

apropos

[Perspektiven auf die Romania]

Sprache/Literatur/Kultur/Geschichte/Ideen/Politik/Gesellschaft

„Mi ricordo“, „je me souviens“: ich erinnere mich
*Sammlungsübergreifende Interviewanalysen in Oral History und
Korpuslinguistik*

Annette Gerstenberg & Cord Pagenstecher

apropos [Perspektiven auf die Romania]

hosted by Hamburg University Press

2022, 9

pp. 213-239

ISSN: 2627-3446

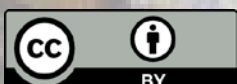
Online

<https://journals.sub.uni-hamburg.de/apropos/article/view/1902>

Zitierweise

Gerstenberg, Annette & Cord Pagenstecher. 2022. „ „Mi ricordo“, „je me souviens“: ich erinnere mich. Sammlungsübergreifende Interviewanalysen in Oral History und Korpuslinguistik.“ *apropos* [Perspektiven auf die Romania] 9/2022, 213-239.
doi: <https://doi.org/10.15460/apropos.9.1902>

Except where otherwise noted, this article is licensed under a Creative Commons Attribution 4.0 International license (CC BY 4.0)



Annette Gerstenberg & Cord Pagenstecher

„Mi ricordo’, „je me souviens’: ich erinnere mich.

**Sammlungsübergreifende Interviewanalysen in Oral History und
Korpuslinguistik**

Annette Gerstenberg

ist Professorin für romanische
Sprachwissenschaft (Italienisch und
Französisch) an der Universität
Potsdam.

annette.gerstenberg@uni-potsdam.de

Cord Pagenstecher

ist Historiker am Center für Digitale
Systeme an der Universitätsbibliothek
der Freien Universität Berlin, Bereich
Digitale Interview-Sammlungen.

cord.pagenstecher@cedis.fu-berlin.de

Keywords

Oral History – Korpuslinguistik – Erinnerung – Zeitzeugen – Interviews

1. Einleitung

Geschichtswissenschaft und Linguistik führen und analysieren Audio- und Video-Interviews, gehen dabei aber sehr unterschiedlich vor und arbeiten selten zusammen. Die soziolinguistischen Interviews der Sprachwissenschaft und die lebensgeschichtlichen Interviews der Oral History entstehen in unterschiedlichen, im Folgenden erläuterten Fragekontexten. Beide sind biographische Zeugnisse und historische Quellen, zugleich aber auch sprachlich gestaltete mündliche Erzählungen. Durch das Entstehen großer digitaler Archive sind die Interviewsammlungen zu nachnutzbaren Forschungsdaten geworden, bei deren Analyse die qualitativ-hermeneutische Oral History und die quantitativ-thesenorientierte Korpuslinguistik fruchtbar zusammenarbeiten können. Dabei stellen unterschiedliche Erkenntnisinteressen, Transkriptionsweisen, Analysemethoden und Terminologien freilich eine Herausforderung dar. Besonders gilt dies bei einer übergreifenden Analyse verschiedener Sammlungen, die bislang in keiner der beiden Disziplinen gängig ist, durch die neue Forschungsumgebung *Oral-History.Digital* nun aber ermöglicht wird.

In der Gegenüberstellung einer Oral-History-Sammlung (*Zwangsarbeit 1939-1945*) und eines linguistisch konzipierten Korpus' biographischer Interviews (*LangAge*) skizziert dieser Beitrag Potenziale der interdisziplinären Analyse. In beiden Sammlungen stehen individuelle und kollektive Erinnerungen im Zentrum; in beiden Sammlungen spielt die lebensgeschichtliche Erfahrung des Zweiten Weltkriegs eine wichtige Rolle. Die hier vorgestellte Interviewanalyse beschränkt sich auf französisch- und italienischsprachige Interviews, ist sammlungsübergreifend angelegt und verbindet Fragestellungen und Methoden aus Korpuslinguistik und Oral History. Um die Sprache der Erinnerung an Weltkrieg und Zwangsarbeit zu erforschen, verfolgen wir einerseits einen thesengeleiteten Ansatz und untersuchen sprachliche Merkmale, deren Verwendungshäufigkeit und Verwendungskontexte Aufschluss geben können über individuelle und kollektive Erinnerungs- und Erzählmuster. Wir fokussieren dabei auf Personalpronomina der ersten Person, Singular und Plural, als erste Annäherung an die Beteiligten (cf. Knowles et al. 2021). In einem weiteren, datengetriebenen Ansatz untersuchen wir andererseits Interviewsammlungen nach viel verwendeten Wortkombinationen (N-Gramme) und unterziehen dann auffällig häufige Wendungen einer genaueren Analyse. Als inspirierend erwies sich in den hier untersuchten drei Beständen – italienische Interviews aus *Zwangsarbeit 1939-1945*, französische Interviews aus *Zwangsarbeit 1939-1945*, sowie *LangAge* als rein französisches Korpus – vor allem die wiederkehrende Wendung *ich erinnere mich*. Die Ergebnisse zeigen teils sprach-, teils sammlungsspezifische Besonderheiten, für deren genaueres Verständnis relevante Interviewpassagen hinzugezogen werden.

2. Interviews aus Linguistik und Oral History

2.1 (Sozio-)linguistische Interviews

Soziolinguistische Interviewtechniken orientieren sich an der Zielsetzung, unter kontrollierten Bedingungen ein Spektrum sprachlicher Ausdrucksweisen zu erfassen, um Phänomene von Sprachwandel und -variation zu untersuchen. Dieses reicht von überwachtem, möglichst spontanem bis hin zum unüberwachtem Sprechen, ergänzt durch Techniken wie Vorlesen von Texten und Wortlisten sowie die Antwort auf Fragen nach Spracheinstellungen und sozioökonomischem Hintergrund (cf. Briggs 2005, 1053; Labov & Auger 1993, 116). In vielen Kontexten sind Interviews auf Grund ihrer hohen Vergleichbarkeit nicht zu ersetzen, wie Wagner und Tagliamonte (2017, 213) feststellen: „Interviews remain the gold standard for eliciting a range of vernacular [maximally unmonitored] and more formal speech styles from informants.“ Parallel zur Weiterentwicklung unterschiedlicher Interviewformate wurde ein breites Spektrum nicht eigens für die Erhebung angebahnter (elizierter), sondern natürlich vorkommender Interaktion für die Untersuchung der gesprochenen Sprache erschlossen. Daraus ergaben sich wiederum Effekte für die Einordnung von Interviews, die in konstruierten Forschungssettings geführt wurden, insofern diese selbst in ihrer interaktionalen Dimension verstanden wurden (cf. Deppermann 2013) und die kooperative Herstellung in die Auswertung einbezogen wird. So zeigen Wagner und Tagliamonte (2017), wie dieses Bewusstsein in die Interviewführung mit einbezogen werden kann, indem

nicht der Versuch unternommen wird, die Rollen zu neutralisieren, sondern sie bewusst einzusetzen.

2.2 Oral History-Interviews

Aus einer ganz anderen Forschungstradition stammen die lebensgeschichtlichen Interviews in der Geschichtswissenschaft. Seit der „Geburt des Zeitzeugen“ (Sabrow & Frei 2012) nach 1945 wurden viele Interviewprojekte nach der Methode der Oral History durchgeführt, besonders mit Holocaust-Überlebenden und Opfern anderer staatlicher Verfolgungen.

Die im angelsächsischen Raum schon länger etablierte Oral History wurde seitens der – lange Zeit auf staatliches Handeln und schriftliche Überlieferungen fixierten – deutschen Geschichtswissenschaft zunächst skeptisch betrachtet. Vor allem außerakademische Initiativen wie Geschichtswerkstätten oder Gedenkstätten begannen, Zeitzeugen-Interviews zu führen. Technisch dominierten zunächst Audioaufnahmen, die seit den 1980er Jahren zunehmend durch verschiedene Videoformate nicht ersetzt, aber ergänzt wurden. Vor allem mit dem Cultural Turn und dem Boom der Memory Studies sind audiovisuell aufgezeichnete Erinnerungen nun aber zu wesentlichen Quellen für die Alltags-, Kultur- und Geschlechtergeschichte geworden (cf. Apel, Leh & Pagenstecher 2022; Eusterschulte, Knopp & Schulze 2016).

Ein Oral History-Interview dauert in der Regel mehrere Stunden. Im Kern steht eine von den Interviewten selbst strukturierte lebensgeschichtliche Erzählung; erst in einer zweiten Phase werden Nachfragen zu dieser Erzählung gestellt, um weitere Narrationen zu evozieren. Anschließend folgen thematische Fragen, schließlich oft eine gemeinsame Betrachtung privater Fotos und Dokumente. Beabsichtigt (wenngleich in der Praxis nicht immer erreicht) sind also im Dialog inspirierte Narrationen der „Erzähler*innen“, keine investigativen Frage-Antwort-Interviews mit „Respondent*innen“. Interessant sind in der Interpretation weniger die berichteten Fakten als ihre erinnernde Deutung.

Darüber hinaus dienen die lebensgeschichtlichen Erzählungen der Geschichtswissenschaft als Quellen für Ereignisse oder Erfahrungen, die in schriftlichen Archivquellen unzureichend oder einseitig dokumentiert sind, etwa von unterdrückten Minderheiten. Gerade im Hinblick auf Massenverbrechen wie den Holocaust ist das Bezeugen des Geschehenen ein zentrales Motiv für und in den Interviews. Ein Zeugnis abzulegen, stellvertretend für die Ermordeten, ist immer auch ein „Versprechen der Wahrheit“ (Eusterschulte, Knopp & Schulze 2016, 21), „eine erlebte, *eingefleischte* Wahrheit“ (Klüger 1996, 405, Herv. im Original). Angesichts der vielfach bewusst vernichteten Akten erweisen sich mündliche Zeugnisse häufig als wertvolle und erstaunlich exakte Quellen für die Ereignisgeschichte oder Topographie bestimmter Lager oder Verbrechenkomplexe. Viele Oral History-Interviews entstehen in Gedenkstätten oder im Rahmen von Erinnerungs-Initiativen, sind also nicht primär Forschungsdaten, sondern entstehen als wichtiger Teil des kulturellen Erbes, genau am Übergang vom kommunikativen zum kulturellen Gedächtnis.

Ein Oral History-Interview ist gewissermaßen als mündliche Autobiographie zu interpretieren; die Zeitzeug*innen sind die Autor*innen dieses Werks. Die Gestaltung dieser Autobiographie ist ein höchst gewagter Prozess, denn der audiovisuelle Text eines Interviews entsteht live am Mikrofon. Es gibt kein Manuskript, keine Korrekturmöglichkeit, kein Lektorat; das Video erlaubt es uns, „den Überlebenden beim Erinnern zuzusehen“ (Nägel 2016, 356). Die meist betagten Autor*innen müssen spontan, möglicherweise noch vor laufender Kamera, und häufig in einer anderen als ihrer Erstsprache, eine erzählerische und sprachliche Form für ihre Lebensgeschichte finden – und dabei mit einer möglichen Reaktivierung traumatischer Erfahrungen umgehen. Zudem wird dieser autobiographische Text nicht nur verfasst, sondern gleich auch zur Aufführung gebracht, denn die Lebensgeschichte wird im Interview mit entsprechender Betonung, Mimik und Gestik in Szene gesetzt. Jenseits der schwierigen Erinnerungsarbeit ist ein Interview also auch eine performative Gestaltungsleistung.

Ferner ist der audiovisuelle Text eines Oral History-Interviews das Ergebnis eines Gesprächs-Settings, eines medial aufgezeichneten Dialogs, einer „Erzählgemeinschaft“ (Nägel 2016, 352). Die Interviewenden sind Co-Autor*innen beim gemeinsamen Verfertigen und Aufführen der lebensgeschichtlichen Erzählung (cf. Pagenstecher & Pfänder 2017). Die schwierige Rolle der Interviewenden ist diesen in der Praxis wohl bewusst, bleibt in der Forschung aber oft unterbelichtet. Als Zuhörende sind sie – und die Kameraleute – die primären Adressat*innen der Erzählung und repräsentieren zugleich doch auch zukünftige Rezipient*innen, vielleicht gar die Nachwelt schlechthin.

Oral History-Interviews sind also historische Quelle, oft auch Zeugnis für anderweitig kaum dokumentierte Verhältnisse und Geschehnisse bis hin zu Verbrechen; zugleich sind sie sprachlich, narrativ und performativ gestaltete Erzählungen und Deutungen der eigenen Lebensgeschichte, die in einem dialogischen Aufnahme-setting entstehen. Und sie sind umfangreiche Daten der gesprochenen Sprache, die auf unterschiedliche Weise analysiert werden können.

3. Digitale Interview-Archive

3.1 Interview-Sammlungen

Neben dem Setting der einzelnen Aufnahme ist auch der breitere Entstehungskontext der Interviews zu berücksichtigen. In der Regel werden Oral History-Interviews nicht isoliert, sondern im Rahmen eines Forschungsprojekts, einer Museumssammlung oder einer erinnerungskulturellen Initiative finanziert und durchgeführt. Am Beispiel des *Visual History Archive* der Shoah Foundation (Michaelis 2013; Shenker 2015; Taubitz 2016; Bothe 2019) wurde untersucht, wie stark sich die Auswahlkriterien, Interviewmethoden oder Kamerarichtlinien des Projekts auf die Gestaltung des Interviewkorpus auswirken.

Als wichtiger Teil unseres kulturellen Erbes sind die in Erinnerungsinstitutionen und Forschungsprojekten entstandenen Oral History-Interviews heute größtenteils digitalisiert. Entsprechende Sammlungen sind aber verstreut und schlecht zugänglich.

An Universitäten wurden die Audio- oder Video-Interviews lange Zeit nur von den Interviewenden selbst für ihre Publikationen genutzt. Erst in letzter Zeit versteht man sie zunehmend auch als Forschungsdaten, die für verschiedene Disziplinen nachnutzbar sind – oder sein sollten (cf. Apel, Leh & Pagenstecher 2022). Große Interview-Archive wie das *Archiv Deutsches Gedächtnis* an der FernUniversität in Hagen (über 3.000 Interviews) oder die *Werkstatt der Erinnerung* in Hamburg (über 2.000 Interviews) sind bislang überwiegend nur vor Ort zu konsultieren. Ähnliches gilt für die meisten Museen und Gedenkstätten mit ihren Interviewsammlungen. Es gibt keinen Katalog, der eine sammlungsübergreifende Recherche in den verstreuten Beständen ermöglichen würde.

3.2. Die Forschungsumgebung *Oral-History.Digital*

An der Freien Universität Berlin entsteht im DFG-Projekt *Oral-History.Digital* (cf. Oh.d 2020–2022) nun eine Informationsinfrastruktur für wissenschaftliche Interview-Sammlungen. Die bis 2023 fertiggestellte Arbeitsumgebung unterstützt Sammlungsinhaber*innen bei der Archivierung, Erschließung und Bereitstellung, Forschungsprojekte bei der Recherche, Annotation und Auswertung von Oral History-Sammlungen. Methodisch und technologisch stützen wir uns bei *Oral-History.Digital* auf Erfahrungen aus früheren Projekten wie den Interview-Archiven *Zwangsarbeit 1939–1945* (ZWAR 2009–2022) oder *Colonia Dignidad. Ein chilenisch-deutsches Oral History-Archiv* (seit 2022).

Die Erschließungs- und Forschungsumgebung *Oral-History.Digital* wird diese und weitere Bestände umfassen, u. a. aus dem *Archiv Deutsches Gedächtnis* in Hagen und der *Werkstatt der Erinnerung* in Hamburg, aber auch aus verschiedenen Museen und Stiftungen. Auch linguistische Interviewprojekte wie *LangAge* werden darüber nutzbar sein. Die versammelten Ressourcen sind thematisch breit gefächert und betreffen zum Beispiel die DDR-, Bergbau- oder Gewerkschaftsgeschichte. Weiterhin ergibt sich dadurch eine intra- und interdisziplinäre Vielfalt, wie folgende Gegenüberstellung einer Oral History-Sammlung (*Zwangsarbeit 1939–1945*) und eines linguistisch konzipierten Korpus' biographischer Interviews (*LangAge*) zeigt.

3.3. Das Online-Archiv *Zwangsarbeit 1939–1945*

Das Online-Archiv *Zwangsarbeit 1939–1945. Erinnerungen und Geschichte* ist der Erinnerung an über zwanzig Millionen Menschen gewidmet, die für das nationalsozialistische Deutschland Zwangsarbeit geleistet haben (ZWAR 2009–2022). Konkret erzählen 249 Zwangsarbeiterinnen und 341 Zwangsarbeiter aus 26 Ländern ihre Lebensgeschichte in ausführlichen Audio- und Videointerviews. Dabei kommen neben der Kerngruppe der „zivilen“, also dem Arbeitsamt oder privaten Firmen untergeordneten Zwangsarbeiter*innen auch die von der SS beaufsichtigten KZ-Häftlinge und die der Wehrmacht unterstellten Kriegsgefangenen zu Wort.

Initiiert von der Stiftung *Erinnerung, Verantwortung und Zukunft*, wurden die Interviews in den Jahren 2005 und 2006 von 32 Initiativen unter der Koordination des Instituts für Geschichte und Biographie der FernUniversität in Hagen in 26

Ländern geführt (cf. Leh, Plato & Thonfeld 2008). Die lebensgeschichtlichen Interviews haben eine durchschnittliche Länge von dreieinhalb Stunden. In einem mehrsprachigen Online-Archiv sind sie mit timecodierten Transkripten bereitgestellt, d. h., mit einer Zeitmarke nach jedem Satz oder nach etwa 120 Zeichen (cf. Apostolopoulos & Pagenstecher 2013, Pagenstecher 2017). Der Zugang erfolgt nach einer Registrierung, die vor der Freischaltung manuell geprüft wird, um die oft sehr persönlichen Lebenserzählungen gegen einen eventuellen Missbrauch zu schützen.

Nach der Anmeldung bietet das Archiv unterschiedliche Rechercheoptionen (Abb. 1a). Über die Kategoriensuche mit Rubriken wie Opfergruppe, Einsatzbereich oder Sprache werden komplette Interviews gefunden. Zudem sind die in 30 Sprachen oder Sprachkombinationen vorliegenden Transkripte im Original oder in ihrer deutschen Übersetzung durchsuchbar und als Untertitel anzeigbar (Abb. 1b). Inhaltsverzeichnisse erleichtern die Orientierung in den lebensgeschichtlichen Erzählungen, editorische Anmerkungen erläutern missverständliche Stellen oder geben weitere Quellenhinweise. Kurzbiographien sowie private Fotos und Dokumente kontextualisieren die Interviews.

ZWANGSARBEIT
1939 - 1945
ERINNERUNGEN UND GESCHICHTE

Interviews
★ Suche speichern 📄 Suchergebnisse exportieren 150 Suchergebnisse

Raster Liste Workflow

K 1
bei: "3.1. Überfall und Verhaftung durch die Deutschen - Verhör im Lager in Marina Gorka"
| Band 1/05 | 00:15:06
dort wie ein **sklave** arbeiten." So also, ich sagte

D 1
Video 3 h 12 min Russisch mit dt. Übersetzung

A 2
Audio 2 h 36 min Russisch mit dt. Übersetzung

A 1
Video 3 h 29 min Russisch mit dt. Übersetzung

E 1
Audio 2 h 41 min Russisch mit dt. Übersetzung

B 2
Video 5 h 01 min Russisch mit dt. Übersetzung

Suche im Archiv

sklav*

zurücksetzen

GRUPPE

- Germanisierte Kinder (0)
- Italienische Militärinternierte (3)
- Jüdinnen/Juden (54)
- KZ-Häftlinge (67)
- Kriegsgefangene (4)
- Ostarbeiter/Innen (aus der Sowjetunion) (33)
- Politisch Verfolgte (28)
- Religiös Verfolgte (1)
- Service du Travail Obligatoire (aus Frankreich) (6)
- Sinti und Roma (6)
- Sonstige (1)
- Weitere Zwangsarbeiter/Innen (27)

EINSATZBEREICH

UNTERBRINGUNG / INHAFTIERUNG

SPRACHE

q

- Bosnisch (0)
- Bosnisch/Romani (0)
- Bulgarisch (1)
- Deutsch (7)
- Englisch (24)
- Französisch (6)
- Hebräisch (11)

1a | Suchmaske in ZWAR

The screenshot shows the website 'ZWANGSARBEIT 1939-1945 ERKENNUNGEN UND GESCHICHTE'. The main content area displays an interview with Anita L. It features a video player at the top, a transcript below it with a highlighted section, and a sidebar on the right with navigation options and a detailed profile for Anita L. The profile includes fields for Name, Birth Date, Gender, Language, Group, Einsatzbereich, Geburtsort, Unterbringung / Inhaftierung, Lager und Haftstätten, Deportationsort, and Wohnorte ab 1945. The transcript text is as follows:

I'm sure they didn't have, when there came sort of thousands of people, probably there wasn't time. But so this girl, who was also a prisoner, asked me various questions, what was my name, where do I come from, and what did I do before. I was arrested and (that I'd always played ?) the cello.

I mean it seems so crazy to even talk about a cello in Auschwitz, you know, and that was very fortuitous, because she said, "That's fantastic, there's an orchestra here, wait a minute."

So she left me there, by that time, I mean, imagine I was naked, with a number on my arm and no hair. I mean, not a particularly beautiful sight. And it was very cold.

Anyhow, I waited and I didn't really know what I was waiting for, and then a lady came into the block where I was and introduced herself as Alma R.

Now, the name Rosé doesn't mean anything to everybody, but she comes from an extremely interesting

1b | Das Interview mit Anita L.: Transkriptansicht

Das Online-Archiv *Zwangsarbeit 1939-1945* unterstützt eine digitale Recherche, aber nur begrenzt eine digitale Analyse. Bestimmte Begriffe und Wortkombinationen können archivweit gesucht und punktgenau angesteuert, aber nur mühsam quantitativ ausgewertet werden. So wird zum Beispiel für eine Recherche zum Wortfeld *Zwangs- und Sklavenarbeit* bei einer Volltextsuche nach *sklav** die Gesamtzahl von 147 Interviews und für jedes dieser Interviews die Anzahl der Segmente mit diesem Wortteil im übersetzten Transkript angezeigt. Durch zusätzliche Filter können die Suchergebnisse nach Gruppe, Geschlecht, Sprache oder anderen Facetten eingegrenzt werden, wodurch sich ein differenziertes Bild ergibt, welche Interviewten welche ihrer Zwangsarbeits-Erfahrungen mit Sklaverei assoziieren. Für im engeren Sinne korpuslinguistische Analysen müssen die Transkripte aber nach entsprechender Genehmigung exportiert und aufbereitet werden (s. u., Abschnitt 5).

517 der 590 Zeitzeugen-Interviews zur Zwangsarbeit sind online zugänglich; die restlichen Interviews sind aus verschiedenen Gründen zugangsbeschränkt oder noch nicht erschlossen. Die meisten Interviews liegen auf Russisch (120) und Polnisch (74) vor, aber auch die romanischen Sprachen sind vertreten: 19 Interviewte sprachen Französisch, 18 Rumänisch, 9 Italienisch, 7 Katalanisch und 2 Spanisch. Darunter finden sich neben weithin unbekanntem Interviewten auch Prominente wie die Auschwitz- und Ravensbrück-Überlebende und spätere italienische Senatorin Liliana Segre, der Buchenwald-Überlebende und spätere Schriftsteller Jorge Semprún oder ein Lagerkamerad und späterer Privatsekretär des Chansoniers George Brassens.

3.4. Das Korpus *LangAge*

Das französischsprachige Korpus *LangAge* entstand in einer 2005, also zeitgleich zu den ZWAR-Interviews, begonnenen Serie biographischer Interviews in Orléans, um

den Sprachgebrauch des höheren Lebensalters zu dokumentieren und der linguistischen Analyse zugänglich zu machen. Die Lebensgeschichten der repräsentierten Generation (Geburtsjahrgänge zwischen 1915 und 1935) sind eng mit dem Zweiten Weltkrieg verbunden, und das Projekt wurde als Initiative der „histoire orale“ vorgestellt. Der thematische Leitfaden von *LangAge* folgt den Lebensgeschichten von der Zwischenkriegszeit über Kriegserfahrungen unter der deutschen Besatzung in die Phase des wirtschaftlichen Aufschwungs, der so genannten *Trente Glorieuses*, und bis zu den gesellschaftlichen Ereignissen und Entwicklungen, die mit dem Jahr 1968 verknüpft sind. Im letzten Abschnitt geht es um die Lebensgestaltung seit dem Ende des Erwerbslebens, wobei zahlreiche Aktivitäten, die Einbindung in familiäre und freundschaftliche Netzwerke und Einstellungen zu sozialen und sprachlichen Fragen angesprochen werden.

Die Methode des biographischen Interviews gestaltete die Interviewsituation aus, insofern die Interviewten die Autorität der Zeitzeugenschaft hatten und die Interviewerin durch aktives Zuhören ihr persönliches Interesse zum Ausdruck brachte. Durch die Adressierung unterschiedlicher Milieus wurde im gegebenen räumlichen (Wohnort Orléans und Umgebung) und zeitlichen (Lebensalter) Rahmen ein breiter historischer Erfahrungsraum einbezogen. Unter den Beteiligten ist ein Shoah-Überlebender, ein weiterer Beteiligter wurde zur Zwangsarbeit verpflichtet. Viele erlebten die deutsche Besatzung und den Krieg in Orléans; Väter und Familienangehörige oder auch sie selbst waren an der Front. Ein häufig thematisierter Zusammenhang ist der so genannte *Exode*, die Fluchtbewegung infolge des deutschen Angriffs ab Mitte Mai 1940, der aus unterschiedlichen Perspektiven berichtet wird. Insofern ergibt sich ein Spannungsfeld zwischen verbindenden historischen Erfahrungen der „älteren Generation“ (ab 70 Jahren) und der Breite der individuellen Biographien.

Zugleich entstanden durch die biographischen Interviews, die monologische Erzählformen entlang einem thematischen Leitfaden bevorzugten, vergleichbare Sprachdaten von Angehörigen der älteren Generation, auf deren Basis sprachliches Altern und seine Dynamik im Längsschnitt untersucht werden konnte. Die französische Stadt Orléans in der Region Centre bot ideale Voraussetzungen, weil hier seit 1968 die groß angelegte *Étude Sociolinguistique sur Orléans* (ESLO) etabliert ist: Der ersten Erhebungsphase (ESLO1) folgte ab 2008 eine zweite Initiative von Sprachaufnahmen (ESLO2), die den Interviews im Format von ESLO1 weitere Module zur Seite stellte. Vor dem Hintergrund dieser sozial breit angelegten Studie konnte ein generationell fokussiertes Projekt wie *LangAge* an Aussagekraft gewinnen. Mit ESLO verbindet *LangAge*, dass ein sozial breites Spektrum von Personen einbezogen wurde. Die Erhebung wurde in weiteren Phasen 2012 und 2015/2016 mit denselben Personen fortgesetzt. Es handelt sich um Audio-Interviews von durchschnittlich einer Dreiviertelstunde Länge in französischer Sprache (Details s.u.), die mit anonymisierten Transkripten derzeit in einer linguistisch erschlossenen Datenbank zugänglich sind (aktuell 129 Transkripte mit 882.353 Tokens) (Abb. 2). Autorisierte Interviews der ersten Serie werden ab 2023 auch in *Oral-History.Digital* nutzbar sein.

_(exode)S

Found 163 results (Total utterance duration: 10:53.293)

Select all results (163) Context: the whole line

1. <input checked="" type="checkbox"/> a001a.tr - a001	mille-neuf-cent-quarante on a fait l'	exode	hein
2. <input checked="" type="checkbox"/> a002a.tr - a002	et euh il y a eu après Nantes ça a été l'	exode	
3. <input checked="" type="checkbox"/> a006a.tr - a006	euh par exemple là quand j' ai fait ma communion j' ai fait ma co() euh j' ai en mille-neuf-cent-quarante au moment de l'	exode	
4. <input checked="" type="checkbox"/>	voulez-vous que je vous parle un peu de de notre de notre équipée de de la guerre enfin de notre départ à l'	exode	là ou
5. <input checked="" type="checkbox"/>	et puis ben si vous voulez que je vous raconte l'	exode ?	
6. <input checked="" type="checkbox"/>		l'	exode euh l' exode chez nous c' éta
7. <input checked="" type="checkbox"/>		l' exode euh l'	chez nous c' était ça a été as
8. <input checked="" type="checkbox"/>		c' était déjà euh l'	rural à l' époque
9. <input checked="" type="checkbox"/>	en gros on euh la façon dont on a vécu l'	exode	
10. <input checked="" type="checkbox"/>	euh vous voulez dire l'	exode	là en quarante ?
11. <input checked="" type="checkbox"/> a007a.tr - a007			
12. <input checked="" type="checkbox"/>	bah d'ailleurs euh l' d'abord le premier c' est l'	exode	
13. <input checked="" type="checkbox"/>	et ben j' ai su quand je suis revenue d'	exode	
14. <input checked="" type="checkbox"/> a008a.tr - a008	euh oui parce qu' on est parti alors en	exode	
15. <input checked="" type="checkbox"/>			ah ben ça l'
16. <input checked="" type="checkbox"/>	on est parti à l'	exode	on l' a fait hein puis ah
17. <input checked="" type="checkbox"/>	enfin c' était c' était l'	exode	avec une voiture que mon pè
18. <input checked="" type="checkbox"/>	d'abord rien qu() nous les les l'	exode	quoi
19. <input checked="" type="checkbox"/> a009a.tr - a009	nous on l' a vécu extérieurement par l'	exode	nous a marqué beaucoup
20. <input checked="" type="checkbox"/> a010a.tr - a010			par tout ça
	ah par contre j' ai fait l'	exode	en mille-neuf-cent-quarante

2 | Ergebnisse für die Suche nach *exode* in *LangAge* Corpora mit LaBB-CAT

4. Korpuslinguistische Herangehensweisen an Oral History-Interviews

4.1 Forschungsstand und Forschungsfragen

Die Geschichtswissenschaft untersucht die Interviews der Oral History vor allem mit qualitativen Analysen, die sich einem Interview als Einzelquelle nähern. Sie zielen auf eine biographische und historische Kontextualisierung und eine hermeneutische Interpretation der individuellen Erfahrung und Erinnerung. Diese Ansätze haben sich bewährt und sind dem Charakter eines lebensgeschichtlich-narrativen Interviews angemessen.

Seit etwa zwei Jahrzehnten werden nun aber mehr und größere Sammlungen von Zeitzeugen-Interviews in digitalen Archiven besser erschlossen, so dass Forschende nun statt weniger selbst geführter Interviews eine größere Anzahl vorhandener Interviews für ihre Fragestellungen nutzen können. Das ermöglicht Vergleiche zwischen verschiedenen Gruppen von Interviewten. Fruchtbar waren komparative Gruppenanalysen etwa zu gemeinsamen Erfahrungsmustern verschiedener Interviewter (Browning 2010), zu den Einflüssen gesellschaftlicher Erinnerungskulturen auf die biographischen Erzählungen (Thonfeld 2014), zu Entstehungs- und Rezeptionskontexten verschiedener Interviewsammlungen (Leh, Plato & Thonfeld 2008; Michaelis 2013; Shenker 2015; Taubitz 2016; Bothe 2019) oder zu Genderaspekten (Pagenstecher & Tausendfreund 2015). Als besonders ergiebig erwies sich der Vergleich mehrerer Interviews mit der gleichen Person in verschiedenen Interview-Sammlungen (Laub & Bodenstab 2008; Kangisser Cohen 2014; Pagenstecher & Pfänder 2017; Schuch 2021).

Oral History-Quellen wurden auch diskursanalytisch (Schiffrin 2000) sowie in einigen korpuslinguistischen Studien verwendet (Grant 2010). Eine breitere wechselseitige Rezeption von Oral History und Linguistik wurde durch das 2015 von

Erich Kasten, Katja Roller und Joshua Wilbur veranstaltete Kolloquium *Oral History Meets Linguistics* angeregt. In diesem Rahmen zeigten Pagenstecher und Pfänder (2017), wie Oral History-Interviews aus einer interaktionalen Perspektive neu gelesen und verstanden werden können. Im vergleichenden Zugriff auf ZWAR-Interviews einerseits und *LangAge*-Daten andererseits untersuchte Gerstenberg (2017), wie sich in französischsprachigen Interviews die Verwendung der Begrifflichkeit für *Zwangsarbeit* bei Betroffenen (ZWAR) von gleichaltrigen Nichtbetroffenen (*LangAge*) unterscheiden: Nur die selbst zur Zwangsarbeit verbrachten Betroffenen verwendeten die personalisierte Abkürzung *un STO* für 'jemand, der den Service du Travail Obligatoire leistete'. Ein weiterer Aspekt des Vergleichs brachte hervor, dass die euphemistischen Slogans der NS-Propaganda, in denen die Zwangsarbeit als ertragreicher Arbeitsaufenthalt im Nachbarland schöngefärbt wurde, von den Betroffenen im Hinblick auf ihren fatalen Beitrag zu einem lange Zeit verharmlosenden Bild der Zwangsarbeit diskutiert wurden. Dagegen reproduzierten die nicht betroffenen Personen aus *LangAge* in einigen Fällen die Slogans unkritisch (cf. Gerstenberg 2017).

Die letztgenannten Beispiele zeigen, wie Oral History-Interviews die Grenze vom alltäglichen Sprachgebrauch zur bewussten Verwendung sozialhistorisch bedeutsamer Begriffe überschreiten (zur Unterscheidung cf. Blank 1997, 29). Als Überlagerung individueller und kollektiver Erfahrungen lassen sich in Lebenserzählungen in *LangAge* unterschiedliche Generationsbegriffe herausarbeiten, die sich auf 1968 als Generationsereignis zwischen Jungen und Alten, und zugleich zwischen Familiengenerationen beziehen (Gerstenberg 2009). Auf das Verbindende der eigenen Altersgruppe zielen auch Kommentare zum Sprachgebrauch ab: „Diachrone Markierungen“ wie *comme on disait à l'époque* 'so sagte man damals' (Gerstenberg 2011, Kapitel 8.2) heben Wörter hervor, die als veraltet dargestellt werden. Diese Markierungen sind weitgehend unabhängig davon, ob ein Wort tatsächlich heute nicht mehr verständlich oder außer Gebrauch geraten ist, sie können vielmehr als Zeichen einer sprachlichen Generationsidentität verstanden werden.

In den engeren Rahmen der historischen Terminologie fällt der Diskurs um die in vielen Nachkriegsgesellschaften erhobenen Kollaborationsvorwürfe gegen die ehemaligen Zwangsarbeiter*innen (Thonfeld 2014). Diese Diskussion ist in den ZWAR-Interviews sehr präsent, entweder explizit oder im Hintergrund, vor allem in den Berichten von eigenen Sabotageversuchen bei der Zwangsarbeit. Auch die Assoziation zur Sklaverei hat hier ihre Funktion. Manche ehemaligen Zwangsarbeiter*innen beschreiben sich selbst als *Sklaven*, vor allem, wenn sie über besonders demütigende Erfahrungen wie ihre Verteilung auf verschiedene Fabriken und über die ihnen vorenthaltene Entschädigung sprechen; möglicherweise verwendeten gerade politisch interessierte Intellektuelle diese Begrifflichkeit häufiger als andere Interviewte (cf. Film „Sklavenarbeit“ 2011, Pagenstecher 2010). In vielen Interviews ordnet sich die Bewertung der eigenen Zwangsarbeit so in ein konfliktbehaftetes Wortfeld zwischen Sklavenarbeit und Kollaboration ein, dessen genauere Analyse vielversprechend erscheint.

4.2 Forschungsmethoden

Um die Interviews nicht nur digital zu durchsuchen, sondern sie auch digital zu analysieren, braucht die Geschichtswissenschaft neue Methoden, die das bewährte Handwerkszeug der Interviewforschung nicht ersetzen, aber ergänzen. In den Digital Humanities werden zahlreiche quantitative Verfahren entwickelt, die auch für die Oral History genutzt werden könnten – von der Netzwerkanalyse über Topic Modeling und Text Embedding bis zu korpuslinguistischen Methoden (cf. Möbus 2020, Knowles et al. 2021). Vielversprechend sind dabei vor allem datengetriebene Ansätze (*corpus driven*), die bemerkenswerte Muster in den vorliegenden Daten zu erkennen suchen, auch jenseits bereits vorab formulierter Hypothesen. In diesem induktiven Vorgehen trifft das Interesse der Digital Humanities auf jenes der Oral History; denn auch diese sucht nicht nach Belegen für vorgefasste Thesen, sondern möchte in den Interviews zunächst immanente, auch unerwartete Deutungsebenen und Sinnstrukturen finden, die dann erst kontextualisierend interpretiert werden. Viele dieser Ansätze verlangen allerdings einen hohen Aufwand in der Datenaufbereitung bei zunächst schwer einschätzbarer Aussagekraft der Ergebnisse, insbesondere wenn einzelne Fundstellen aus den lebensgeschichtlichen Erzählungen herausgelöst und ohne deren Kontext untersucht werden.

Das berührt die zentrale Frage der Herangehensweise an die Interviews als Texte. Verstehen wir die Oral History-Interviews als autobiographische Texte, ist ein literaturwissenschaftlich informiertes Verständnis des autobiographischen Genres erforderlich, gerade auch um die Spezifika eines mündlichen Erinnerungstextes herauszuarbeiten (cf. Michaelis 2013). Blicken wir eher auf das dialogische Aufnahmesetting der Interviews, helfen die Methoden der Gesprächsanalyse und interaktionalen Linguistik, um die zentrale Rolle der Interviewenden oder die performativen Erzählstrategien der Interviewten zu beleuchten (cf. Pagenstecher & Pfänder 2017). Soweit die Interviews filmisch aufgezeichnet wurden, kann die Videoanalyse Erkenntnisse über multimodale Erzählmuster, das Körpergedächtnis oder das spontane Auftauchen von Erinnerungsfragmenten beisteuern (cf. Gülich & Pfänder 2022, Freyburger & Jäger 2021, 189). Diese Ansätze fokussieren auf einzelne Sätze, Wörter, Gesten, die sehr genau transkribiert und analysiert werden, sind also nicht korpusbasiert und nicht biographisch orientiert.

Für diesen Beitrag besonders relevant ist eine gesprächsanalytische Fallstudie zu den italienischen Interviews in *Zwangsarbeit 1939–1945* (Pfänder et al. 2022). Danach dient die Wendung *mi ricordo* ‘ich erinnere mich’ manchmal dazu, eine vorangegangene Passage zu beglaubigen, also epistemische Autorität zu behaupten. Vor allem aber bereitet *mi ricordo* einen Wechsel zwischen Textsorten vor und leitet etwa eine Erzählung eines konkreten, selbst erlebten oder gestalteten Ereignisses ein. Oft bringt *mi ricordo* dann einen Wechsel vom beschreibenden Imperfekt zum erzählenden Perfekt oder Präsens. Dieser Wechsel betrifft auch die Mimik; mit der Äußerung von *mi ricordo* wandert der Blick häufig kurz nach unten, weg von der interviewenden Person, der danach beim Erzählen wieder ins Gesicht geblickt wird.

In korpuspragmatischer Perspektive werden Möglichkeiten gesucht, quantitative Verfahren mit einer kontextsensitiven Herangehensweise zu verbinden, indem Frequenzen über verschiedene Texttypen bzw. Gruppen (Lernersprache vs. Erstsprache) hinweg oder im Übersetzungsvergleich hinzugezogen werden, wie Aijmer (2015) am Beispiel von *I think* zeigt. Grant (2010) zeigt die unterschiedlichen Funktionen von *I don't know – I dunno* in Texttypen gesprochener Sprache – beide Formen nehmen Spitzenplätze der Frequenz von N-Grammen ein („3-word chunk“, Grant 2010, 2290). Sie konstatiert in den von ihr manuell annotierten Interviews, darunter auch Oral History-Interviews, zu einem großen Teil die Funktionen des distanzierenden Heckenausdrucks und der Markierung von Unsicherheit.

Dieser kurze Überblick zeigt, dass es zwischen den hermeneutischen Methoden der Oral History und korpuslinguistischen Ansätzen durchaus Verbindungslinien gibt, insofern Funktionen und Verteilungsmuster aus den Daten herausgefiltert werden. Die im nächsten Absatz behandelte Beschaffenheit der Daten spielt dabei eine fundamentale Rolle.

4.3 Unterschiedliche Transkriptionsstandards

Sammlungs- und disziplinübergreifende Analysen stehen vor der Herausforderung unterschiedlicher Traditionen und Regeln bei der Interviewführung, -aufzeichnung, -transkription und -erschließung. Problematisch ist das insbesondere bei verschiedenen Transkriptionsrichtlinien.

Die audiovisuellen Medien bilden den Kern der Interviews; für Suche, Annotation, Analyse und Zitation sind aber manuell erstellte, zunehmend auch automatisch generierte Transkriptionen oder Indexierungen von zentraler Bedeutung. Diese folgen jedoch den unterschiedlichen Transkriptionsgewohnheiten und -regeln der jeweiligen Disziplin und Sammlung.

Oral History-Transkripte sind meist orthographisch geglättete und mit Interpunktion versehene Texte, die oft als unstrukturierte Word-Dateien vorliegen und nur manchmal Timecodes und Auszeichnungselemente enthalten, z. B. für Sprecherwechsel oder Pausen. Wortwiederholungen und nonverbale Äußerungen werden nicht immer transkribiert. Sprachwissenschaftliche Transkriptionssysteme (GAT2 2009) sind unüblich und wären aufgrund der begrenzten Ressourcen bei den mehrstündigen Interviews kaum umsetzbar.

Digitale Interview-Archive wie *Zwangsarbeit 1939–1945* koppeln die oft bis zu 100-seitigen Transkripte mit den mehrstündigen Audio- oder Videoaufnahmen, indem Timecodes automatisch oder manuell in die Texte eingefügt werden (cf. Nägel 2016, 360-363). Diese Segmentierung durch regelmäßig nach bestimmten Zeitabständen (z. B. eine Minute), Zeichenzahlen (100 Zeichen) oder Sinneinheiten (Wort oder Satz) eingefügte Zeitmarken erlaubt eine synchrone Untertiteldarstellung und eine sekundengenaue Volltextsuche oder Annotation. Aus diesem Sprung in die alignierten Videos oder Audioaufnahmen ergeben sich weiterführende Perspektiven für die Kontrolle der Transkription und die Analyse der multimodalen, para- oder nonverbalen Facetten der Interviews, die auch Mimik und Gestik einbezieht. Standardisierte Im- und Exportformate wie das tabellarische

bzw. komma-getrennte Format (CSV) oder der Untertitelstandard Video Text Tracks (VTT) stehen bei entsprechender Berechtigung zur Verfügung; die in Editionen gängigen Kodierungen der Text-Encoding Initiative (TEI) können in *Oral-History.Digital* generiert werden (ISO 2016, cf. TEI Guidelines).

Das Korpus *LangAge* liegt in Tondateien (*.wav) und zeitalignierten Transkripten vor. Diese wurden mit dem Werkzeug Transcriber (Transcriber 1998–2008) erstellt, das Metadaten und Zeitkodierungen in ein XML-Format umsetzt. Die Transkription folgt dem orthographischen Prinzip (Gerstenberg, Hekkel & Kairat 2018) und erfasst Wortabbrüche, Wortwiederholungen, Häsitationen und einige redegleitende Ereignisse. Dieser auf die Mündlichkeit ausgerichteten Transkription gemäß wird auf Interpunktion und andere Schriftlichkeitsstandards verzichtet. Das ist allerdings nicht leicht: Wenn nicht explizit dafür geschult wird, verschriften auch gerade geübte Personen „richtige“ Formen statt der im Gesprochenen üblichen Varianten.

Eindrücklich zeigt sich dies am Beispiel der französischen Negation. Wie das Beispiel *je veux pas* (gesprochen) vs. *je ne veux pas* (Standardsprache, ‘ich will nicht’) zeigt, wird das Element *ne* im Gesprochenen sehr oft ausgelassen. Dies ist auch in den ZWAR-Interviews der Fall. Mit der Archivsuche und dem Anmerkungswerkzeug wurden dort 24 Vorkommen des Strings *je ne me* geprüft, der vor verschiedenen Verben vorkam, wie *je ne me souviens pas* ‘ich erinnere mich nicht’. In zwei Fällen war nicht zu hören, ob das verneinende *ne* realisiert wurde, in den verbleibenden 22 Fällen konnte die Transkription nicht bestätigt werden, *ne* war nicht zu hören. Diese ungenaue Transkription erschwert den quantitativen Vergleich mit den *LangAge*-Interviews; die bei den ZWAR-Interviews nur scheinbar häufigere Verneinung erweist sich so als ein Artefakt der Transkriptionsweise.

Dieses Bild bestätigte sich in weiteren Kontrollen und weist eine eigene Systematik auf: Die Unterschiedlichkeit von Gesprochenem und Geschriebenem wird so weitgehend vorausgesetzt, dass in der Verschriftung unmittelbar „korrigiert“ wird, ohne weitere Markierung. Die Geschichtswissenschaft fragt stärker danach, was erzählt wird, als wie es erzählt wird. Auch spielt die erinnerungskulturelle und würdigende Funktion des Online-Archivs eine Rolle: Durch eine lautgenaue Transkription hätten sich die Interviewten als unbeholfen, ja stotternd bloßgestellt fühlen können. Im Ergebnis geben die Oral History-Transkripte in ZWAR eher die Form wieder, die der sprachlichen Norm der Standardgrammatik entspricht, als das, was tatsächlich gesprochen wurde.

Dagegen folgen die linguistischen Transkriptionen in *LangAge* dem Wortlaut in allen Details, motiviert durch ein deskriptives Selbstverständnis: Alle Varianten werden ohne Bewertung erfasst und eingeordnet. Unterschiedliche Ausdrucksweisen für das Gleiche werden nicht im Sinne der normativen Grammatik als „richtig“ und „falsch“ beurteilt. Stattdessen ermöglicht die Verschriftung „wie gesprochen“, die Formulierungsarbeit nachzuvollziehen und daraus Rückschlüsse auf die Sprecherinnen und Sprecher zu ziehen.

Solche Transkriptionsunterschiede müssen bei sammlungsübergreifenden Analysen also stets beachtet und – durch eine multimediale Forschungsumgebung

wie *Oral-History.Digital* unterstützt – direkt an den Mediendateien überprüft werden. Für korpuslinguistische Analysen schränkt dies die Vergleichbarkeit aber ein: Einfache Häufigkeitsvergleiche führen in die Irre. Basisangaben zu einer Sammlung wie die Gesamtlänge (üblicherweise in Tokens) sind unterschiedlich zu bewerten, je nachdem, ob Wortwiederholungen, Diskursmarker und Wortabbrüche mitgezählt werden. Daher wird die Länge der Interviews hier in Minuten angegeben.

5. Sammlungsübergreifende Analyse

In den folgenden Fallstudien werden die in der Oral History etablierten Methoden der explorativen Herangehensweise in der Anwendung korpuslinguistischer Methoden quasi nahtlos weitergeführt. Zunächst beschreiben wir die Eckdaten der in die Analyse einbezogenen italienischen und französischen Interviews aus *Zwangsarbeit 1939–1945* und aus *LangAge*. Das italienische Teilkorpus aus *Zwangsarbeit 1939–1945* (im Folgenden: ZWAR-it) umfasst neun Interviews mit zwei Frauen (einer als Jüdin und einer politisch Verfolgten) und sieben Männern (vier Militärinternierten, zwei als Juden und einem politisch Verfolgten). Die neun Video-Interviews dauern im Durchschnitt 4 Stunden und 20 Minuten. Eine Interviewerin hat fünf, die andere vier Interviews geführt (cf. Felsen & Frenkel 2008). Zum Zeitpunkt der Interviews 2005/2006 waren die Interviewten zwischen 75 und 85 Jahre alt.

Das französische Teilkorpus aus *Zwangsarbeit 1939–1945* (im Folgenden: ZWAR-fr) umfasst 18 Interviews mit ausschließlich männlichen Interviewten, die alle im Service du Travail Obligatoire (STO) Zwangsarbeit leisten mussten. Die 14 Video- und 5 Audio-Interviews dauern im Durchschnitt 2 Stunden und gut 20 Minuten. Eine Interviewerin hat zehn, die andere fünf Interviews geführt (cf. Granet-Abisset 2008). Zum Zeitpunkt der Interviews 2005/2006 waren die Interviewten zwischen 80 und 85 Jahre alt. Die Interviewten stammen aus verschiedenen Regionen, etwa die Hälfte aus der Umgebung Lyons.

Das hier ausgewertete Teilkorpus aus *LangAge* (im Folgenden: LangAge-oh.d) umfasst 32 Interviews. Die Interviews wurden 2005 sämtlich von einer Person, der Ko-Autorin, in Orléans und Umgebung geführt, die beteiligten 14 Männer und 18 Frauen waren zu diesem Zeitpunkt zwischen 71 und 94 Jahre alt. Die durchschnittliche Länge liegt bei 50 Minuten.

Für die sammlungsübergreifende Suche wurden die Transkripte aller drei Teilkorpora so bereinigt, dass jeweils nur die Äußerungen der Interviewten erhalten blieben. Analyse und Annotation wurden in SketchEngine (Kilgariff et al. 2014) vorgenommen. Diese korpuslinguistische Software ermöglicht es, eigene Daten in eine geschützte Cloud zu laden und programmgestützt zu lemmatisieren und nach Wortarten (Part-of-Speech, POS) zu taggen. Das bedeutet, dass im Hintergrund jedem Wort die Grundform (das Lemma) zugeordnet wird (also *werden* für *wird*), und die Wortart wie Verb oder Substantiv sowie morphologische Merkmale wie Plural oder Person ermittelt werden.

5.1. Schritt 1: Schlüsselwörter

In einer ersten korpuslinguistischen, induktiven Skizze wurden in SketchEngine die beiden französischen Korpora auf ihre Schlüsselwörter (Keywords) hin verglichen, nämlich solche Wörter erfasst, die signifikant häufiger in ZWAR-fr als in LangAge-oh.d vorkommen (in SketchEngine: Keywords mit Referenzkorpus LangAge-oh.d; lemma, ohne Unterscheidung von Groß- und Kleinschreibung). Zu den Schlüsselwörtern in ZWAR-fr gehören auf den ersten 10 Rängen überwiegend Orts- und Personennamen sowie Nationalitäten: *russes*, *breslau*, [Familiennamen], *bunzlau*, *grenoble*, *juifs*, *romans*, [Vorname]; Ausnahmen betreffen die Fabrikproduktion: *ciment* (Rang 5), *contremaître* (Rang 7). Darin zeichnet sich ab, dass die beiden Korpora viele Gemeinsamkeiten haben, Verben (*fumer* ‘rauchen’, Rang 17; *affecter* ‘zuweisen’, Rang 20) oder Pronomina kommen erst auf hinteren Rängen. Kennzeichnend für ZWAR-fr ist etwa das Verb *affecter*: Es beschreibt – stets im Passiv – wie die einzelnen oder Gruppen ihren Einsatzorten ‘zugewiesen wurden’. Die 30 Okkurrenzen dieses Verbs kommen in über der Hälfte der ZWAR-fr-Interviews vor, was die Bewertung als Schlüsselwort unterstreicht.

In der Analyse der Schlüsselwörter wurden die Häufigkeiten in ZWAR-fr mit den Häufigkeiten in *LangAge* als so genanntes Referenzkorpus verglichen. An der Tatsache, dass hauptsächlich Eigennamen auf den ersten Plätzen vorkommen, zeigt sich, dass im allgemeinen Wortschatz keine signifikanten Unterschiede zu finden sind. Diese lassen sich an wenigen Verben wie *fumer* und *affecter* festmachen.

5.2. Schritt 2: Personalpronomina

Für den zweiten Schritt war die Frage leitend, ob der Gebrauch der Personalpronomina spezifische Verteilungsmuster aufweist; dabei gingen wir von der Hypothese aus, dass die Singularform der Personalpronomina in ZWAR-fr und ZWAR-it seltener verwendet wird als in LangAge-oh.d, weil die zur Zwangsarbeit Verschleppten ein stark im Kollektiv geprägtes Erleben berichten. Zudem gingen wir geschlechtsspezifischen Unterschieden nach, wie sie Knowles et al. (2021) fanden. Die Personalpronomina sind ausreichend häufig, um aussagekräftige Werte zu erreichen.

Dabei ist zu berücksichtigen, dass das Italienische grundsätzlich keine Personalpronomen in Subjektposition erfordert. Die vergleichende Herangehensweise an die erste Person Singular und Plural ermöglicht es aber dennoch, relative Frequenzen zu betrachten. Dadurch wird auch der unterschiedliche Umgang mit Wortwiederholungen ausgeglichen, die in LangAge-oh.d wie gesprochen transkribiert werden, in ZWAR aber, wie oben dargestellt, nicht vollumfänglich.

Der Vergleich der Häufigkeit pro Teilkorpus (Tab. 1a) zeigt, dass in ZWAR-it und in LangAge-oh.d die Singularform ‘ich’ öfter vorkommt, während in ZWAR-fr die Pluralform ‘wir’ dominiert. Diese Angaben sind aber problematisch, denn Wortwiederholungen wurden in ZWAR-fr und LangAge-oh.d unterschiedlich transkribiert, und das Pronomen der ersten Person Singular *je* wird häufiger wiederholt als das *nous* der ersten Person Plural. Außerdem hat die im Gesprochenen übliche Form *on* für ‘wir’ zugleich die Bedeutung ‘man’, so dass sie für weiterführende

Untersuchungen zuerst disambiguiert werden müsste. Schließlich wird *noi / nous* auch häufig in Präpositionalphasen wie ‘für uns’, ‘von uns’ verwendet, was den Vergleich verzerrt.

Daher wurden im zweiten Schritt Kombinationen des ‘ich’/‘wir’-Pronomens und der Verbform ‘habe(n)’ ausgewertet; auf diese Weise werden Phrasen mit wiederholtem Pronomen nur einmal erfasst. Es handelt sich bis auf wenige Ausnahmen im einstelligen Bereich (*on a espoir* ‘wir haben Hoffnung’, *on a juste le titre* ‘wir haben nur den Titel’) dabei um Formen der zusammengesetzten Vergangenheit. Die Bedeutung von *on* ‘man’ ist in der zusammengesetzten Vergangenheitsform zwar möglich, aber weniger häufig, weil mit generalisierender Verwendung das eine Dauer beschreibende Imparfait üblich ist. Die Wortverbindung *on a* kann daher mit geringen Unschärfen zur ersten Person Plural gerechnet werden. Die deutlich selteneren Vorkommen von *io sono* bzw. *je suis* sowie *noi siamo* bzw. *nous sommes/on est* wurden nicht einbezogen: Dadurch würde die Vergleichbarkeit weiter eingeschränkt, weil diese Konstruktionen andere Verben bevorzugen und häufig im Passiv sind.

Pronomen	ZWAR-it absolut	ZWAR-it pro Std.	ZWAR-fr absolut	ZWAR-fr pro Std.	LangAge-oh.d absolut	LangAge-oh.d pro Std.
<i>io</i> <i>je / j'</i>	2094	54	9119	207,61	8603	325,7
<i>noi</i> <i>nous</i>	1094	28	2104	47,9	1307	49,9
<i>on</i>			8525	194	5291	200,3
Quotient 1.Sg./1.Pl.	1,9	1,9	0,9	0,9	1,3	1,3

Tab. 1a | Personalpronomina der ersten Person in den drei Teilkorpora mit Zahl der Vorkommen und Vorkommen pro Stunde (Lemma, Konkordanz-Tool SketchEngine)

Pronomen	ZWAR-it absolut	ZWAR-it pro Std.	ZWAR-fr absolut	ZWAR-fr pro Std.	LangAge-oh.d absolut	LangAge-oh.d pro Std.
<i>io ho</i> <i>j'ai</i>	134	3,4	1374	31,3	2416	91,5
<i>noi abbiamo</i> <i>nous avons</i>	46	1,2	64	1,5	105	4,0
<i>on a</i>			1121	25,5	909	34,4
Quotient 1.Sg./1.Pl.	2,9	2,9	1,2	1,2	2,4	2,4

Tab. 1b | Personalpronomina der ersten Person mit Auxiliar im Perfekt, in den drei Teilkorpora mit Zahl der Vorkommen und Vorkommen pro Stunde (Phrase, Konkordanz-Tool SketchEngine)

Diese konkreten, auf die Vergangenheit bezogenen Verwendungen (Tab. 1b) bestärken – bereinigt – die beschriebene Tendenz (Tab. 1a). Das Personalpronomen der ersten Person wird in ZWAR-it dreimal so häufig im Singular wie im Plural verwendet, wie der Quotient zeigt.

Die italienischen Interviewten berichten also häufiger in der Ich-Form als in der Wir-Form. Individuelle Unterschiede sind allerdings groß: Die beiden jüdischen Überlebenden Piero T. und Liliana S. sagen deutlich öfter *io* als die politisch Verfolgten oder die Militärinternierten. Wie lässt sich das interpretieren? Spiegelt die Ich-Form die Einsamkeit eines mit 14 Jahren allein in Auschwitz überlebenden Mädchens? Können politisch Verfolgte ihre Lebensgeschichte eher in einem kollektiven Schicksal verorten? Für eine Antwort sind weitere Analysen auf breiterer Basis nötig.

Die französischen Korpora unterscheiden sich deutlich: In LangAge-oh.d wird die Singularform mehr als doppelt so oft verwendet wie die Pluralform; in ZWAR-fr ist das Verhältnis ausgeglichener – freilich bei großen individuellen Unterschieden. So entfallen zwei Drittel aller *nous avons*-Vorkommen auf einen Teilnehmer (Roger C., za076; entspricht 9,9 Okkurrenzen pro Stunde). Mehr als dreimal pro Stunde verwendet Jean A. (za074) *j'ai* ‘ich habe’ in seinem Interview, damit entfallen knapp 12 Prozent der gesamten Okkurrenzen auf ihn.

Damit bestätigt sich die Hypothese, dass die zur Zwangsarbeit Verschleppten, übrigens alles Männer, mit *on* und *nous* viel mehr über gemeinsam Erlebtes berichten als die während des Krieges im Heimatort Verbliebenen. Interessant ist die Beobachtung, dass die in Anknüpfung an Knowles et al. (2021) formulierte Erwartung geschlechtsspezifischer Unterschiede nicht bestätigt wurde. Liliana S. verwendete die erste Person Singular sehr häufig, die zweite Frau in ZWAR-it weniger als halb so oft. Für eine belastbare quantitative Aussage dazu müssten mehr Interviews ausgewertet werden. Besonders häufig benutzen die ehemaligen Zwangsarbeiter das *wir* in Wendungen wie *on nous a dit* oder *ils nous ont mis*, schildern also die Wir-Gruppe als Objekt fremder Befehle (s.u.).

In der Analyse der Häufigkeiten persönlicher Fürwörter wurden die Schwierigkeiten des direkten Korpusvergleichs deutlich, die sich aus den unterschiedlichen Transkriptionsstandards ergeben, d. h. der Berücksichtigung oder Nicht-Berücksichtigung von Wiederholungen. Diesem Problem wurde dadurch begegnet, dass nicht isolierte Formen ausgewertet wurden, sondern Wortfolgen von Pronomen und Verb. Darin zeigten sich deutliche Unterschiede im Gebrauch der Personalpronomen, mit einer klaren Präferenz für die erste Person Singular in ZWAR-it und in LangAge-oh.de, während in ZWAR-fr die erste Person Plural dominiert.

5.3. Schritt 3: N-Gramme

Der nächste Schritt beruht nicht auf einer Hypothese, sondern folgt einer explorativen Bottom-up-Herangehensweise. Nun verglichen wir die für die drei Teilkorpora – ZWAR-it, ZWAR-fr und LangAge-oh.d – charakteristischen, also häufig wiederholten, Wortkombinationen, zunächst im Rahmen von 3–4 Wörtern. Häufige Wortkombinationen oder N-Gramme bilden typische Muster im Sprachgebrauch ab (cf. Bubenhofer 2009). Dabei wurden die ersten drei Ränge der Frequenzlisten

verglichen und danach N-Gramme mit lexikalischem Element, die einen neuen Aspekt brachten (Tab. 2).

Im Italienischen nimmt die Verbphrase *mi ricordo* 'ich erinnere mich' die erste Position ein und widerspiegelt somit das zentrale Motiv der Interviewten, aber auch des in der Erinnerungskultur positionierten Interviewprojekts. Diese auch gesprächsanalytisch von Pfänder et al. (2022) untersuchte Wendung wird im nächsten Abschnitt genauer betrachtet.

Bezeichnend ist auch die häufige Verbphrase *ci hanno portato* 'sie brachten uns': Das als Kollektiv der Zwangsarbeiter*innen referenzierte 'uns' wird zum Objekt, das von einem ungenannten Subjekt – meistens den Deutschen – von einem Ort zum anderen gebracht wird. In dieser syntaktischen Konstruktion bilden sich die Machtverhältnisse von handelnden Personen (Agens) und den ihnen Ausgelieferten (Patiens) ab. Deutlich wird diese ohnmächtige Passivität auch durch die anderen, zusammen mit *ci hanno* 'sie haben uns' genutzten Begriffe des Verladens (*caricato/i*) und Verlegens (*trasferiti*), Verräumens (*messo*), Verschickens (*mandato*) oder Befehlens (*fatto* + Verb im Infinitiv).

Im Französischen dominieren auf Rang 1 die im Gesprochenen sehr häufigen Präsentativa *il y avait* 'es gab' (ZWAR-fr) und *il y a* 'es gibt' (LangAge-oh.d). Darin zeigen sich die Charakteristika des gesprochenen Französisch; der Unterschied beider Teilkorpora besteht in der Verwendung der Vergangenheitsform 'es gab', die in ZWAR-fr häufiger ist, während die Präsensform 'es gibt' in LangAge-oh.d dominiert.

ZWAR-it (Okkurrenzen)	ZWAR-fr (Okkurrenzen)	LangAge-oh.d (Okkurrenzen)
<i>mi ricordo che</i> 'ich erinnere mich, dass' (161 Okk., Rang 1)	<i>il y avait</i> 'es gab' (1.666 Okk., Rang 1)	<i>il y a</i> 'es gibt' (910 Okk., Rang 1)
<i>un po' di</i> 'ein bisschen' (125 Okk., Rang 2)	<i>il y a</i> 'es gibt' (608 Okk., Rang 2)	<i>il y avait</i> 'es gab' (667 Okk., Rang 2)
<i>un certo momento</i> 'ein bestimmter Moment' (102 Okk., Rang 3)	<i>il y en</i> 'es [gibt/gab] davon' (520 Okk., Rang 3)	<i>c'est c'</i> 'das ist, das' (255 Okk., Rang 3)
<i>non lo so</i> (100, Rang 4)	<i>je ne sais</i> 'ich weiß nicht' (412 Okk., Rang 5)	<i>je sais pas</i> 'ich weiß nicht' (213, Rang 9)
<i>ci hanno portato</i> 'sie brachten uns' (42 Okk., Rang 20)	<i>à l'époque</i> (203, Rang 24)	<i>un petit peu</i> (189, Rang 16)

Tab. 2 | Auswahl häufiger Wortkombinationen (N-Gramme, 3–4 Wörter) in den drei Teilkorpora mit Zahl der Vorkommen und Rang (ohne Unterscheidung von Groß- und Kleinschreibung, SketchEngine)

Die Wendung *il y en* kommt 312 mal im Präsens vor (*il y en a* 'es gibt') und knapp 200 mal in der Vergangenheit (*il y en avait* 'es gab'). Beide Formen werden aber mit wenigen Ausnahmen im Kontext von Vergangenheitsformen verwendet, wie die Überprüfung ergibt: *Mais il y en a un qui a réussi à se sauver* 'aber es gibt einen, dem es gelungen war sich zu retten' (Robert W., za092, Band 2, 0:05:32). Zu dieser

sich nach Häufigkeit als dominant abzeichnenden Vergangenheitsperspektive passen adverbiale Angaben wie ‘damals’, frz. *à l’époque* und, auf dem dritten Platz in der italienischen Rangfolge, die unbestimmte Zeitangabe [*a(d) un] certo momento* ‘[in einem] gewisse[n] Moment’, die unten genauer betrachtet wird.

In allen drei Teilkorpora vertreten sind Einschränkungen in Bezug auf das eigene Wissen wie *je (ne) sais pas* und *non lo so* ‘ich weiß (es) nicht’ („epistemische markierungen“, Gülich & Pfänder 2022, 35). Je nach Kontext ist damit tatsächlich eine Abschwächung der epistemischen Autorität gemeint oder wird damit eine Planungspause überbrückt (cf. die Analyse in Grant 2010). Die Mengenangaben *un petit peu* bzw. *un po’ di* ‘ein bisschen (von)’ heben geringe Quantitäten hervor; in ZWAR-fr ist dieses Triple nicht auf den vorderen Plätzen (Rang 71, 108 Okkurrenzen). Damit kann einerseits ganz wörtlich eine geringe Menge angesprochen sein, aber auch eine Abschwächung des Gesagten, wie zum Beispiel in der Kombination mit ‘vielleicht’: *C’était peut-être un petit peu patriotique peut-être* ‘das war vielleicht ein bisschen patriotisch vielleicht’ (André D., za077, Band 1, 0:11:11). In allen Gruppen sind sich die Interviewten also der Grenzen ihres Erinnerungsvermögens bewusst und formulieren dies auch.

Die programmgestützte Ermittlung besonders häufiger Wortkombinationen führt zu einer weiteren Profilierung der drei Teilkorpora: die Vergangenheitsperspektive tritt deutlich hervor, weiterhin die Thematisierung des Erinnerns – und der Grenzen der Erinnerung. Dieser Punkt soll im nächsten Schritt differenziert und unter Berücksichtigung der jeweiligen Kontexte analysiert werden.

5.4. Schritt 4: ‘ich erinnere mich’

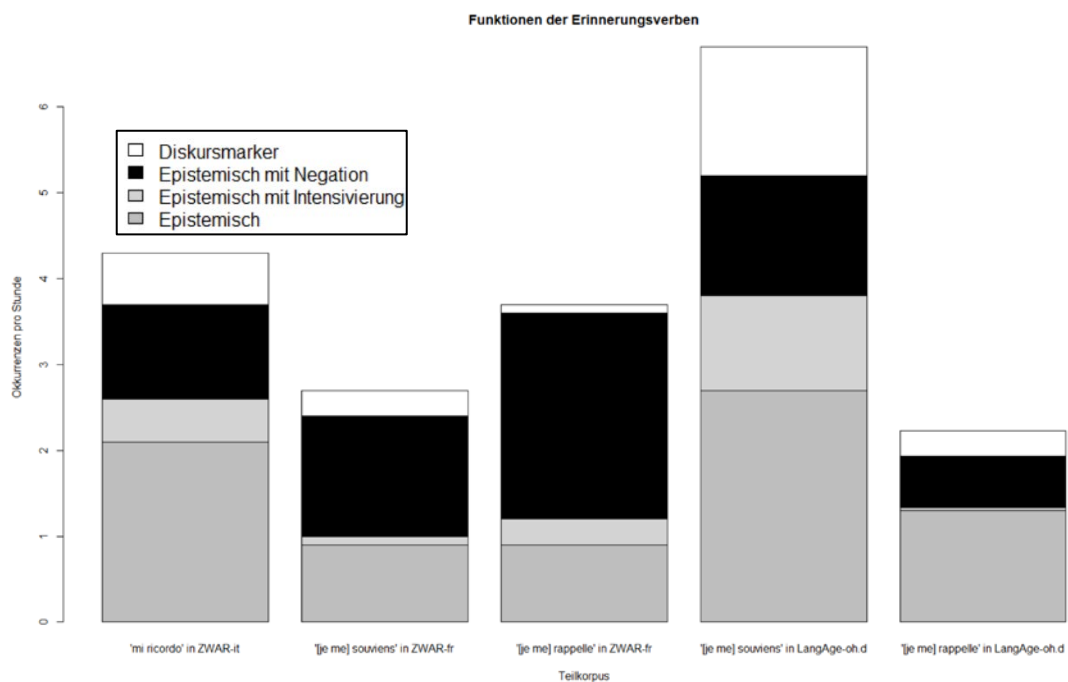
Aufbauend auf dem explorativ und datengetrieben ermittelten Ergebnis der hohen Frequenz von ‘ich erinnere mich’ wird im Folgenden gefragt, wie sich dies im Vergleich der Teilkorpora darstellt. Wie oben dargestellt, sind die programmgestützt ermittelten Frequenzen der Wortkombination ‘ich erinnere mich, dass’ zwar kennzeichnend für das Teilkorpus ZWAR-it, aber das Gesamtbild verlangt nach Differenzierung. Dazu gehört, dass die Verteilung über die einzelnen Interviews berücksichtigt wird, und dass unterschiedliche Verwendungsweisen sichtbar gemacht werden. Dafür wurde eine manuelle Annotation in SketchEngine umgesetzt, d. h. die einzelnen Vorkommen wurden manuell im Kontext geprüft und Kategorien zugeordnet.¹

Um die Verteilung über die einzelnen Interviews im Korpus gleichmäßig zu beobachten, wurden die ersten 20 Okkurrenzen pro Interview annotiert, bzw. alle, die überhaupt vorkamen. Diese Zahl wurde so gewählt, dass in den französischen Teilkorpora die höchsten Werte ausgeschlossen wurden, ohne dass auf eine zu große Anzahl Okkurrenzen verzichtet werden musste. In ZWAR-it wurden die Vorkommen von *ricordo* mit linkem und rechtem Kontext in SketchEngine erfasst (Konkordanz-Tool, nur Verbform), in ZWAR-fr und LangAge-oh.d wurden die französischen

¹ Für die Belegannotation von *ricordo* und *souviens/rappelle* danken wir Elena Bandt. Die Annotation wurde mit Screenshots gesichert; SketchEngine erlaubt keinen Export der annotierten Daten.

Äquivalente [*je me*] *souviens* bzw. [*je me*] *rappelle* (wiederum Verbform, erfasst wurde nur die 1. Person).

Ausgehend von den N-Gramm-Ergebnissen wurden dabei nicht die weiteren textuellen Funktionen analysiert (cf. Pfänder et al. 2022), sondern die Verwendung benachbarter Modifikatoren wie Verneinung oder Verstärkung berücksichtigt. Auf diese Weise wurde die epistemische Funktion von *ich erinnere mich* unterschieden von einer ebenfalls epistemischen, aber durch intensivierende Adverbien wie italienisch *bene* bzw. *benissimo* 'gut', 'sehr gut' und französisch *très bien* '(sehr) gut', *toujours* 'immer', *surtout* 'vor allem', *vraiment* 'wirklich' und *effectivement* 'tatsächlich' verstärkten Form und von der ebenfalls auf das eigene Wissen und Erinnern bezogenen, aber verneinten Form. In Anlehnung an die Studie zu engl. *I don't know* 'ich weiß nicht' (Grant 2010) wurden weiterhin Okkurrenzen unterschieden, in denen *ich erinnere mich* als Diskursmarker fungiert, also nicht inhaltlich bedeutsam ist, sondern in seiner Funktion, eine Planungspause zu füllen bzw. Zögern und Unsicherheit zu markieren. Um ein klares Kriterium zu formulieren, wurden Diskursmarker annotiert, die aus weiteren Hinweisen wie Satzabbrüchen oder Pausen im Kontext als Häsitiation eingeschätzt werden konnten.



3 | Funktionen der Erinnerungsverben in der ersten Person Singular: Ergebnisse der manuellen Annotation von bis zu 20 Okkurrenzen pro Interview (n=603)

Durch die Grenze von 20 berücksichtigten Fällen werden sehr hohe Vorkommen bei einzelnen Personen statistisch ausgeglichen, und eine Verzerrung der zusammengefassten Werte wird vermieden. Im Vergleich der Teilsammlungen zeigt sich dann, dass die Relevantsetzung des Erinnerns, also die epistemische Funktion mit und ohne Intensivierung, in LangAge-oh.d am häufigsten ist, es folgt ZWAR-it und dann ZWAR-fr. Umgekehrt ist die Reihenfolge bei den verneinten

Formen: das Nicht-Erinnern wird am häufigsten in ZWAR-fr erwähnt, dann in LangAge-oh.d und am seltensten in ZWAR-it.

Die beiden französischen Teilkorpora unterscheiden sich also deutlich: Die zwar auch biographisch, aber spezifisch zu ihrer lange zurückliegenden Zwangsarbeit Interviewten verwenden häufiger die verneinte Form der Erinnerungsverben. Sie schränken damit häufiger die Aussagekraft der erinnerten Inhalte ein als die allgemein-biographisch (und kürzer) interviewten *LangAge*-Zeitzeug*innen. Die verneinte Form ‘ich erinnere mich nicht’ ist im italienischen ZWAR-Teilkorpus deutlich seltener als im französischen ZWAR-Teilkorpus, wohingegen die positiven Formen häufiger verwendet werden. ZWAR-it und LangAge-oh.d zeigen in diesem Bereich also Gemeinsamkeiten, während ZWAR-fr sich davon abgrenzt. Dass sich somit Gemeinsamkeiten über die Sprachgrenzen hinweg ausprägen, unterscheidet das Bild der Erinnerungsmarker von der oben ausgewerteten Verteilung der Pronomina.

Wenn dagegen keine Obergrenze von 20 Okkurrenzen eingesetzt wird, treten die Besonderheiten einzelner Interviews hervor. Insbesondere in ZWAR-it, wo die Standardabweichung viel höher ist als in den anderen beiden Teilkorpora, zeigen die neun Interviews sehr unterschiedliche Häufigkeiten von *mi ricordo*. Die Spanne reicht von 8 bei Claudio S. bis zu 153 bei Liliana S. Die Auschwitz-Überlebende Liliana S. sagt auch im Vergleich aller Interviews aller Sprachen in *Zwangsarbeit 1939-1945* – ausweislich der deutschen Übersetzungen – am häufigsten *ich erinnere mich*.

Sie erzählt sehr anschaulich, im häufigen Wechsel zwischen konkret geschilderten Erlebnissen und deren historischer Einordnung, und betont ihre epistemische Autorität als Überlebende der Shoah. Als aktive Zeitzeugin in Schulen und Öffentlichkeit, seit 2018 als Senatorin, ist Liliana S. zudem eine geübte Rednerin. All dies sind individuelle Faktoren, die bei einer korpuslinguistischen Interpretation der Wendung ‘ich erinnere mich’ berücksichtigt werden müssen.

5.5. Schritt 5: Der Moment

Wie die N-Gramm-Analyse ergab, ist *in un certo momento* ‘in einem bestimmten Moment’ in den italienischen Interviews eine viel verwendete Wortkombination. In einzelnen Interviews ist dieser ‚Moment‘ besonders häufig, allein 42 mal etwa bei dem Militärinternierten Claudio S., fast in Art eines Füllwortes, das die Erzählung in Art eines ‚und dann‘ vorantreibt. Hier ist der ‚bestimmte Moment‘ meist gerade nicht genau bestimmt, sondern bezeichnet eine nicht präzise datierbare Veränderung nach einer längeren Phase von Wiederholung und Alltäglichkeit, etwa wenn die jugendliche Begeisterung für Mussolinis Faschismus der Ernüchterung weicht (za126, Bd. 1, 00:34:42), die Offizierslaufbahn vom Sturz Mussolinis in Frage gestellt wird (za126, Bd. 2, 00:07:32) oder der ausgehungerte Überlebenskampf im Kriegsgefangenenlager in lebensmüde Gleichgültigkeit umschlägt (za126, Bd. 3, 00:31:41).

Liliana S. dagegen nutzt die Wendung *un certo momento* gar nicht. Ihr ‚Moment‘ ist oft präziser (*in quel momento*); mit ihm beschreibt sie ihre Gefühle oder bewertet

ihr Verhalten in der zuvor geschilderten Situation, häufig einem emotionalen Erlebnis oder einem zentralen Wendepunkt ihrer Biographie. Solch ein *momento cruciale* war etwa die Entscheidung der 13-jährigen Liliana, auf der Flucht im Moment der Zurückweisung durch die Schweizer Grenzpolizei mit ihrem Vater umzukehren – trotz der nun drohenden Deportation. Im Rückblick bekräftigt sie die Bedeutung dieses Moments durch eine Wiederholung: „non mi sono mai pentita di questo, è stato un momento cruciale della mia vita, ma non mi sono mai pentita di questo“ – ‘ich habe das niemals bereut, das war ein entscheidender Moment in meinem Leben, ich habe das niemals bereut’ (za124, Bd. 2, 00:31:33). Die Wiederholung „un momento ... un momento“ verwendet Liliana S. auch im Moment der endgültigen Trennung vom Vater bei der Selektion in Birkenau (za124, Bd. 4, 00:07:55) oder nach dem Krieg bei der Geburt ihres ersten Kindes (za124, Bd. 4, 00:30:48).

In diesen verschiedenen Momenten der Interviewten bündeln sich also unterschiedliche Zeitlichkeiten, manche unerträglich lange, manche plötzlich überraschend. Der ‚Moment‘ zeigt die individuellen Bedeutungen dieser Zeitlichkeiten – in der erinnerten Erfahrung, der rückblickenden Erinnerung und der aktuell zu strukturierenden Erzählung.

6. Fazit

Diese Fallstudie zur Sprache des Rememberns in den lebensgeschichtlichen Interviews des Archivs *Zwangsarbeit 1939-1945* und des *LangAge*-Korpus zeigte in ihren fünf einzelnen Arbeitsschritten, wie die in der Oral History-Forschung etablierte Herangehensweise, Hypothesen aus den Daten heraus zu entwickeln, korpuslinguistisch umgesetzt werden kann. Statistische Verfahren wie die Ermittlung von Schlüsselwörtern (Keywords) und häufigen Wortkombinationen (N-Gramme) sowie die Extraktion von Personalpronomina (*ich* und *wir*) führten zu einer ersten Profilierung der drei Teilsammlungen. Diese quantitativen Merkmale bringen Gemeinsamkeiten und Unterschiede hervor, die teils die Teilkorpora der gleichen Sammlung (ZWAR-fr und ZWAR-it vs. LangAge-oh.d), teils die Teilkorpora der gleichen Sprache (ZWAR-fr und LangAge-oh.d vs. ZWAR-it) verbinden.

Auch quer dazu wurden Ähnlichkeiten festgestellt, so in der Analyse von ‘ich erinnere mich’, wo ZWAR-it und LangAge-oh.d sich näher waren in Bezug auf die positive und verstärkte Wendung. Diese Muster sind linguistisch aussagekräftig für die registerbezogene Einordnung der Interviews (cf. Biber, Egbert & Keller 2020) und die jeweils genutzten sprachlichen Möglichkeiten, um die eigene Position zum Ausdruck zu bringen.

Die zu diesen Befunden oben vorgeschlagenen historischen Interpretationen sind – auch aufgrund der heterogenen und im Umfang begrenzten Datengrundlage – nur als erste Annäherungen zu verstehen, die zu weiteren Analysen anregen wollen.

Zudem ergeben sich aus diesen Beobachtungen verschiedene Anschlussfragen. Sie betreffen zum einen die Rolle der „Zeitzeug*innen“ in den unterschiedlichen Interviewsettings, d. h. die Frage, ob die Personen ihr individuelles Zeugnis ablegen

oder eher ihrer Rolle als Vertretung einer Gruppe entsprechen wollen. Zum anderen zeigt sich daran die Notwendigkeit, persönliche, fallbezogene Informationen auszuwerten. Der biographische Zugang der Oral History und das geschichtswissenschaftliche Kontextwissen müssen dazu mit den Ergebnissen der korpuslinguistischen Analyse zusammengeführt werden.

Die methodischen und forschungspraktischen Verschiedenheiten zwischen Oral History und Linguistik erfordern eine hohe Sensibilität für die Standards der Datenaufbereitung. Dies wird am Beispiel unterschiedlicher Transkriptionsstandards deutlich, die an vielen Stellen die Vergleichbarkeit einschränken.

Die Perspektiven der sammlungsübergreifenden Analyse sind vielfältig; sie können zum Beispiel auf weitere Phänomenbereiche wie den Wortschatz bezogen werden, um generationstypische Erfahrungen etwa des Bombenkriegs besser zu verstehen. Der Aufbau der Forschungsumgebung *Oral-History.Digital* wird in Zukunft solche sammlungsübergreifenden Analysen unterstützen. Mit diesem digitalen Blick auf die Interviews ist freilich eine stärkere Distanz zu den lebensgeschichtlichen Erzählungen der Zeitzeug*innen verbunden, die aber stets eine sorgsame und respektvolle Interpretation erfordern.

Bibliografie

- Aijmer, Karin. 2015. „Pragmatic Markers.“ In *Corpus Pragmatics*, ed. Aijmer, Karin & Christoph Rühlemann, 195–218, Cambridge: Cambridge University Press.
- Apel, Linde, Almut Leh & Cord Pagenstecher. 2022. „Oral History im digitalen Wandel. Interviews als Forschungsdaten.“ In: *Erinnern, erzählen, Geschichte schreiben. Oral History im 21. Jahrhundert*, ed. Apel, Linde, 193–222, Berlin: Metropol.
- Apostolopoulos, Nicolas & Cord Pagenstecher (ed.). 2013. *Erinnern an Zwangsarbeit. Zeitzeugen-Interviews in der digitalen Welt*. Berlin: Metropol.
- Biber, Douglas, Jesse Egbert & Daniel Keller. 2020. „Reconceptualizing register in a continuous situational space.“ *Corpus Linguistics and Linguistic Theory* 16 (3), 581–616.
<<https://doi.org/10.1515/cllt-2018-0086>>.
- Blank, Andreas. 1997. *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Tübingen: Niemeyer.
- Bothe, Alina. 2019. *Die Geschichte der Shoah im virtuellen Raum. Eine Quellenkritik*. Berlin, Boston: De Gruyter Oldenbourg.
- Briggs, Charles L. 2005. „Sociolinguistic Interviews/ Soziolinguistisches Interview.“ In *Sociolinguistics: An International Handbook of the Science of Language and Society*. HSK 3.2, ed. Ammon, Ulrich et al., 1052–1062, Berlin, New York: De Gruyter.
- Browning, Christopher R. 2010. *Remembering Survival. Inside a Nazi Slave-Labor Camp*. New York: Norton.
- Bubenhofer, Noah. 2009. *Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse*.
<<https://doi.org/10.5167/uzh-111287>>.
- Deppermann, Arnulf. 2013. „Interview als Text vs. Interview als Interaktion.“ *FQS Forum: Qualitative Sozialforschung Social Research* 14(3).
<<https://doi.org/10.17169/fqs-14.3.2064>>.

- Eusterschulte, Anne, Sonja Knopp & Sebastian Schulze (ed.). 2016. *Video-graphierte Zeugenschaft: Ein Interdisziplinärer Dialog*. Weilerswist: Velbrück Wissenschaft.
- Felsen, Doris & Viviana Frenkel. 2008. „Die italienischen Deportationen 1943-45.“ In *Hitlers Sklaven. Lebensgeschichtliche Analysen zur Zwangsarbeit im internationalen Vergleich*, ed. Leh, Almut, Alexander von Plato & Christoph Thonfeld, 285-297, Wien: Böhlau.
- Freyburger, Philipp & Frank Jäger. 2021. „Emergentes Erinnern. Sensorische, kognitive und mediale (Spiel-)Räume in Oral-History-Interviews und literarischen Erinnerungstexten.“ *Romanische Forschungen* 133, 176-205.
- GAT 2. 2009. „Gesprächsanalytisches Transkriptionssystem 2 (GAT 2).“ *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 10, 353-402.
<<http://www.gespraechsforschung-ozs.de/heft2009/px-gat2.pdf>>.
- Gerstenberg, Annette, Valerie Hekkel, & Julie Kairet. 2018. *Corpus LangAge: Transcription Guide*. University of Potsdam: Romance Linguistics.
<<https://doi.org/10.5281/zenodo.6444538>>.
- Gerstenberg, Annette. 2009. „The Multifaceted Category of ‘Generation’: Elderly French Men and Women Talking about May ’68.“ *International Journal of the Sociology of Language* 200, 153–170.
<<https://doi.org/10.1515/IJSL.2009.049>>.
- Gerstenberg, Annette. 2011. *Generation und Sprachprofile im höheren Lebensalter: Untersuchungen zum Französischen auf der Basis eines Korpus biographischer Interviews*. Frankfurt am Main: Vittorio Klostermann.
<<https://doi.org/10.5281/zenodo.7315875>>.
- Gerstenberg, Annette. 2017. „A Difficult Term in Context: The Case of French STO.“ In *Oral History Meets Linguistics*, ed. Kasten, Erich, Katja Roller & Joshua Wilbur, 159–184, Fürstenberg: Kulturstiftung Sibirien.
<https://dh-north.org/siberian_studies/publications/orhili_gerstenberg.pdf>.
- Granet-Abisset, Anne M. 2008. „Berichte aus ‚Vichy-Frankreich‘. Über die Schwierigkeit, Erinnerungen zu sammeln und Geschichte zu schreiben.“ In *Hitlers Sklaven. Lebensgeschichtliche Analysen zur Zwangsarbeit im internationalen Vergleich*, ed. Leh, Almut, Alexander von Plato & Christoph Thonfeld, 103-112, Wien: Böhlau.
- Grant, Lynn E. 2010. „A Corpus Comparison of the Use of ‘I don’t Know’ by British and New Zealand Speakers.“ *Journal of Pragmatics* 42 (8), 2282–2296.
<<https://doi.org/10.1016/j.pragma.2010.01.004>>.
- Gülich, Elisabeth & Stefan Pfänder. 2022. „Erinnerungsmarkierungen in Zeitzeugenerzählungen. Episodische Rekonstruktion und epistemische Authentifizierung in Gesprächen mit Überlebenden der NS-Zwangsarbeitslager“. In *Romanistisches Jahrbuch*. Im Erscheinen.
- ISO 24624. 2016. Language resource management – Transcription of spoken language.
<<https://www.iso.org/standard/37338.html>>.
- Kangisser Cohen, Sharon. 2014. *Testimony and Time: Holocaust Survivors Remember*. Jerusalem: Yad Vashem Publications
- Kilgarriff, Adam et al. 2014. „The Sketch Engine: ten years on.“ *Lexicography* 1. 7–36.
<<https://doi.org/10.1007/s40607-014-0009-9>>.
- Klüger, Ruth. 1996. „Zum Wahrheitsbegriff in der Autobiographie.“ In *Autobiographien von Frauen. Beiträge zu ihrer Geschichte*, ed. Heuser, Magdalene, 405-410, Tübingen: Niemeyer.
- Knowles, Anne K. et al. 2021. „Mind the Gap: Reading across the Holocaust

- Testimonial Archive.“ In *The Holocaust in the 21st Century: Relevance and Challenges in the Digital Age*, ed. Cole, Tim & Simone Gigliotti, 216–241. United States: Northwestern University Press.
- Labov, William & Julie Auger. 1993. „The Effect of Normal Aging on Discourse: A Sociolinguistic Approach.“ In: *Narrative Discourse in Neurologically Impaired and Normal Aging Adults*. ed. Brownell, Hiram H. & Yves Joanette, 115–133, San Diego (CA): Singular.
- Laub, Dori, & Johanna Bodenstab. 2008. „Wiederbefragt. Erneute Begegnung mit Holocaust-Überlebenden nach 25 Jahren.“ In *Hitlers Sklaven. Lebensgeschichtliche Analysen zur Zwangsarbeit im internationalen Vergleich*, ed. Leh, Almut, Alexander von Plato & Christoph Thonfeld, 389–401, Wien: Böhlau.
- Leh, Almut, Alexander von Plato & Christoph Thonfeld (ed.). 2008. *Hitlers Sklaven. Lebensgeschichtliche Analysen zur Zwangsarbeit im internationalen Vergleich*. Wien: Böhlau.
- Michaelis, Andree. 2013. *Erzählräume nach Auschwitz: Literarische und videographierte Zeugnisse von Überlebenden der Shoah*. WeltLiteraturen Bd. 2. Berlin: Akademie Verlag.
- Möbus, Dennis. 2020. „Holleriths Vermächtnis – ein Beitrag zur Geschichte von Frauen in der EDV. Topic Modeling als Methode digitaler Sekundäranalyse lebensgeschichtlicher Interviews.“ *BIOS – Zeitschrift für Biographieforschung, Oral History und Lebensverlaufsanalysen* 2, 162–180. <<https://doi.org/10.3224/bios.v33i2.01>>.
- Nägel, Verena Lucia. 2016. „Zeugnis – Artefakt – Digitalisat. Zur Bedeutung der Entstehungs- und Aufbereitungsprozesse von Oral History-Interviews.“ In *Videographierte Zeugenschaft: Ein interdisziplinärer Dialog*, ed. Eusterschulte, Anne, Sonja Knopp, & Sebastian Schulze, 347–68, Weilerswist: Velbrück Wissenschaft.
- oh.d = *Oral-History.Digital, 2020-2022. Informationsinfrastruktur für die Erschließung, Recherche und Annotation von audiovisuellen narrativen Interviews*. Berlin: Freie Universität Berlin, Universitätsbibliothek. <www.oral-history.digital>.
- Pagenstecher, Cord & Stefan Pfänder. 2017. „Hidden Dialogues. Towards an Interactional Understanding of Oral History Interviews.“ In *Oral History Meets Linguistics*, ed. Kasten, Erich, Katja Roller & Joshua Wilbur, 185–207, Fürstenberg: Kulturstiftung Sibirien.
- Pagenstecher, Cord & Doris Tausendfreund. 2015. „Interviews als Quellen der Geschlechtergeschichte. Das Online-Archiv ‚Zwangsarbeit 1939-1945‘ und das ‚Visual History Archive‘ der USC Shoah Foundation.“ In *Geschlecht und Erinnern im digitalen Zeitalter. Neue Perspektiven auf ZeitzeugInnenarchive*, ed. Bothe, Alina & Christina Isabel Brüning, 41–67, Berlin: Lit.
- Pagenstecher, Cord. 2010. „‘We were treated like slaves.’ Remembering forced labor for Nazi Germany.“ In *Human Bondage in the Cultural Contact Zone. Transdisciplinary Perspectives on Slavery and Its Discourses*, ed. Mackenthun, Gesa & Raphael Hörmann, 275–291, Münster: LIT.
- Pagenstecher, Cord. 2017. „Oral History und Digital Humanities.“ *BIOS – Zeitschrift für Biographieforschung, Oral History und Lebensverlaufsanalysen* 1-2, 76–91.
- Pfänder, Stefan et al. 2022. „Doing remembering as a multimodal accomplishment. On the use of *mi ricordo* (‘I remember’) in Oral History Interviews.“ *Interactional Linguistics* 2 (1), 110–136.
- Sabrow, Martin, & Norbert Frei. 2012. *Die Geburt Des Zeitzeugen Nach 1945*. Göttingen: Wallstein.
- Schiffrin, Deborah. 2000. „Mother/daughter discourse in a Holocaust oral history.“ *Narrative Inquiry* 10 (1), 1–44. <<https://doi.org/10.1075/ni.10.1.01sch>>.

- Schuch, Daniel. 2021. *Transformationen der Zeugenschaft. Von David P. Boders frühen Audiointerviews zur Wiederbefragung als Holocaust Testimony*. Göttingen: Wallstein.
- Shenker, Noah. 2015. *Reframing Holocaust Testimony*. Bloomington: Indiana University Press.
- „Sklavenarbeit“: War die NS-Zwangsarbeit Sklaverei? Themenfilm mit Interview-Ausschnitten. 2011. In *Zwangsarbeit 1939-1945*. Webseite, Freie Universität Berlin.
<<https://www.zwangsarbeit-archiv.de/zwangsarbeit/ereignisse/sklavenarbeit/index.html>>.
- Taubitz, Jan. 2016. *Holocaust Oral History und das lange Ende der Zeitzeugenschaft*. Göttingen: Wallstein Verlag.
- TEI Guidelines, P5, chapter 8: Transcription of Speech.
<<http://www.tei-c.org/release/doc/tei-p5-doc/de/html/TS.html>>.
- Transcriber = Geoffrois, Edouard, Mark Liberman & Zhibiao Wu (ed.). 1998–2008. *Transcriber 1.5.1. A Tool for Segmenting, Labeling and Transcribing Speech*. DGA: Sourceforge.
- Thonfeld, Christoph. 2014. *Rehabilitierte Erinnerungen? Individuelle Erfahrungsverarbeitungen und kollektive Repräsentationen von NS-Zwangsarbeit im internationalen Vergleich*. Essen: Klartext.
- Wagner, Suzanne E. & Sali A. Tagliamonte. 2017. „What Makes a Panel Study Work? Researcher and Participant in Real Time.“ In *Panel Studies of Variation and Change*, ed. Wagner, Suzanne & Isabelle Buchstaller, 213–232, New York: Routledge.
- ZWAR = *Zwangsarbeit 1939–1945. Erinnerungen und Geschichte*, 2009–2022. Berlin: Freie Universität/Centrum für Digitale Systeme.
<<https://www.zwangsarbeit-archiv.de/>>.
Zugang zum Archiv: <<https://archiv.zwangsarbeit-archiv.de/de>>.

Zusammenfassung

Im interdisziplinären Spannungsfeld zwischen Geschichtswissenschaft und Linguistik werden in diesem Beitrag Oral History-Interviews als zentrale Ressource beider Fächer ausgewertet. Die Fallstudie untersucht die Sprache des Erinnerns in italienischen und französischen lebensgeschichtlichen Interviews des Archivs *Zwangsarbeit 1939-1945* und des *LangAge*-Korpus. In fünf Arbeitsschritten setzt sie die in der Oral History-Forschung etablierte Herangehensweise, Hypothesen aus den Daten heraus zu entwickeln, mit korpuslinguistischen Methoden um. So werden die Vorkommen der Personalpronomina *wir* und *ich* verglichen, und Schlüsselwörter (Keywords) und häufige Wortkombinationen (N-Gramme) ermittelt. Diese quantitativen Analysen werden durch die exemplarische Analyse häufiger Wendungen wie ‘ich erinnere mich’ oder ‘in einem bestimmten Moment’ vertieft. Dabei werden der biographische Zugang der Oral History und das geschichtswissenschaftliche Kontextwissen mit den Ergebnissen der korpuslinguistischen Analyse zusammengeführt. Als problematisch erweisen sich dabei unterschiedliche Transkriptionsstandards zwischen Oral History und Linguistik. Die neue Forschungsumgebung Oral-History.Digital unterstützt in Zukunft solche sammlungsübergreifenden und doch quellennahen Analysen lebensgeschichtlicher Interviews.

Abstract

Situated in the interdisciplinary fields of historiography and linguistics, this article evaluates oral history interviews as a resource valuable to both disciplines. The case study analyses the language of remembering in Italian and French life story interviews from the archive *Forced Labour 1939-1945* and the *LangAge* corpus. In five working steps, it implements the oral history approach of developing hypotheses from the data itself with methods from corpus linguistics. Occurrences of personal pronouns *we* and *I* are compared and keywords as well as frequent word combinations (*n-grams*) are identified. These quantitative approaches are complemented by exemplary analyses of frequent phrases such as ‘I remember’ or ‘at a certain moment’. In doing so, the biographical approach of oral history and historical contextual knowledge can be combined with the results of corpus linguistic analysis. Although different transcription standards of oral history and linguistics prove to be problematic, in the future the new research environment Oral-History.Digital will simultaneously support cross-collection and close-to-source analyses of life history interviews.