketten, usw.) oder Informationen unter eine digitalen Form (wie auf den Audiominikassetten, den VHSKassetten, usw.). Ein Teil dieser Träger ist dafür bestimmt, auf Hardwaren gebraucht zu werden, ein anderer Teil auf audio/digitalem Material. Auch hier ist es wie für jeden Träger : sie verfallen unweigerlich im Laufe der Zeit. Diese Träger

bleiben jedoch sehr beliebt, da der Zugang zu den Werkzeugen, die ihr Schreiben und Lesen ermöglicht, einfach bleibt.

- *Optische Träger*

Die optischen Träger erschienen als letzte ; sie sind hauptsächlich in ihrer Compact-DiscForm bekannt (Audio-CD, CD-ROM, usw.). Die Technologie beruht auf den optischen Eigenschaften der Komponente, nämlich zum Beispiel auf die Eigenschaft der Zellen, aus denen die Audio-CD bestehen, das Licht eines Laserstrahles widerzuspiegeln. Diese Träger werden hauptsächlich benutzt, um digitale Daten zu lagern (mit Ausnahme von einigen wenig populären Laserplatten und den für die Tonaufnahme wenig gebrauchten analogen Filmen). Die Mehrzahl dieser Träger ist für eine Benutzung auf der Hardware bestimmt, was den Zugang, die Übertragung und die Ausarbeitung der Daten erleichtert. Die Bewahrungsprobleme sind dieselben für jeden Träger, selbst wenn sie nicht gegen dieselben Angriffe (Licht, Wärme, Magnetfelder, Feuchtigkeit, usw.) empfindlich sind. Wie die Magnetträger haben sie den Vorteil, jüngst und populär zu sein, was ihre Benutzung heute erleichtert.

Es gibt andere Trägertypen, die zum Beispiel optische und magnetische Techniken mischen.

DIE WAHLSKRITERIEN

Da die Bewahrung der Träger sowieso dieselben Probleme stellt, egal welche Träger gewählt worden sind, werden die Wahlskriterien des richtigen Aufnahme- und dann Bewahrungsträgers eher auf der Qualität der Codierung, der Zugangs- und Bearbeitungsleichtigkeit, wie auf der Möglichkeit, seinen Inhalt ohne Informationsverlust wiederzugeben, beruhen. Zu begünstigen werden also die digitalen Träger sein im Vergleich zu den analogen Trägern, denn sie können identisch und ad infinitum vervielfältigt werden. Man wird auch die Datenträger begünstigen, und das wegen der Reihe von Werkzeugen, die die Informatik für die Verwaltung, den Zugang zum Lesematerial, die Verbreitung und die Bearbeitung der Daten (Verschlüsselung, Anonymisierungstechniken, usw.) anbietet, indem man zugleich berücksichtigt, dass diese Werkzeuge noch viele Standardisierungsprobleme stellen (zum Beispiel was die Wahl der Softwaren, des Formats, der Kompressionscodecs betrifft). Schließlich ist für die Bewahrung ein nicht auszulöschender Träger auch vielleicht eine gute Garantie für die Vermeidung ärgerlicher Zwischenfälle.

Die Wahl eines Formats, das eine wesensgleiche Vervielfältigung erlaubt, garantiert den Daten eine Art Fortbestand. Diese bezweifelt den Begriff „Original“, der sich dann weniger auf den Träger als auf die Daten selbst bezieht.

STANDARDISIERUNG DER VERMERKE

Die Korpora der gesprochenen Sprache bestehen meistens aus Audio- oder Videoaufnahmen und aus deren Vermerken.

- *Primär- vs Sekundärdaten*

Es wird im allgemeinen zwischen Primär- und Sekundärdaten unterschieden :

– die *Primärdaten* bestehen aus den *Aufnahmen*, die ein so enges Band wie möglich mit dem dokumentierten Ereignis haben. Sie enthalten auch die anderen Objekte, die im Rahmen des Tates gesammelt worden sind, wie die während der aufgenommenen Tat gelesenen oder geschriebenen Dokumente, die manipulierten Gegenstände, die nachgeschlagenen Bilder, usw. Sie enthalten auch die durch die Aktivität hintergelassenen Computerspuren.
– die *sekundären Daten* bestehen aus der Reihe von Beschreibungen, Transkriptionen, Vermerken, die die Primärdaten anreichern und oft im nachhinein auf der Grundlage der Primärdaten geliefert werden. Sie enthalten auch die Metadaten, die Transkriptionskonventionen, die Erlaubnisse der Teilnehmer, usw.

Die Unterscheidung zwischen Primär- und Sekundärdaten ist notwendig, vor allem um Interpretationsebenen zu differenzieren und die Wichtigkeit der Rückkehr zu den Primärdaten und ihrer Zugänglichkeit zu betonen. So konzentriert sich eine Analyse auf das Audio- oder Videoband und nicht nur auf die Transkription, selbst wenn diese ein wichtiger Zusatz ist, ohne den die Analyse wahrscheinlich unmöglich wäre. In diesem Sinne werden die Ausrichtungswerkzeuge zwischen die Ton-/Bildquelle und den Text der Transkription entwickelt. Jedoch ist diese Unterscheidung zwischen Primär- und Sekundärdaten begrenzt : sie darf die Tatsache nicht vergessen lassen, dass jede Aufnahme das Ergebnis zugleich techniker und praktischer Entscheidungen ist - die zum Beispiel die Wahl des aufzunehmenden Augenblicks und die Abgrenzung des aufgenommenen Abschnittes, die Wahl der Bildeinstellung und der Optik für die Video, der Positionierung und der Orientierung des Mikrofons für die Audio betreffen - , Entscheidungen, die auf einer Vorkenntnis der aufgenommenen Aktivität beruhen. Die „Daten“ werden nie „geschenkt“ oder „gesammelt“, sie werden von den Forschern aktiv hervorgebracht (Mondada, 2006).

- *Erläuterung der Datenstruktur*

Um die Vermerke zu schreiben, werden Ausdrucksformalismen gebraucht, die zugleich erlauben, den Inhalt der Kommentare zu verschlüsseln und zu erläutern, um welche Art Kommentar es sich handelt. In den relationalen Datenbanken wird man zum Beispiel Tabellen gebrauchen, die Felder mit Namen umfassen, (d. h. *pos* für „*part of speech*“) die dazu dienen werden, Werte zu lagern, (z. B. „*Verb*“) die in besonderen Strukturentypen (Charakterketten, Nummern, usw.) ausgedrückt sind.

Ein alternativer und im Gebiet der Textvermerkung breit gebrauchter Formalismus wird von der umfangreichen Familie der Textmarkierungssprache gebracht. Dieser Formalismus grenzt die Kommentare durch formale Zeichen (d. h. Markierungen) ab, die anzeigen, um welchen Typ Kommentar es sich handelt. Es besteht heute ein ziemlich breiter Konsens von allen Disziplinen über die Annahme der jüngsten Textmarkierungssprache XML als Strukturierungs- und Austauschformalismus der Dokumente.

- *Standardisierung/ Normalisierung*

Während die Wahl eines Formalismus, der es erlaubte, die gesamten Vermerke auszudrücken und ihre Struktur zu erläutern, notwendig ist, reicht sie aber nicht, um den Austausch oder die Bewahrung eines Dokumentes zu erlauben. Um ein Dokument auszutauschen oder zu bewahren, muss die für die Verschlüsselung seiner Struktur und des Inhalts seiner Kommentare gebrauchte Sprache den Teinehmern (im Rahmen eines Austausches) gemein sein oder im Laufe der Zeit bekannt bleiben (im Rahmen eines langfristigen Gesprächs). Im Kontext eines Dokumentes, das eine Textmarkierungssprache gebraucht, müssen die Namen der Strukturelemente (Markierungen, Attribute) bekannt sein und ihre Begriffsbestimmung angenommen und geteilt werden, so wie die gesamten Zwänge (Markierungverkettung, kontrollierte Wortschätze, optionaler oder zwangsläufiger Charakter bestimmter Strukturen...).

Wenn viele Leute oder eine ganze Gemeinschaft sich auf eine gemeinsame Sprache einigen können, wird von Standardisierung gesprochen. Das war zum Beispiel der Fall mit der internationalen Lautschrift. Während die Standardisierung für den Austausch notwendig ist, erfordert die langfristige Bewahrung Garantien über die Übertragbarkeit und den Zugang zu der Dokumentation der gemeinsamen bereitgestellten Sprachen. Als solche müssen die Normalisierungsinstitutionen einen bestimmten Fortbestand zu den von ihnen bereitgestellten Normen bringen können, so wie eine Unabhängigkeit den privaten Interessen gegenüber. Sie müssen auch für das Allgemeinwohl repräsentativ sein. Unter diesen Bedingungen wäre es günstig, Normen für die Verschlüsselung und die Formatierung der Daten zu gebrauchen, überall wo sie bestehen. Als Beispiel kann man die Charakterverschlüsselung ISO-10646, die unter dem Namen Unicode bekannter ist, zitieren, eine Charakterverschlüsselung, die sich als universal gibt und die meisten Schreibarten der Welt berücksichtigt, die internationale Lautschrift inbegriffen. Für die Codierung der linguistischen Analyse wird es interessant sein, die Empfehlungen der Text Encoding Initiative (TEI) zu lesen, die Analysen für Datenstrukturen wie Wörterbücher, Gedichte oder die Transkription der Sprache vorschlägt. Es wird auch sehr nützlich sein, den Fortschritten der Arbeitsgruppe der ISO über die Verwaltung der linguistischen Ressourcen TC37 SC4 zu folgen.

So können die Prinzipien, die die Wahl der einen oder der anderen Technologie für die Vermerkung lenken müssen, können in vier Fragen zusammengefasst werden :

- Ermöglicht diese Technologie, alle Vermerke *eindeutigerweise zu verschlüsseln* ?
- Weist diese Technologie einen *vorrangigen Charakter* oder eine *gesetzliche Begrenzung* auf, die daran hindern könnten, die Vermerke mit anderen zu teilen (Besitzformate, auf Patenten ruhende Techniken, usw.) ?
- Wird diese Technologie *von der Gemeinschaft angenommen*, mit der der Datenaustausch geplant wird ?
- Wurde diese Technologie der Gegenstand einer *Normalisierung* ?

3.2.2 BEFRAGUNGSTECHNIKEN

Die linguistischen Befragungen haben aus technischen Gründen nicht immer zu Aufnahmen geführt (die ersten Aufnahmewerkzeuge gehen kaum bis zu mehr als einem Jahrhundert zurück), aber auch aus methodologischen und theoretischen Gründen. So sind noch die geschriebenen Fragebogen, das Notieren, das Zurückgreifen auf die Intuition und/oder die Beobachtung des Forschers Beschreibungswerkzeuge, die von den Sprachwissenschaftlern gebraucht werden. Die Möglichkeit, die Sprache aufzunehmen, und die Entwicklung der Techniken (Miniaturisierung der Geräte, Qualität des aufgenommenen Zeichens, Digitalisierung und Computerbearbeitungen der Ton- und Videodaten) haben es trozdem den Terrainbefragungen ermöglicht, Methodologien zu entwickeln, die doch sehr unterschiedlich sind, es sei nur durch die Verschiedenartigkeit der betroffenen wissenschaftlichen Gebiete (Dialektologie, Soziolinguistik, konversationelle Analyse, Psycholinguistik, Linguistik des Mündlichen, automatische Bearbeitung der Sprache, Ethnolinguistik...). Die Forschungen über die Methodologie der Befragung haben jedoch die Forscher dazu geführt, die aufgenommenen Daten als das Ergebnis der Befragungssituation zu betrachten, ganz anders als eine Auffassung vorhandener Daten, die einfach gesammelt wären (Cameron et *al.*, 1991).

Die Befragungstechniken haben zum Schluss eine wichtige Rolle, indem sie der befragten Person ermöglichen (oder nicht), die den Forschern gegebenen Daten kontrollieren zu können. Der Rest dieses Kapitels ist einer Bestandaufnahme der verschiedenen Befragungstechniken, die bei dem Aufbau der Korpora der gesprochenen Sprache angewendet worden sind, gewidmet.

DER FRAGEBOGEN

Der aufgenommenen mündliche Fragebogen kann verschiedene Formen annehmen; er besteht meistens aus geschlossenen oder halb offenen Fragen und aus Listen lexikalischer Wörter oder von dem Forscher vorbereiteter Texte. Nur der Fall der Texte, die vor dem Fragebogen vorhanden sind, kann die Frage des geistigen Eigentums stellen (wie zum Beispiel das Lesen eines vom Urheberrecht geschützten Textes oder eine Produktion, deren Inhaltsbesonderheit geschützt wäre). In den anderen Fällen geht es darum, unter anderem die Schwankungen, die Regelmäßigkeiten zu bekommen und wie der Befragte diese Regelmäßigkeiten einem gemeinsamen linguistischen System entsprechend wahrnimmt.

Der Empfindlichkeitsgrad der gesammelten Informationen ist meistens vorauszusehen, da der Forscher selbst den Fragebogen herstellt und so die Risiken je nach der Natur der Fragen schätzen kann. Jedoch kann bei anscheinend harmlosen Fragen auch etwas auf dem Spiel stehen, das vom Forscher unvermutet wurde und aus dem besonderen Kontext der Befragung auftaucht. Dazu muss betont werden, dass der Fragebogen mehr als jede andere Technik das Kennzeichen der Befragungshandlung und des *Ergreifens* des Forschers trägt (Encrevé, 1983) und infolgedessen potentiell ein Einschätzungsgefühl, selbst wenn dieses durch die explizit angebotene Möglichkeit, alle oder einige Fragen nicht zu beantworten, oft gemildert wird (für eine Analyse der Fragebogensituationen vgl Achard 1991). Unterstreichen wir zum Schluss einen Punkt, der manche Befragungssituationen betrifft, der aber mit dem Fragebogen besonders gebunden ist : dieser enthält oft

einen Teil, der dem Sammeln der persönlichen Daten (Alter, berufssoziologische Kategorie...) zu dem Zweck, das soziologische Profil des Befragten aufzustellen, gewidmet ist.

DAS GESPRÄCH

Das Gespräch besteht aus offenen Fragen, da das Ziel hauptsächlich das Sammeln einer großen Menge linguistischer Daten ist. Das Gespräch setzt immer eine mehr oder weniger starke Führung von dem Forscher voraus (vom autoritären bis zum semi-autoritären, sogar nichtautoritären Gespräch ; von dem am meistens genormten bis zu dem am wenigtens genormten) und bringt es dabei der mündlichen Befragung oder der weniger gezwungenen Interaktion näher (Maynard *et al.,* 2002, Houtcoop-Streenstra, 2000). Obwohl der Forscher die Kategorien und die Themen ins Gespräch oft einführt, von denen er wünscht, dass die Informanten sich damit befassen, kann die Methodologie der Forscher auch verlangen, dass das Thema der Forschung vor dem Gespräch nicht ausführlich mitgeteilt wird, im Bemühen um ein Sammeln der so natürlichen Produktionen wie möglich, und stellt also das Problem des Wahls des Augenblickes und des Inhalts der Informationen, die den Befragten gegeben werden (vgl Mondada 2001).

DIE MÄRCHEN-, LIEDER- ...SAMMLUNG

Die Märchen-, Liedersammlung und die Sammlung der mündlichen Produktionen der traditionellen Kulturen ist eine häufige Praktik vor allem in den Bereichen der Beschreibung der Sprachen mit mündlicher Tradition und der Ethnolinguistik. Außer dass es wichtig ist, diese Märchen, Lieder und Erzählungen in ihren Kontext zu bringen (Bedeutungen, die in einem kulturellen Kontext implizit sind, können in einem anderen entgehen oder unwichtig scheinen), müssen zwei Elemente hauptsächlich berücksichtigt werden : das geistige Eigentum traditioneller Produktionen einer Gemeinschaft und die Sammelnbedingungen, die oft mit sozialen Aktivitäten in einem öffentlichen oder privaten Rahmen gebunden sind.

DIE LEBENSGESCHICHTEN

Die Lebensgeschichten werden häufig bei Anthropologie-, Geschichte-, Ethnolinguistik-, aber auch Dialektologieforschungen und in vielen anderen Gebieten in Anspruch genommen (Guillaumou *et al.* 1997). Diese Art Aufnahmen bedeuten zwangsläufig eine wichtige Quelle von persönlichen Daten, die den Verfasser der Erzählung und Dritte betreffen, die eventuell mit einem sozialen oder historischen besonders heiklen Kontext verbunden werden können, insbesondere wenn die persönliche Erzählung ein von einer oder mehreren Gemeinschaften erlebtes Ereignis widerhallt.

So können selbst in der Forschung über exklusiv linguistische Phänomene die in den Lebensgeschichten enthaltenen Worte und die Frage über den Einfluss ihrer Verbreitung im öffentlichen Raum der Verantwortung des Forschers, der sie ersucht und ausnutzt, nicht entkommen.

Dazu können die Auswertungs- und Verbreitungsbedingungen dieser Erzählungen sich in einem sehr unterschiedlichen Kontext machen von dem sehr merkwürdigen,

der die Sammlung geprägt hat und der oft in einem präzisen Rahmen und dank einer privilegierten Beziehung zwischen dem Forscher und dem Zeugen stattfand.

Zum Schluss kann sich die Frage des geistigen Eigentums einer Lebensgeschichte und des unveräußerlichen moralischen Rechtes im Fall Originalerzählungen als besonders relevant erweisen.

DIE AUFNAHME IN EINEM LABOR

Die Aufnahmen in einem Labor gemäss einem Versuchsprotokoll werden in den Sprachwissenschaften, besonders in den Bereichen der Psycholinguistik, der Phonetik und der automatischer Sprachbearbeitung gebraucht. So interessieren bestimmte Korpora direkt die angewandte Forschung und die vom linguistischen Engineering betroffenen Unternehmen und sie werden manchmal teilweise oder ganz mit privaten Geldern finanziert.

So wie die Fragebogen, außer in den ganz bestimmten Fällen der Texte, die urheberrechtpflichtig sind, scheinen die Produktionen der Teilnehmer, die nach einem von den Forschern erarbeiteten Versuchsprotokoll gebracht wurden, von dem Urheberrecht (außer in Einzelfällen) nicht betroffen werden zu müssen. Die typische Situation der aufgenommenen Person muss jedoch allen Fällen von experimentellen Forschungen auf einem Menschen nähergebracht werden.

DIE AUFNAHME VERURSACHTER AKTIVITÄTEN

Es handelt sich hauptsächlich um Aktivitätenaufnahmen im üblichen Kontext der betroffenen Sozialarbeiter, selbst wenn die Vorschriften vom Forscher gegeben werden (Aktivitäten, die Kindern im schulischen Rahmen angeboten werden, simulierte Aufgaben in beruflicher Situation, usw.). Diese Situation kombiniert zugleich die Aufnahmedaten nach einem Versuchsprotokoll (wofür der Forscher die Verantwortung trägt) und die Daten des üblichen Kontextes in ökologischem Milieu; sie bietet infolgedessen einen vom Forscher kontrollierbaren Doppelrahmen. Diese eindeutige Intervention des Forschers (dessen Rolle von den Teilnehmern klar identifiziert werden kann) erleichtert die Erhaltsbedingungen einer aufgeklärten Einwilligung ; jedoch soll eine ganz besondere Aufmerksamkeit dem beruflichen Milieu, das die Einwilligung zwingen kann (Vertraulichkeit), geschenkt werden.

DIE AUFNAHME VON AKTIVITÄTEN IN IHREM ÜBLICHEN KONTEXT

Die Forschungen in der Soziolinguistik, in der konversationellen Analyse und in der Analyse der Technologiengebräuche (Computer Supported Cooperative Work ; Dialog Mensch-Maschine) interessieren sich für das Sammeln von Daten in einer Tätigkeitssituation, die nicht vom Forscher organisiert wird und nicht durch ihre Anweisungen verursacht wird. Es handelt sich um Tätigkeiten, so wie sie üblicherweise, auch ohne die Anwesenheit des Forschers, stattfinden. Diese Tätigkeiten können sehr unterschiedlich sein : Versammlungen, berufliche Aktivitäten, Einholung von Auskünften, telefonische Interaktionen, usw. Die Sammelntechniken sind auch sehr verschieden. Sie gehen von der teilnehmenden Beobachtung bis zu der erlaubten Aufnahme über den Gebrauch von „Ressource-Leuten“, die innerhalb der beobachteten Berufskollegen gewählt worden sind, und

insbesondere damit beauftragt sind, die Aufnahmegeräte zu tragen (Mikrofon, eventuell Kamera).

Der von diesen Techniken geteilte Zweck ist die Forschung, in einer natürlichen Situationen, nach Angaben, und nimmt eine Methodologie an, die sich bemüht, die von den Aufnahmeeinrichtungen hervorgebrachten Folgen zu mindern (Heath, 1997; Jordan & Henderson, 1995). Es ist also höchst wahrscheinlich, dass die Daten empfindliche Informationen, was den Schutz des Privatlebens betrifft, enthalten. Die Modalitäten der Einwilligungssammlung müssen es berücksichtigen und sich darauf anpassen.

DIE WIEDERAUFNAHME

Einige Korpora, die aus Aufnahmen bestehen, die für andere Zwecke als die wissenschaftlichen oder als die bei der Einwilligungssammlung erwähnten, von anderen Akteuren als die Forscher hervorgebracht wurden, können zu dem Zweck linguistischer Forschungen zu einer Wiederaufnahme oder zu patrimonialen, memorialen oder politischen Zwecken zu einer Bereitstellung führen (so stellte zum Beispiel die Brigade der Feuerwehrleute von New-York im August 2005 die Aufnahmen der Radiomitteilungen während des Attentats vom 11. September 2001 zur Verfügung). Diese Korpora kennzeichnen sich also durch die Anwesenheit der Einwilligung für den neuen Zweck und durch die Tatsache, dass die archivierten Worte ohne die Kenntnis dieses Zweckes hervorgebracht wurden, sondern in einem anderen Rahmen und mit anderen Zwecken. So kann bei Interviews oder Seminaren - die zum Beispiel mit dem Zweck einer Ausbreitung der transkribierten Inhalte aufgenommen wurden - die Verbreitungserlaubnis die transkribierten und bestätigten Worte und keine spätere Wiederaufnahme betreffen.

DIE WIEDERAUFNAHME VON MEDIENAUFNAHMEN

Die Wiederaufnahme von Medienaufnahmen ist ein Einzelfall der vorigen Kategorie, die die Besonderheit anbietet, sich auf Daten zu beziehen, die in einem Rahmen öffentlicher Verbreitung hervorgebracht wurden.

Auch in diesem Fall ist die Einwilligungssammlung die Vorbedingung jeder Auswertung, wenn der Inhalt der Aufnahmen vom Urheberrecht geschützt ist (zum Beispiel im Fall einer originellen Produktion). Es existiert jedoch eine Ausnahme für einen bestimmten Zeitraum, wenn es sich um Reden handelt, die an das Publikum gerichtet sind und vor einem Publikum gehalten werden, wie in den folgenden Zeilen festgelegt :

> *Französisches Urheberrechtsgesetz, Art. 122.5*[6] *:*
> *Die Verbreitung der Reden, die an das Publikum gerichtet sind und in den politischen, administrativen, richterlichen oder akademischen Versammlungen so wie in den öffentlichen Versammlungen mit politischem Charakter und in den amtlichen Feierlichkeiten gehalten werden, selbst wenn sie*

[6] Artikel 122.5 des französischen Urheberrechtsgesetzes.

ungekürzt ist, durch die Presse oder Fersehübertragung, als Information des Tagesgeschehens.

Erinnern wir daran, dass die persönliche Aufnahme einer Sendung für diese Kopie, die dem privaten Nutzen streng vorbehalten ist, einer gesetzlichen Lizenz entspricht. Die Aufführung dieser Kopie kann sich erst im „Familienkreis" vollziehen. Ebenso verleiht das Vervielfältigungsrecht (mit seinen Grenzen) einer Kassette oder einer DVD aus dem Handel kein Verbreitungsrecht.

Ein letztes Detail zum Schluss : der öffentliche Charakter des Kontexts der Fernsehverbreitung bedeutet keine Einschränkung des Schutzes der persönlichen Daten.

Die Verschiedenartigkeit der für das Datensammeln gebrauchten Techniken bestimmt genauso viele *Situationen*, die *Teilnehmer* klar herausstellen, deren Rolle das erste Kategorisierungselement ist.

3.2.3 ROLLE DER TEILNEHMER

Die Teilnehmer an die Befragung und an die aufgenommenen Aktivitäten kann man auf verschiedene Weisen kategorisieren, die alle spezifisch erläutern, was sie machen und was sie sagen (Sacks, 1972). So können die Teilnehmer an eine Aufnahmesituation zugleich für Befragte (wenn die Situation einem Forschungsobjekt zugeordnet wird) und für Sozialarbeiter gehalten werden - deren genaue Karakterisierung vom Kontext, von der Aktivität, von den Engagement- und Teilnahmeformen abhängt - , was zugleich die soziale Geschichte der Personen und die lokale Erfüllung ihrer Rolle, aber auch ihrer Identität während des Treffens voraussetzt. Je nach der Art und Weise, wie die Forscher selbst diese viele Kategorien behandeln, können verschiedene Konsequenzen zugleich für das Objekt der Forschung und für die Bewertung des mehr oder weniger empfindlichen Charakters der Aktivität erscheinen.

TEILNEHMERKATEGORIEN

Die in der Literatur sehr abwechslungsreiche gebrauchte Terminologie, um die Kategorien der Teilnehmer an eine Befragung zu definieren, macht verschiedene ethische und theoretische Auswirkungen bekannt (Cameron et *al.*, 1991). Hier eine nicht erschöpfende Liste der gebrauchten Worte, die in verschiedenen Forschungszusammenhängen gebraucht werden, um die Teilnehmer vom Standpunkt ihres Engagement in die Forschung zu charakterisieren :

– Informanten,
– Sprecher,
– Subjekte,
– „Versuchskaninchen",
– Einheimische,
– Sozialarbeiter,
– Teinehmer,
– Mitarbeiter,
– Partner,
– Befragte,

– Zeugen.

Diese terminologischen Entscheidungen sind meistens das Ergebnis theoretischer und politischer Betrachtungen, die die Art der existierenden, gebildeten, oder zwischen den Befragten und den Forschen entwickelten Beziehungen erkennen lässt.

Wenn wir hier nicht entwickeln können, was bei diesen theoretischen Betrachtungen auf dem Spiel steht, ist es jedoch bedeutend, die Kennzeichen einer besonderen *Beziehung* ausfindig zu machen, die verschiedene Verwirklichungen des *Paares* Befragter/Meinungsforscher begründet und dabei je nach den Daten dieser Beziehung verschiedene Rechte impliziert (Sacks, 1972).

Zwei Elemente kennzeichnen vor allem diese Beziehung : die Nähe/Entfernung der Teilnehmer und die Rollen in Aktion und in der Situation.

- *Nähe/Entfernung*

Die Frage der Zugänglichkeit der Befragungssituationen für den Forscher hat sich immer gestellt und hat verschiedene Formen des *fieldwork* (Terrainarbeit) rechtfertigt, Arbeit, die vom Eintauchen in eine dem Forscher ganz fremde Gemeinschaft bis zur Auswertung seiner Gehörigkeitsbande zu seiner Gemeinschaft geht.

Diese Probleme wurden mit den Worten Paradox des Beobachters bearbeitet - nach dem das Phänomen der Befragung sich auflöst, sobald es beobachtet wird (so das Einheimische für Labov, 1972) - so wie auch mit den Worten symbolischer Gewalt zwischen dem Befragten und dem Forscher (Bourdieu, 1993). Sie wurden auch mit dem Wort *Reflexivität* bearbeitet - durch Forscher, die ihre Anwesenheit und die der Forschungseinrichtung in die Analyse des befragten Objekts integrierten (vor allem in der Anthropologie, Clifford & Marcus, 1986 ; Mondada, 1998).

Die Befragungen bei den „Verwandten“ des Forschers erleichtern die Kontaktaufnahmen und den Zugang zum Terrain, und stellen zugleich Undeutlichkeitsprobleme zwischen die durch die Befragung diktierten Beziehungen und die persönlichen Beziehungen, wenn der Forscher seine eigenen Netze für eine Befragungsarbeit ausnutzt. Diese Fragen stellen sich im Fall der Befragungen bei „Entfernten“ nicht (beobachtete Gemeinschaft, bemusterte Panel, Zeugen, die nicht vom Forscher ausgesucht wurden,...), wo die Zugangschwierigkeiten größer sein können, wo der Forscher aber einen oft klareren und besser anerkannten Status als solcher haben kann, wenn das Vertrauen einmal erlangt und eine Beziehung hergestellt wird (Beaud & Weber, 1977). Die Forschung in den Sozial- und Humanwissenschaften hat außerdem oft „gefangene Bevölkerungen“ gebraucht, im Sinne, dass der Forscher einen von Institutionen erleichterten Zugang (die Schule, das Krankenhaus...) genießt, und dass diese Bevölkerungen begrenzte Möglichkeiten haben, die Mitarbeit abzulehnen (Kinder, Schüler...). Von nun an soll eine besondere Aufmerksamkeit der Annäherung von Bevölkerungen gewidmet werden wie :

– den benachteiligten Personen,
– den behinderten Personen,
– den Kindern,
– den Schülern und Studenten,
– den Angestellten von Unternehmen oder Institutionen, mit denen man durch ihre Hierarchie in Kontakt gesetzt wurde,

– usw.

Es ist in diesem Fall üblich, die Erlaubnisse für die Personen durch die Erlaubnis einer gesetzlichen Verantwortungsperson (die Kinder werden von den Erwachsenen ersetzt) zu verdoppeln.

Dieser Einzelfall zeigt, dass die unterzeichnete Erlaubnis nicht immer als eine genügende Handlung betrachtet werden kann und dass einige Befragten über dem, was sie unterzeichnet haben, geschützt werden müssen (Verantwortung des Forschers).

- *Rollen in Situation*

Während der Befragungssituation und je nach den gebrauchten Techniken kann die Beziehung Meinungsforscher-Befragter sehr unterschiedliche Formen annehmen und mehr oder weniger direkte Verpflichtungen voraussetzen.

- *Rollen des Forschers*

– *äußerlicher Beobachter,*
– *teilnehmender Beobachter,*
– *engagierter Beobachter* (verteidigt die Gemeinschaft),
– *Mitglied der Gemeinschaft*, das an eine Aktion-Forschung teilnimmt (ein Projekt, das von der Gemeinschaft auskommt oder ihre Probleme und Ziele berücksichtigt und spezifisch unter der Form einer „Aktion-Forschung" eingreift),
– *verkleideter Beobachter* (*cross-dressing* in der ethnographischen Tradition), der sich durch Beziehungen, eine Arbeit oder ein Amt in die Gemeinschaft eingliedert, seine Fragenderidentität aber nicht erklärt,
– „*Zauberer von Oz*" : Meinungsforscher, der sich hinter eine technologische Einrichtung versteckt, die dem Informanten antworten sollte.

- *Rollen der Befragten*

– *Befragter*/Informant/fokussierter Sprecher
– „*Peripherleute*" : die Techniker, die Passanten, die Schauspieler...
– Leute, die mit denjenigen, die für die Befragung genehmigt wurden, *gebunden* sind (z. B. Kunden, die ein Call-Center anrufen oder Ehemann der befragten Frau)
– der „*Kumpan*" : privilegierter Informant, der das Aufnahmegerät trägt und dem Meinungsforscher ermöglicht, eine Gruppe zu infiltrieren, zu der der Kumpan gehört oder Zugang hat.

Diese Rollen legen in die widerspiegelte Spannung zwischen Teilnahme und Beobachtung, die durch das Wort „teilnehmende Beobachtung" (Becker, 1960 ; Platt, 1983 ; Spradley, 1980) widerspiegelt ist, Rechenschaft über die möglichen Schwankungen zwischen beide ab. Je nach diesen Rollen (Adler, 1987) wird das Engagement der Befragung und den Aufnahmen gegenüber sehr unterschiedlich sein, so wie die Kontaktmodalitäten für das Erhalten einer aufgeklärten Einwilligung.

3.2.4 ORTE

Die Information über den Ort des Sammelns bedingt durch seine eigenen Merkmale und durch die Rolle, die er in der Befragungssituation hält, besondere juristische Antwortselemente.

So kann man zuerst die *öffentlichen Orte*, wo die wissenschaftliche Betätigung der Audio-videoaufnahme keine andere Erlaubnis verlangt als die der befragten Person, von den *privaten Orten* unterscheiden, wo man sich der Vorerlaubnis des Eigentümers/Verantwortlichen beugen muss, die vom Erhalten der Einwilligung des Befragten verschieden ist.

Der Ort kann auch von der Beziehung abhängen, die die Teilnehmer hergestellt haben. Handelt es sich um einen Ort, wo die Anwesenheit der befragten Person dem Meinungsforscher zuzuschreiben ist (Labor, Aufnahmeraum...) oder soll sich dieser an Ort und Stelle begeben und so den eigenen Raum des Befragten einnehmen ?

Schließlich kann der Aufnahmeort in die Daten integriert werden (Audio- oder Videomerkmale, die in den Daten vorkommen) oder nur von einer Information abhängen, die in den Metadaten eventuell anwesend ist.

3.3 TERRAINPRAKTIKEN

Dieser Kapitel hat die Absicht, die Allgegenwart der ethischen und juristischen Einsätze in den Etappen des Terrainverfahrens zu zeigen, das den Aufbau von Korpora von mündlichen, interaktiven und multimodalen Daten bezweckt. Wir werden unter anderem die *Vorphasen* der Befragung betonen, die der Aufnahme der Daten vorherig sind und wo es unter anderem um die Herstellung einer Beziehung zwischen die betroffenen Personen geht : diese Herangehensweisen sind nicht nur mit den Befragungsmethodologien (vgl oben 3.3.2) eng gebunden, sondern auch mit den Möglichkeiten und technischen Grenzen der gewählten Aufnahmeeinrichtung, von denen die spezifischen Zwänge abhängen um die Erlaubnisse, eine Aufnahme zu machen. Wenn die Meinungsforschung einmal fertig ist und die Daten analysiert sind, heißt es, die *Rückkehr auf das Terrain* zu verschiedenen Formen der „Zurückgabe" der Ergebnisse und der Erfahrungen zu organisieren - Zurückgabe, die man lieber voraussehen sollte und die die Art des Engagements den betroffenen Leuten gegenüber gestaltet.

3.3.1 HERANGEHENSWEISEN

Die Meiningsforschungen, die das Sammeln aufgenommener Daten bezwecken, hängen zwangsläufig von der Qualität der Beziehung mit den Ressourcen-Leuten ab - ob man sie Informanten oder Partner nennt (vgl oben 3.2.3). Der Einsatz dieser Personen variiert nach der gewählten Befragungsmethode : unten werden wir die Zeitlichkeit der verschiedenen Herangehensweisen bei den direkt betroffenen Personen oder ihrer Hierarchie betonen, so wie die Frage, wie man die Rückkehr, die Gegengabe, eventuell die Bezahlung dieser Personen organisieren kann.

BEZIEHUNGS- UND HERANGEHENSWEISENTYPOLOGIE

Die Art der Annäherung auf dem Terrain - wie eine persönliche und soziale Beziehung entsteht - mag als eine Tatsache gelten, die sofortige ethische und juristische Folgen hat. Die Herstellung der Beziehung mit den Informanten wirkt einerseits auf die Qualität ihrer Mitarbeit, also infolgedessen auf die Qualität der hergestellten Daten : andererseits wirkt sie auf die Vertrauens-, Annahme-, sogar Interessebeziehungen oder die wissenschaftliche Neugierde, die die Informanten gegen die Meinungsforscher hegen werden.

Man kann eine Typologie der mit den Informanten hergestellten Beziehungen skizzieren, indem man sie mit dem Augenblick ihrer Annäherung im Befragungsverfahren verbindet :

- Wenn die Meinungsforschung mit *namentlicher Einladung* der Informanten im Labor verfährt, werden die Modalitäten ihres Einsatzes meistens *vorher* verdeutlicht, das heißt im Augenblick, wo die Personen annehmen, an die Aufnahme teilzunehmen, die in vorher vereinbarten Orten und zu vorher vereinbarten Momenten gemacht wird. Da sind die Personen entweder vom Forscher (oder durch eine Einrichtung, die für ihn arbeitet) ausgesucht und mit ihm in Verbindung gesetzt oder sie antworten auf eine „Aufforderung für Freiwillige". Die Aufforderung oder die Rekrutierungsbekanntgabe ist die erste Kommunikationshandlung, die die Form des Kontakts, sogar des Vertrags, der sich mit dem Forscher herstellt, zum Ausdruck bringt (oder Erwartungen verursacht, was den Kontakt, sogar den Vertrag betrifft).
- Wenn die Meinungsforschung unter der Form eines *fieldwork* (Terrainarbeit) vorgeht, die die mehr oder weniger lange Anwesenheit des Forschers auf dem Terrain und Formen einer *teilnehmenden Beobachtung* voraussetzt, - Formen, die ganz gewöhnlich innerhalb der von den Sprachwissenschaftlern wie von anderen Forschern in Sozial- und Humanwissenschaften übernommenen ethnographischen Methoden diskutiert wurden (Depperman 2000, Duranti, 1997, Hammersley, Atkinson 1995 & Moerman 1988) - entwickelt sich die Beziehung mit den Informanten in der *Dauerhaftigkeit* dieser Anwesenheit und sie ist oft mit dem Aufbau persönlicher Beziehungen geknüpft, die unter anderem ein gegenseitiges Vertrauen voraussetzen. Auf bestimmten Terrains ist der Forscher nicht der erste, der eingreift und andere sind vielleicht vorangegangen. Nach dem Verhalten seiner Vorgänger wird der Empfang durch die Gemeinschaft mehr oder weniger leicht sein und insbesondere die Anspüche, was die Rückkehr betrifft, werden mehr oder weniger groß sein.
- Wenn die Befragung mit *Gesprächen*, „*Aufnahmen auf dem Bürgersteig*", *Tätigkeitenaufnahmen* vorgeht, die *zufälligerweise* in öffentlichen Orten gemacht werden, ohne besondere Zeugen zu erzielen, sondern Passanten, die nur wegen ihrer Anwesenheit am Ort der Aufnahme gewählt wurden, so ist definitionsgemäß ein Vortreffen mit den

Informanten unmöglich. Dann findet die Erklärung der Zwecke und die Bitte um die Erlaubnis *gerade vor*, *während* oder *gerade nach* der Aufnahme.

– In bestimmten Fällen ist es möglich, einen der Aufnahme *folgenden* Kontakt zu planen : so ist es für Aufnahmen, die von einerm Teil der Teilnehmer unbemerkt gemacht wurden, deren aufgenommener Auftritt nicht voraussehber war (es ist zum Beispiel der Fall der Telefonate, wo ein Teil an der Forschung mitarbeitet und der andere nicht immer Bescheid weiß ; man wird sich später mit ihm in Verbindung setzen, um seine Erlaubnis zu bekommen).

Die Form des Kontakts, des Einsatzes, der Glaubwürdigkeit, des Vertrauens ändert sich sehr je nachdem, ob die Beziehung zwischen dem Forscher und dem Befragten zunächst oder während der Arbeit auf dem Terrain, dauerhafterweise oder im Moment der Aufnahme, oder noch nach dieser aufgenommen wird.

DIE KONTAKTIERTEN PERSONEN

In der eben gemachten Vorstellung haben wir aus Einfachheitsgründen bedacht, dass der Kontakt sich mit der oder den von der Aufnahme direkt angesprochenen Personen hergestellte ; es handelt sich aber oft um Personen, die zu einer Gruppe oder einer Einrichtung gehören – was voraussetzt, dass man sich mit mehreren Leuten in Verbindung setzt. Man soll auch folgende Fälle unterscheiden :

– Den Fall, wo der Informant *in seinem Namen* individuell handelt.

– Den Fall, wo man sich mit dem Informanten im Rahmen seiner beruflichen oder institutionellen Tätigkeiten in Verbindung setzt, so dass er als Mitglied einer Gemeinschaft eingreift. Man setzt sich zunächst in Verbindung mit der Hierarchie der durch die Befragung betroffenen Personen : so kann es für die Führung eines Unternehmens sein, einen Häuptling oder Schülereltern. Man soll bemerken, dass die Beziehung zwischen der Person und ihrer Hierarchie oft nicht eindeutig ist und zu unterscheiden einlädt, was den Personen und ihrer Hierarchie versprochen, erklärt, gezeigt, usw wird.

BEZAHLUNGEN

Wenn man sich mit den durch die Befragung getroffenen Personen in Verbindung setzt, mögen Versprechungen gemacht werden, echte Verträge, Entschädigungen, Bezahlungen, Rückerstattungen vorgeschlagen werden. Diese Verpflichtungen können zugleich ethisch und juristisch, sozial, materiell und sogar finanziell sein. Die Frage der Form einer „Entschädigung“ der Informanten stellt sich sowieso - die sehr verschieden ist, wenn man sie als „Gegengabe“, „Bezahlung“, „Entschädigung“, „geleisteter Dienst“ bezeichnet...

Mehrere Möglichkeiten sind denkbar :

– *während, sogar vor der Meinungsforschung :* finanzielle Bezahlungen, die schon bei der Bearbeitung des ersten Vertrags versprochen werden, Gegengaben in Naturalien, symbolische Gegengaben, Leistungen für die betroffene Gemeinschaft ;

– *nach der Meinungsforschung :* Dankbarkeit des Informanten unter Formen, die vom Dank oder Zitat bis zur Erwähnung als Mitverfasser oder Mitarbeiter, sogar als Partner der Forschung gehen, Rückgabe der Ergebnisse, Rückgabe der Daten/Korpora unter der Form eines Archivs, Verbreitung vom Know-how, günstige Rückgaben, die für die breite Gemeinschaft langfristig erwartet sind (nach der Vorlage der Vorteile, die nach einer ärztlichen Forschung erwartet sind).

Für eine Diskussion über diese „*Rückgabe*formen" verweisen wir (oben, 3.3.4) auf die Diskussion über die „*Rückkehr*" auf das Terrain. Die Frage ist, was man den Informanten versprechen kann/soll, wenn man die Beziehung aufnimmt, mit Berücksichtigung der folgenden Tatsachen :

– Diese Beziehung kann in Partizipationsprojekte, wo der „Eingeborene" mehr als seine eigenen Leistungen bringt (zum Beispiel beim Mitarbeiten an die Transkriptionen, Übersetzungen, Glossen der Daten), den Informanten mehr oder weniger als einen *Partner* des Forschungsprojektes anerkennen (und nicht nur als ein „*Objekt*").
– Die *finanzielle Bezahlung* kann weniger problematisch sein für (manchmal sogar durch spezialisierte Einrichtungen) rekrutierte Informanten im Rahmen eines ausdrücklichen Vertrags ; sie kann auf dem Terrain problematischer sein, wo sie voraussetzt, nicht nur die möglichen Informanten aber auch die Forscher, die dazu Zugang haben können, in Konkurrenz zu stellen (es ist zum Beispiel der Fall für Sprachwissenschaftler aus Universitäten, die mit weniger Mitteln ausgestattet sind, gegenüber Forschern aus Universitäten, die besser ausgestatttet sind - und daher von den Informanten begünstigt werden können oder bei ihnen schwer zu befriedigende Ansprüche erzeugen können). Die Praktiken der Anthropologen und der Sprachwissenschafler unterscheiden sich in diesem Punkt. Im Fall einer teilnehmenden Beobachtung kann es für einen Anthropologen heikel sein, die Personen zu entlohnen, die ihm die Informationen liefern, mit dem Risiko, ein Mehrangebot des Informationskostens zur Folge zu haben. Dagegen ist die Bezahlung des Sprechers und/oder des Übersetzers, der mehrere Stunden täglich mit dem Sprachwissenschaftler verbringt, eine gerechte Entschädigung für eine richtige Arbeit und behindert nicht zwangsläufig die Vertrauensbeziehung, die zwischen die zwei Personen zustande kommen konnte.
– Die finanzielle Bezahlung ist ein Beispiel unter den anderen der „*Rückkehr*" (oder der Entschädigung, des Lohnes...), die sich für die Terrainforschungen immer auf eine mehr oder weniger implizite Weise macht, im Laufe des Alltags und der Verhandlung der gegenseitigen Verhältnisse.

3.3.2 AUFNAHMEEINRICHTUNG

Die Wahl der Einrichtung der Korporaaufnahme hat Folgen auf die Art und Weise, wie die betroffenen Personen behandelt werden, wie man ihre Einwilligung gewinnt, wie über die Akzeptanz oder über die Akzeptabilität verhandelt wird.

Wir wollen hier einige Aspekte erötern, die sich als relevant erweisen können, von der Wahl der Rahmen der Aufnahme bis zu den Aufnahmemodalitäten.

RAHMEN DER AUFNAHME

Definitionsgemäß ist es unmöglich, *alles* aufzunehmen und die Forscher sind gezwungen, Entscheidungen zu treffen. Diese hängen von dem gezielten Forschungsobjekt, von den technischen Zwängen ab (zum Beispiel ist es schwer, nachts mit dem Video oder mit dem Audio in sehr lauten Plätzen aufzunehmen), und auch von dem Respekt der aufgenommenen Personen.

Was besonders eine Rolle spielt :

- die Wahl der *Weile* der Aufnahme : es handelt sich darum, ein Gleichgewicht zwischen die für die Forscher interessanten Momente und den Respekt des Privatlebens des Befragten zu finden ;
- die Wahl der *Tätigkeiten*, die man aufnehmen wird : diese können eher öffentlich und sozial sein oder intim und privat ;
- die Wahl des *Ortes*, wo man aufnehmen wird : da gibt es auch eine Spannung zwischen öffentlichen Orten, die vom Alltag der Personen gelöst sind, und intimen Orten : das *Labor* ist ein vom Lebensraum der Informanten völlig gelöster Ort – und eben deswegen vermeiden es die Forscher, die über die gelegenen sozialen Sitten arbeiten wollen ; der *Wohnort* der Personen ist ihr Lebensort, der selbst aus „öffentlicheren“ oder “intimeren“ Orten besteht (ein Essen, das im Esszimmer, in der Küche oder im Bett genommen wird, hat nicht denselben Inhalt, so wie ein Gespräch im Wohnzimmer oder um den Küchentisch) : die *Arbeitsräume* sind auch, wenn auf eine andere Weise, durch einzuhaltende Vertraulichkeitsfragen strukturiert ; ihre Nichtbeachtung kann für die gesammelten Daten vielleicht eine Vertraulichkeitspflicht mit sich bringen, die die Unmöglichkeit ihrer Auswertung bedeutet (vgl 3.4) ; die heiligen und/oder Tabus unterwürfigen *religiösen Räume* müssen auch respektiert werden. Im allgemeinen ist eine gute Kenntnis des Ortes und seiner geographischen und sozialen Organisation notwendig, bevor man jede Aufnahme (Bild oder Ton) plant.

Das zu findende Gleichgewicht steht also zwischen dem kontextuellen und natürlichen Aspekt der aufgenommenen Daten im Rahmen und der Schaulust - die Wahl der aufzunehmenden Momente kann wichtige Folgen auf den Rest der Forschung haben (auf die Erlaubnisse, um die Daten auszunutzen, und auf das Widerrufsrecht *post hoc* der Subjekte).

AUFNAHMEMODALITÄTEN

Die Aufnahmemodalitäten spielen oft eine Rolle in die Wahl der aufzunehmenden Kontexte (vgl oben), der gezielten Tätigkeiten und in die Akzeptanz- oder

Widerstandsmodalitäten der betroffenen Personen. Verschiedene technische Aspekte können in die Akzeptabilität durch die aufgenommenen Personen eine Rolle spielen :

- Die Tatsache, dass die Aufnahme mit *Audio-* oder *Video*geräten gemacht wird : für bestimmte Tätigkeiten können die betroffenen Personen dem Video - für eindringender gehalten - das Audio bevorziehen, auch auf die Gefahr hin, dass man in einer zweiten Phase vom Audio zum Video übergeht, wenn die Modalitäten und die Wirkungen der Aufnahme auf die Tätigkeit einmal festgestellt sind.
- Die Tatsache, dass die Aufnahme von dem *anwesenden Forscher*, von *Technikern* oder durch eine *vorher aufgestellte Einrichtung*, die in Anwesenheit des Forschers funktioniert, gemacht wird, hat Folgen auf ihre Akzeptanz : selbst wenn die Kamera oder das Mikrofon oft als „Prothesen" oder Verlängerungen des Forschers angesehen werden (zum Beispiel, wenn die Teilnehmer sie direkt ansprechen), kann die Anwesenheit des Forschers von einigen Teilnehmern bevorzugt werden.
- Die Tatsache, dass die Aufnahme von dem *Forscher* oder von den *Teilnehmern selbst* gemacht wird : einerseits kann das Delegieren der Aufnahme an die Teinehmer als eine Kontrolleform von ihnen auf das Aufgenommene erscheinen : andererseits kann dieses Delegieren abgelehnt werden, als eine zu weite Mitarbeitsform, die den Teilnehmer von der Tätigkeit ablenkt.
- Die Tatsache, dass die Aufnahme durch eine *auffallende oder diskrete*, sogar versteckte *Einrichtung* gemacht wird : es gibt manche Debatten über den Gebrauch eines versteckten Microfons und über seine Folgen auf die möglichen Beziehungen mit den Teilnehmern (Mitchell, 1991, Mondada, erscheint demnächst, Welland & Pugsley, 2002) ; außerdem kann der Gebrauch einer auffälligen Einrichtung wie ein Transparenzbeweis oder ebenso wie eine Störung angesehen werden, selbst wenn die Teilnehmer Bescheid wissen. Oft ermöglicht die Miniaturisierung der Einrichtungen, sie so aufzustellen, dass sie nicht versteckt werden sondern oft einen Teil des Dekors bilden.
- Die Tatsache, dass die Aufnahme von *technischen Mitteln* abhängt, die *auf kurze Sicht ein Eingreifen brauchen* (das sich zum Beispiel auf die Dauer der Batterie oder der Kassette bezieht), verursacht Störungen der Tätigkeit von dem Forscher (oder von den Teilnehmern, die die Kassette ersetzen würden), was andere Einrichtungen mit mehr Autonomie vermeiden (indem sie zum Beispiel auf Festplatten direkt aufnehmen). Es kann wegen der veursachten Störung das Verhalten der Zeugen beeinflussen, insbesondere für bestimmte Aktivitäten (wie einen Patienten zu operieren, eine Therapieberatung zu halten, einen heiklen Vertrag zu diskutieren, in einen Schöpfungsprozess hineinzugeraten).
- Die Tatsache, dass die Aufnahme den Teilnehmern, die sich ihr für einen Augenblick entziehen möchten, *tote Winkel* vorschlägt oder nicht :

der Rahmen und das Feld, die durch eine Kamera begrenzt werden, lassen unbedeckte Zonen ausfolgern, während die vorgestellte Stärke eines Mikrofons oder die Anwendung mehrerer Kameras auf der selben Szene als Überwachungsanlage wirken können, der man sich nicht entziehen kann.

– Die Tatsache, dass man *die Aufnahme* stoppen oder *kürzen* kann, kann als Realisierung der Widerrufsmöglichkeit eingreifen ; die Tatsache, dass das Löschen oder die Abstellung der Aufnahme von den Teilnehmern gemacht werden können, wenn sie es wünschen oder später, oder von anderen gemacht werden müssen, kann den Eindruck einer bestimmten Handlungsfreiheit des Eingreifens auf die Daten geben und setzt andere Vertrauensbeziehungen voraus. Diese Frage - wie manch andere - ist auch mit den technischen Voraussetzungen der Aufnahme und mit dem hohen Entwicklungsstand der Einrichtung gebunden. Man kann in die Wahl der Träger berücksichtigen, die ein unmittelbares Löschen der Daten erlauben oder nicht, oder das an Ort und Stelle anzusehen, was aufgenommen worden ist, erlauben oder nicht.

Diese Betrachtungen (Mondada 2006) zeigen ja die Überschneidung der technischen Fragen und der juristischen Fragen, während der zugleich persönliche, ethische und juristische Respekt der Teilnehmer in den eingesetzten technischen Entscheidungen in die Tat umgesetzt wird.

3.3.3 GENEHMIGUNGSANTRAG UND AUFGEKLÄRTE EINWILLIGUNG

Die Begriffsbestimmung der „aufgeklärten Einwilligung“ und ihre Übertragung in Formen einer sozialen Beziehung (der Kontakt mit den Informanten) und in materialisierte Formen (die ausgetauschten und unterzeichneten Dokumente) sind empfindlich für den Rahmen und die Objekte der Forschung, so wie für die soziokulturellen Bedingungen der Gruppe, wo sie sich abspielt. Wir deuten hier einige Überlegungen an, indem wir von der Begriffsbestimmung der „aufgeklärten Einwilligung“ abgehen, indem wir die Idee des Augenblicks, wo sich diese Fragen stellen, wieder aufnehmen, so wie die Frage der Personen, die man informiert und um die Einwilligung bittet, die Frage der Formen dieser Information, der Objekte, worüber man zu informieren unterscheidet, und der Formen der Einwilligung selbst.

DEFINITION DER „AUFGEKLÄRTEN EINWILLIGUNG“

Es wird oft von Einwilligungformularen gesprochen, die den Informanten vorgelegt werden sollen ; jedoch ist es wichtig, diese Einwilligung von der Information anhängen zu lassen, die den betroffenen Personen gegeben wurde : ohne *Information* hat der *Genehmigungsantrag* weder Objekt noch Sinn. Deswegen spricht man von *aufgeklärter Einwilligung (informed consent),* im Sinne, dass die Akzeptanz der Aufnahme von dem Verstehen der gezielten Zwecke sehr eng abhängt. Auf bestimmten Terrains soll die Schwierigkeit, die Zwecke der Forschung verstehen zu lassen, jedoch den Forscher nicht dazu verleiten, den Genehmigungsantrag zu ignorieren und dieser soll im Einklang mit der Art der Gesellschaft, wo sie sich abspielt, formuliert werden (wie kann man sich zum Beispiel eine individuelle Einwilligung

vorstellen, die in einer Gesellschaft mit verbalen Tradition unterzeichnet wird, wo das private Recht nichts bedeutet ?).

MOMENT DER INFORMATION UND DES ANTRAGS

Der Genehmigungsantrag hängt von der Art, wie man den aufgenommenen Personen herangekommen ist. Er kann je nach dem Moment, wo er stattfindet, unterschiedlich sein :

- Information und *im voraus vorbereiteter* Antrag, der während eines Aufenthaltes auf dem Terrain gemacht wurde und von der Gegenkenntnis- und Vertrauensbeziehung mit dem Forscher abhängt,
- Information und *genau vor* der Aufnahme gemachter Antrag,
- Information und *genau nach* der Aufnahme gemachter Antrag,
- Information, mündlicher Antrag *vor* der Aufnahme und schriftlicher Antrag *nach* der Aufnahme (mit Widerrufsmöglichkeit).

Die Information ist reichhaltiger, wenn der Forscher lange auf dem Terrain bleibt : sie ist begrenzter, wenn der Genehmigungsantrag kurz vor oder nach der Aufnahme stattfindet, ohne andere Kontaktform zwischen den Forschern und den Befragten.

Wann die Information und der Genehmigungsantrag stattfinden, kann in Beziehung mit den geplanten Wirkungen auf die Strukturierung der aufgenommenen Tätigkeit bestimmt werden : das Moment der Information und des Genehmingungsantrags wird oft so bestimmt, dass es die Tätigkeit nicht stört, was die Teilnehmer betrifft (ein Genehmigungsantrag an einen Kunden im Moment des Verkaufs kann zum Beispiel eine mögliche Störung des Verkaufs für den Verkäufer verursachen und kann also dem Forscher, der diese Forschung dokumentieren möchte, verweigert werden), oder was die Forscher betrifft (ein Genehmigungsantrag am Anfang eines Gesprächs kann zum Beispiel die Organisation des sequentiellen Verlaufes dieses Anfangs verändern).

Wenn die Information und der Antrag *nach* der Aufnahme stattfinden, kann die Information als eine „Enthüllung", eine „Offenbarung" erscheinen, die die Aufnahme *nachträglich* als „Verstellung" bezeichnet : dies kann zu Wiederkategorisierungen der Teilnehmer und der Tätigkeiten führen (derjenige, der sich als ein Tourist vorgestellt hatte, der sich in der Stadt verlaufen hat und sich nach dem Weg erkundigt, wird ein Forscher, der über die Raumbeschreibungen in den Erkundigungen nach dem Weg arbeitet) (Mondada erscheint demnächst). Dazu ist diese Technik für manche Forschungsterrains nicht denkbar. So sind diese Verheimlichungsfälle in einigen Gemeinschaften besonders unangebracht und sie schaden der gesamten wissenschaftlichen Gemeinschaft und den folgenden Forschern wirklich sehr.

STATUS DES ANTRAGSSTELLERS

Selbst wenn der Forscher derjenige ist, der informiert und nach der Aufnahmegenehmigung fragt, sind verschiedene Möglichkeiten denkbar :

- Der klassischste Fall ist der *des Forschers*, der sich um die Information und um den Genehmigungsantrag kümmert.

– Oft schickt jedoch der Forscher *Studenten* oder *Mitarbeiter* auf das Terrain, die so viele Sprecher des Projekts sind.
– In einigen Fällen ist es denkbar, dass die Teilnehmer selbst *die Sprecher des Projekts* werden : es ist der Fall, wenn der Forscher einen Teilnehmer darum bittet, andere Teilnehmer zu informieren (zum Beispiel den Gastgeber, der bei sich Freunde zu einem Essen einlädt, das aufgenommen wird ; der Verkäufer, der seine Kunden darum bittet, sich aufnehmen zu lassen ; der Lehrer, der seine Schüler oder Studenten um die Genehmigung bittet, usw.). Diese Delegation ist eine der Formen der Zusammenarbeit zwischen Befragten und Forschern auf dem Terrain ; sie kann aber zu Missverständnissen und Schwierigkeiten führen.

Die Erlaubnis kann ebenso die Unterzeichner selbst betreffen oder Personen, die von ihnen abhängen (Untergebene, Kinder, Studenten, usw.). In diesem letzten Fall ist es wichtig zu berücksichtigen, dass *Erlaubnis* nicht immer mit *Einwilligung* übereinstimmt. In den Gesellschaften ohne Privatrecht sind oft die Meinung und die Erlaubnis der gesamten Gruppe oder einiger ihrer (politischen oder religiösen) Verantwortlichen unvermeidlich.

WAS HEIßT *INFORMIEREN* ?

Mitten in der aufgeklärten Einwilligung ist es verlangt, die aufgenommenen Teilnehmer zu informieren. Sobald man dieses Verlangen berücksichtigt, tauchen aber Fragen plötzlich auf. Was ist „informieren“ ? „Worüber“ informieren ? Unter welchen Bedingungen kann man sagen, dass diese Information ihrem Empfänger einen „aufgeklärten“ Status hergibt ?

Der Begriff „Information“ kann auf eine einfache Botschafts- und Inhaltsübertragung deuten ; er beseitigt letzlich die Prozesse, die Kontexte und die Belanglosigkeiten, die diese Kommunikationsaktivität charakterisieren, durch die ein Forscher seinen Partnern auf dem Terrain das Objekt seiner Forschung erklärt. Sobald es mit den Worten eines Aktivitätstyps nachgedacht wird, stellt die „Information“ der Befragten eine Menge zu lösender Probleme :

– die *Anpassung an den Empfänger* : damit die Erklärung des Forschungsprojekts verstanden und geteilt wird, muss sie an die Kompetenzen, an das Sprach- und Auffassungsgabeniveau des Empfängers angepasst werden, und diese Anpassung betrifft auch den Forschungskontext und die Forschungsmodalitäten beim Berücksichtigen der Anpassung zwischen dem, was die Partner auf dem Terrain geschehen sehen und den gegebenen Erklärungen ;
– die *Explizitheit der Finalitäten* der Forschung soll dieser nicht *schaden* : dies stellt die Frage des zu findenden Gleichgewichts zwischen der Durchsichtigkeit der Forschung und den Verwandlungen, die sie auf das Verhalten der Teilnehmer herbeiführen kann ;
– die Erklärung des Forschungsprojekts kann auf *verschiedene Allgemeinheitniveaus* gemacht werden (von „es ist eine Forschung über die

Sprechweisen der Leute“ bis „es ist eine Forschung über die Frequenz und die Kontexte der nicht obligatorischen Bindung im Französisch“).

Die Information der Befragten besteht nicht nur aus Erklärungen über das wissenschaftliche Projekt sondern auch aus genauen Informationen, die zum Beispiel folgendes betreffen :

- die *Verantwortlichen* der Forschung und ihre institutionelle Mitgliedschaft, so wie diejenigen, die finanzieren,
- eine *Kontakt*adresse,
- die Personen, die *Zugang zu den Daten haben werden* und über diesen Daten arbeiten werden,
- *wie* die Befragten gewählt wurden und zu welcher Population sie gehören,
- wie die Daten anonymisiert werden,
- die Tatsache, dass die Daten nach *bestimmten Konventionen* transkribiert werden (Möglichkeit, ein Beispiel zu geben),
- wie die Daten nach dem Ende der Forschung *archiviert* werden (Bewahrung oder Zerstörung am Ende der Forschung, Bewahrung bei welchem Garanten, Modalitäten einer eventuellen Wiederverwendung, Übertragung auf andere Forscher),
- die *Zugangsmodalitäten* zu den Informationen, die das Projekt betreffen und ganz besonders die Daten/Analysen betreffen, die sich auf die Person beziehen (Möglichkeit eines Zugangs zu den Dateien und Informationen, die die Person ganz besonders betreffen),
- die Rechte der Person, vor allem das *Widerrufs*recht,
- die eventuellen *Risiken* so wie die positiven, moralischen oder materiellen Nebenwirkungen der Forschung.

Auch die Informationsmodalitäten können nach der Kultur der Empfänger variieren, insbesonders :

- die Information kann mündlich gegeben werden : individuell in *üblichen Gesprächen*, kollektiv in *Informationsversammlungen*...
- sie kann schriftlich (durch eine Broschüre, einen Faltprospekt...) oder per E-mail gemacht werden.

Im Rahmen einer schriftlichen Kultur ist es empfohlen, einen Text zu überlassen ; auch kann der Hinweis auf eine Webseite, wo man der Entwicklung des Projekts folgen kann, nützlich sein (eventuell mit besonderen Zugangsweisen).

DER GEGENSTAND DES GENEHMIGUNGSANTRAGS

Erst nach dieser Informationsphase kann der Genehmigungsantrag für das Sammeln der Daten stattfinden. Die Frage wird gestellt, wie man den Gegenstand dieses Antrags umgrenzen kann.

Die Genehmigung betrifft ja folgende Aspekte, die Wechselwirkungen haben können und sich überlagern können :

- die *Aktionen*, die von den Forschern im Rahmen des Projekts gemacht wurden : die Aufnahme, die Vorbereitung des Korpus (Transkription, Übersetzung, Anmerkungen, usw.), die Bedingungen der Archivierung

(Ort des Depots, vorgesehene Bewahrungsdauer, haftende Institutionen...), die Analyse im Rahmen der angekündigten Zwecke, den Gebrauch gekürzter oder ungekürzter Daten, die Verbreitung der Ergebnisse der Analyse, die Bewahrung/Zerstörung der Daten am Ende der Forschung ;

- die *Formate* und die Bedingungen der Aufnahme : audio/video, mit mehreren Kameras/Mikrofonen, in Momenten, die von den Befragten bekannt oder nicht bekannt sind, gut umgrenzt oder von langer Dauer ; jede technische Wahl, die in der Art, wie die Person in den Daten stehen wird, eine Rolle spielt, kann erläutert und sogar verhandelt werden ;
- die *Verbreitung*sbedingungen der Daten und der Ergebnisse : unter einer ungekürzten oder lückenhaften Form (kurze Auszüge, deren maximale Länge vorgesehen werden kann), unter nur textueller Form (Transkripte) oder audiovisueller (zum Beispiel in Powerpoint Dokumenten) ;
- die Verbreitungsrahmen der Daten und der Ergebnisse : Forschungskontexte (Workshops, Kolloquien, Kongresse), Rahmen des Hochschulwesens, breitere Ausbildungs- und Popularisierungsrahmen, Rahmen, die mit dem Terrain gebunden sind (man muss zum Beispiel explizit um die Erlaubnis bitten, die Daten wiederzuverwenden, wenn man im Rahmen einer Ausbildung ist, die in der selben Einrichtung stattfindet, als die, wo die Daten gesammelt wurden - wo sie sich als sehr empfindlich erweisen können) ;
- weite Verbreitungskontexte : in Form einer CD oder auf einer Webseite.

Die Explizitheit dieser Kontexte überdeckt sich mit der der Aktivitäten, in denen die Daten verbraucht werden ; in beiden Fällen ist der Einsatz derjenige der Personen, die im Rahmen dieser Aktivitäten Zugang zu den Daten haben werden. Man kann die Verbreitungsrahmen, die einer bestimmten Kontrolle des Forschers unterworfen sind (zum Beispiel durch Übereinkünfte) von denen unterscheiden, die definitionsgemäß unkontrollierbar sind (zum Beispiel eine Webseite).

Denkbar ist, dem Befragten die Möglichkeit zu überlassen, persönliche Zwänge hinzufügen ; diese Eventualität stellt jedoch das Doppelproblem ihrer Gesetzlichkeit und ihrer Interpretierbarkeit. Eines der Hauptprobleme für den Genehmigungsantrag - wie für die Information - betrifft die immer mögliche Entwicklung der Forschungszwecke, die am Anfang nicht ganz festgestellt werden mögen und sich vor allem im Laufe der Arbeit auf dem Terrain und über die Korpora verändern mögen. Deswegen ist es wichtig, die Zwecke ziemlich generell zu formulieren, um eventuelle Zwecksentwicklungen aufzunehmen, die im Laufe der Forschungsarbeit auftauchen können. Jede Zwecksänderung muss dagegen Gegenstand von einem neuen Antrag sein (vgl unten).

DIE FORMEN DER GENEHMIGUNG

Der Genehmigungsantrag kann verschiedene Formen annehmen, die auch vom soziokulturellen Rahmen der Forschung abhängen : so ist es zum Beispiel nur in den Schriftkulturen sinnlos, die Unterschirft des Befragten zu verlangen, wo dieses Verfahren einen Sinn hat, nicht erschrickt und mit keinen anderen Praktiken gebunden ist, mit denen sie verwechselt werden könnte (wie die Unterschrift von Bankanweisungen).

Man kann also die Formen des Antrags nach dem Träger differenzieren, unter denen sie eingetragen sind :

- geschriebener und unterzeichneter Antrag,
- mündlicher Antrag,
- möglich und nützlich ist vorherzusehen, dass die mündliche Genehmigung selbst unter einer Audio- oder Videoform aufgenommen wird, was erlaubt, ihre Rückverfolgbarkeit zu sichern : es ist die zu begünstigende Lösung bei der Arbeit in Gesellschaften mit mündlicher Tradition, indem man je nach den Bedürfnissen den von den Sprachpraktiken der betroffenen Geinschaft erforderlichen Formalitätsgrad und die Wahl der Sprache achtet (zum Beispiel eine individuelle Aufnahme mit dem Sprecher für eine punktuelle Erlaubnis oder eine aufgenommene Erlaubnis auf einer formelleren Versammlung mit den Behörden).

Im Fall des schriftlichen Antrags kann dieser verschiedene Formen annehmen - in einem vorgeformten Text (Formular) :

- Ein kompakter Text, der die verschiedenen Aspekte des Gehnemigungsantrags zusammenfasst und eine globale Zustimmung (oder Ablehnung) verlangt.
- Ein Text mit anzukreuzenden Feldern, also mit *Wahlmöglichkeiten* : diese Form ist der ersten überlegen, indem sie echte Wahlmöglichkeiten für den Befragten materialisiert, also ihm partielle Ablehnungen ermöglicht (er kann zum Beispiel die Audioaufnahme annehmen aber die Videoaufnahme ablehnen), sogar zusätzliche Zwänge ermöglicht (er kann zum Beipiel die Anonymisierung der Video und Audioaufnahme verlangen). Die Frage ist also, wie die Alternativen formuliert werden, so dass sie nicht redundant sind und auch weder zu kompliziert noch zu lang zu bearbeiten für den Befragten sind.

Ein Problem mag bei kollektiven Anträgen entstehen, wenn Gruppen betroffen sind (zum Beispiel im Rahmen von Aufnahmen von Versammlungen) : wenn zu viele Alternativen zur Auswahl der Teilnehmer gelassen werden, können die Antworten zu gegensätzlichen Ergebnissen führen, wo kein gemeisamer Nenner auftaucht ; so stellen die Gruppenanträge andere Probleme und Zwänge als die individuellen Anträge.

3.3.4 *NACH DER FORSCHUNG : RÜCKKEHREN, DEBRIEFINGS*

Die Vorbereitung auf dem Terrain wird oft betont, es ist aber auch wichtig, die Abfahrt und die Rückkehr aufs Terrain vorherzusehen. Es ist von wissenschaftlicher und ethisch-juristischer Wichtigkeit : die Rückkehr auf das Terrain kann sich irgendwann für eine Nachprüfung, eine zusätzliche Information, eine neue Kontaktaufnahme mit den Informanten als nötig erweisen. Wenn die Abfahrt vom Terrain schlecht verlaufen ist, wird die Rückkehr unmöglich sein. Die Anwesenheit auf dem Terrain verursacht amsonsten nicht nur Vertrauensbeziehungen sondern auch Erwartungen, die dauerhafte Verpflichtungen schaffen : vom Terrain einfach verschwinden, nachdem man ein Untertauchen praktiziert hat, das oft enge Beziehungen mit den Teilnehmern einführt und sie um Hilfe und Leistungen bittet, kann große Enttäuschungen hervorbringen. Wenn man Wissen, Antworten, Korpora auf dem Terrain „entnommen" hat, heißt es also zu wissen, wie man den Leuten etwas „zurückgeben" kann, ohne die die Forschung unmöglich gewesen wäre (vgl auch die Bezahlungsprobleme oben). Es ist zum Beispiel jetzt unmöglich, auf bestimmten Terrains zu arbeiten (im Falle der bedrohten Sprachen), ohne mit einer Rückgabe an den Sprecher und an die Gemeinschaft, sogar mit einer Zusage des Forschers zu rechnen, egal was sie ist (Implikation in Erziehungsprojekte, Schreibprojekte, usw.)[7].

Es muss auch darauf hingewiesen werden, dass die „Feedbacks", die „Debriefings", die Erfahrungsrückkehre schon während der Arbeit auf dem Terrain gemacht werden können, zum Beispiel unter der Form von Berichten über partielle Ergebnisse. Die Unterscheidung zwischen „während des Terrains" und „nach dem Terrain" kann so relativisiert werden.

Mehrere Praktiktypen sind denkbar, um eine „Rückkehr" bei den befragten Bevölkerungen zu sichern. Wir zählen einige davon auf, die von der Darstellung der Ergebnisse gehen, die am nächsten vom akademischen Kontext ist, bis zur Formulierung von Wissen und Know-how, die am nächsten vom Terrain ist. Zweifellos liegen die Wahlmöglichkeiten der „Terrainpolitik" in der Abschätzung des Abstands zwischen der „Rückkehr" und der Akademie oder dem Terrain :

- Darstellung der Ergebnisse am Ende des Prokekts : die Formulierung der Ergebnisse kann mehr oder weniger popularisiert werden, mehr oder weniger mit den Besorgnissen der Befragten verwandt sein, die Darstellung der Ergebnisse kann zum Beispiel Elemente für die *Transkription* und die Analyse von Transkripten aufweisen : die Teilnehmer reagieren auf sehr verschiedener Weise (sie sind manchmal erstaunt, manchmal schockiert) auf die Vorstellung ihrer Stimme.
- *Empowerment* Verfahren (Rückgabe) : es besteht darin, die „Rückkehr" nicht nur als „Information" zu denken sondern auch als eine Wissens-

[7] Bericht der UNESCO (2001), *Language vitality and Endangerment* : « Any research in endangered language communities must be reciprocal and collaborative. Reciprocity here entails researchers not only offering their services as a quid pro quo for what they receive from the speech community, but being more actively involved with the community in designing, implementing, and evaluating their research projects ».

und Know-howseinlage für die Gemeinschaft der Befragten (Cameron et *al.*, 1991) : es ist also denkbar, nicht nur Analysen vorzustellen aber auch den Teilnehmern zu erlauben, Daten weiterzu*sammeln* und ihre eigenen Daten für ihre eigenen Zwecke zu analysieren, man kann die *Auswirkungen der Analyse* nach dem Wortlaut des Kalenders, der Themen, der Besorgnisse der Täter formulieren, man kann im Rahmen der Kompetenzen des Forschers auf die *Untersuchungsanträge* antworten, die oft von den Gemeinschaften ausgedrückt werden (zum Beispiel Überlegungsateliers über den Übergang zum Schreiben oder über die Übersetzung offizieller Unterlagen, Verwicklung in zweisprachige Erziehungsprogramme), man kann das von der Forschung *erzeugte Wissen* zur Verfügung der Gemeinschaft stellen, indem man es unter eine andere Form als die der traditionellen Universitätsschriften verwirklicht (zum Beispiel unter der Form einer Ausstellung oder anderer abgeleiteter kultureller Produkte), es kann eine auf den Ergebnissen/Methoden der Forschung ruhende *Ausbildung* angeboten werden ; im allgemeinen kann man beabsichtigen, Analysewerkzeuge zu übermitteln, Kompetenzen umzulagern, die auf dem Terrain nützlich sein könnten.

– Die Frage der „Rückkehr" der Daten selbst als Korpora oder Archive kann sich als heikel erweisen : sie kann sich in einigen Fällen aufdrängen (so sollten[8] für die bedrohten Sprachen patrimoniale, an die Gemeinschaft weitergegebene Archive angelegt werden) aber auch vermeidet werden sollen, um die Informanten zu schützen (so könnten im Fall der Forschungen in Unternehmen oder Einrichtungen die gesammelten Daten bestimmte Stufen der Hierarchie interessieren aber Untergebenen schaden). Die Rückkehr der Archive, wenn sie relevant ist, stellt also oft Fragen : begrenzter Zugang der Leute, die in diesen Archiven nachschlagen können, mit Berücksichtigung der Risiken und Vorteile, die die Bereitstellung der Zugangsweisen und -technologien zu den Archiven auf dem Terrain vorbereitet : wenn die Archive so formatiert werden, dass die betroffene Bevölkerung Zugang zu ihnen haben kann, müssen die Technologien den Gebräuchen und Möglichkeiten dieser Bevölkerungen angepasst werden (es ist sinnlos, eine DVD zu machen, wenn keiner einen DVD-Spieler besitzt, oder eine Webseite zu bauen, wenn niemand Zugang zur Informatik hat ; hier stellt sich die Frage, wie man mit der Asymmetrie zwischen der „akademischen Arbeit" und dem „Terrain" zurechtkommt), die Garantie *des Zugangs zu den Veröffentlichungen* stellt ähnliche Fragen wie die des Zugangs zu den Daten, obwohl es oft nicht so schwierig ist.

[8] Es ist sogar nach den Empfehlungen der UNESCO eine Pflicht auf diesem Gebiet.

3.4 ANONYMISIERUNG

Die Möglichkeit oder die Garantie (die wir später relativieren werden), die gesammelten Daten zu anonymisieren, ist für den Schutz des Privatlebens der von der Forschung betroffenen Personen wichtig und für die Gesetzlichkeit der von den Forschern gesammelten Korpora. Die Anonymisierung der Daten ist jedoch weder ein einfaches Verfahren noch eine nicht-problematische Garantie, denn sie lässt viele zugleich technische, wissenschaftliche und soziologische Probleme auftauchen.

Die Anonymisierung der Daten ist eine wichtige Garantie, was die Gesetzlichkeit der Daten und ihren Gebrauch betrifft ; in einigen Fällen kann sie einen Gebrauch der Daten auch ohne vorangehenden Genehmigungsantrag ermöglichen, wenn sie die Nicht-Identifizierung der betroffenen Personen wirklich sichert und wenn außerdem die Daten nicht vom Urheberrecht geschützt sind. Man soll aber mit diesem Punkt vorsichtig sein, indem man alle Begrenzungen und Schwierigkeiten berücksichtigt, auf die man bei der Anonymisierung stößt (vgl unten).

3.4.1 BEGRIFFSBESTIMMUNG

Obwohl es oft von Anonymisierung die Rede ist, ist die gesetzliche Frage, die sich stellt, die der *Unmöglichkeit, die Personen zu identifizieren* : der Einsatz ist, dass die betroffenen Personen auf der Basis von den gesammelten Daten und ihren Vorstellungsformen (zum Beispiel Transkript) nicht identifiziert werden können. Die Identifizierungsverfahren werden durch die aktuellen Technologien völlig verändert, die die Lagerung und die Verbreitung der Daten erleichtern, aber auch starke Werkzeuge für die Bearbeitung der Informationen (Sortieren, Gegenüberstellungen, gekreuzte Ersuchen...) anbieten.

Diese Betrachtungen betreffen :

- alles, was erlaubt, eine Person *direkt* zu identifizieren : dem Sprecher oder einem Dritten und seiner Privatsphäre entsprechend, auf der Basis von den Äußerungen des Sprechers, wie seine Stimme oder sein Aussehen ;
- alles, was ihm schaden kann ;
- alles, was indirekt durch Informationsgegenüberstellung ermöglichen kann, bis zum betroffenen Sprecher zu gelangen.

Die Prozesse, die jeden Bezug oder jede Äußerung streichen, werden „Anonymisierungsprozesse“ der Daten genannt.

3.4.2 BETROFFENE DATEN

Die Anonymisierung betrifft nicht nur die Aufnahmen oder die Transkripte aber auch eine Reihe von Daten, die in den Korpora enthalten sind und sich je nach diversen Trägern und Formaten, von denen die Anonymisierungstechnicken abhängen werden, differenzieren :

- die Video- Primärdaten,
- die Audio- Primärdaten,
- die textuellen Primärdaten : offizielle ode unoffizielle Dokumente, die auf dem Terrain gesammelt wurden,

– die Sekundärdaten : Transkripte, Notizen auf dem Terrain, Metadaten, Analysen, ethnographische Beschreibungen,
– die visuellen Sekundärdaten : *screenshots*, sogar Darstellungen der Stimme (Oszillogramme, Spektrogramme...).

Man wird bemerken, dass bestimmte persönliche Daten der Anonymisierung entgehen : es ist der Fall für öffentliche Männer und Frauen in Beiträgen mit öffentlichem Charakter (zum Beispiel Politiker im Fernsehen), wo sie sich in Kenntnis der Sachlage, was die Ausstrahlung ihres Images betrifft, einschalten und wo ihre Worte selbst als eine öffentliche Rede betrachtet werden. In diesem Fall werden die Worte, wenn sie als „original" angesehen werden, mit einem Toleranzbereich für einen von dem „Zeitgeschehen" bestimmten Zeitraum den Verbreitungszwängen, die das Urheberrecht regeln, unterworfen sein. Sobald diese Beiträge als nicht mehr mit dem Zeitgeschehen gebunden angesehen werden, entziehen sie sich dieser Bezeichnung[9].

3.4.3 WANN SOLL MAN ANONYMISIEREN ?

Verschiedene Momente, wo die Anonymisierung stattfinden kann, sind zu unterscheiden. Je nach den Zwecken der Studie und den Rahmen der Forschung kann man meinen, dass die Anonymisierung so *früh* oder so *spät* wie möglich stattfinden muss. Die erste Lösung vermehrt die Vertraulichkeitsgarantien für die Person, die zweite maximiert die Analysemöglichkeiten für den Forscher. Die Zeitlichkeiten können auch nach der Art der Daten unterschiedlich sein :

– die Anonymisierung der primären originellen Referenzdaten wird vermieden, denn sie könnte die Daten selbst beschädigen ; die so nicht anonymisierten Daten müssen jedoch in einem sicheren Ort aufbewahrt werden,
– die Daten können/müssen/müssen nicht (nach den eingenommenen Politiken) bei ihrer Abgabe für die Bewahrung anonymisiert werden ; die Rolle des Garanten der Institutionen, der die Bewahrung sichert, ist hier in Frage,
– man kann (in einer gut begrenzten Forschungsgruppe, die sichert, dass die Daten nicht außerhalb in Umlauf gesetzt werden) auf nicht anonymisierten Daten arbeiten und jedoch die Anonymisierung jedes Auszuges sichern, der in einem Schriftstück oder in einer mündlichen Darstellung steht,
– die Anonymisierung wird immer für die Kopien gemacht, die für den Umlauf unter Forschern bestimmt sind, die außerhalb des Projekts sind (es ist vor allem der Fall für große Forschungskonsortien oder für Projekte, die wichtige Mannschaftsnetze miteinander verbindet).

3.4.4 WIE SOLL MAN ANONYMISIEREN ?

Die Anonymisierungsverfahren betreffen zugleich die Datenträger und -formate und stellen also technologische Möglichkeiten und Zwänge auf das Spiel ; sie betreffen

[9] Vgl den Artikel 122 des französischen Urheberrechtsgesetzes.

auch symbolische Formen und Äußerungen der Identität der Personen und setzen also Analysefragen aufs Spiel.

FORMEN UND ELEMENTE, DIE DURCH DIE ANONYMISIERUNG BETROFFENEN SIND

Wie wir es sehen werden, ist es schwierig - sogar unmöglich - , eine abgeschlossene Liste der durch die Anonymisierung betroffenen Formen zu machen. Man kann jedoch die Hauptformen betonen :

- Namensformen (Name, Vorname, Spitzname, Abkürzung des Unternehmens...),
- persönliche Daten (Adresse, Telefonnummer, Passnummer, Kontonummer, Alter, Geburtsort...),
- Beruf, Status, Titel,
- soziale Tätigkeiten,
- Verwandtschaft, Netze,
- Bezug auf Orte (Ortsnamen, Einrichtungen, Dienste...),
- Bezug auf typische (physische, kulturelle, ärztliche...) Merkmale der Person, die in seiner identifizierten Umgebung einzigartig oder selten sind,
- physische Merkmale : Stimme, Gesicht, körperliche Merkmale,
- usw.

Das „usw" am Ende dieser Liste betont die Tatsache, dass jedes Element nach den Kontexten der Aufnahme und des Empfangs dieser Aufnahme Information über die Identität der Personen geben kann. Die Identifizierung der durch die Anonymisierung betroffenen Formen setzt also eine soziologische und kulturelle Kompetenz voraus, die den Forscher befähigt, sich die Gebräuche, Kenntnisse und Verbindungen vorzustellen, die es ermöglichen könnten, eine Person auf der Basis einer gegebenen Form zu identifizieren.

ERSATZFORMEN

Wenn die Formen einmal identifiziert sind, die zur Identifizierung der Personen leiten könnten, muss man sie umwandeln, um die Anonymisierungsoperationen zu machen.

Man kann merken lassen, dass die radikalste Anonymisierungsform die einfache *Streichung* der Daten ist - obwohl oft andere Mittel, die die Daten besser bewahren könnten, gesucht werden, um die Anonymisierung zu sichern. Wohlgemerkt kann die Streichung partiell sein (man kann daran denken, die Auszüge zu zestören, die zu viele problematische und vertrauliche Elemente tragen würden, um in diesem Zustand gebraucht werden zu können).

Die im allgemeinen eingenommene Anonymisierungsform geht mit dem Ersetzen von vertraulichen Elementen durch neutrale Formen vor. Diese Formen sind nach den betroffenen technischen Trägern unterschiedlich : wir unterscheiden hier zwischen dem Text, dem Audio und dem Video.

- *Text*

Die betroffenen Texte sind zuerst das Transkript und all seine Erwähnungen in Artikeln, Beispielsammlungen, Kursen, Konferenzen... Andere zu anonymisierende Texte sind die textuellen Primärdaten (auf dem Terrain gesammelte Dokumente). Diese können auch eine textuelle Form oder die Form eines Bildes haben (so ist es der Fall für einen Brief, ein behördliches Dokument, ein unter fotokopierter oder digitalisierter Form aufbewahrtes Manuskript).

Das Prinzip des Austauschens besteht darin, das ersetzte Textstück sichtbar zu machen und so allgemeine Informationen über dieses Stück (die wenigstens seine Dauer betreffen) zu geben.

- *Ersetzen durch eine „Lücke"* : es ist die am wenigstens informative und vor allem die am wenigstens sichtbare Lösung.
- *Ersetzen durch ein Hyperonym* oder eine Abkürzung, wie NN, SN oder KN für Name, Stadtname oder Krankenhausname, usw. Diese Lösung kann informativ bleiben (präzisiert wird die Art der Referenz der anonymisierten Form). Sie ist nützlich, wenn das Austauschen gegen ein Pseudonym (vgl unten hier) unmöglich, schwer oder unwahrscheinlich ist. Diese Lösung setzt die Entwicklung spezifischer Konventionen für das Notieren dieser Hyperonyme voraus, die nicht dieselbe Natur wie der ersetzte Text haben (deswegen wird manchmal der Gebrauch von Großbuchstaben vorgeschlagen, wenn er anderen Gebräuchen der Großbuchstaben nicht widerspricht, die in den Konventionen für die Transkription vorgesehen wurden).
- *Ersetzen durch einen Decknamen* : es ist die meist gebrauchte Lösung, wenigstens für die Personennamen, denn sie ermöglicht eine gute Integration der Ersetzungsform in den Sprachfaden, lenkt die Aufmerksamkeit nicht auf sich, ist wahrscheinlich und behält einige Hinweise der ursprünglichen Form. Dies ist aber nur möglich, wenn die Wahl der Decknamen überlegt ist und folgenden Problemen antwortet : der Deckname wird im selben paradigmatischen Feld ausgesucht wie die ersetzte Form (zum Beispiel „Ahmed" wird eher durch „Mustapha" als durch „Albert" ersetzt, da der Deckname einige Ethnizitätszüge aufzubewahren versucht) ; in einigen Fällen, besonders wenn die aufgenommene Interaktion den Namen treffend macht, wird man darauf aufpassen, die möglichen Konnotationen des Wortes zu bewahren (wenn es z. B. die Grundlage für Scherze oder Witze ist), so wie die Zahl der Silben und bestimmte phonetische und prosodische Merkmale (wenn sie in der Interaktion verwertet werden) ; der Deckname wird so gewählt, dass der ursprüngliche Name nicht wieder hergestellt werden kann (in diesem Sinne muss man vermeiden, einen Decknamen zu wählen, der mit denselben Buchstaben wie der ursprüngliche Name anfängt, selbst wenn er Vorteile für seine Speicherung bietet) ; der Deckname wird so ausgesucht, dass er die Person nicht lächerlich macht (in diesem Sinne sollen die Decknamen vermeidet werden, die auf Merkmale der Person verweisen würden -

zum Beispiel „Herr Dick“) ; die Straßennamen, Telefonnummer, usw. können auf dieselbe Art wie die Personennamen ersetzt werden.

Zu bemerken ist, dass es leichter ist, für Personen als für Städte Decknamen zu finden (man kann den Namen einer Kleinstadt, eines Viertels oder einer Straße finden, aber schwerer den Namen einer Großstadt oder einer Hauptstadt) ; es ist manchaml denkbar aber nicht immer möglich, Decknamen für institutionelle Dienste zu finden (es hat keinen Sinn, „Chirurgieabteilung“ durch „Dermatologieabteilung“ im Fall eines Krankenhauses zu ersetzen). Wenn die Wahl eines Decknamens schwer oder unwahrscheinlich ist, wird man zum Hyperonym greifen.

- *Audio*

– *Ersetzen durch die Stille* ; diese Lösung bietet den Nachteil, dass man das Ersetzen mit einer Pause verwechseln kann,
– *Erstezen durch ein Tonzeichen oder ein anderes Geräusch*, das sich mit keinem anderen Signal verwechselt, das in der Aufnahme vorkommt,
– *Ersetzen durch das ursprüngliche Signal, das gefiltert und verstellt wird* ; diese Technik wird besonders in den Medien verbraucht, damit die Stimme nicht identifizierbar ist ; wenn sie von Nichtexperten gebraucht wird, kann sie Probleme stellen, was ihre Unumkehrbarkeit betrifft (die Möglichkeit, das ursprüngliche Signal wieder herzustellen).

- *Bild*

Das betroffene Bild ist vor allem das dynamische Bild der Videoaufnahmen. Man kann aber auch an feste Bilder denken, zum Beispiel an die Fotos auf Dokumenten und an Screenshots in den Transkripten. Man kann auch an die Anonymisierung einer visuellen Darstellung des Tonflusses denken (zum Beispiel in einem Spektrogramm), wenn sie die Aussprache eines Namens oder einer Nummer erkennbar machen könnte.

– für diese Daten ist die *Streichung* unter der Form von Unterbrechungen bei der Montage denkbar ; in diesem Fall wird beraten, die Dauer des geschnittenen Ausschnittes auf das Band zu kennzeichnen und nicht den Eindruck einer Kontinuität zu geben ;
– *Ersetzen durch eine Signalstörung* : durch Verschwommung, durch Pixelisierung ou durch Randabgrenzung des Bildes oder Anwendung anderer Filter (diese Verarbeitung kann *das gesamte Bild oder nur eine Einzelheit* betreffen) ; in diesem Fall ist es technisch schwieriger, wenn diese Einzelheit sich bewegt ;
– Setzung einer schwarzen Binde auf die Augen der Person.

3.4.5 DIE GRENZEN DER ANONYMISIERUNG

Selbst wenn die Anonymisierung eine grundsätzliche Aktion für die Sicherung der gesetzlichen Verbreitung der Daten ist, muss man vorsichtig sein, was die Versprechungen und Garantien betrifft, die den Befragten gemacht wurden und sich auf die Anonymisierung der Daten beziehen.

Die Beschränkungen sind von zwei sehr unterschiedlichen Arten, die erste betrifft die Kontexte, die die Erkennbarkeit der Personen verstärken oder vermindern, die zweite betrifft die Zwänge, die die Anonymisierung auf die Objekte der Forschung selbst lasten lässt.

HERSTELLUNGS- UND UMLAUFSKONTEXTE

Die Anonymisierung wird durch verschiedene Faktoren relativisiert, die entweder bei dem Hervorbringen der Daten – und gemäß den spezifischen Besonderheiten des Geschehens der Aufnahme – oder bei dem Empfang dieser Daten zustande kommen :

- Die Anonymisierung greift zuerst auf einer Serie von Formen, die die wichtigsten Hinweise, die die Identifizierung der Person erlauben, enthalten sollen ; irgendwelche Referenz oder Form kann jedoch nach den Kontexten zur Identifizierung der Person führen und oft auf einer Weise, die auf den ersten Blick für den Forscher nicht auffällt. So kann sich zum Beispiel in der Interaktion die Erwähnung einer seltenen Einzelheit (einer seltenen Pathologie der Person, eines außergewöhnlichen Attributes, eines einzigartigen und in der Gegend der Person bekannten Merkmals...) als bedeutsam für einige erweisen (in einigen Fällen, ohne dass der Forscher dessen bewusst wird).
- Der erkennbare Charakter dieser Einzelheiten hängt entscheidenderweise vom Kontext des Empfangs ab und spezieller von dem Publikum, das in den Korpora nachschlagen wird oder sie zur Kenntnis nehmen wird. So werden die Mitglieder einer Betäubungsabteilung sehr leicht einen ihrer Kollegen auf der Basis von typischen Ausdrücken, spezifischen Untersuchungen oder eigenen Sprech- oder Handelnarten erkennen ; dagegen werden die Fachleute eines anderen Krankenhauses oder erst recht Studenten in Sprachwissenschaften diesen Einzelheiten kaum Beachtung schenken. Auch da hängt aber die Erkennungsmöglichkeit nicht nur von der geographischen oder sozialen Entfernung des Kontexts der Datenaufnahme : die Personen bewegen sich im Raum und in den sozialen Kontexten und der Sohn eines Patienten mag seinen Vater in einer Vorlesung über therapeutische Sprechstunden erkennen. Das Identifizierungswert einer Einzelheit hängt also vom Kontext des Empfangs der Daten ab.
- Nach den Fällen kann der Bezug auf eine Einrichtung oder auf eine Institution die Anonymisierung notwendig oder nicht notwendig machen : der Bezug auf ein bekanntes Schild muss anonymisiert werden, wenn es sich um den Arbeitsort eines Beamten handelt, muss es aber nicht, wenn er als Teil der Landschaft in die Angabe eines Weges erwähnt wird, und muss erneut anonymisiert werden , wenn er in verleumderischen Worten zitiert wird.
- Andere Aspekte sind mit der *Überschneidung* von Informationen gebunden, die aus mehreren Quellen kommen (es kann zum Beispiel die Beziehung zwischen anonymisierten Daten und Metadaten betreffen).

Die Anonymisierung mag wegen anderer Betrachtungen, die mehr mit den Analysepraktiken der Forscher gebunden sind, begrenzt werden.

Das Grundproblem wird durch den eventuellen Widerspruch zwischen Anonymisierung und Verfügbarkeit der Einzelheiten für die Analyse (auf dem Verfügbarkeitsprinzip Mondada, 2003) gestellt. Die Aufnahmen und die Transkripte zielen zwar darauf, zur Verfügbarkeit der zu beobachtenden Einzelheiten zu führen, damit sie durch die Analyse ausgenutzt werden können ; dagegen kann die Anonymisierung einige dieser Einzelheiten nicht verfügbar machen, indem sie sie verwischt oder verwandelt.

Dies kann zum Beispiel der Fall der Anonymisierung durch die Anwendung eines Pieptons für einen Namen sein, der in Überlappung mit einem anderen Sprechturnus ausgesprochen wird, und die Analyse dieser Überlappung unmöglich macht.

Dies kann der Fall der Anonymisierung der Telefonnummern im Notfallanrufen sein, die die Art, wie der Anrufende seine Telefonnummer in einer Stress- oder Aufregungssituation gibt, nicht verfügbar macht und diese Information also entscheidenderweise betreffen kann.

Dies kann der Fall der Anonymisierung der Gesichter auf einem Videoband sein, die die Analyse der Blicke unmöglich macht.

Ähnlicherweise ist die Filterung der Stimme (wie sie in den Medien angewendet wird) nicht denkbar für die meisten sprachwissenschaftlichen Studien, die sich auf die wahren Eigenschaften des Tonsignals stützen.

Deshalb betonen die Forscher die Notwendigkeit und erheben Anspruch auf das Recht, mit nicht anonymisierten Daten zu arbeiten - indem die Sicherheit und die Unzugänglichkeit der Daten garantiert werden - , sie unter dieser Form aufzubewahren, und die Anonymisierung so spät wie möglich zustande kommen zu lassen und auf eine Art, die das berücksichtigt, was für die Analyse treffend ist.

3.5 TRANSKRIPTION

Die Transkription ist eine Praktik, die sich nicht gar nicht auf eine technische Vervielfältigungsübung beschränkt und zahlreiche theoretische und interpretative Einsätze aufnimmt (schon Ochs, 1979). Im Übergang vom Mündlichen zum graphisch-visuellen Schriftlichen werden mehrere Kategorisierungsaktionen ausgeführt, was entweder die linguistischen, optisch in Einheiten geteilten Formen angeht (Blanche-Benvéniste & Jeanjean, 1987 ; Mondada, 2000) oder die Identität der Sprecher selbst (Mondada, 2003). Aus der Sicht des Bild- und Identitätschutzes der befragten und aufgenommenen Personen sollen diese Wirkungen geschätzt werden, um die Überinterpretation, die Stereotypisierung (Jefferson 1996) und die Stigmatisierung der Sprecher und ihrer Art zu sprechen zu vermeiden. Wir werden uns hier auf diese Umsätze der Transkription beschränken ; in dem folgenden Kapitel werden wir einen ganz verschiedenen Aspekt berücksichtigen, den der Standardisierungsfragen der Transkriptionen und ihrer Konventionen.

3.5.1 *DIE ETHNOGRAPHISCHEN BESCHREIBUNGEN*

Die Transkription wird oft von einer kurzen ethnographischen Beschreibung begleitet, die den Kontext skizziert, wo sie gesammelt wurde, wie den Betätigungstyp und die Identität der Teilnehmer. Diese Beschreibung, die Elemente der Korpusmetadaten integriert, kann mehrere Wirkungen auf die Lesart (oder auf die Aufnahme eines mündlichen Vortrags) haben :

- sie kann Informationen enthalten, die die Identifizierung der Person ermöglichen und im Widerspruch zu den Anonymisierungsprinzipien stehen.
- sie kann Hinweise enthalten, die die Lesart oder die Interpretation der Daten beeinflussen. Indem sie die Zugehörigkeit zu einer Kategorie oder einem anderen sinnvollen Aspekt des Befragten restaurieren, können diese Hinweise ein besonderes Bild der Tätigkeit und der Sprecher geben.
- sie kann insbesondere Anspielungen enthalten, Folgerungen erlauben, die bestimmte Stereotypen verstärken (oder sie sogar benutzen, um komische Effekte zu bewirken, um das Publikum zu gewinnen - es ist in den mündlichen Vorträgen häufig).

Diese Bemerkungen betreffen nicht nur die Beschreibung der Daten, sondern auch die Namen der Korpora, die manchmal vertrauliche Elemente integrieren können. In dieser Hinsicht sollten die Namen der betroffenen Täter in den Namen des Korpus vermieden werden, selbst wenn es oft eine Erinnerungsfunktion besitzt.

3.5.2 *DIE IDENTIFIZIERUNG DER SPRECHER*

Das Transkript soll die Ergebnisse der Anonymisierung integrieren. Da wo das Versehen mit Anmerkungen eine Codierung der Sprechturnusse vorhersieht, können Teile des Transkripts deutlichen, auf verschiedene Weisen identifizierten Sprechern zugeschrieben werden. Der Gebrauch von Decknamen ist ziemlich verbreitet, andere Möglichkeiten, die dennoch unterschiedliche Wirkungen auf die Interpretation des ihnen folgenden Textes haben, sind aber denkbar. Jede Entscheidung auf diesem Gebiet stellt das Problem der Art, wie der Sprecher behandelt wird. Zum Beispiel :

- A, B, C…: die am wenigstens konnotierte Lösung, die aber, indem sie die alphabetische Reihenfolge einnimmt, die Sprecher als erster, zweiter, dritter... ordnet.
- S1, S2, S3... (für Schüler): eine Wahl, die die Personen innerhalb einer mit einer einzigen Kategorie bezeichneten Klasse homogen macht. Dasselbe gilt für SP1, SP2, SP3, wo SP auf den Sprecher verweist : wenn der Sprachwissenschaftler finden kann, dass alle Sprecher gleich sind und dass die sozialen Akteure ihn vor allem als Sprechende interessieren, dann nehmen sie zuerst unter anderen Kategorien an die laufende Tätigkeit teil, es sei denn als Forscher/Befragter, Vater/Sohn, Arzt/Patient, usw.

– M, F (für Mann und Frau) : diesmal wird wieder die Kategorie des Geschlechts/Genus jeder anderen Kategorie bevorzugt und so die allgemeine Richtigkeit dieser Kategorie für die Verständlichkeit der laufenden Tätigkeiten postuliert.

Diese Bemerkungen laden dazu ein, sich zu fragen, zu welchen interpretativen Wirkungen die Wahlmöglichkeiten der Identifizierungsmerkmale führen. In dieser Hinsicht soll man sich fragen, welches die treffenden Identifizierungsmerkmale für die Teilnehmer sind - besonders in analytischen Schritten, die sich (wie die Konversationsanalyse) um die Perspektive der Teilnehmer kümmern. Deshalb können die alternativen Lösungen folgende sein :

– EVA, MAR, ROB, AND...: Angabe der drei ersten Buchstaben der Decknamen, ob Vorname oder Name - nach der Grundstimmung des Gesprächs,
– ANR/TEL für Anrufender/Telefonist oder ARZ/PAT für Arzt/Patient, oder auch BEF/INT für Befragter/Interviewer, wenn die institutionnelle Tätigkeit durch kategoriale Paare dieser Art geregelt wird. Über diese Fragen kann man auf H. Sacks Betrachtungen über die Kategorisierungen der Personen und über die Richtigkeit der Kategorien nach den laufenden Tätigkeit und Zusammenhang verweisen (eine Person, die in einem Zusammenhang Arzt ist, kann in einem anderen Zusammenhang Vater sein ; wie er identifiziert wird, hängt also von der laufenden Tätigkeit ab) (Sacks, 1972, 1992).

Der Begriff des Privatlebens und der Intimität hat nicht den selben Wert in allen Gesellschaften, der Forscher sollte sich also über die Wünsche der Sprecher in bezug auf die Anonymisierung der Daten erkundigen. In bestimmten Gemeinschaften wird die Tatsache, dass man die Namen der Personen nicht erwähnt, als ein Mangel an Respekt für den Autor des Gesprächs oder für die teilnehmenden Personen angesehen, während die Erwähnung in anderen eine Einschränkung des Privatlebens ist. In diesem Punkt scheint es zum Beispiel grosse Unterschiede zu geben zwischen bestimmten Terrains in Afrika (wo die Sprecher wünschen, erwähnt zu werden) und anderen wie denen in Amazonien, besonders in Französisch-Guayana.

3.5.3 EINSÄTZE

Beim Transkribieren werden ständig Entscheidungen getroffen in bezug auf die Art, wie man die Sprecher und ihre Ausdrucksweisen darstellt. So schleichen sich sofort die Analyse - und manchaml das Urteil - in die Transkriptionspraktik hinein. Wir werden einige Einsätze der in der Transkription gemachten Wahlen unterstreichen.

(ORTHO)GRAPHISCHE EINSÄTZE

Seit mehr als zwanzig Jahren gab es viele Diskussionen über den Gebrauch der Standardrechtschreibung, der angepassten Rechtschreibung und der internationalen Lautschrift in den Transkripten (vgl 2.1.3). Die phonetischen Transkriptionen (internationale Lautschrift oder sonstige) können lauter von den Fachleuten gelesen werden, und nur für kurze Texte. So haben viele europäische Sprachwissenschaftler die Standardrechtschreibung gewählt, um grosse Korpora zu lesen, aber sie schlagen

auch vor, andere Anmerkungen übereinander legen zu können, wenn es sich darum handelt, bestimmte Phänomene ausführlicher zu beobachten.

Im Gegenteil dazu kann in einigen Fächern wie die Phonetik eine orthographische Transkription nicht treffend sein (zum Beispiel für die Transkription von Logatomen, von Pseudowörtern, usw.).

Das geschriebene Transkript der Sprache erstaunt oft die Sprecher und kann ihnen sogar erheblich missfallen. Es kommt vor, dass sie das Bild ihrer durch das Transkript weitergegebenen Sprache ablehnen, dass sie den Forscher verleugnen und seine Arbeit ablehnen.

DIE DARSTELLUNG DER EXOSPRACHLICHEN SPRACHE

Die Entscheidung, bestimmte Auszüge oder nur die von bestimmten Sprechern eher als andere mit der internationalen Lautschrift zu transkribieren, erlaubt sicher eine grössere Genauigkeit in der Darstellung der Einzelheiten ihrer Sprache, es kann aber auch unbeherrschte Asymmetriewirkungen verursachen.

So kann die Anwendung von der internationalen Lautschrift und von der angepassten Rechtschreibung „Nicht-Eingeborenen" gegenüber Stigmatisierungs- und Asymmetriewirkungen verursachen - wenn diese anders als die „Eingeborenen" dargestellt werden (diese mit Standardanmerkungen, und die „Nicht-Eingeborenen" mit speziellen Schreibungen, die nicht nur den Unterschied, sondern auch die „Abnormalität", die „Abnormativität" hervorheben).

Vergleichbarerweise führt das explizite Notieren der verschiedenen Sprecharten der Sprecher (Differenzierung durch Zeichensätze, Formatvorlagen, Alphabete, die den verschiedenen im zweisprachigen Gespräch gebrauchten Sprachen spezifisch sind, oder der Zwischensprache des Lernenden in einem exosprachlichen Gespräch) durch stillschweigende Übereinkunft eine Vorkategorisierung dieser Mannigfaltigkeit durch : diese Mannigfaltigkeit ist aber oft ein durch die Teilnehmer verhandeltes Element, das sich im Laufe des Gesprächs verändert (wo bestimmte Formen manchmal als „fremd" oder „seltsam" kennzeichnet werden und wo ihr Unterschied in anderen Momenten überhaupt nicht berücksichtigt wird).

Die selben Fragen stellen sich für die Übersetzung des Transkripts :

- die Tatsache, dass man die Wörter bestimmter Sprecher eher als anderer übersetzt, kann als Werturteil betrachtet werden ;
- die Art, wie man mehr oder weniger wortgetreu übersetzt, kann dazu führen, eine verarmte Version der Rede des Sprechers zu bringen und deren Unterschied auszulöschen oder im Gegenteil zu unterstreichen ;
- es existieren verschiedene Formate für die Übersetzung (die als Notiz, nach dem Original, Zeile per Zeile, gegeben wird ; oder so, dass es der Originalfassung entspricht, so dass es ein fast wortgetreues Band mit dem Original bewahrt, so dass es eine grammatikalische Glosse davon gibt), die für jedes ein verschiedenes Bild der Kultur und der Sprache des anderen geben (Traverso, 2003).

Genauer gesagt handelt es sich hier um Übersetzung im spezifischen Rahmen der Korpora der gesprochenen Sprache. Diese Übersetzung ist unentbehrlich für die Arbeit über andere Sprachen als das Französische, sie bleibt aber oft ein Instrument

für den Forscher und in diesem Fall darf er nicht versuchen, das Abbild der Rede des Sprechers zu sein. Sie soll mit metalinguistischen Hinweisen begleitet werden, die es ermöglichen, die für eine gründliche Analyse der Sprache nötigen Nuancen besser zurückzugeben. So soll eine richtige Übersetzungsarbeit aus völlig anderer Sicht als der des Datensammeln für die Sprachanalyse geplant werden, wenn eine zweisprachige Veröffentlichung des Korpus vorgesehen wird.

EINSÄTZE DER MULTIMODALEN UND DER DETAILLIERTEN TRANSKRIPTION

Die Tatsache, dass nur mündliche Tätigkeiten notiert und die anderen Kommunikationshinweise ignoriert werden - wie es zur Zeit in den meisten Transkriptionen der Fall ist - kann ein gestörtes Bild des Verhaltens der Sprecher geben. Es mag vor allem der Fall Aphasiker oder Kinder sein, die sich durch andere Mittel als die linguistischen Standardmittel ausdrücken : die gesamten Mittel, die von diesen Sprechern mobilisiert werden, nicht zu berücksichtigen, heißt ein reduziertes Bild von ihnen geben, das ihr Benehmen pathologisiert oder anormal macht.

So können auch verschiedene Granularitätsgrade der Transkription (Jefferson, 1985) der Darstellung von nicht standardisierten Benehmen schaden (zum Beispiel kann die von einem Aphasiker ausgesprochene Vokalisierung sinnvoll sein und eine angepasste Transkription verlangen ; sie kann aber auch in einer oberflächigen Transkription zu einem einfachen „Geräusch“ ohne Sinn reduziert werden).

Der mehr oder weniger gründliche Charakter der Transkription entspricht also nicht nur wissenschaftlichen Vorschriften ; er entspricht auch ethischen und juristischen Vorschriften, die es erlauben, das von den Sprechern gegebene Bild zu differenzieren und zu komplizieren - indem man sich um so mehr von der Gefahr entfernt, ihn zu karikieren und durch stereotypisiertes Verhalten zu brandmarken.

4 SIND DIE KORPORA DER GESPROCHENEN SPRACHE ERBOBJEKTE ?

EINE LÖSUNG FÜR DIE ERHALTUNG UND DEN ZUGANG ZU DEN KORPORA DER GESPROCHENEN SPRACHE ?

4.1 ERINNERUNG AN DIE SITUATION

DIE VON DEN FORSCHERN IN DEN EINRICHTUNGEN ERZEUGTEN KORPORA DER GESPROCHENEN SPRACHE

Die Aufnahme von Korpora der gesprochenen Sprache fügt sich in eine schon hundertjährige Geschichte ein, die durch die Möglichkeit, die Stimme festzuhalten, eine neue und eigenartige Bedeutung bekommt. Schon 1986 halten Gelehrte, Forscher (Anthropologe, Ethnomusikwissenschaftler, Sprachwissenschaftler) ihre Sammlungen auf Zylindern, dann auf Platten fest. Da die Forscher dessen bewusst sind, dass sie neue Sammlungen schaffen, die an die zukünftigen Generationen weitergegeben werden müssen, finden die bei den „ethnographischen Aufträgen" aufgenommenen Produktionen natürlich ihren Platz in Institute unter staatlicher Schirmherrschaft. Das Spracharchiv, ein Konservatorium der Sprachen und Mundarten Frankreichs, entsteht innerhalb der Pariser Universität im Jahre 1911, die Phonotek des *Musée de l'Homme* 1932, die Nationale Phonothek 1938 und sie wird 1977 innerhalb der audiovisuellen Abteilung der BnF übernommen. Die grossen ethnographischen Sammelkampagnen, die von dem Museum für die volkstümlichen Künste und Überlieferungen[10] geführt wurden, das auch das Ethnologiezentrum Frankreichs ist, betreffen zum Beispiel die Bretagne im Jahre 1939 und die wichtige multidisziplinäre Forschung über dem Aubrac, die zwischen 1964 und 1968 vor allem beinahe vier tausend Tonaufzeichnungen und ungefähr zwölf Filme produzierte. Die Sprachwissenschaftler und dann die Ethnologen kümmern sich in erster Linie um das Werden ihrer Aufnahmen, inbegriffen um ihre Nutzung von anderen Forschern. In den siebziger Jahren führen bestimmte Soziologen wir Daniel Berteaux[11] die „Lebensgeschichte" in ihren Methoden ein. Diese Spur bereitet den Weg für multidisziplinäre Forschungen, von denen die „Ethnotexte" eine von Jean-Claude Bouvier und Philippe Joutard ausprobierte Spur sind.

FRANKREICH IST ABER EIN LAND GESCHRIEBENER TRADITION UND DAS MÜNDLICHE BEKOMMT KEINEN KULTURELLEN WERT UND NOCH MEHR KEINEN PATRIMONIALEN STATUS.

Die Universitätsforscher haben also keine spezifische kritische Methodologie, die seiner Problematik angepasst wäre. Dass es keinen normalisierten Wortschatz gibt, um die verschiedenen Formen der Korpora der gesprochenen Sprache zu

---

[10] Das MNATP (*Musée National des Arts et Traditions Populaires*) wurde im Juni 2005 das MCEM, *Musée national des Civilisations de l'Europe et de la Méditerranée* (Nationalmuseum der europäischen und Mittelmeerkulturen).

[11] Daniel Berteaux, « L'approche biographique. Sa validité méthodologique, ses potentialités » *Cahiers internationaux de sociologie*, 1980.

beschreiben, lässt erkennen, dass die Korpora der gesprochenen Sprache keinen wissenschaftlichen oder patrimonialen Status haben. Jede Disziplin gebraucht ihre eigene Terminologie und verleiht ihr einen präzisen Sinn. Claude Martel[12] erinnert an die Vielfalt der bekannten Definitionen von Wörtern wie Lebensgeschichten, Aussagen, Gespräche, je nach dem Fachgebiet der Gebrauchers.

Die Historiker haben sehr lange nur mit Vorbehalt die mündliche Aussage als eine sichere beachtungswürdige Quelle angesehen. Philippe Joutard, einer der Initiatoren der mündlichen Geschichte, erinnert an die Abgeschiedenheit Frankreichs den anderen europäischen Ländern wie zum Beispiel Großbritannien, Italien, Spanien, Argentinien gegenüber, die innerhalb der Universität eine dynamische und überschäumende Entwicklung dieser Disziplin kennen. Zahlreiche Zeitschriften beweisen diese Vitalität.

Die vorzügliche Forschung[13], die auf Antrag des Forschungsministers von Françoise Cribier und Elise Feller geführt wurde, hat bewiesen, dass die französischen Forscher in den dreißig letzten Jahren in allen Gebieten der Human- und Sozialwissenschaften außer der Geschichte unheimlich viel aufgenommen haben. Ihre Aufnahmen sind aber ohne offizielle Anerkennung oder Empfangsort in den Laboren geblieben. Sie wurden vor allem weder beschrieben noch kommentiert und die Bewilligungen der Zeugen, wenn sie bestehen, beschränken sich bestenfalls auf den Gebrauch der Forscher, die sie gemacht haben.

Die Bestände werden oft in privaten Händen aufbewahrt, denn die Sammlungen der gesprochenen Sprache, die während der offiziellen Aufnahmekampagnen realisiert wurden, bringen meistens die Behörden in Verlegenheit. In dieser Hinsicht ist das wichtige Unternehmen, das durch die DGRST am Anfang der sechziger Jahren um Plozévet, ein bretonisches Dorf, koordiniert wurde, ganz exemplarisch. Diese sehr wichtige Untersuchung wurde durch das *Musée de l'Homme* geführt, hat beinahe fünf Jahre gedauert, und sie hat Historiker, Geographen, Soziologen, Wirtschaftsexperten, Ethnologen mobilisiert. Viele hatten Tonbandgeräte. Diese Untersuchung hat aber, anstatt eine multidisziplinäre Arbeit zu liefern, nur eine Reihe von Einzeldarstellungen gebracht und niemand hat sich um die realisierten Aufnahmen gekümmert, außer um die, die vom Ethnologen Donatien Laurent produziert wurden. Er ist einer der wenigen Forscher, die ihre gesamte Sammlung nicht nur dokumentiert hat sondern im Zentrum für Forschung und keltische und bretonische Kultur an der Universität von Brest abgegeben hat. Heute sind ihre Aufnahmen digitalisiert und man kann im Universitätsrahmen in ihnen nachschlagen. Die anderen Aufnahmen wurden verloren oder die Bänder wurden aus Kreditmangel wiederaufgenommen.

---

[12] Siehe den Artikel von Claude Martel « La recherche et les sources orales, les mots pour le dire » in : *Bulletin de liaison des adhérents de l'AFAS* 10, 1998.

[13] Cribier F. & Feller E. (2003) *Projet de conservation des données qualitatives des sciences sociales recueillies en France auprès de la « société civile »* , Bericht, der der Ministerin für die Forschung und die neuen Technologien vorgelegt wurde. getippt. 2 Bände
http://www.iresco.fr/labos/lasmas/rapport/Rapdonneesqualita.pdf Eine andere kurze Befragung wurde auf Wunsch des CNRS von Dubar C. realisiert.

DIE ÄRA DER DIGITAL- UND NETZARBEIT : DIE 80ER JAHRE

So stellen noch einige dieser Sammlungen ohne wissenschaftlichen Status juristische Fragen, die immer noch nicht gelöst sind, was ihre Erhaltung und ihr Nachschlagen betrifft.

Man kann in den digital aufgenommenen Tondokumenten nur in realtime nachschlagen. Ihre Indexierung reicht nicht immer, um schnell davon Kenntnis zu nehmen. Diese Arbeit ist den meisten Forschern zuwider.

In den achtziger Jahren kündigen die Digitalisierungstechniken[14] ein neues Interesse für die gesprochene Sprache an, die eine empfindliche Angabe mit einem oft einzigartigen Inhalt ist. Die digital erzeugten Aufnahmen, die der Forscher selbst im Augenblick ihrer Realisierung indexiert hat, erlauben in der Tat, im Ton schnell zu „blättern", wie es im Schriftlichen möglich ist.

Wenn die Digitaltechniken aber wie für das Schriftliche und das Bild, den Zugang zu den Korpora der gesprochenen Sprache revolutioniert haben, so haben sie durch den einwandfreien Charakter der realisierten Kopien eine andere viel wichtigere intellektuelle Revolution gebracht, unter allem für den weiteren Gebrauch. *Indem sie den Begriff der Originalfassung ausstreichen, haben sie die Markierungen verschlossen*, die bis damals das Feld der Sammlungen begrenzten. Die Korpora der gesprochenen Sprache werden von ihrem Erzeuger innerhalb einer Institution beigelegt und werden Sammlungsobjekte aber es wird von daher unmöglich, die erste als „original" bekannte Aufnahme von den aufeinanderfolgenden Kopien eines Korpus der gesprochenen Sprache zu unterscheiden.

Wer wird es also beschließen, *den* Augenblick der Fassung auszuwählen und festzulegen, die beim Eintreten in eine Institution die Forschung seines Erzeugers bezeugen wird, da der Träger es nicht mehr erlaubt, die verschiedenen Elemente voneinander zu unterscheiden ? Welche Art *Metadaten* werden gleichzeitig den Sammlungen integriert werden ?

SAMMLUNGEN DER GESPROCHENEN SPRACHE OHNE STATUS

Die Korpora der gesprochenen Sprache stehen nicht im Urheberrecht als geschützte Werke, außer wenn sie eine identifizierte und als solche schützbare Form haben : *Aussagen, Interviews, Gespräche, Radiosendungen.*

Im allgemeinen werden die Sammlungen der gesprochenen Sprache so wie die Tondimension in dieser grossen kulturellen Unternehmung nicht berücksichtigt, die André Malraux 1964 eingeführt hatte : *das generelle Inventar der Gebäude und künstlerischen Schätze Frankreichs*. Keines der Verfahren, die ein Erbe[15] begründen, kann ihnen zugeschrieben werden. Sie werden nirgendwo klassifiziert noch eingetragen, dementsprechend kümmert sich kein spezialisierter Ausschuss „des

---

[14] « Musique et son : les enjeux de l'ère numérique. Création musicale, recherche, archivage, transmission », *Culture et Recherche* 91-92, 2002.

[15] Über dem sehr abgedroschenen Wort « Erbe » kann man Jean-Pierre Babelon und André Chastel lesen, *La notion de patrimoine*, Liana Levi, 1994, und die ausführliche historische Analyse von André Desvallées, « Emergence et cheminement du mot Patrimoine » in *Musées et collections publiques de France* 208 : 6-29,1995.

Erbes“ darum. Allein die UNESCO hat Initiativen in diese Richtung ergriffen. Die in den achtziger Jahren innerhalb des Kultus- und Kommunikationsministeriums geschaffene Mission des ethnologischen Erbes wird die Korpora der gesprochenen Sprache bald als Objekte einschreiben. Diese Sorge wird sehr schnell aus den Programmen beseitigt werden.

4.1.1 DIE SAMMELWERKE DER KORPORA DER GESPROCHENEN SPRACHE

PRAKTIKEN UND GEBRÄUCHE DER PATRIMONIALEN INSTITUTIONEN

Da die Korpora der gesprochenen Sprache, die einmalig von individuellen Erzeugern oder Institutionen produziert wurden, keine besondere Kategorie vom Standpunkt des Erbes und des Urheberrechtsgesetzes darstellen, hat der Gesetzgeber keine spezielle Einrichtung für ihr Sammeln und die Gestaltung ihrer Erhaltung vorgesehen.

Die Universität hat das Interesse an dieser reichen und wimmelnden Gesamtheit verloren, die zu diversifizierten Fachgebieten gehörte. *Es gibt also keinen Depotzwang für die Korpora der gesprochenen Sprache.*

Die Korpora der gesprochenen Sprache können innerhalb einer patrimonialen Institution lediglich durch eine *freiwillige Initiative* (Spende oder Depot) des Sammlers oder durch den *Entschluss der Institution* geschützt werden, die sich darum besorgt, Sammelwerke der gesprochenen Sprache über eigene Themen anzulegen. Die patrimonialen Institutionen können also simultan oder nacheinander Erzeuger von Korpora der gesprochenen Sprache sein und Verwalter von mündlichen Dokumenten, die von anderen erzeugt wurden. Die Institutionen, die für diese Art Sammlungen zuständig sind, leiten Forschungen über die Erhaltung der Tondokumente ein.[16] Sie setzen auch selektive Kriterien für die Gründung der Bestände ein.

- Im allgemeinen werden die Sammlungen innerhalb der patrimonialen Institutionen (Archiv, patrimoniale Bibliotheken, Museen) nach *dem Prinzip der Kohärenz der Bestände* angelegt. Eine einzelne Aufnahme wird nur für sich relevant sein. Die einzige Aufnahme der Stimme eines Schriftstellers in dem ihm gewidmeten Museum bleibt anekdotisch.
- Das heißt, dass die Anlage eines kohärenten Bestands das Ergebnis einer *anspruchsvollen Sortierungs- und Auswahlpolitik* ist : sie wird nach den allgemeinen Richtungen gemacht, die als vorrangig von der Institution (gesprochener Bestand für die BnF, Bestand über die Verschleppung für das Nationalarchiv) bestimmt worden sind. Diese Richtungen sollen breit und vollständig sein, damit sie für die Zukunft bedeutsame Bezugsquellen darstellen. In den Unternehmensmuseen, die die Nachfolger der in den siebziger Jahren auf Georges-Henri Rivière Initiative definierten Ökomuseen sind,

[16] Calas, M.-F. Fontaine, J-M. (1996) *La conservation des documents sonores*, Paris : CNRS-Verlag.

zielt das Sammeln von mündlichen Befragungen darauf, die Abwesenheit von Objekten zu schließen oder ihre Schwierigkeit, von der menschlichen Dimension innerhalb einer Gemeinschaft zu zeugen. In Fecamp lässt die Aufnahme der Arbeiterinnen der ehemaligen Fischfabriken eine Gesellschaftsorganisationsform der Stadt in der ersten Hälfte des 20. Jahrhundert erkennen, worüber weder ein Objekt noch ein Schriftstück berichten kann[17]. Es geht desgleichen im Museum der Tabakmanufaktur in Morlaix, im Ökomuseum der Städtegemeinschaft in Creusot-Montceau-les-Mines (Saône et Loire).

- Das Sammeln wird nicht immer als Sammlungsobjekt oder als Werk angeshen. In der BnF, im Nationalarchiv ist die dokumentarische Bearbeitung nicht durch den Träger der Sammlung bestimmt. Nichts Derartiges in den Museen. Mit Ausnahme des Nationalmuseums der europäischen und Mittelmeerkulturen (ehemaliges Nationalmuseum der volkstümlichen Künste und Überlieferungen), des Museums der Dauphiné, das die Befragungen von Charles Joïsten sehr früh wie die Objekte ins Museumsinventar aufgenommen hat, tragen die meisten Museen wie zum Beispiel das Museum von Salagon die Korpora der gesprochenen Sprache auf Inventare mit dem Typ Bibliothek ein. Da Ökomuseum von Saint-Quentin-en-Yvelines hat sich ebenso dafür entschieden, die mit den politischen Akteuren und den Bewohnern geführten Gesprächen auf ein separates Verzeichnis, wo die Studiensammlungen aufgenommen werden, einzutragen. Am Ende der neunziger Jahre genossen die Suche nach der Identität und die Gedächtnispflicht einem sehr starken, sogar übertriebenen Interesse. Dieses Archiv der gesprochenen Sprache genießt noch keine allgemein anerkannte Anerkennung.
- Die Sammeln der gesprochenen Sprache kann man nicht auf die Aufnahme der Stimmen zurückführen. Sie gewinnen ihren vollen Sinn nur in die Bereitstellung der zeitlichen, technischen, wissenschaftlichen Daten ihrer Herstellung. Diese gesamten Kontextualisierungselemente (Metadaten), die für das aufgenommene Korpus spezifisch sind, bilden mit ihm ein nicht voneinander zu trennendes Ganzes, ohne das die Aufnahme Zeitlichkeit und Sinn verlieren würde. Man könnte ihr also alles und sein Gegenteil bedeuten lassen.
- Wie jedes patrimoniale Objekt kann man das Dokument der gesprochenen Sprache, obwohl es datiert, identifiziert wurde, nicht nur auf den Gebrauch seines Erzeugers reduzieren, wie etliche Forscher es sehr lange geglaubt haben. Die Forschungen über die

[17] Diese Gesprächsserie wurde in Zusammenarbeit von dem Museum und von dem Dienst des Stadtarchivs gemacht und gab den Anlass zu einer CD mit Heftchen *Femmes de marins, compagnes de pêche*, Fécamp, Museum der Terre-Neuvas, 2003.

gesprochene Sprache gehen oft über das Projekt ihrer Führung hinaus. Sie können im Rahmen anderer Disziplinen gebraucht werden.

> *„Eine neue Lesung führt dazu, einen anderen Blick auf das, was gesagt wurde, zu werfen, weil die Zeit vorbeigegangen ist und die Fragen, die man sich stellt, anders geworden sind“* [18].

- Sie sollen im Laufe der Zeit von verschiedenen Forschern durch ihre persönliche Analyseschema analysiert werden können. Das Programm für die Digitalisierung der Tondokumente, das Ende 1999 vom Kultus- und Kommunikationsministerium eingeführt wurde, hat das Defizit an Informationen ans Licht gebracht, was sich auf diese Sammlungen der gesprochenen Sprache bezieht. Einige Bestände, die als historisch angesehen werden, konnten aus Mangel an unentbehrlichen Kontextualisierungsdokumenten ihre Bedeutung nicht ordentlich beweisen. Dazu besass keine der Sammlungen, die auf die Digitalisierungsaufforderung geantwortet haben, die Nutzungsrechte, die erlauben konnten, das Nachschlagen vom Publikum vor allem durch das Internet zu organisieren.
- Welche Schutzart genießen die Korpora der gesprochenen Sprache in den öffentlichen und privaten Einrichtungen ? Das Spenden einer Sammlung zu einer Institution hat keinen *Probewert*. Kann das Spendendatum einen eventuellen Prioritätsbeweis sein im Vergleich zu einer Aufnahme, die sich als eine Fälschung der Ersten erweisen würde? Mit Ausnahme der Spenden die, von Natur her, immer widerruflich sind, gehen die Sammlungen (unter der Form der Träger oder digitalisierter Daten) endgültig und unverjährbar in die Bestände der Institution. Diese Abtretung zieht nicht die Abtretung der Nutzungsrechte mit sich, wie eben gesagt, vorbehaltlich spezifischer Bestimmung. Die Institutionen verpflichten sich a priori dazu, den physischen Fortbestand zu sichern und das Nachschlagen in die Korpora der gesprochenen Sprache in der Einhaltung der Rechte derjenigen, die an die Anfertigung teilgenommen haben, zu organisieren, aber es ist unvermeidlich, dass die Abtretenden wenigstens die Nachschlagenerlaubnisse abtreten. Seit den achtziger Jahren hat das Nachschlagen von weitem in gewisser Weise das Interesse für die Korpora der gesprochenen Sprache „geweckt“ und früher unvorstellbare Zugangsmöglichkeiten ahnen lassen. Der Zugang zu den digitalen Korpora der gesprochenen Sprache stellt das Problem mit den Worten der Bewahrung und der Identifizierung der Quellen vor der Digitalisierung. Er beziffert sich auch in Geld- und Personalmittel.

[18] Françoise Cribier & Elise Feller, *op. cit.*

4.1.2 *DIE NATIONALBIBLIOTHEK FRANKREICHS*

KONSERVATORIUM DER MÜNDLICHKEIT

Die audiovisuelle Abteilung der Nationalbibliothek Frankreichs ist der Nachfolger des Spracharchivs, das 1911 von Ferdinand Brunot gegründet wurde, des 1928 darauf folgenden Museums des Wortes und der Geste, dann des 1938 geschaffenen nationalen Tonbandarchivs. Ihre Aktion steht im Einklang mit der dieser Institutionen. Heute ist also mehr als ein Jahrhundert Mündlichkeitsgedächtnis bewahrt und zur Verfügung des Publikums gestellt.

Die audiovisuelle Abteilung führt aber gleichzeitig eine aktive Politik für die Entwicklung ihrer Sammlungen, vor allem im Gebiet der Mündlichkeit. Die Nationalbibliothek Frankreichs hat ja neben dem Sammeln der Pflichtabgabe die Aufgabe und die Mission, ihre Sammlungen durch Errungenschaften, Schenkungen, Spenden, Vermächtnisse, Übergaben, usw zu erweitern. Es ist also der Fall für die audiovisuelle Abteilung die, zu der Pflichtabgabe der Ton-, Video-, Multimedia- und Informatikdokumente, wofür sie verantwortlich ist, die Grundlinien einer Erweiterungspolitik ihrer Sammlungen, was unveröffentlichte Tonaufnahmen betrifft, bestimmt hat. Am Ende dieser Einleitung sind einige der vor kurzem in die audiovisuelle Abteilung gekommenen Bestände, die für den Platz der gesprochenen Sprache in ihren Sammlungen charakteristisch sind.

DIE GRUNDLINIEN DER ERWEITERUNG DER SAMMLUNGEN DER AUDIOVISUELLEN ABTEILUNG

Die audiovisuelle Abteilung beschreibt als „unveröffentlichte“ Dokumente „Quellen“dokumente, die einzig vorhanden sind, nicht mehrmals verbreitet werden und von keiner besonderen Auflageform bestimmt werden. Dazu hat die audiovisuelle Abteilung gewisse Grundprinzipien bestimmt, die angesichts der unbegrenzten Erweiterung des Feldes und der Vielfalt der Inhalte (Sprachwissenschaft, Ethnologie, mündliche Geschichte...), der Vielfalt der möglichen Quellen (institutionnellen, unabhängige Forscher...), der notwendigen Komplementarität mit anderen Institutionen so wie vor den auszugleichenden Lücken, was die Bewahrung, die Verbreitung und die Aufwertung betrifft, ihre Erweiterungspolitik auf diesem Gebiet leiten können.

DAS DOKUMENTARISCHE UND PATRIMONIALE KRITERIUM

Die Politik der Abteilung beruht zuerst auf einem Auswahlprinzip. Was grundsätzlich dazu führt, eine Spende unveröffentlichter Dokumente anzunehmen oder abzulehnen, ist vor allem das dokumentarische und/oder patrimoniale Interesse des vorgeschlagenen Bestands. Dieses Kriterium kann als dies des „nationalen Gedächtnisses“ betrachtet sein. Welches sind mit anderen Worten die unveröffentlichten Aufnahmen, die man als nichts anderes als Gedächtnisstück, als Teil eines nationalen Erbes betrachten kann ? Dieses Kriterium begrenzt das Feld der Dokumentarpolitik nicht auf den französischen „Boden“ sondern es gibt den Beständen, die entweder als Quelle (der Sammler, die Einrichtung...) oder im Inhalt einen Zusammenhang mit Frankreich haben, den Vorrang. Illustrierung davon sind das Spenden des Bestands von Deben Bhattacharya, eines indischen

Musikwissenschaftlers, der in der ganzen Welt Aufnahmen gemacht hat, in Paris aber von 1954 bis 2001 lebte, oder das der pygmäischen Sammlungen von Simha Arom (Lacito-CNRS).

Die audiovisuelle Abteilung ist in enger Beziehung mit diesem Kriterium des dokumentarischen und/oder patrimonialen Interesses, dadurch eng eingegrenzt und sie schenkt Dokumenten oder Beständen, für die es a priori keinen festgestellten Konservierungs- und/oder Nachschlagenort gibt, eine privilegierte Aufmerksamkeit. Es ist zum Beispiel der Fall in einigen Laboren für bestimmte Privatarchive oder Bestände ohne Erben, aus Mangel an angepassten Einrichtungen.

DIE AKZEPTABILITÄT DES BESTANDS UND DAS DOKUMENTARISCHE PRINZIP

Wenn dieses Prinzip der Auswahl und die dokumentarische und patrimoniale Prinzipien einmal gesichert sind, werden Akzeptabilitätsbedingungen gestellt, was den Erhalt eines Bestands betrifft. Zuerst handelt es sich um dokumentarische Bedingungen. So müssen die unveröffentlichten Quellen dokumentiert sein oder in dokumentarischer Weise ausgenutzt werden können, um erhalten oder gekauft werden zu können. Geplant sein kann entweder, dass die dokumentarische Bearbeitung unter der Form von Metadaten gleichzeitig mit dem Archiv besorgt wird oder eventuell, dass alle Informationen der BnF unter der einen oder der anderen Form besorgt werden, damit sie dokumentarisch bearbeitet werden können.

DIE AKZEPTABILITÄT DES BESTANDS UND DAS GERICHTSPRINZIP

Die gerichtlichen Bedingungen sind eine andere Komponente der Akzeptabilitätsbedingungen. Die - physische oder moralische - Person, die das Spenden macht, soll vor allem sich dessen vergewissern :

- dass er die physischen Träger besitzt, worauf die Aufnahmen gemacht worden sind, und dass diese Aufnahmen der BnF gegeben werden können ;
- dass er die Urheberrechte auf den realisierten Werken und die Nebenrechte des Herstellers der Tonaufzeichnungen und eventuell der Interpreten der Musik besitzt oder garantieren kann.

Für die BnF bedarf das Erhalten der Träger auch der Verfügung über die Urheberrechte und über die für ihre Vervielfältigung und Mitteilung zu den Lesern erforderlichen Nebenrechte, da die Tonaufzeichnungen vervielfältigt und dargestellt werden müssen, um aufbewahrt und nachgeschlagen werden zu können. Aber die - physische oder moralische – Person, die das Spenden macht, hat nicht immer die gerichtliche Fähigkeit, diese Vervielfältigungs- und Mitteilungserlaubnisse auszufertigen.

Es müssen der BnF abgegeben können :

- das Vervielfältigungsrecht des Dokuments, das heißt, die Möglichkeit, aus Gründen der Konservierung des Signals seinen Inhalt auf einen geeigneten (digitalen) Träger zu übertragen ;
- das Darstellungsrecht. Dieses Recht versteht sich wenigstens als die Möglichkeit für das Forscherpublikum, das Dokument im Raum P

(Niveau „Forschung“ der Bibliothek) nachzuschlagen. Zugegeben wird das Prinzip eines Mitteilungserlaubnis von Fall zu Fall. Es wird auch akzeptiert, dass eine Vorbehaltszeit für die Mitteilung verlangt werden kann für andere Gründe als die, die auf die Urheberrechte (Vertraulichkeit der Daten, die das Privatleben betreffen...) zurückzuführen sind.

EINIGE BEISPIELE UNTER DEN LETZTEN UNVERÖFFENTLICHTEN BESTÄNDEN, DIE VON DER AUDIOVISUELLEN ABTEILUNG ALS SPENDE BEKOMMEN WURDEN

(in den Sammlungen nach der Ankunftsreihenfolge geordnet) :

- Bestand der regionalen Sprachatlasse (1979 und folgende) ;
- Bestand des Zentrums für geschichtliche Recherchen, EHESS/CNRS (1979) : mündlich gelieferte Geschichte, Lebensgeschichten, Jahrgänge 1970-1980 ;
- Bestand Félix Quilici (1981) : korsische Musik mit mündlich gelieferten Tradition, 1959-1963 ;
- Bestand Geneviève Massignon (1985) : ethno-linguistische Sammlungen, Acadie, Westen von Frankreich, Korsica..., 1946-1963 ;
- Bestand Nicole Revel (1995) : Epen Palawan, Philippinen, Jahrgänge 1980 ;
- Bestand Gilles Deleuze (1997) : Vorlesungen, Universität Paris VIII, 1979-1984 ;
- Bestand Deben Bhattacharya (2003) : ethnomusikologische Sammlungen, Asien, Europa..., 1954-2000 ;
- Programm „Archivierung“, LACITO/CNRS : seltene Sprachen, Transkripte, Vermerke, http://lacito.vjf.cnrs.fr/archivage/

4.1.3 DIE ARCHIVE FRANKREICHS

Im Buch II des französischen Erbrechtes werden die Archive im Artikel L 211-1 wie folgt definiert :

„Als Archive werden alle Dokumente bezeichnet, die von jeder physischen oder moralen Person und durch jedes öffentliche oder private Dienst oder Einrichtung im Rahmen seiner/ihrer Tätigkeit hervorgebracht werden, egal wann, unter welcher Form und auf welchem Träger sie produziert wurden. Die Bewahrung dieser Dokumente wird sowohl für die Bedürfnisse der Verwaltung und der Rechtfertigung der Rechte der physischen oder moralischen, öffentlichen oder privaten Personen wie für die geschichtliche Dokumentation der Forschung im öffentlichen Interesse organisiert.“ Die Archive bilden zwei Kategorien : die öffentlichen Archive, die von der Tätigkeit des Staates, der Gebietskörperschaften und der öffentlichen Unternehmen kommen, und die privaten Archive.

Die Produktionsweise und nicht der Typ des Trägers bestimmt die Zugehörigkeit zu der einen oder der anderen Kategorie. Die Aufnahme einer Sitzung des *Conseil Général* (das oberste Exekutivorgan eines Departements) ist ein Dokument des

öffentlichen Archivs, wenn die Aufnahme eines Politikers am Radio ein Dokument des privaten Archivs ist.

EINIGE BEISPIELE VON KORPORA DER GESPROCHENEN SPRACHE IN ARCHIVBESTÄNDEN

- *Das Nationalarchiv*

Es steht unter der Verantwortung der Leitung der Archive Frankreichs und legt fünf Zentren zusammen.

- Das Historische Zentrum der Nationalarchive (CHAN) in Paris. Innerhalb der Abteilung XX. wurde in den achtziger Jahren eine Zelle für die mündlichen Archive gegründet. Diese Zelle bekommt Spenden, zum Beispiel die von der Stiftung für die Erinnerung an die Deportierten, sie produziert aber Aussagen als Ergänzung der geschriebenen Archive, „indem gesagt wird, was nicht geschrieben werden kann, indem die Ereignisse auf menschlicher Ebene wieder dimensioniert werden, und indem existierende geschichtliche Lücken gegebenfalls durch das Erzählen verschleierter Einzelheiten geschlossen werden“.[19] So geht es auch für die Videoaufnahmen der Gerichtsarchive (Gerichtsverfahren von Klaus Barbie, Paul Touvier, oder über die Affäre des „verseuchten Blutes“) und die Archive des Staatspräsidenten : Reden und Pressekonferenzen der Staatspräsidenten Georges Pompidou und Valéry Giscard d’Estaing.
- Die von den Verwaltern geschaffenen Quellen entsprechen zwei Herangehensweisen : das autobiographische Erzählen trägt zu der Geschichte der Eliten bei und die thematischen Korpora können es erlauben, mehrere Erzählungen über ein selbes Ereignis zu kreuzen (zum Beispiel das Grundschullehreramt in den fünfziger Jahren).
- Das Zentrum für die zeitgenössischen Archive (CAC) in Fontainebleau. Dahin werden zum Beispiel die 400 Stunden der Aufnahmen beigelegt, die im Rahmen des Programms realisiert wurden, das von dem Komitee der Geschichte der staatlichen Sozial- und Krankenversicherung von Dominique Aron-Schnapper eingeführt wurde.
- Das Zentrum für die Archive der Arbeitswelt (CAMT) in Roubaix, das jedes Archiv über sein Gebiet sammelt, darunter Aufnahmen.
- Von den zwei anderen Zentren bewahrt das von Esperran nur Mikrofilme und das für die Archive aus Übersee bewahrt vor allem einen geschlossenen gedruckten Bestand.

[19] Agnès Callu, « Aux Archives nationales, une politique raisonnée en faveur des témoignages oraux » *Colonnes : archives d’architecture du XX^e siècle*, 20, décembre 2002.

- *Die Einrichtungen für die Archive in den Departements*

Diese Einrichtungen wurden viel früher als andere in die Provinz verlegt, sie sammeln oft Kopien von Radiosendungen, Amateurfilmen, Dokumentarfilmen, und führen allein oder mit Vereins- und Universitätshilfe Befragungsprogramme. Ihre Situation ist sehr diversifiziert und die Wichtigkeit der Bestände der gesprochenen Sprache hängt mit den bearbeiteten Thematikbereichen und vor allem mit der Motivierung und dem Interesse des Leiters zusammen.

Das Dienst der Stadtarchive hat oft in der Bewegung, die aus allen Erinnerungen Erbe macht, Jugendliche, die als „Gedächtnishüter" rekrutiert wurden (Beispiele : Martigues, Lille), mit der Realisierung von Archiven der mündlichen Sprache betraut.

4.1.4 *PLATZ DER KORPORA DER GESPROCHENEN SPRACHE IN DEN MUSEEN*

Als Museum im breitesten Sinne des Wortes wird jede dauerhafte Sammlung betrachtet, die aus Gütern besteht, deren Bewahrung und Vorführung ein öffentliches Interesse annehmen, und die für das Wissen, die Erziehung und die Freude des Publikums organisiert ist.

Die Sammlungen bestehen aus Gegenständen und Werken jeder Art, dessen Tatsächlichkeit greifbar ist.

Die Aufnahmen der gesprochenen Sprache stellen definitionsgemäß Immaterielles dar. Die ICOM, internationale Organisation der Museensammlungen, die innerhalb der UNESCO die Entwicklung aller Museenformen bestimmt, hat jedoch 2004 die Debatte über die immaterielle Dimension des ungreifbaren Erbes eingeführt. Das von den westlichen Museen im allgemeinen empfundene Unbehagen vor der Eingliederung der Tondimension, der audiovisuellen und landschaftlichen Dimension in die Museen gibt dieser Art Widerspruch des Museums zwischen Gegenstände und Mündliches ganz genau kund.

Die Geschichtemuseen, Ökomuseen, Gesellschaftsmuseen gebrauchen dagegen seit sehr lange (zum Beispiel das Museum der Dauphiné in Grenoble) die Aufnahmen der mündlichen Erinnerungen als eines der wesentlichsten Elemente des kulturellen und wissenschaftlichen Projekts, worum das Museum sich organisieren wird. Die Tonaufzeichnungen sind wie die anderen Sammlungen in der Inventarliste des Museums aufgeschrieben, der Fall ist aber weit davon, verbreitet zu sein, und viele Ton- und Videoaufnahmen sind bestenfalls im Buch der Studien- oder Dokumentarsammlungen eingeschrieben.

Wenn die Korpora der gesprochenen Sprache wie im Museum de Dauphiné als Werke anerkannt wären, die in der Inventarliste aufgeschrieben sind und deren Redaktionsbedingungen durch gesetzgebende Texte bestimmt werden, dann wären sie unveräußerlich und unverjährbar.

4.1.5 *DIE „KORPORA DER GESPROCHENEN SPRACHE" IN DEM INA*

Durch das Nachschlagen in die Werke der Pflichtabgabe des Rundkunks und des Fernsehens gibt in der Tat das INA Zugang zu einer grossen Vielfalt von Korpora

der gesprochenen Sprache, die von verschiedenen aufgenommenen Aussagen, Worten, Ansprachen, Reden gebildet sind und zum Verbreitungszweck aufgenommen wurden.

Die Forscher, die das Nachschlagenzentrum der Inathek besuchen, bilden für ihre spezifischen Bedürfnisse Korpora aus den Quellen des Rundfunks und des Fernsehens, die sich in verschiedene disziplinarische Logikarten der Auswertung der Korpora der gesprochenen Sprache einschreiben : Sprachwissenschaft, Soziologie, Geschichte...

Bei der Beschäftigung mit diesen Korpora kann es um die Redeverfahren in dieser oder jener Sendungsart (das Fernsehinterview, die Rundfunkkommentare...) gehen, um verschiedene (politische, journalistische...) Analysearten der Rede, um die Schaffung lexicographischer Verzeichnisse, um soziolinguistische Analysen (das Wort des Tänzers, Worte von Arbeiterinnen) usw.

Bestimmte Sendungensammlungen, die im INA archiviert sind, bilden auf Anhieb „geschlossene Korpora" von mündlichen Produktionen.

Um nur einige zu erwähnen : „Die Archive des zwangzigsten Jahrhunderts", die von Jean-José Marchand produziert wurden. Es sind Zusammenstellungen von Gesprächen mit Persönlichkeiten aus der Welt der Literatur und der Kunst ; „Die Erzähler", eine von André Voisin realisierte Sammlung, die vom Forschungsdienst der ORTF produziert wurde. Es ist eine Zusammenstellung persönlicher, regionaler Geschichten (*Ceux de La Hague, Au coeur de l'Aubrac*...).

Das INA macht sonst seit seiner Gründung bei der Produktion von Sammlungen patrimonialer Aufnahmen und Zusammenstellungen von Aussagen mit.

Diese Gespräche sind von unterschiedlicher Dauer (bis 15 Stunden Unterhaltung) und über „@propos", eine Schnittstelle, die die Navigation im Programm erleichtert, für ein aktives Nachschlagen zugänglich.

- So beruht die Sammlung „Musique Mémoires" auf einer Archivierungskampagne, die dazu strebt, die Aussagen von Komponisten, Interpreten, Dirigenten und Persönlichkeiten aufzunehmen, deren Schaffungen und Aktionen das Musikleben der sechzig letzten Jahre geprägt haben. Diese Unterhaltungen, die von Bruno Serrou geführt werden, erforschen den eigenen Lauf jedes Künstlers : Herkunft, Ausbildung, Einflüsse, Begegnungen, Ausübung des Berufes... Unterhaltungen, die schon ausgeführt wurden : mit François Bayle, Claude Helffer, Betsy Jolas, Claude Ballif, Pierre Boulez, Marius Constant, Antoine Duhamel, Luis de Pablo, Yvonne Loriod, Michel Fano, Ivo Malec.
- « Geschichte von Geschichtswissenschaftlern » bietet eine Sammlung von Selbstporträts zeitgenössischer Historiker an ; wenn die Geschichte ihres Lebens so erzählt wird, erlaubt sie ein besseres Verständnis ihres Werkes. Unterhaltungen, die schon ausgeführt wurden : mit Maurice Agulhon, Pierre Chaunu, Emmanuel Le Roy Ladurie, Claude Nicolet, Pierre Nora, Robert

Paxton, Madeleine Rebérioux, René Rémond, Zeev Sternhell, Jean Tulard.

- « Télé notre histoire » ist eine Sammlung langer Unterhaltungen, die eine Erinnerung an das Fernsehen bietet, die von denen erzählt wird, deren persönlicher Lauf und berufliche Praxis die Geschichte dieses Mediums erklären : Schriftsteller, Künstler, Produzenten, Programmplaner, Ingenieure, Techniker, Entscheidungsträger, Pioniere oder jüngste Praktiker. Unterhaltungen, die schon ausgeführt wurden : mit Igor Barrère, Marcel Bluwal, Yves Jaigu, Jacques Krier, Claude Santelli, Pierre Tchernia…
- Andere Unterhaltungen, die sich nicht in eine Sammlungslogik einschreiben, bieten jedoch die Aussagen wesentlicher Akteure des kulturellen, wissenschaftlichen und künstlerischen zeitgenössischen Lebens an. Unterhaltungen, die schon ausgeführt wurden : mit Françoise Gilot, K.S. Karol, Claude Lévi-Strauss.
- „Mémoires de la Shoah » ist im Produzieren, es ist eine Sammlung von 110 Unterhaltungen (ungefähr drei Stunden) von Zeugnissen des Holocausts : Deportierte, Waisenkinder, „Gerechten".

Zu all diesen Sammlungen wird auf Dauer im Nachschlagenzentrum der Inathek in Frankreich Zugang gegeben werden.

4.2 DIE PRIVATEN INITIATIVEN

Die Aufnahme von Aussagen kennt seit 1972 (Datum der Schaffung der ständigen Geschichtekommission des Schulwesens) eine erhebliche Entwicklung innerhalb von Programmen der „*Comités d'Histoire orale*" [Komitees der mündlichen Geschichte], die von den öffentlichen Einrichtungen eingerichtet wurden, die darum besorgt sind, das Gedächtnis ihrer Einrichtungen aufzuwerten.

Heute werden 67 Komitees und Dienste[20] einer Einrichtung (Geschichtekomitee des Kultus- und Kommunikationsministeriums, Geschichtekomitee der BnF) integriert.

Die AHICF, Verein für die Bahngeschichte in Frankreich, nimmt einen besonderen Platz ein. Sie stellt sich zur Verfügung der Einrichtungen, deren Geschichte zu schreiben sie die Absicht hat. Die AHICF wurde 1987 geschaffen und hat zwei Aufträge : Forschung und Sicherung des Erbes. Sie begünstigt die Erhaltung der Quellen, ist aber nicht dazu beauftragt, sie selber zu gewährleisten. Es bestehen Dienste nach Wunsch (Historiker), um dabei zu helfen, das Gedächtnis im industriellen Gebiet zu schaffen.

Diese Komitees sehen generell die realisierten Aufnahmen wie private Archive an, die vom Urheberrechte gehaftet sind. Die Erbfolgeklausel der erzeugten Korpora

[20] Handbuch der Komitees für die Geschichte und die Geschichtsdienste. Paris, Komitee für die wirtschaftliche und finanzielle Geschichte Frankreichs.

der gesprochenen Sprache zugunsten eines Archivdienstes, falls die Vereine aufgelöst werden, ist eine ziemlich verbreitete Regel.

Unter den aktiven Partnern eines Netzes der „lokalen Archive" können die assoziierten Zentren der BnF wie die FAMDT, DASTUM, die MMSH (Südländisches Haus der Humanwissenschaften) in Aix-en-Provence (vgl BnF, assoziierte Zentren) zitiert werden. Diese Zentren verfügen nur selten über die vollständigen Rechte der aufbewahrten Korpora.

4.3 DER ZUGANG ZU DEN SAMMLUNGEN

Es besteht keinen Sammelkatalog der Korpora der gesprochenen Sprache. Mehrere Initiativen haben es erlaubt, Institutions- und Vereinsstrukturen zu identifizieren, die zum Erhaltungs- und Nachschlagenszweck Korpora der gesprochenen Sprache erzeugen oder sammeln. Allerdings machen sie eher eine globale Beschreibung als eine ausführliche Aufstellung der Inhalte, da die meisten von diesen Korpora von ihren Erzeugern sehr wenig beschrieben werden. Die Veröffentlichung[21] von diesem Jahr resultiert aus dem Studium einer umfangreichen Umfrage über die in Frankreich bewahrten mündlichen Quellen in den Sozialwissenschaften. Sie wird vielleicht den Anfang der Einrichtung einer gemeinsamen Quelle für die Mündlichkeit signalisieren, wenn der computerisierter Katalog eine Aktualisierung durch das Netz der Erzeuger erlaubt.

Die Bedingungen des Zugangs werden durch den Vertrag bestimmt. Es existiert aber keinen Vertragstyp.

In den Einrichtungen werden die mündlichen Aufnahmen überwiegend durch das Urheberrecht behandelt. Im allgemeinen verfügt der Zeuge um ein Aufsichtsrecht über den Gebrauch seiner Stimme (Gesetz vom 17. Juli 1970). Es kann keiner die Worte und das Bild einer Privatperson, die sich in einem privaten Ort befindet, ohne ihre Erlaubnis festhalten, aufbewahren oder verbreiten. Der Artikel 9 des Zivilgesetzbuches und der Artikel 226-1 des Strafgesetzbuches zwingen dazu, die Bewilligung der Person zu bekommen. Wenn der Zeuge Eigenständigkeit in seinen Worten beweist, kann er als Autor betrachtet werden und als solcher die damit zusammenhängenden moralischen Rechte genießen. Die Verwendung seiner Aufnahme kann durch die Pflicht einer im Rahmen eines Vertrags festgesetzten Bezahlung bedingt sein. Der Sammler muss die breiteste Nachschlagenerlaubnis erzielen.

> *Die Zugänglichkeit stellt Rechts- und Ethikfragen (Respekt des Privatlebens, Recht auf seine Stimme für einen Zeugen, Lebensgeschichten, heikle Aussagen, Worte, die verleumderisch werden können...). Aber aus Gründen, die mit der Natur des Inhalts zusammenhängen (Lebensgeschichten und Aussagen, die andere Personen beschuldigen, Gespräche in psychiatrischer Umgebung) können diese Korpora der gesprochenen Sprache an*

[21] Callu, A., Lemoine, H. (2004) *Patrimoine sonore et audiovisuel français : entre archive et témoignage : guide de recherche en sciences sociales*, Paris, Belin, 7 Bände, 1 CD-Rom, 1 DVD-Rom.

Ort und Stelle nicht zum Nachschlagen gegeben, und noch weniger auf dem Internet verbreitet werden.

Jeder Fall ist also ein besonderer Fall und die Anerkennung der Rechte der einen und der anderen untersteht einer feinen und heiklen Analyse, bei der folgenden Fragen eine Antwort gegeben werden muss : wer besitzt die Rechte ? Ist der Besitzer mit der Übertragung der Rechte einverstanden, unter welchen Bedingungen und für welchen Zweck ? Für wie lange ? Unmittelbarerweise ? Mit zeitlicher Verzögerung ?

Der Forscher-Sammler, für den die Aufnahme der Korpora ein Augenblick in einer gründlichen Forschung darstellt, sollte als Autor geschützt werden können. Er wird meistens *Sammler* genannt. Um dem Fragenden ein Urheberrecht anerkennen zu können, sollte die originale Form seiner Worte klar herausgestellt werden können.

Die Einrichtungen können infolgedessen nur in ihren eigenen Räumen das Nachschlagen erlauben und die Digitalisierungsarbeiten, wozu die Forscher die Initiative ergreifen können, werden oft ohne Erlaubnis der echten Rechtebesitzer gemacht (Nationaler Digitalisierungsplan).

Es bleiben viele Schwierigkeiten.

Die Frage des unselbständigen Sammlers, der im Rahmen seiner öffentlichen Aufträge handelt und seine Rechte zugunsten vom Staat aufgeben sollte, unterstreicht das Problem der Rechte der unselbständigen „Autoren", das im öffentlichen Dienst auf finanzielle nicht gelöste Fragen stößt.

Was soll über die Rechte gesagt werden, die Studenten fordern könnten, die an die Technik des Interviews wohl wenig gewöhnt sind und dazu bezahlt werden, die Fragen in die Reihenfolge eines vorher festgelegten Fragebogens zu stellen ?

Der Status der Sammlungen der mündlichen Archiven ist nicht belanglos.

Der Fall der ersten wichtigen Befragung über die Geschichte der staatlichen Sozial- und Krankenversicherung, die zwischen 1973 und 1975 von Dominique Schnapper nach Bedarf des Geschichtekomitees dieser 1973 geschaffenen Einrichtung geleitet wurde, hat das Aufnehmen 200 Zeugen erlaubt, die zu 400 Stunden Interviews und Aussagen führen. Es handelte sich definitionsgemäß um Privatarchive. Es wurde aber vor dem Anfang der Befragung entschlossen, dass die gesamte Forschung als öffentliches Archiv eingeteilt werden sollte und dass infolgedessen nach sechzig Jahren darin nachgeschlagen werden konnte. Dieser Beschluss hat wichtige Folgen gehabt. Philippe Joutard hat mehrmals dieses Beispiel erwähnt, wo er einen der möglichen Gründe für die mangelnde Dynamik der Entwicklung der mündlichen Geschichte in Frankreich sieht.

Florence Descamps teilt auch diese Analyse, indem sie diese innovativen mündlichen Archive, die vom Anfang an „gefroren" wurden, prangert.

Da die Forscher wenig dazu neigten, ihre Korpora institutionalisiert zu sehen, haben sie diese für sich behalten. Sie wurden von den Einrichtungen wenig ermutigt, wie von dem CNRS und der Universität, die bis vor kurzem nie konstruktive Initiativen ergriffen haben, um Korpora der gesprochenen Sprache, die sich jeder akademischen Begriffsbestimmung entzogen, zu bewahren (eine Ausnahme ist die Übereinkunft, die zwischen dem CNRS und der BN 1979 für den Schutz der linguistischen Atlasse unterschrieben wurde), während die mündliche Geschichte in Großbritannien, wo sie geboren ist, wie in den anderen romanischen Ländern als in Frankreich einen großen Wuchs kennt.

4.3.1 WELCHES NETZ FÜR MORGEN ?

EIN VERWALTUNGS-, SCHUTZNETZ DER SAMMLUNGEN DER KORPORA DER GESPROCHENEN SPRACHE, DAS VON DEN UNIVERSITÄTEN UND DEN FORSCHUNGSINSTITUTEN ORGANISIERT WIRD, ODER PATRIMONIALE EINRICHTUNGEN ?

Wie es in vielen europäischen Ländern der Fall ist, könnten die Universitäten und die Forschungsinstitute außer den patrimonialen Einrichtungen die Fähigkeit und den Willen haben, ein umfangreiches Netz der Human- und Sozialwissenschaften zu schaffen, wodurch die zur Verfügung der anderen Forscher gestellten Korpora geschützt und anderen zugänglich werden könnten.

Der von Françoise Cribier in Zusammenarbeit mit Elise Feller verfasste Bericht[22] hat die Situation und die Netze studiert, die für die Sicherung und den Zugang zu den qualitativen Daten der Sozialwissenschaften in sechs europäischen Ländern existieren. Zwei Initiativen werden als eventuelle Vorbilder für die französischen Forscher dargestellt : Qualidata (Großbritannien) und SIDOS (in der Schweiz).

> In Großbritannien wurde Qualidata[23] 1994 geschaffen. Sie ist in Colchester in der Soziologieabteilung der Universität von Essex angesiedelt. Diese Initiative hat sich in einen universitären Rahmen gepasst, der für die Erhaltung der mündlichen Daten breit sensibilisiert war, vor allem durch die von Paul Thomson auf die Anregung des ESRC (Rat der wirschaftlichen und sozialen Forschung vom Vereinigten Königreich) geleitete Befragung. Sie könnte als Beispiel dienen. Die

[22] *op.cit.*

[23] Qualidata, UK Data Archive, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK. www.qualidata.ac.uk.

Einrichtung ist sehr selektiv, was die nach 1995 erzeugten Bestände betrifft (unter den Kriterien : genau identifizierte Themen, dokumentierte Korpora, digitalisierte Tondokumente, die in einem guten Zustand sind, und deren Merkmale juristisch festgestellt sind).

Die Einrichtung gebraucht Kriterien, die zu dem Zweck einer zu kommenden sekundären Analyse nützlich sind. Interessant zu bemerken ist das Engagement der Einrichtung in die Ausbildung der Forscher, die zukünftig Daten erzeugen werden.

So kann sie die Forschung in einigen Gebieten besser beherrschen und die Redundante vermeiden.

Das SIDOS ist ein schweizerisches Informations- und Archivierungsdienst der Daten für die Sozialwissenschaften, wurde 1992 durch die schweizerische Akademie der Human- und Sozialwissenschaften geschaffen und stellt auch eine Art Verwaltungsagentur der qualitativen oder quantitativen Daten dar, die von den Forschern erzeugt werden.

Das SIDOS hält den Datenträger für einen Autor und jede Archivierungsbeschäftigung für eine Ausgabe der Daten und der Dokumentation.

Die Archivierung lenkt sich auf den Datenaustausch zwischen Forschern. Sie ist ein Instrument der Bereicherung der wissenschaftlichen Tätigkeit, unter der Bedingung, dass diese Daten später ordentlich verbreitet werden.

Die Verwirklichung solcher Netze wäre von unbestreitbarem Interesse für die Forschung. Wir sind nicht davon überzeugt, dass der patrimoniale Status und der Fortbestand dieser mündlichen Sammlungen dadurch besser garantiert wären.

WELCHE MÜNDLICHE QUELLEN FÜR MORGEN ?

Seit dem Beginn der Digitalisierung ist der langfristige Fortbestand immer noch problematisch, vor allem wegen dem Veraltetsein der Systemestandards und Systemekompatibilität. Der zukünftige Zusammenhang der Sammlungen wird aber durch die Modalitäten der Datenarchivierung umgestoßen. Seine Bestände zu spenden heißt für den Forscher eine echte *Ausgabearbeit* der Korpora und ihrer Dokumentation zu machen, um immer verständliche und zusammenhängende Dokumente zugänglich zu machen. Diese Arbeit sollte immer der Forscher machen. Wann wird er sich Zeit dazu nehmen ? Welches Bild seiner Arbeiten wird er

spenden wollen ? Welche Form sollte behalten werden ? Welches ist das Interesse für den zukünftigen Forscher ? Es besteht keine einzigartige Antwort.

Der Forscher, der Korpora der gesprochenen Sprache, die von anderen geschaffen wurden, gebrauchen möchte, braucht eine Vermittlung, das heißt eine Dokumentation, die die Variable beschreibt, aber auch das Sammeln der Daten und den Kontext des Projekts.

In diesem letzten Fall ist der Erzeuger-Forscher nicht derjenige, der am besten seine Daten beschreiben kann, deren Gebrauch von Personen gemacht wird, die seinem Gebiet nicht gehören. Es kommt dem Fachmann der dokumentarischen Ausarbeitung, dem Biblithekar, dem Dokumentalisten, dem Archivar zu, *mit Hilfe normalisierter und von allen verständlicher Mittel Korpora zu beschreiben, die für andere bestimmt sind.*

Eine zu genaue Beschreibung, wie *„nachträgliche", Aussagen, „Lebensgeschichten im nachhinein"*, die auf dem Begriff der Zeitlichkeit beruht, und ja für den Analysebedarf des Forschers nützlich ist, ist für die Verwaltung dieser Sammlungen innerhalb einer Bewahrungseinrichtung nicht wirksam. Diese Kriterien gehören ja zu der objektiven Beschreibung des mündlichen Dokuments, es kommt aber der Einrichtung zu, *sie in zu enge Kategorien zu ordnen, die schon von der Analyse kommen und die Freiheit der zukünftigen Gebraucher beschränkt, indem es ihren Standpunkt zwingt.*

Alles in allem ist der Korpuserzeuger sicher der einzige, der seine Korpora der gesprochenen Sprache dokumentieren kann. Ihr Gebrauch von Dritten wird erst möglich sein, wenn die Beschreibung von Fachleuten der Dokumentation gemacht wird.

4.3.2 *AUF DEM WEG ZU EINER ANERKENNUNG EINES STATUS DES MÜNDLICHEN ERBES*

Die Zukunft der mündlichen Quellen ist keine exklusiv juristische Frage. Diese Dimension kann durch vertragliche pragmatische Lösungen gemeistert werden. Dieses *Handbuch* hat keine andere Absicht, als es zu zeigen.

Was bei der Frage der mündlichen Quellen auf dem Spiel steht, hat mit Kultur und Politik zu tun. Ihre Anerkennung erfordert zugleich die Erarbeitung anspruchsvoller Auswahlkriterien, ohne die kein wohl benanntes Erbe existieren kann, und auch eine Erkenntnis der Gesellschaft, dass diesen wissenschaftlich erzeugten Dokumenten *ein Status von Erbeteil* verleiht werden soll.

Ihre Integration in die Einrichtung, die die Erbobjekte bestimmt, wird dann eine natürliche Sache sein.

Es ist zu bemerken, dass Frankreich erstaunlich hintergeblieben ist, was das immaterielle Erbe betrifft.

5 ABKÜRZUNGEN

- AFAS : (Association Française des détenteurs de documents Audiovisuels et Sonores), französischer Verein der Besitzer von audiovisuellen und Tondokumenten.
- AHICF : (Association pour l'HIstoire des Chemins de fer en France), Verein für die Bahngeschichte in Frankreich.
- API : (Alphabet Phonétique International), internationale Lautschrift.
- BnF : (Bibliothèque nationale de France), Nationalbibliothek Frankreichs.
- CAC : (Centre des Archives Contemporaines), Zentrum der zeitgenössischen Archive.
- CDF : (Centre de Droit Fondamental), Zentrum des Grundrechtes.
- CECOJI : (Centre d'Etudes sur la Coopération Juridique Internationale), Zentrum für das Studium über die juristische internationale Zusammenarbeit.
- CAMT : (Centre des Archives du Monde du Travail), Zentrum der Archive der Arbeitswelt.
- CELIA : (Centre d'Etude des Langues Indigènes d'Amérique), Zentrum für das Studium der einheimischen Sprachen Amerikas).
- CHAN : (Centre historique des Archives nationales), Historische Zentrum der Nationalarchive.
- CHILDES : (CHIld Language Data Exchange System), System für Datenaustausche über die Kindersprache.
- CLAPI : (Corpus de LAngue Parlée en Interaction), Korpus der gesprochenen Sprache in Interaktion.
- CNRS : (Centre National de la Recherche Scientifique), Nationales Forschungszentrum
- CORAL : (Cente Orléanais de Recherche en Anthropologie et Linguistique), Forschungszentrum von Orléans in Anthropologie und Linguistik.
- CPI : (Code de la Propriété Intellectuelle), Urheberrechtsgesetz.
- DASTUM : (organisme de sauvegarde du patrimoine oral et musical de la Bretagne, « dastum » heißt « zusammenstellen » in Bretonisch), Einrichtung für das Erhalten des bretonischen mündlichen Erbes.
- DGLFLF : (Délégation Générale à la Langue Française et aux Langues de France), oberster Rat für die Überwachung des Französisches und der Sprachen Frankreichs.
- DGRST : (Délégation Générale à la Recherche Scientifique et Technique), Allgemeine Delegation für die wissenschaftliche und technische Forschung.
- DMF : (Direction des Musées de France), Leitung der Museen Frankreichs.
- EHESS : (Ecole des Hautes Etudes en Sciences Sociales), Hochschule für Sozialwissenschaften.
- ELRA ELDA : (European Language Resources Association/ Evaluations and Language resources Distribution Agency), Europäischer Verein für die Sprachressourcen/ Agentur für die Sprachauswertungen und -ressourcen.
- ENS-LSH : (Ecole Normale Supérieure Lettres et Sciences Humaines), Hochschule für Geistes- und Humanwissenschaften.

- EPHE : (Ecole Pratique des Hautes Etudes), Praktische Hochschule für die Forschung.
- FAMDT : (Fédération des Associations de Musiques et Danses Traditionnelles), Verband der volkstümlichen Musiken und Tänze.
- ICAR : (Interactions, Corpus, Apprentissages, Représentations), Interaktionen, Korpora, Lernen, Darstellungen.
- ICOM : (International Council Of Meseums), internationale Organisation der Museensammlungen.
- INA : (Institut National de l'Audiovisuel), Nationales Institut des Audiovisuellen.
- IRD : (Institut de Recherche pour le Développement), Forschungsinstitut für die Entwicklung.
- JO : (Journal Officiel), Gesetzblatt.
- LACITO : (LAngues et CIvilisations à Tradition Orale), Sprachen und Kulturen mit mündlicher Tradition.
- LIMSI : (Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur), Informatiklabor für Mechanik und Ingenieurwissenschaften.
- MMSH : (Maison Méditerranéenne des Sciences de l'Homme), südländisches Haus der Humanwissenschaften.
- MNATP : (Musée National des Arts et Traditions Populaires), Museum für die volkstümlichen Künste und Überlieferungen.
- ORTF : (Office de la Radio-Télévision Française), französisches Rundfunk- und Fernsehamt.
- SIDOS (Service suisse d'information et d'archivage de données pour les sciences sociales), schweizerisches Dienst für die Information und die Datenarchivierung in den Humanwissenschaften.
- TAL : (Traitement Automatique des Langues), automatische Bearbeitung der Sprachen.
- TEI : (Text Encoding Initiative), Initiative für die Codierung der elektronischen Texte.
- UMR : (Unité Mixte de Recherche), Gemischte Forschungseinheit.
- Unicode : norme informatique, Informatische Norm.