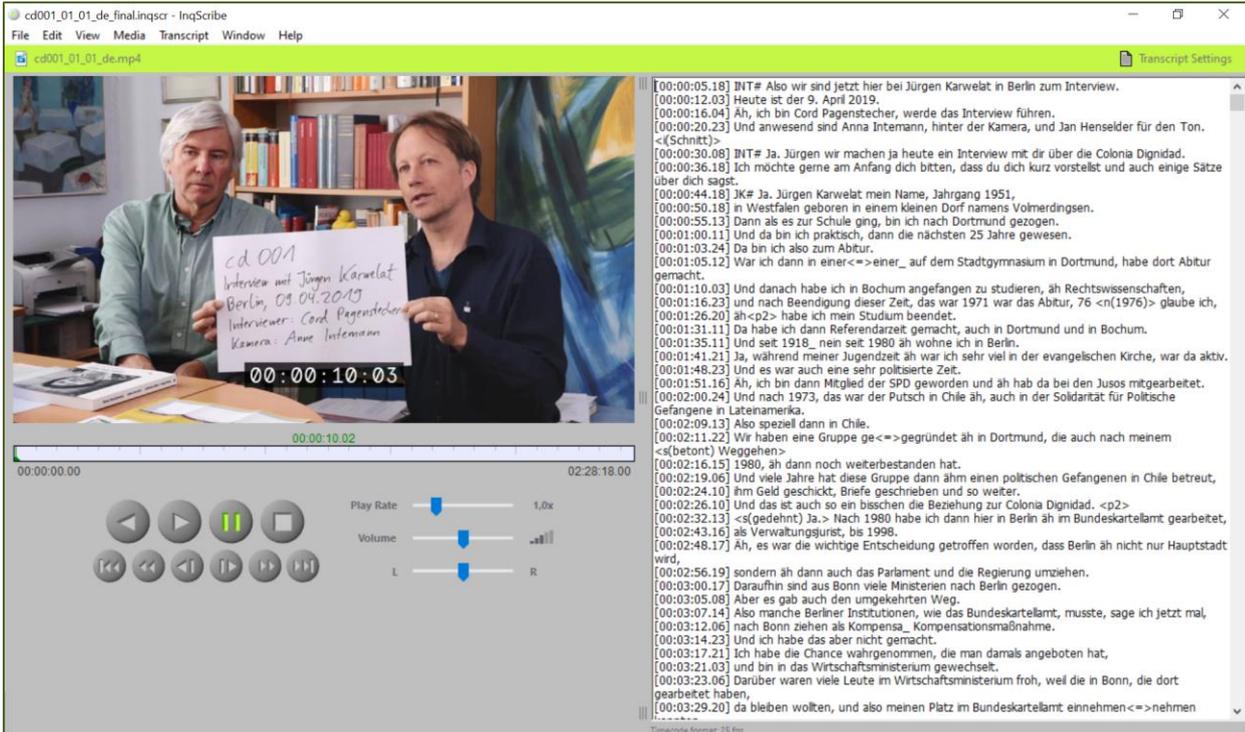


Richtlinien für die Transkription und Segmentierung narrativer Video-Interviews im Projekt „Colonia Dignidad. Ein chilenisch-deutsches Oral History-Archiv“

Freie Universität Berlin, Universitätsbibliothek, Center für Digitale Systeme (CeDiS)/
Universitätsbibliothek, Digitale Interview-Sammlungen, Projekt „Colonia Dignidad. Ein
chilenisch- deutsches Oral History-Archiv“

Dorothee Wein und Lydia Dollmann, unter Mitarbeit von Sophie Metzler, Peter Kompel,
Philipp Kandler und Cord Pagenstecher

Aktualisierte Fassung vom 20.12.2021



Beispieltranskript, Jürgen K., Interview cd001, 9.4.2019, Colonia Dignidad. Ein chilenisch-deutsches Oral History-Archiv,
<https://archiv.cdoh.net/de/interviews/cd001>

Inhaltsverzeichnis

Einführung	3
Struktur und Inhalt der Interviews.....	3
Zweck der Transkription	3
Wortgetreu, nicht lautgetreu	4
Transkriptionszeichen	4
Namen und Zahlen.....	4
Markierung der Transkriptionszeichen	5
Timecodes ([hh:mm:ss.ff]).....	5
Sprecher*innen-Zuordnung (INT#, KAM#).....	5
Auslassungen im Transkript (<res>)	5
Unterbrechung der Aufnahme (<i>).....	6
Pausen (<p2>).....	6
Undeutliche Äußerungen (<? xyz>).....	6
Unverständliche Äußerungen (<?5>)	6
Zitate, wörtliche Rede, Buch- und sonstige Titel („xyz“).....	6
Abgebrochene Wörter oder Sätze (_).....	6
Schneller Anschluss der Wörter (<=>)	6
Fremdsprachige Wörter (<l(es)>/ <l(es-de)>).....	6
Begriffe und Redewendungen im Dialekt (<ld>)	6
Nonverbale lautlich wahrnehmbare Äußerungen (<v>).....	7
Sprechweise (<s(lachend) xyz>, <s(laut) xyz>, <s(leise) xyz>, <s(leise) xyz>)	7
Simultansprechen (<sim>)	7
Interjektionen, Füllwörter und Verzögerungslaute	8
Anredeformen.....	8
Anmerkungen (<n>)	8
Ereignisse und Geräusche (<nl>)	8
Gestik (<g>)	8
Mimik (<m>)	9
Empfohlene Arbeitsweise	9
1. Arbeitsgang: Vorbereitung, Gesamteindruck	10
2. Arbeitsgang: Rohtranskript, Schwerpunkt Inhalt.....	10
3. Arbeitsgang: Feintranskript, Grunderschließung.....	10
4. Arbeitsgang: Feintranskript, Schwerpunkt Sprechweise.....	10
5. Arbeitsgang: Timecodes	10
6. Arbeitsgang: Rechtschreibung und Zeichensetzung.....	10
7. Arbeitsgang: Schlusskorrektur	10
Transkriptionszeichenliste	11

Einführung

„Colonia Dignidad – Ein chilenisch-deutsches Oral History-Archiv“ sammelt und erschließt Interviews mit Zeitzeug*innen einer deutschen Sektensiedlung im südlichen Chile. Zwischen 1961 und 2005 wurden die Sektenmitglieder und ihre Kinder isoliert, indoktriniert, ausgebeutet, gequält und sexuell missbraucht. Während der chilenischen Diktatur 1973 bis 1990 wurden Oppositionelle dort gefoltert und ermordet. Diese mit Wissen der deutschen Botschaft begangenen Verbrechen wurden bislang nur ungenügend aufgearbeitet.

Im Rahmen des Projekts sollen etwa 60 lebensgeschichtliche Interviews mit Bewohner*innen, politischen Gefangenen, Angehörigen und weiteren Zeitzeug*innen geführt werden. Die deutsch- oder spanischsprachigen Video-Interviews werden transkribiert, übersetzt und wissenschaftlich aufbereitet. Das auf drei Jahre angelegte Projekt der Freien Universität Berlin wird vom Auswärtigen Amt gefördert. Unter der Leitung von Prof. Stefan Rinke wird es von der Abteilung Geschichte des Lateinamerika-Instituts und dem Bereich Digitale Interview-Sammlungen des Centers für Digitale Systeme/ Universitätsbibliothek der Freien Universität Berlin umgesetzt.

Struktur und Inhalt der Interviews

Die Interviews, die offen und nicht standardisiert über das Leben der Befragten, ihre Erfahrungen in und mit der Colonia Dignidad sowie ihren Umgang mit dieser Geschichte danach geführt werden, folgen der narrativ-biographischen Methode. Die Zeitzeug*innen wählen als Interviewwort meist die vertraute häusliche Umgebung. Viele sprechen zum ersten Mal in dieser Ausführlichkeit über ihre Erinnerungen und Erfahrungen.

Die Interviews sind durchschnittlich ca. drei Stunden lang, das bisher längste Interview ist acht Stunden lang. Die Kamera ist meist statisch und nichtinterpretativ auf Kopf und Oberkörper des Interviewten gerichtet. Nur ganz zu Beginn ist die interviewende Person zusammen mit der Interviewten zu sehen. An einem Interview sind in der Regel nur zwei Sprecher*innen beteiligt, wobei die meiste Zeit der Interviewte erzählt. Die interviewende Person stellt in der Regel nur wenige Leit- und Zwischenfragen und erst im letzten Teil des Interviews Nachfragen. Hinzu kommen wenige Sätze der Kameraperson. Am Schluss werden meist historische und aktuelle Familienfotos, private Dokumente und ähnliche Objekte gezeigt und von den Interviewten kommentiert. Es gibt einzelne Interviews, in denen Zeitzeug*innen am Ende durch historische Orte und Schauplätze des Geschehens führen.

Zweck der Transkription

Der Bereich der Digitalen Interview-Sammlungen des Centers für Digitale Systeme der Freien Universität Berlin werden Interviews erschlossen und für Wissenschaftler*innen unterschiedlicher Disziplinen zugänglich und nutzbar gemacht. So ist bei der sozialwissenschaftlichen Interpretation narrativer Interviews „nicht nur der Wortlaut des Interviews wichtig wie bei einem journalistischen Interview, sondern auch die Art wie etwas gesagt wird.“¹¹ Die Transkriptionen sollen die Bearbeitung einer Vielzahl von Forschungsfragen erleichtern.

Aufbauend auf den Erfahrungen aus frühere Projekten des Bereichs der Digitalen Interviewsammlungen, bei denen Interview-Transkriptionen für den deutschsprachigen Bestand des Visual History Archive der „USC Shoah Foundation“², für die Interview-Archive „Zwangsarbeit 1939-1945“ und „Erinnerungen an die Okkupation in Griechenland“³ angefertigt wurden, bieten die vorliegenden Regeln eine stellenweise Vereinfachung und eine Erleichterung für eine Übersetzung in maschinenlesbare Formate. Um in Zukunft das Austauschformat TEI⁴ nutzen zu können, wurden die Transkriptionszeichen so gewählt, dass sie inhaltlich den oben aufgeführten Anforderungen entsprechen und leicht in ein TEI-kompatibles Format übersetzt werden können.

Bei der Transkription werden Zeitstempel (Timecodes) in den Text eingefügt, sodass der Text in Segmente geteilt wird. Diese Segmentierung dient dazu, den geschriebenen Text an die Videodatei zu koppeln. Dadurch können Transkript und Übersetzung als Untertitel im laufenden Video angezeigt werden.

¹ Ivonne Küsters, Narrative Interviews. Grundlagen und Anwendungen, Wiesbaden 2006, S. 73.

² „Die Transkription und Übersetzung der Interviews der USC Foundation“. Vgl. Woggon, Helga: Transkription und Übersetzung. Video-Interviews als Lesetexte, in: Zeugen der Shoah. Die didaktische und wissenschaftliche Arbeit mit Video-Interviews des USC Shoah Foundation Institute, Berlin 2012, https://www.zeugendershoah.de/unterrichtsmaterialien/vha_broschuere/VHAS_Broschuere_Web.pdf, S. 24-28

³ Transkriptionsrichtlinien für das Interview-Archiv „Erinnerungen an die Okkupation in Griechenland“

⁴ TEI ist ein von der Text Encoding Initiative entwickeltes Austauschformat zur maschinenlesbaren Kodierung von Texten. Zur Sprachtranskription gibt es die TEI Guidelines, P5, chapter 8: Transcription of Speech (<http://www.tei-c.org/release/doc/tei-p5-doc/de/html/TS.html>). Darauf basiert der Standard ISO 24624:2016 (<https://www.iso.org/standard/37338.html>), vgl. Schmidt, Thomas/Hedeland, Hanna/Jettka, Daniel: Conversion and annotation web services for spoken language data in CLARIN. Selected papers from the CLARIN Annual Conference 2016. Linköping Electronic Conference Proceedings 136, 2017, 113-130, www.ep.liu.se/ecp/article.asp?issue=136&article=009.

Außerdem kann die Nutzer*in dadurch per Volltextsuche an eine spezifische Stelle des Video-Interviews springen.

Die Transkriptionen dienen folgenden Zwecken:

- Sie dienen als Grundlage für die Übersetzungen ins Deutsche bzw. ins Spanische.
- Sie ermöglichen im Online-Archiv die Volltextsuche in den Interviews.
- Sie können als spanische und deutsche Untertitel eingeblendet werden.
- Sie werden im Online-Archiv als mitlaufende Transkription/Übersetzung angezeigt.
- Sie werden als gesonderter Lesetext (PDF) bereitgestellt.

Auf der Interviewplattform werden unterschiedlich ausführliche Versionen des Transkripts angezeigt: Die detaillierte Auszeichnung von Wörtern und Textstellen (insbesondere die Kennzeichnung nonverbaler Phänomene) wird in der ausführlichen PDF-Version angezeigt, damit die Video-Interviews nach allem Hör- und Sehbareren durchsucht und mediengerecht interpretiert werden können.

In der Untertiteldarstellung dagegen wird die Detailauszeichnung ausgeblendet, da hier die Transkription zusammen mit dem Video angesehen wird und dieses über die auditive und die visuelle Ebene alle Informationen vermittelt. Damit diese Versionen im Online-Archiv automatisch generiert werden können, ist die genaue Anwendung der Timecodes und Transkriptionszeichen mit Spitzklammern unabdingbar.

Wortgetreu, nicht lautgetreu

Die Interviews werden wortgetreu und nicht lautgetreu in phonetischer Umschrift transkribiert. Alle Eigenheiten der gesprochenen Sprache (auch Füllwörter, Interjektionen, Wiederholungen, Fehler) bleiben jedoch im Transkript erhalten. Die Reihenfolge und der Wortlaut des Transkripts entsprechen exakt dem, was die interviewte Person sagt. Markiert werden Wort- und Satzabbrüche, Pausen sowie Dehnlaut und laut und leise Gesprochenes. Zudem werden in der Transkription auch Mimik und Gestik des Interviewten vermerkt, jedoch nur dann, wenn diese sehr auffällig ist.

Für eine bessere Lesbarkeit der Transkription und ihre Darstellung als Untertitel sollen eine normale Interpunktion und grammatische Regeln angewendet werden, d. h. Fragesätze, Ausrufesätze usw. werden mit Frage- oder Ausrufezeichen gekennzeichnet.

Als Richtlinie gilt die aktuelle Standardrechtschreibung und Standardaussprache. Es gilt die Groß- und Kleinschreibung und Interpunktion nach Duden. Maßstab für die Standardaussprache sind die Laut- Buchstaben-Zuordnungen nach der amtlichen Regelung der deutschen Rechtschreibung⁵ und das Aussprachewörterbuch, z. B. ist „wia habn wieda“ die Standardaussprache für die Schreibung „wir haben wieder“ und wird also „wir haben wieder“ geschrieben.⁶

Regionalsprachliche Eigenheiten in Grammatik, Syntax etc., z. B. lateinamerikanisches „ustedes“ für „vosotros“, werden unverändert transkribiert. Regionalspezifische Aussprachen werden nicht verschriftlicht, da die Transkription wortgetreu, aber nicht lautgetreu erfolgt. Das betrifft z. B. die argentinische Aussprache bei „calle“, „yo“ etc. Hier wird in der Standardschreibweise transkribiert.

Das Interviewtranskript sollte als Lesetext, unabhängig von dem Video, verständlich und gut zu lesen sein. Für die Lesbarkeit des Transkripts ist darauf zu achten, dass die Sätze eine überschaubare Länge haben. Falls die Sprechenden zu sehr langen Sätzen neigen, müssen Transkribierende versuchen, an geeigneter Stelle einen Punkt, ein Frage- oder ein Ausrufezeichen zu setzen. Mitunter beginnt der nächste Satz dann mit „Und“ oder „Denn“ usw. Wichtig ist, dass der Satz syntaktisch unabhängig ist.

⁵ https://www.rechtschreibrat.com/DOX/rfdr_Regeln_2016_redigiert_2018.pdf.

⁶ Duden, Bd.6, 6. Aufl., 2005, bes. S.34 ff.

Transkriptionszeichen

Namen und Zahlen

Alle Eigennamen (von Orten, Straßen, Personen, Organisationen) werden in korrekter Form geschrieben. Wenn diese nicht bekannt sind, bitte im Internet und ggf. mit den zur Verfügung gestellten Hilfsmitteln die korrekte Schreibweise recherchieren. Ist die Schreibweise nicht feststellbar, wird der Name nur nach Gehör transkribiert und entsprechend der Regel für unsichere Transkription markiert (<? hier steht die unsichere Transkription>). Bei unvollständig ausgesprochenen Jahreszahlen wird die korrekte Jahreszahl als Anmerkung ergänzt und mit dem Kürzel <n> für notes gekennzeichnet. Ein Beispiel ist: „Der wohnte ab 77<n (1977)> am Kudamm<n (Kurfürstendamm)>“.

Zahlen, Ordnungszahlen und Daten werden grundsätzlich nach Duden geschrieben. Im Fließtext werden Zahlen bis zwölf sowie volle (runde) Zehner, Hunderter, Tausender ausgeschrieben, wenn es sich nicht um eine Maßangabe, Datumsangabe oder andere rechnerische Angaben handelt, z. B. „Da waren vielleicht dreißig Leute vor dem Haus. Ich war damals gerade 13 geworden.“ Jahreszahlen werden beispielsweise wie gesprochen geschrieben: 1980er, 80er oder Achtzigerjahre. Werden Jahreszahlen nicht vollständig ausgesprochen, wird die vollständige Zahl mit der Anmerkungs-Kennzeichnung ergänzend hinzugefügt, z. B. „Das war so dreiun-, nee, 34<n (1934)>.“ Wenn Zahlen Teil eines Begriffs/Namens sind, müssen sie in der Standardschreibweise auftauchen, z. B.: Block 10, 196. Grundschule, Drei-Zimmer-Wohnung.

Markierung der Transkriptionszeichen

Fast alle Transkriptionszeichen werden mit den als Klammern verwendeten ASCII-Zeichen „Kleiner als“ < begonnen und mit „Größer als“ > beendet (im Folgenden als „spitze Klammern“ bezeichnet). Diese Markierung des gekennzeichneten Textes ermöglicht eine maschinenlesbare Verarbeitung der Anmerkungen. **Jede geöffnete Klammer muss innerhalb eines Segments, also vor dem nächsten Timecode, geschlossen werden.**

Innerhalb einer Klammer können neue Klammern geöffnet und geschlossen werden. Sie dürfen sich also verschachteln, aber nicht überlappen, z.B.:

[hh:mm:ss.ff] INT# <s(lachend) so war das also <s(laut) Echt?>>

[hh:mm:ss.ff] MM# <s(lachend) Ja so war das. <s(laut) OhneWitz!>>

An diesem Beispiel ist zudem ersichtlich, dass Spezifizierungen in der Transkription immer in runden Klammern „()“ stehen.

Im Folgenden sind alle Auszeichnungen detailliert erklärt, am Ende dieses Dokumentes befindet sich eine Übersichtstabelle.

Timecodes ([hh:mm:ss.ff])

Vor jedem Sprecher*innenwechsel, nach jedem Satz oder jeder Sinneinheit, nach maximal 120 Zeichen muss ein *Timecode* gesetzt werden. Die Angabe von maximal 120 Zeichen (ohne Transkriptionszeichen) ist als Schätzwert zu verstehen. [hh:mm:ss.ff] ist das empfohlene Format des Zeitstempels, das Transkriptionsprogramme wie InqScribe automatisch setzen. ff entspricht den 25 Frames pro Sekunde Video. Um die Genauigkeit der Timecodes zu gewährleisten soll das Video zum Setzen eines Timecodes angehalten werden. Aufeinanderfolgende Timecodes dürfen dabei nie exakt identisch sein. Zur besseren Lesbarkeit werden Timecodes in eine neue Zeile gesetzt und danach ein Leerzeichen eingefügt, z. B.:

[hh:mm:ss.ff] MM# <s(laut) Ja, echt!>

Sprecher*innen-Zuordnung (INT#, KAM#)

Jeder Sprecher*innenwechsel wird mit dem jeweiligen Initialen und einer Raute (Hashtag) markiert. Dabei steht *INT* immer für die Interviewer*in und *KAM* immer für die Kameraperson. Für die interviewte Person wird immer der erste Buchstabe vom Vor- und Nachnamen verwendet, z. B. für Maria Malta *MM*. Für jede weitere an dem Interview beteiligte Person wird ebenso der erste Buchstabe vom Vor- und Nachnamen verwendet. Die Namen und Kürzel der Beteiligten stehen im Interviewprotokoll, das die Transkribierenden erhalten.

Die Raute wird anstatt eines Doppelpunktes eingesetzt, da dieser auch an anderen Stellen im Transkript als Satzzeichen vorkommen und daher für die Kennzeichnung nicht genutzt werden kann.

Vor jedem Sprecher*innenwechsel steht ein Timecode:

[hh:mm:ss.ff] KAM# immer für die Kameraperson

[hh:mm:ss.ff] INT# immer für die Interviewer*in

[hh:mm:ss.ff] MM# erster Buchstabe vom Vor- und Nachnamen für die interviewte Person, z. B. hier MM für Maria Malta.

[hh:mm:ss.ff] DT# jeweils erster Buchstabe vom Vor- und Nachnamen für jede weitere am Interview beteiligte Person, z. B. DT für Doris Tausendfreund.

Unterbrechung der Aufnahme (<i>)

Eine Unterbrechung (*interruption*) des Interviews bzw. der Aufnahme wird mit <i> gekennzeichnet. Ist der Grund der Unterbrechung bekannt, wird dies in einer Klammer angegeben, z. B.

„<i(Batteriewechsel)>“. Da die Kamera in der Zeit der Unterbrechung ausgestellt wurde bzw. das Interview in einer neuen Datei fortgeführt wird, kann die Länge der Unterbrechung von den Transkripteur*innen nicht erkannt werden. Auch der Grund für eine Unterbrechung ist nicht immer bekannt, sollte aber in dem von den Interviewenden erstellten Protokoll stehen, das neben der Transkription online gestellt wird.

Pausen (<p2>)

Kommt es zu Pausen während der Erzählung, in denen nicht gesprochen wird, werden diese folgendermaßen dargestellt: <p2> für eine Pause von ca. zwei Sekunden, <p3> für eine Pause von ca. drei Sekunden usw. Sprechpausen bitte nur kennzeichnen, wenn sie deutlich als Sprechpausen auszumachen sind, also in der Regel erst ab ca. zwei Sekunden.

Undeutliche Äußerungen (<? xyz>)

Ist der Inhalt der Äußerung nur schwer zu verstehen bzw. bei unsicherer Transkription, werden in spitze Klammern ohne Leerzeichen ein Fragezeichen und der vermutete Text eingesetzt: <? hier steht die unsichere Transkription>.

Unverständliche Äußerungen (<?5>)

Ist der Inhalt der Äußerung komplett unverständlich, wird dies mit einem Fragezeichen und der Zeitangabe kenntlich gemacht. Wie bei der Pause wird die Länge der unverständlichen Äußerung mit der Angabe von Sekunden markiert, z. B. bei einem unverständlichen Abschnitt von etwa vier Sekunden: <?4>.

Zitate, wörtliche Rede, Buch- und sonstige Titel („xyz“)

In Anführungsstriche werden die in der Erzählung genannten Zitate, wörtliche Rede, Buch- und sonstigen Titel gemäß den in den Originalsprachen geltenden Regeln (in Deutsch „unten - oben“) gesetzt, z. B.: Die Frau sagte: „Das war eine schlimme Zeit.“ oder bei Quellenangaben: „Erinnern an Zwangsarbeit. Zeitzeugen-Interviews in der digitalen Welt“.

Abgebrochene Wörter oder Sätze (_)

Ein Unterstrich kennzeichnet ein in der Erzählung abgebrochenes Wort oder einen abgebrochenen Satz. Dabei wird direkt im Anschluss an den gesprochenen Teil des Wortes oder Satzes der Unterstrich gesetzt (z. B. Ich wollte arbeiten, aber_ Können Sie mir bitte etwas zu trinken geben?). Das gleiche Vorgehen gilt bei einem abgebrochenen Gedankengang oder einer Formulierung.

Schneller Anschluss der Wörter (<=>)

In sehr auffälligen Fällen werden schnell und ohne Pause hintereinander gesprochene Wörter mit Gleichheitszeichen markiert: Aber<=>aber.

Fremdsprachige Wörter (<l(es)>/ <l(es-de)>)

Aussagen oder einzelne Wörter in einer anderen Sprache als der (Haupt-)Sprache des Interviews werden mit „l“ für *lexical phenom*en und dem Sprachkürzel nach ISO 639-1, z. B. „l(de)“ für Deutsch, „l(es)“ für Spanisch oder „l(en)“ für Englisch, gekennzeichnet. Ein Beispiel ist „There was the <l(de) Pisspott> in the corner“. Dies betrifft nicht die durchgängig zweisprachigen Interviews.

Manchmal kommen in den Interviews auch Begriffe vor, die aus dem Spanischen stammen, aber eingedeutscht verwendet werden, diese werden mit <l(es-de)> oder im umgekehrten Fall mit <l(de-es)> ausgezeichnet.

z.B. MM# Und dann hat er sich schnell wieder <l(es-de) rekuferiert>

Auch spanischsprachige Begriffe, die in die Sprache der Colonia Dignidad übernommen wurden und im Glossar erläutert werden, werden mit <l(es)> ausgezeichnet, dann aber groß geschrieben.

Das sind: Pito, Fundo, Fichas, Jeraquía, Galpón.

AUSNAHME: Colono/ Colona als Bezeichnung für die Bewohner der (ehemaligen) Colonia Dignidad gilt als Eigennamen und wird daher nicht ausgezeichnet.

Umgekehrt werden auch die deutschsprachigen Begriffe, die aus der Sprache der Colonos stammen und in den spanischsprachigen Interviews verwendet werden im Transkript mit <l(de)> ausgezeichnet.

Das sind u.a.: Kinderhaus, Freihaus, Zippelhaus, Keile

Begriffe und Redewendungen im Dialekt (<ld>)

Bei Erzählungen, in denen größtenteils Hochsprache gesprochen wird, werden im Dialekt gesprochene Wendungen mit einem kleinen L und D (*lexical phenomén, dialect*) gekennzeichnet, z. B. „<ld Nachtigall ick hör dir trapsn.>“ Wenn die Art des Dialektes erkannt wird, soll sie vermerkt werden, z. B.: <ld(Berliner Dialekt) Nachtigall ick hör dir trapsn.>

Nonverbale, lautlich wahrnehmbare Äußerungen (vocal but not verbal) (<v>)

Damit die für die Auswertung meist sehr wichtige nonverbale Kommunikation im Transkript berücksichtigt wird, werden auffällige nonverbale, aber lautlich wahrnehmbare Phänomene (*vocal, but non-lexical phenomén*) mit <v> markiert. Die Art der nonverbalen Äußerung wie Weinen und Lachen wird in Klammern ohne Leerzeichen direkt hinter dem v markiert, z. B.: <v(Lachen)>.

<v>
<v(schweres Atmen)>
<v(Schlucken)>
<v>Weinen)>
<v(Lachen)>
<v(Räuspern)>
<v(Husten)>
<v(Seufzen)>
<v(Gähnen)>

Sprechweise (<s(lachend) xyz>, <s(laut) xyz>, <s(leise) xyz>, <s(weinend) xyz>)

Wird während eines nonverbalen Phänomens gesprochen, kann dieses als Veränderung in der Art zu sprechen charakterisiert werden (*shift in vocal quality*) und wird mit <s> markiert. Dann werden die Wörter oder Erzählsequenzen, die die interviewte Person beispielsweise weinend oder lachend äußert, wie folgt gekennzeichnet: <s(weinend) Ich wollte das nicht> oder <s(lachend) Ich wollte das nicht>. Die Art der Sprechweise steht in Klammern ohne Leerzeichen direkt hinter dem <s, der gesprochene Text folgt dann nach einem Leerzeichen.

Besonders laut und/oder betont gesprochene Wörter oder Sequenzen werden mit <s(laut)> kenntlich gemacht, z. B. MM# <s(laut) Ich wollte das nicht>.“ Auch besonders gedehnt gesprochene Wörter werden mit <s> markiert, z. B. „<s(gedehnt) Ja>, das war so.“ Besonders leise gesprochene Wörter oder Sequenzen werden mit <s(leise)> kenntlich gemacht, z. B. „<s(leise) Ich wollte das nicht>.“

Sprechweisen sollen nach der folgenden Liste bezeichnet werden:

<s>
<s(betont)>
<s(gedehnt)>
<s(leise)> <s(tief)>
<s(laut)> <s(hoch)>
<s(lachend)>
<s(lächelnd)>
<s(stockend)>
<s(traurig)>
<s(flüsternd)>
<s(weinend)>
<s(zitternde Stimme)>

Simultanes Sprechen (<sim>)

Sprechen mehrere Beteiligte gleichzeitig, wird dies mit <sim> als Kürzel für *simultan* gekennzeichnet. Vor jedem Sprecher*innenwechsel wird ein Timecode gesetzt. Dadurch entsteht ein neues Segment, das, so lange gleichzeitig gesprochen wird, erneut mit <sim> gekennzeichnet wird. Aufeinanderfolgende Timecodes dürfen dabei nie exakt identisch sein, sondern müssen sich mindestens im Bereich der Framerate oder der Sekundenangabe unterscheiden.

Beispiel: [01:23:45.67] INT# Wo waren Sie <sim während dieser Zeit?> [01:23:46.12] MM# <sim Ich war bei Maria Malta,> als das passierte. In dem Beispiel sind „während dieser Zeit?“ und „Ich war bei Maria Malta“ simultan gesprochen. Bei längeren Passagen, in denen wiederholt mit kurzen Unterbrechungen gleichzeitig gesprochen wird, sodass eine Transkription nach dieser Regel schwer lesbar wäre, kann vor der Passage eine entsprechende Anmerkung vermerkt werden. Erstreckt sich das gleichzeitige Sprechen über mehrere Segmente oder Minuten, werden Anfang und Ende notiert: <sim(Anfang)> xyz <sim(Ende)>.

Interjektionen, Füllwörter und Verzögerungslaute

Interjektionen, Füllwörter und Verzögerungslaute gehören auch in die Transkription. Grundsätzlich sollte jedes gesprochene Wort mit in die Transkription aufgenommen werden. Zumal diese Wörter beispielsweise Empfindungen zum Ausdruck bringen, den Sprachrhythmus wiedergeben oder aufzeigen, wenn die Person nachdenkt. Am häufigsten sind Verzögerungslaute wie äh, ähm, die die Überlegungspausen füllen, und Interjektionen der Aufmerksamkeitsbekundung (hm, hmhm).

Interjektionen sind im Duden verzeichnet und werden nach Duden geschrieben. Nur falls es keine festgelegte Schreibweise gibt, wird nach Gehör transkribiert.

Anredeformen

Zum besseren Verständnis sollten die Anredeformen Sie, Ihnen, Ihre, Ihres, Ihrem und Ihren großgeschrieben werden. Kleingeschrieben werden (außer am Satzanfang) – alle von „du“ und „ihr“ abgeleiteten Anredeformen („du, dir, dich, dein/e, deiner/m/n, ihr, euch, euer, eure/m/n) – alle von der 3. Person Sg. fem. und 3. Person Pl. abgeleiteten Formen des Personalpronomens „sie“.

Anmerkungen (<n>)

Anmerkungen und Ergänzungen (*notes*) erfolgen beispielsweise bei der Korrektur offensichtlicher Irrtümer werden mit <n(Erläuterung)> gekennzeichnet, z. B.: Der wohnte ab 77 <n(1977)> am Kudamm <n(Kurfürstendamm)>! Oder: nach dem Mauerbau 51 <n(gemeint ist 1961)>. Die Anmerkung steht also in runden Klammern direkt hinter dem <n(.

Ereignisse und Geräusche (<nl>)

Ereignisse und Geräusche (*non-linguistic incidents*) wie Telefonklingeln, Straßengeräusche, der Eintritt einer anderen Person oder das Herunterfallen eines Gegenstands usw. werden beispielweise mit <nl(Telefonklingeln)> kenntlich gemacht. Wenn dabei gesprochen wird, wird das Gesagte innerhalb der spitzen Klammern angegeben, z. B.: <nl(Telefonklingeln) Wir haben uns immer verstanden, aber>“. Erstreckt sich ein Geräusch über mehrere Segmente, werden nur der Anfang des Geräuschs und sein Ende notiert: <nl(Telefonklingeln Anfang)> transkribierte gesprochene Sprache <nl(Telefonklingeln Ende)>.

Auch wenn sich die interviewte Person beispielsweise ein Glas Wasser eingießt oder trinkt, wird dieses Ereignis und Geräusch mit <nl>ausgezeichnet, z.B. <nl(MM trinkt Wasser)>.

Für die Geräusche sollen die unten angegebenen Begriffe verwendet werden. Die Liste ist nicht vollständig. Geräusche, die das Interview beeinflussen und dort nicht aufgeführt sind, sollen von den Transkribierenden den Projektverantwortlichen mitgeteilt werden.

<nl>
<nl(Türklingeln)>
<nl(Straßengeräusche)>
<nl(Gesprächsgeräusche)>
<nl(Hundebellen)>
<nl(undefinierbares Geräusch)>
<nl(MM trinkt Wasser)>
<nl(Kamerageräusch)>

<nl(Fernsehgeräusch)>
<nl(Handklappe)>

Gestik (<g>)

Die Gestik (*gesture*) beim Interviewten wird nur gekennzeichnet, wenn sie sehr auffällig ist, z. B.

<g(ablehnende Handbewegung)>, oder, wenn sie das Gesagte verändert: <g(ablehnende Handbewegung) der dachte, dass ich das machen würde>.

Gestik soll mit den folgenden Begriffen ausgezeichnet werden. Diese gelten als Standard. Bedeutungstragende Gestik, die damit nicht abgedeckt werden kann, soll den Projektverantwortlichen mitgeteilt werden.

<g>
<g(Gestikulieren)>
<g(Achselzucken)>
<g(ablehnende Handbewegung)>
<g(hebt den Finger)>
<g(zählt mit den Fingern auf)>
<g(Kopfschütteln)>
<g(darstellende Handbewegungen)>
<g(Anführungszeichengebärde)>

Mimik (<m>)

Die Mimik (*mimic*) beim Interviewten wird ebenfalls nur gekennzeichnet, wenn sie sehr auffällig ist. Z. B.: <m(Stirnrunzeln) Was soll ich dazu sagen?>.

Sehr auffällige Mimik soll mit den folgenden Begriffen ausgezeichnet werden. Diese gelten als Standard. Bedeutungstragende Mimik, die damit nicht abgedeckt werden kann, soll den Projektverantwortlichen mitgeteilt werden.

<m>
<m(Stirnrunzeln)>
<m(lächelt)>
<m(rümpft die Nase)>
<m(traurig)>
<m(kämpft mit Tränen)>
<m(verzieht das Gesicht)>
<m(schließt die Augen)>

Empfohlene Arbeitsweise

Zu jedem Interview werden weitere Informationen wie die Kurzbiografie des Interviewten zur Verfügung gestellt. Dies dient dem besseren Verständnis des Gesprochenen und der korrekten Schreibweise von Namen.

Die transkribierende Person muss sehr genau hinhören können, was die im Interview aufgenommenen Menschen sagen. Gute Kopfhörer und eine ruhige Umgebung sind hilfreich.

In der gesprochenen Sprache reden wir nicht in grammatisch strukturierten, ausformulierten Sätzen, sondern formen die Gedanken nach und nach zu einem „Satz“, in dem immer wieder Denk- und Formulierungspausen auftreten und Sätze nicht eindeutig und korrekt beendet werden. Dies sollte in den Transkripten berücksichtigt werden.⁸ Dennoch gilt die Form grammatischer Sätze, d. h. jeder Satz beginnt mit einem Großbuchstaben und endet mit einem Satzschlusszeichen.

Die Transkription von der gesprochenen in die geschriebene Sprache, von Lauten in Buchstaben, ist in gewisser Weise vergleichbar einer „Übersetzung“. Wie bei Übersetzungen, kommt es auch bei der Transkription darauf an, zuerst den Gesamtzusammenhang (Inhalt und Tenor der individuellen Sprech- und Erzählweisen) zur Kenntnis zu nehmen, bevor man beginnt zu transkribieren. Zudem müssen mehrere unterschiedliche Phänomene der gesprochenen Sprache in das Transkript aufgenommen werden, die man beim Zuhören schwer alle gleichzeitig beachten kann, daher werden mehrere Arbeitsdurchgänge durch das Interview empfohlen.

Allgemein empfiehlt es sich, jeweils ein Interview möglichst in kompletten Arbeitsdurchgängen zu transkribieren, anstatt sich zu lange an einzelnen Abschnitten aufzuhalten. Dadurch behalten Sie den Gesamtzusammenhang des Interviews und ihr Arbeitstempo im Blick. Wollten Sie für jede Interviewminute alle Arbeitsgänge jeweils segmentweise bis zur Schlusskorrektur erledigen, wäre es ein frustrierend langsames Vorankommen.

Teilen Sie die Gesamtaufgabe jeweils in überschaubare Arbeitsgänge und konzentrieren Sie sich bei jedem Arbeitsgang jeweils nur auf eine genau begrenzte Aufgabe. Wir empfehlen Ihnen die folgenden Arbeitsgänge. Im Laufe der Arbeit werden Sie das für Ihre individuelle Arbeitsweise effektivste Vorgehen herausfinden.

1. Arbeitsgang: Vorbereitung, Gesamteindruck

Machen Sie sich mit dem Gesamt-Interview vertraut und lesen Sie die Kurzbiografie. Diese Informationen helfen Ihnen beim Verständnis und bei der Schreibweise einzelner Namen, Orte oder Begriffe aus dem Interview.

2. Arbeitsgang: Rohtranskript, Schwerpunkt Inhalt

In vielen Fällen finden Sie zu Beginn Ihrer Arbeit bereits ein automatisch durch ein Spracherkennungsprogramm erstelltes – und entsprechend fehlerhaftes – Rohtranskript vor. Diese Rohversion überarbeiten Sie bei laufendem Video von Anfang bis Ende. Wenn Sie bereits einen Eindruck von den unverfälschten Stimmen der Sprechenden haben, können Sie mit Hilfe der Transkriptions-Software die Sprechgeschwindigkeit reduzieren, z. B. auf 85%, was die Stimmen relativ unverzerrt belässt. Damit können Sie beim Mitschreiben besser dem Sprechtempo folgen. Bleiben Sie dem gesprochenen Text immer möglichst nah „auf den Fersen“.

Ist der Inhalt der Äußerung nur schwer zu verstehen bzw. bei unsicherer Transkription, wird diese als unsicher <?> gekennzeichnet. In den weiteren Arbeitsgängen können diese Stellen sukzessive mit dem bislang unverstandenen Text vervollständigt werden. Wenn der Inhalt auch bei mehrmaligen Hören nicht verstanden wird, wird die Kennzeichnung beibehalten.

Falls nicht schon Timecodes im Text vorhanden sind, sollten Sie in größeren Abständen einen Timecode setzen, etwa alle zehn Minuten oder bei Arbeitspausen, damit Sie rasch wieder an die gewünschte Stelle Ihrer Transkription zurückfinden.

3. Arbeitsgang: Timecodes

Die Segmentierung setzt Zeitstempel (Timecodes, TC) in den Text.

Setzen Sie bei jedem Sprecher*innenwechsel und nach jedem Satz einen Timecode. In der

gesprochenen Sprache ist nicht immer eindeutig, wann ein Satz beendet ist. Achten Sie dann auf die Sprachmelodie und wählen Sie eine geeignete Sinneinheit. Setzen Sie **nach maximal 120 Zeichen** auf jeden Fall einen Timecode, damit die daraus erzeugten Untertitel nicht mehr als zwei Zeilen umfassen. Das Setzen eines TC sollte mit einem Absatzwechsel kombiniert werden, das erhöht die Übersichtlichkeit des Transkripts.

4. Arbeitsgang: Feintranskript, Grunderschließung

Nun ergänzen und korrigieren Sie nach und nach die weiteren Details, die die Richtlinien vorgeben. Prüfen Sie dabei in gesamten Arbeitsdurchgängen die folgenden Arbeitsschritte:

- Ergänzen Sie nun nach erneutem Anhören alle Stellen, die Sie beim ersten Durchgang als unverständlich übersprungen hatten.
- Achten Sie auf Wortabbrüche, Satzabbrüche, Pausen etc. und notieren diese mit den vorgegebenen Zeichen (s. o.).
- Vermerken Sie starke Betonungen, nonverbale Phänomene und besonders auffällige Gesten oder Mimik sowie Simultansprechen entsprechend den Richtlinien.

5. Arbeitsgang: Rechtschreibung und Zeichensetzung

Nun prüfen Sie ihr Transkript auf Rechtschreibung und Zeichensetzung.

Überprüfen Sie insbesondere die Namen von Orten, Personen und Institutionen und fügen Sie ggf. die korrekten Schreibweisen als Anmerkung hinzu (<n(xyz)>). Halten Sie das Ergebnis Ihrer Recherchen und offene Fragen in einer kurzen Anmerkung an das Projektteam fest.

6. Arbeitsgang: Schlusskorrektur

Prüfen Sie ihr Transkript noch einmal im Ganzen in zwei Schritten:

- Lesetext: Lesen Sie den ausgedruckten Gesamttext des Transkripts wie einen selbständigen Text, ohne das Video dabei anzuschauen. Markieren Sie beim Lesen Wörter, Sätze oder Passagen, die Ihnen unverständlich vorkommen.
- Übereinstimmung: Vergleichen Sie den Gesamttext des Transkripts mit dem Video, prüfen Sie die exakte Übereinstimmung des Wortlauts und achten Sie dabei besonders auf die als unverständlich markierten Stellen. Es kann sein, dass ein Satz unverständlich erscheint, weil Sie vergessen haben, den Abbruch eines Wortes oder eines Satzes zu markieren, eine nonverbale Besonderheit zu kommentieren, wichtige Pausen anzugeben, oder weil andere Transkriptions- oder einfache Tippfehler vorkommen.
- Überprüfen Sie die Passgenauigkeit der Timecodes.
- Unklarheiten: Überprüfen Sie alle verbliebenen Unklarheiten und notieren Sie sie für Ihre Besprechung mit dem Qualitätsmanagement.

Transkriptionszeichentabelle

Transkriptionszeichen	Beispiel	Erklärung
[hh:mm:ss.ff]	[01:23:45.67]	Vor jedem Sprecher*innenwechsel, nach jedem Satz, jeder Sinneinheit, nach maximal 120 Zeichen muss ein Timecode gesetzt werden. [hh:mm:ss.ff] ist das empfohlene Format, ff entspricht den 25 Frames
INT#, KAM#, MM#	[01:23:45.47] INT# Wo waren Sie während dieser Zeit? [01:23:45.67] MM# Ich war bei Maria Malta, als das passierte.	INT steht immer für die Interviewer*in, KAM für die Kameraperson. Für die interviewte Person wird immer der erste Buchstabe vom Vor- und Nachnamen verwendet, z. B. für Maria Malta MM. Für weitere an dem Interview beteiligte Person wird ebenso der erste Buchstabe vom Vor- und Nachnamen verwendet. Raute/Hashtag # wird statt Doppelpunkt gesetzt, damit Sprecher*innenmarkierungen automatisch
<n>	Der wohnte ab 77 <n(1977)> am Kudamm <n(Kurfürstendamm)>.	Anmerkungen und Ergänzungen (notes) wie die Korrektur offensichtlicher Irrtümer werden mit <n(Erläuterung)> gekennzeichnet.
<i>	Wo waren Sie während dieser Zeit? [01:23:45.67] <i(Batteriewechsel)> [01:23:45.67] MM# Ich war	Eine Unterbrechung (interruption) des Interviews bzw. der Aufnahme wird mit <i> gekennzeichnet. Ist der Grund der Unterbrechung bekannt, wird er in einer Klammer angegeben, z. B.
<p2>, <p3> usw.	Ich war <p3> bei Maria Malta, als das passierte.	Sprechpausen werden mit der Länge in Sekunden gekennzeichnet,
<? xyz>	Ich war bei Maria <? Meier>, als das passierte.	Eine undeutliche, schwer verständliche Passage wird mit <? hier steht die unsichere
<?1>, <?2>, <?3> usw.	Ich war bei Maria <?4>, als das passierte. Ich war <?3>.	Unverständliche Passagen werden mit <? und ihrer Dauer in Sekunden markiert, <?1>, eine Sek., <?2> zwei Sek., <?3> drei Sek. usw.
„xyz“	Die Frau sagte: „Das war eine schlimme Zeit.“ Oder bei Quellenangaben: „Zeitzeugen-Interviews	Zitate, wörtliche Rede, Buch- und sonstige Titel werden in Anführungszeichen gesetzt (in Deutsch „unten und oben“).
–	Ich wollte arbeiten, aber_. Können Sie mir bitte etwas zu trinken	Ein Unterstrich kennzeichnet abgebrochene Wörter oder Sätze.
<=>	Ja<=>nein<=>vielleicht.	Der schnelle Anschluss von Wörtern wird mit
<l(es) xyz> <l(es-de)> <l(de-es)>	Wo waren Sie <l(es) en este tiempo? MM# Und dann hat er sich schnell wieder <l(es-de) rekuperiert>	Fremdsprachige Begriffe werden mit „l“ für <i>lexical phenomen</i> und dem Sprachkürzel, z.B. „l(de)“ für Deutsch, l(es)“ für Spanisch oder „l(en)“ für Englisch, gekennzeichnet.“. Begriffe, die aus dem Spanischen stammen, aber eingedeutscht verwendet werden, werden mit <l(es-de)> oder, im umgekehrten Fall, mit <l(de-es)>

		ausgezeichnet, und groß geschrieben: z.B. <(es) Pito>, in den spanischen Interviews: z.B. Kinderhaus (vgl. Glossar CDOH).
<ld xyz>	<ld(Berliner Dialekt) Ick kanns nich glob'n>, aber ich war bei Maria Malta.	Bei Erzählungen, in denen größtenteils Hochsprache gesprochen wird, werden im Dialekt gesprochene Wendungen mit einem kleinen L und D (<i>lexical phenomen, dialect</i>) gekennzeichnet. Wenn die Art des Dialektes erkannt wird, soll sie vermerkt werden.
<v>	<v(Lachen)>	Nonverbale, aber lautliche Äußerungen wie Lachen, Gähnen etc. (<i>vocal</i>) werden mit <v(Art der Äußerung)> gekennzeichnet. Für deren Spezifizierung gibt es eine Liste von Standardeinträgen (vgl. Tabelle).
<s(Sprechweise) xyz>	Wo waren Sie <s(lachend) während dieser Zeit?>	Sprechweisen wie Lautstärke, Tempo (<i>shift in vocal quality</i>) etc. werden mit <s(Art der Sprechweise) xyz> gekennzeichnet. Diese Kennzeichnung wird vorgenommen, wenn besonders laut, leise, langsam, gedehnt, lachend, weinend usw. gesprochen wird
<sim>	[01:23:45.47] INT# Wo waren Sie<sim während dieser Zeit?> [01:23:45.67] MM# <sim Ich war bei Maria Malta,> als das passierte.	Beim Simultansprechen, wenn Interviewer*in (INT) und Interviewte gleichzeitig sprechen, wird beides mit <sim> (<i>simultan</i>) gekennzeichnet. Vor der Sprecher*inneninitiale muss ein Timecode stehen. Erstreckt sich das gleichzeitige Sprechen über mehrere Minuten, werden Anfang und Ende notiert: <sim(Anfang)> xyz <sim(Ende)>.
<nl(Geräusch) xyz>	Wo waren <nl(Telefonklingeln) Sie während dieser> Zeit?	Geräusche oder Ereignisse, die das Interview beeinflussen, werden mit <nl> (<i>nonlinguistic incidents</i>) kenntlich gemacht. Erstreckt sich ein Geräusch über mehrere Segmente, werden Anfang und Ende des Geräusch oder Ereignisses notiert: <nl(Telefonklingeln Anfang)> xyz <nl(Telefonklingeln Ende)> (vgl. Tabelle).
<g(Gestikart) xyz>	Wo waren <g(Fingerzeigen) Sie> während dieser Zeit?	Besonders auffällige Gestik (<i>gesture</i>) wird mit <g(Art der Geste)> gekennzeichnet. Für deren Spezifizierung gibt es eine Liste mit Standardbezeichnungen (vgl. Tabelle).
<m(Mimikart)>	<m(stirnrunzelnd) Hat er mich wirklich verstanden?>	Besonders auffällige Mimik (<i>mimic</i>) wird mit <m(Art der Mimik)> gekennzeichnet. Für deren Spezifizierung gibt es eine Liste mit Standardbezeichnungen (vgl. Tabelle).

