

## Tagungsbericht über das 10. Treffen des Netzwerks für Doktorand\*innen der Gesprächsforschung (DokGF) im Dezember 2024

**Veronika Sahlbach / Laura Schleicher / Regina Warmers<sup>1</sup>**

Am 13. und 14. Dezember 2024 konnte das Netzwerk für Doktorand\*innen der Gesprächsforschung (DokGF) sein Jubiläum in Form des 10. Treffens begehen.

Bei dem Netzwerktreffen der DokGF handelt es sich um ein selbstorganisiertes Format, das Promovierende der Gesprächsforschung in unterschiedlichen Stadien ihrer Qualifizierungsprojekte die Möglichkeit bietet, sich in einem offenen und unterstützenden Rahmen kennenzulernen und über aktuelle Forschungsthemen, Promotions-Herausforderungen, methodische Zugänge in der Gesprächsforschung und Vieles mehr austauschen zu können.

Das 10. Netzwerktreffen wurde von den Organisator\*innen Helena Konstanze Budde (Mannheim), Simon Arvo Titze (Heidelberg) und Samuel Lo Presti (Heidelberg) online ausgerichtet und moderiert. Rund 25 Doktorand\*innen trafen sich, um sich bei einer spielerisch moderierten Kennenlernrunde über ihre Promotionsprojekte und gemeinsame Interessen auszutauschen. Anschließend blickten sie gemeinsam auf die Entstehungsgeschichte der DokGF zurück. Neben dem informellen Austausch wurde sich durch einen Hands-On Workshop dem fachlichen Thema der 'automatischen Transkription' genähert. Spannende Datensitzungen, die Einblicke in die jeweiligen Promotionsprojekte boten, standen im Zentrum des Treffens. Im Folgenden wird ein Überblick über die inhaltlichen Ergebnisse der einzelnen Agenda-Punkte gegeben.

### **Maximilian Krug (Universität Duisburg-Essen) Beitrag zur Geschichte des Netzwerks anlässlich des 10. Treffens im DokGF**

Zum Jubiläumstreffen des Netzwerks hielt Mitgründer Maximilian Krug einen Vortrag, in dem er die Entstehung und Entwicklung des Doktorand\*innennetzwerks für Gesprächsforschung vorstellte.

Die Idee zur Gründung des Netzwerkes entstand auf der 5. ICCA-Konferenz<sup>2</sup> im Jahr 2018, wo sich die Gründer\*innen Elena Becker und Maximilian Krug kennenlernten. Der Wunsch, andere Promovierende des gleichen Fachgebietes kennenzulernen, konkretisierte sich bei der IEMCA-Konferenz<sup>3</sup> 2019, damals noch mit dem ungelenkten Namen *Doktorand\*innen der Konversationsanalyse und Interaktionalen Linguistik* (DNKAIL). Ziel war es, ein Netzwerk für Promovierende in der Konversationsanalyse und Interaktionalen Linguistik im deutschsprachigen Raum zu schaffen. Insbesondere sollten jene unterstützt werden, die an ihrer eigenen Universität keine etablierte Fachgemeinschaft haben.

Die zentralen Ziele des Netzwerks sind die Stärkung des informellen Austauschs unter Promovierenden und die gegenseitige Unterstützung bei fachlichen Fragen. Daneben werden gemeinsame Datensitzungen und niedrigschwellige Alternativen

<sup>1</sup> Die Autorinnen haben zu gleichen Teilen zur Ausarbeitung des Tagungsberichts beigetragen.

<sup>2</sup> International Conference on Conversation Analysis in Loughborough

<sup>3</sup> International Institute for Ethnomethodology and Conversation Analysis Konferenz in Mannheim



zu großen fachwissenschaftlichen Konferenzen organisiert. Ein wesentliches Strukturmerkmal des Netzwerks ist das Prinzip der "rollierenden Verantwortung": Die Organisation weiterer Treffen wird kontinuierlich an neue Mitglieder übergeben, um eine dynamische und nachhaltige Struktur zu gewährleisten.

Das erste Treffen des Netzwerks fand im November 2019 in Halle mit 15 Teilnehmenden statt. Sie diskutierten verschiedene Themen, darunter auch Herausforderungen, die insbesondere Frauen in der Promotion erleben. Aufgrund des erfolgreichen Auftaktes findet das Netzwerktreffen seither regelmäßig statt.

Seit 2019 besteht das Doktorand\*innennetzwerk der Gesprächsforschung in Abstimmung mit dem Verein für Gesprächsforschung.<sup>4</sup> Im Jahr 2022 wurde das Netzwerk in die *International Society for Conversation Analysis*<sup>5</sup> (ISCA) integriert, die seither die strukturelle Infrastruktur bereitstellt, unter anderem durch eine Mailingliste, eine Forschungswerkstatt und das Nachwuchsjournal *fokus:interaktion*.

Die Mitgliederzahl zeigt kontinuierliches Wachstum. Im Dezember 2024 waren 162 Mitglieder registriert, davon sind etwa 9 bis 10% mittlerweile Post-Docs. Innerhalb des Netzwerks gibt es verschiedene Schwerpunktbereiche, die sich aus Interaktionaler Linguistik, Konversationsanalyse und multimodalen Ansätzen zusammensetzen. Viele Mitglieder forschen zudem im Bereich der Angewandten Linguistik.

Die Präsentation von Maximilian Krug unterstrich die Bedeutung des Netzwerks für die Forschungslandschaft im deutschsprachigen Raum und zeigte die Erfolge sowie die wachsende Relevanz dieses Austauschs für Nachwuchswissenschaftler\*innen auf.

### **Thomas Schmidt / Renat Shigapov (Universität Mannheim) Hands-On Workshop: Automatische Transkription**

Renat Shigapov ist Referent für Data Science im Forschungsdatenzentrum (FDZ) der Universitätsbibliothek Mannheim; Thomas Schmidt arbeitet dort als Referent für Forschungsdatenmanagement für Data Literacy. Seit April 2024 arbeitet das Team am FDZ an einem eigenen Tool (*whisply*) mit dem Ziel, automatische Transkriptionen erzeugen zu können, um diese beispielsweise als Videountertitel zu verwenden oder in diversen Projekten Annotationen von Sprecher\*innen zuzuführen.

Mit der Veröffentlichung von ChatGPT und anderen Anwendungen von OpenAI im Herbst 2022 eröffneten sich im Bereich der automatischen Transkription von Audiodateien neue Möglichkeiten. Zunächst waren die Zugänge vor allem denjenigen vorbehalten, die über Programmierkenntnisse verfügten, mittlerweile gibt es jedoch Oberflächen,<sup>6</sup> die die Verwendung der *large language models* anwendungsfreundlicher anbieten. Das bietet neben kommerziellen und privaten Anwendungen auch der Forschung neue Möglichkeiten. Transkriptionen von Audiodateien waren bisher eine rein händische Leistung (mit technischer Unterstützung durch Programme wie MAUS<sup>7</sup> (Poerner/Schiel 2016) und EXMARaLDA<sup>8</sup> (Nolda 2024) oder andere Anwendungen), die zeitliche Ressourcen und Arbeitskapazitäten bindet.

<sup>4</sup> Die Onlinepräsenz des Netzwerkes ist abrufbar unter: <https://dokgf.conversationanalysis.org/>

<sup>5</sup> Mehr Informationen zur ISCA sind hier zu finden: <https://www.conversationanalysis.org/>

<sup>6</sup> Neben dem vorgestellten *noScribe* kann an dieser Stelle auch *aTrain* (Haberl et al. 2024) genannt werden.

<sup>7</sup> Abrufbar unter: <https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface>

<sup>8</sup> Abrufbar unter: <https://exmaralda.org/de/>

Eine Verbesserung dieses methodischen Vorgehens wünschen sich viele, besonders in der Gesprächsforschung, wo Transkriptionen ein fester Bestandteil der Untersuchungsabläufe sind. Deswegen war die Freude sehr groß, sich mit den neuen Möglichkeiten im Netzwerktreffen auseinanderzusetzen.

Der Workshop sollte den Teilnehmenden einen Einblick in den aktuellen Entwicklungsstand der automatischen Spracherkennung (*automatic speech recognition*, ASR) und einen Überblick zu Modellen, Ressourcen und Leaderboards geben. Ein weiteres Ziel war die Testung des vorgestellten Tools sowie die anschließende Evaluation der Transkriptionsergebnisse, um die Herausforderungen und Möglichkeiten herauszuarbeiten. Unter ASR definieren Malik et al. ein System, das die Umwandlung eines Inputs (bspw. eine Audiodatei) in einen Text ermöglicht (2021:9412). An dieser Idee wird seit den 1950er Jahren geforscht; in den letzten 20 Jahren beschleunigte sich das Verfahren durch die Untersuchung künstlicher neuronaler Netze und wissenschaftliche Durchbrüche im Bereich *deep learning*.

Mit dem 2022 von OpenAI publizierten ASR-Transformer-Modell *Whisper* (Radford et al. 2022) ging ein Tool online, das jedem und jeder die automatische Transkription ermöglicht. *Whisper* ist ein kostenfreies Open Source Tool, das lokal auf jedem Rechner genutzt werden kann. Das Besondere an diesem Modell sind der besondere Umfang von auditiven Trainingsdaten sowie die Sprachauswahl, in denen die Daten vorliegen. Je nach zur Verfügung stehenden Rechenkapazitäten können verschiedene Modellgrößen für die Transkription ausgewählt werden. Dabei überzeugen die größten Modelle mit der größten Genauigkeit. Die Exaktheit des besten *Whisper*-Modells entspricht in etwa den Ergebnissen einer menschlichen Transkription (Radford et al. 2022:6). Dabei gilt es aber zu beachten, dass diese Ergebnisse nur für die indo-europäischen Sprachen gelten. Neben der Genauigkeit überzeugt *Whisper* mit einer hohen Transkriptionsgeschwindigkeit (in Abhängigkeit von der verfügbaren Rechenkapazität), der Eignung für lange und umfangreiche Daten sowie einer relativen Robustheit gegen Umgebungsgeräusche. Um das zu belegen, zeigte Thomas Schmidt die Demonstration auf der Website von *Whisper* und wies dabei auch das Problem der Normalisierung, also der fehlenden Akkuratheit, hin.

Das vorgestellte Programm *noScribe* (Dröge 2025) ist eine auf den *Whisper*-Modellen basierende Open Source Software, die eine Oberfläche für die automatische Transkription von eigenen Daten anbietet. Der Vorteil dieser Lösung ist die lokale Installation der von *noScribe* verwendeten Modelle (small oder large-v2), die Erkennung verschiedener Sprecher\*innen sowie die anwendungsfreundliche Oberfläche. Daneben ist ein Editor Teil der Anwendung, mit dem die Daten nachträglich bearbeitet werden können. Thomas Schmidt betonte besonders den Aspekt der Datensicherheit, da mit der lokalen Anwendung ein hoher Schutz von sensiblen Audiodaten einherginge. Die transkribierten Daten können anschließend in verschiedenen Exportformaten heruntergeladen werden. Für die wissenschaftliche Weiterverwendung ist die Bereitstellung des VTT-Formates interessant, da damit ein Direktimport zu EXMARaLDA möglich ist. Als Hinweise bekamen die Teilnehmenden die Rechenzeiten der Transkriptionsprozesses erläutert, die sich auf das Zwei- bis Fünffache der Länge der Aufnahme belaufen. Die Genauigkeit der Transkription unterliegt der Qualität des Audio-Inputs. Des Weiteren gilt es zu beachten, dass auch hier die üblichen rechtlichen Rahmenbedingungen für Forschungsdaten

gelten. Nach dieser Einführung erhielten die Teilnehmenden die Möglichkeit, die Software unter Anleitung selbst zu testen.

Schritt für Schritt wurden die einzelnen Felder der Eingabemaske gezeigt und an einem exemplarischen Transkriptionsbeispiel erläutert. Nachdem die Transkription abgeschlossen war, zeigte Thomas Schmidt die Ergebnisse und stellte mögliche weitere Vorgehensweisen vor. So kann man beispielsweise mit dem Klicken auf einzelne Abschnitte die entsprechenden Audiosegmente gezielt anhören und eventuelle Korrekturen vornehmen. In einem nächsten Schritt wurde die Exportfunktion nach EXMARaLDA vorgestellt. Dabei wurde darüber diskutiert, welche Programme neben EXMARaLDA noch verwendet werden und wie der Datenimport in diese funktionieren könnte. Auch Erfahrungen von Workshopteilnehmenden mit ähnlichen Tools wurden besprochen und ausgetauscht. Dabei ging das Gespräch in die Erörterung der Anwendbarkeit auf die unterschiedlichen Projekte über. Dabei wurde auch *whispily* kurz gezeigt, das für Anwender\*innen weitere technische Möglichkeiten bietet. Die Teilnehmenden berichteten weiter, dass besonders bei mehrsprachigen Daten und Sprachwechsel in einer Aussage sowie dialektal eingefärbten Audiodateien die Ergebnisse nicht überzeugten, viele berichteten allerdings von ihrer Begeisterung über die vorgestellten Möglichkeiten.

Ein weiterer Schwerpunkt des Workshops war *verbatim speech recognition*. Diese Modelle versuchen, Transkriptionen ohne Normalisierungen (Weglassen von Verzögerungslauten, Wiederholungen) zu ermöglichen. Thomas Schmidt stellte dazu *CrisperWhisper* vor, welches genau diese Anforderungen erfüllt und frei zur Verfügung steht. Testen kann man das Tool unter anderem auf *huggingface*. Inwieweit die Versprechungen der fehlenden Normalisierung eingehalten werden, muss überprüft werden; eine erste Einschätzung gibt es nur vom Entwicklungsteam. Ebenfalls auf *huggingface* wurde auf das *Speech Recognition Leader Board* hingewiesen, auf dem verschiedene *large language Modelle* (LLM) hinsichtlich ihrer Genauigkeit sortiert sind.

Für die daran anschließende Diskussion brachten die Vortragenden folgende Leitfragen ein:

- Wie brauchbar sind die automatischen Transkriptionsergebnisse für die Gesprächsforschung?
- Was braucht es, um automatische Spracherkennung für Gesprächsforschung nutzbar zu machen?

Zuerst wurde die notwendige Überarbeitung der ausgegebenen Transkripte thematisiert. Für Forschungszwecke müssen die Daten meist genauer transkribiert werden, als sie derzeit noch ausgegeben werden können. Die Nutzung von ASR-Modellen kann hingegen einen ersten Überblick über die Daten ermöglichen, an den sich eine manuelle Transkription anschließt. Zur Diskussion stand dabei, ob durch die maschinelle Transkription ein tieferes Datenverständnis eingeschränkt werden könnte. Je mehr Gesprächsdetails für eine Untersuchung notwendig sind, desto weniger Nutzen bringt die Anwendung der ASR für den Transkriptionsprozess. Damit ist die Brauchbarkeit stark themen- und anwendungsspezifisch und lässt sich nicht einfach einschätzen. Die Anwendung in universitären Lehrveranstaltungen wurde in Frage gestellt, da dort das Erlernen der manuellen Transkriptionsmethode als besonders wichtig erachtet wird und ASR dies eher verhindern

würde. Trotz der geäußerten Skepsis über die Anwendbarkeit überraschte die aktuelle Detailgenauigkeit der ASR. Der Nutzen wurde insbesondere für eine Übersicht über große Datenmengen antizipiert. Thomas Schmidt merkte dazu an, dass Whisper erst zwei Jahre alt ist und dementsprechend Weiterentwicklungen wahrscheinlich sind. Auch der Ablauf des Datentrainings wurde diskutiert, da dieser einen Einfluss auf die Genauigkeit haben kann. Mit Hinweis auf den Artikel zu Whisper (Radford et al. 2022) wurde thematisiert, dass genau dieser Punkt nicht veröffentlicht wurde und dementsprechend keine Rückschlüsse möglich sind. Mit der amüsanten Überlegung, als Gesprächsforschung ein eigenes LLM zu erstellen, endete der informative und spannende Workshop.

**Rebekka Gallusser (Universität Paderborn)**  
**Gesprächsstruktur in Pflege-Patient\*in-Gesprächen (Datensitzung)**

Rebekka Gallusser ist im Arbeitsbereich *Deutsch als Zweitsprache, Deutsch als Fremdsprache und Mehrsprachigkeit* am Institut für Germanistik und Vergleichende Literaturwissenschaft an der Universität Paderborn tätig. Sie arbeitet im interdisziplinären Projekt *Sprachliche Teilhabe in der Pflegeausbildung stärken – schulische und pflegerische Praxis als Lerngelegenheit* (StePs)<sup>9</sup>, in dem unter anderem an Fragen zu Bedürfnissen und sprachlichen Bedarfen für migrationsbedingt mehrsprachige Pflegeauszubildende, Möglichkeiten zur Partizipation und Gestaltungsideen für die sprachbildende Ausbildung geforscht wird. Zurzeit beschäftigt sich Rebekka Gallusser mit der Gesprächsstruktur in Pflege-Patient\*in-Gesprächen; aus dieser Arbeit sind auch die Daten für die gemeinsame Sitzung. Damit soll das Desiderat der mehrsprachigen Sprachkonstruktionen ausgearbeitet und eine grundlegende Forschung zu Gesprächsstrukturen angestrebt werden, die der funktionalen Pragmatik zugeordnet wird.

In einem kurzen Input führte Rebekka Gallusser die Teilnehmenden der Datensitzung in die Ideen des Forschungs- und Projektumfeldes ein. Für die Untersuchung von Gesprächsstrukturen und mehrsprachigen Konstruktionen werden migrantische Pflegefachkräfte in ihrem Arbeitsumfeld betrachtet. Die Pflegekräfte haben in ihrem Herkunftsland eine Ausbildung abgeschlossen und müssen für die Anerkennung in Deutschland einen Prüfungsprozess bestehen. In enger Verzahnung durchlaufen die Betroffenen praktische Tätigkeiten in Pflegesituationen, die daneben noch schulisch begleitet werden. Die vorliegenden Daten stammen aus einer pflegepraktischen Situation in einem Raum mit zwei Patient\*innen. Vor Ort stand ein Mikrofon für die Aufzeichnung, dazu wurde die Situation von der anwesenden Forscherin beobachtet. Für die Datensitzung lag ein Transkript vor, das im HIAT-Format (Ehlich/Rehbein 1976) angefertigt wurde. Auf die Audiodaten konnte aus Datenschutzgründen während der Sitzung nicht zugegriffen werden. Rebekka Gallusser bat nach einem ersten Blick auf die Daten um die Fokussierung auf sprachlich-kommunikative Auffälligkeiten jeder Art.

In einem ersten Schritt sammelten die Teilnehmenden alle ersten Eindrücke, die sich zeigten. So wurde der Ablauf des Gesprächs besprochen, der abweichend von den erwartbaren Routinen verlief. Auffällig war die augenscheinliche Zweiteilung der Situation. In einem ersten Schritt war eine klassische Visite-Situation ablesbar, in der nach dem Wohlbefinden und Wünschen der Patient\*innen gefragt wurde. In

---

<sup>9</sup> Mehr Informationen zum Projekt: <https://www.stepsprojekt.de/>

der zweiten Hälfte folgte eine Vorstellung der Pflegekraft. Sie erklärte, dass sie bereits eine Pflegeausbildung in ihrem Herkunftsland absolviert hatte, jetzt aber für die Arbeit in Deutschland ein Programm durchlaufen und Dinge erlernen muss. Damit endete der Transkriptauszug.

Der erste Abschnitt der Situation wurde von den Teilnehmenden eher kurz thematisiert. Besonders wurde hervorgehoben, dass dieser Teil von der Pflegekraft sprachlich sehr gut umgesetzt wurde und bereits ein großes Vokabular vorhanden sein musste. Ursächlich dafür war wahrscheinlich die schulische Begleitung der Ausbildung, in der die Pflegekräfte auch Deutschunterricht besuchen. Überrascht waren die Teilnehmenden von den gewählten Fragen der Pflegekraft, die eher unüblich für die professionelle medizinische Situation waren. Als Erklärung wurden kulturelle Unterschiede oder das Herstellen einer freundlichen Beziehung zu den Patient\*innen diskutiert.

Der Fokus der Diskussion lag auf dem zweiten Abschnitt des Transkriptes und der Selbsterklärung der Pflegekraft. Der Teil wirkte losgelöst und nachgeschoben zum ersten Abschnitt. Nach dem erneuten Betrachten des Transkripts fiel die lange Pause zwischen den Abschnitten auf. Eine besprochene Interpretation war die vorgegriffene Rechtfertigung für ihre Sprachschwierigkeiten. Als alternative Erklärung wurde die Möglichkeit diskutiert, dass der Teil als Abschluss der Untersuchungssituation dienen könnte. Ebenso könnte die Pflegekraft etwas über sich als Person erzählt haben, um die Situation aufzulockern und eine emotionale Verbindung zu den Patient\*innen aufzubauen. Daraus ergab sich die Diskussion über die Rolle von Empathie in Pflegegesprächen. Aus der Gruppe wurde zur weiteren Recherche Maxi Kupetz (2015: Empathie im Gespräch. Eine interaktionslinguistische Perspektive) empfohlen.

Durch den gemeinsamen Austausch der Ideen innerhalb der Datensitzung erarbeitete die Gruppe neue Aspekte für die Analyse und alle konnten neue Perspektiven für ihre eigene Arbeit mitnehmen.

**Myriam Goll (RPTU Kaiserslautern-Landau)**  
**Explizite Sanktionierungen in (Polit-)Talkshows (Datensitzung)**

Myriam Goll ist als wissenschaftliche Mitarbeiterin an der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau (RPTU) tätig und bearbeitet dort ihr Promotionsprojekt *Sprecherwechselorganisation in (Polit-)Talkshows – eine medienlinguistische und multimodal interaktionale Analyse*. Forschungsgegenstand ihrer Untersuchung sind 'turbulente' Sequenzen aus (Polit-)Talkshows des öffentlich-rechtlichen Fernsehens, welche sich u.a. durch Unterbrechungen, Simultansprechphasen, (implizite und explizite) metapragmatische Strukturierungen charakterisieren (Birkner/Auer/Bauer 2020; Schegloff 2000). Diese werden hinsichtlich der Fragestellung, wie der Sprecher\*innenwechsel im Kontext des Mediums (Polit-)Talkshow organisiert wird, näher beleuchtet. Dabei bedient sich Myriam Goll sowohl der konversationsanalytischen als auch multimodal-interaktionsanalytischen Methodik und bezieht sowohl multimodale Ausdrucksressourcen der Gesprächsbeteiligten (z.B. Körperhaltung, Mimik) als auch die Gesprächsdarstellung nach außen (bspw. die Kameraperspektive) mit ein.

Für die gemeinsame Datensitzung wurde ein besonderer Fokus auf 'explizite Sanktionierungen' gelegt. Hierzu wurden von Myriam Goll exemplarisch zwei Sequenzen aus unterschiedlichen Polit-Talkshows mit Transkripten nach GAT2-

Konventionen (Selting et al. 2009) zur Diskussion gestellt. Beide Sequenzen charakterisieren sich durch Formen von "Quid pro Quo-Sanktionierungen" der Sprechenden, die in besonderem Erkenntnisinteresse der Analyse stehen. Im ersten Fallbeispiel handelt es sich um eine Sequenz der Sendung *Markus Lanz* (vom 3.4.2024), in der unter anderem Kai Ambos (KA) und Michael Wolffsohn (MW) eingeladen wurden. In der ausgewählten Sequenz wird der Turn von KA durch MW unterbrochen. Es folgen wenige Sekunden einer turbulenten Abfolge von simultanem Sprechen, Unterbrechungen, Überlappungen und Turn-Taking-Versuchen. Diese turbulente Phase wird durch eine explizite Sanktionierung des ursprünglich Sprechenden KA beendet, indem er darauf hinweist, dass er MW doch auch habe aussprechen lassen. MW räumt KA lächelnd das Rederecht ein; *Ja natürlich*, und es folgt eine (für die Daten ungewöhnliche) Pause, bevor MW die Argumentation fortführt. Trotz der Rüge findet keine Entschuldigung statt, obwohl die Regeln ratifiziert wurden, was als charakteristisch für Polit-Talkshows gewertet werden kann.

Im zweiten diskutierten Fallbeispiel, einer Sequenz aus der Sendung *Maybrit Illner* (vom 21.09.2023), kann die turbulente Sequenz nicht eindeutig beendet werden. Hier führt ein Interventionsversuch durch die Moderation, einen ausführlichen Turn von Jens Spahn zu unterbrechen (*Ich habe Ihnen zugehört...*), zu keiner eindeutigen Turn-Übergabe. Mehrere Gesprächsparteien versuchen, sich das Rederecht zu erkämpfen. Vielmehr wird deutlich, dass selbst die Moderation, die formal durch ihre Funktion das Rederecht innehaben sollte, mehrmalige Anläufe braucht um, gestützt von deiktischen Mitteln und dem Kamerafokus auf ihr, das Rederecht weiter zu geben, während im Off (abseits des Kamerafokus) die turbulente Gesprächssequenz fortgeführt wird. Eine Besonderheit für die Analyse der Daten ist die Gesprächsdarstellung durch die Kameraperspektive, die sowohl die holistische multimodale Analyse einschränkt als auch selbst Gegenstand der Untersuchung sein kann. So entziehen sich durch Kameraschnitt und Porträtaufnahmen bestimmte multimodale Hinweise der Sprechenden (insbesondere Handgesten und Körperhaltung) der Betrachtung durch Forschende. Gleichwohl scheint aber die Kameraeinstellung selbst als Methode der Sprecher\*innenorganisation zu fungieren.

Insgesamt bieten sich in der Detailbetrachtung der Daten spannende Anhaltspunkte zur Analyse der (nicht-)sprachlichen Gestaltung von "Quid Pro Quo-Sanktionierungen", die im Zusammenspiel von Mimik, Gestik, Prosodie, Kameraeinstellung und Konditioneller Relevanz der Unterbrechungen in turbulenten Gesprächssequenzen wirken und mehr oder weniger erfolgreich eine Turn-Übernahme einleiten. Die gewonnenen Eindrücke der Datensitzung lieferten den Teilnehmenden wertvolle Einblicke in die diskursiven Praktiken in Medienformaten.

#### **Tianyi Bai (Universität Mannheim)**

#### **Sprachprobleme in Freizeitaktivitäten von SprachpartnerInnen (Datensitzung)**

Tianyi Bai promoviert an der Universität Mannheim. In der Datensitzung präsentierte er seine Forschung zu "Sprachproblemen und ihre Lösungen in der freizeitaltivistischen Interaktion von SprachpartnerInnen". Die vorgestellten Daten stammen aus zwei unterschiedlichen Settings: einer Videospiele-Interaktion sowie einer Koch-Interaktion. Beide Interaktionen fanden zwischen einer Person mit Deutsch als Erstsprache und einer Person mit Deutsch als Fremdsprache statt.

In der Videospiele-Interaktion gibt die Deutschlernende Anweisungen zur Verwendung des Controllers. Auch in der Koch-Interaktion liegt das Wissen über den

Kochvorgang bei der Lernenden, während der Erstsprecher die epistemische Autorität über die Sprache innehat. Durch die Diskrepanz in den Sprachkenntnissen der Teilnehmenden sind die Interaktionen von häufigen Missverständnissen geprägt, deren Reparatur den Spiel- und Kochprozess häufig unterbrechen. Die Gesprächspartner\*innen stehen also vor der Herausforderung, zwei *joint projects* gleichzeitig bearbeiten zu müssen: das Spielen bzw. Kochen und gleichzeitig den Sprachlernprozess.

Die Interaktionen werden jeweils durch sprachliche Korrekturen des Erstsprechers begleitet. Die Wissensstände müssen also in der Interaktion ausgehandelt und vermittelt werden, weswegen *embodied interaction* im Fokus der Datensitzung stand. Dabei wurde insbesondere auf räumliche und objektbezogene Ressourcen in der Interaktion geachtet. Innerhalb der Datensitzung wurde herausgearbeitet, wie die beiden Gesprächspartner\*innen prosodische Markierungen und Gesten zur Überbrückung von Verständnisproblemen nutzen. Auch die Einleitung von Reparaturen, der Prozess der Wortsuche und die Closing Sequence wurden näher beleuchtet. Insgesamt wurde der erhöhte Kooperations- und Koordinationsaufwand der Gesprächspartner\*innen in Interaktionen mit unterschiedlichen Sprachkompetenzen anerkannt.

Bei der Diskussion der Daten wurde auch über die besondere Herausforderung der Transkription gesprochen: durch den Sprachlern-Kontext war es stellenweise schwierig, Worte der Gesprächspartner\*innen eindeutig zu identifizieren. Durch die gemeinsame Betrachtung des Datenmaterials konnten unterschiedliche Deutungsansätze diskutiert werden. Dies verdeutlicht die Relevanz von Austausch und Diskussion in Datensitzungen in der Gesprächsforschung.

Tianyi Bai reflektierte nach der Datensitzung, dass er ein besseres Verständnis seiner Daten erlangt hat und sich ermutigt fühlt, sein Projekt weiter zu verfolgen. Hier bestätigt sich das Ziel des Netzwerktreffens, durch gemeinsame gesprächslinguistische Datensitzungen Austauschmöglichkeiten für Forschende bieten zu können. Insgesamt wurde deutlich, dass die Interaktion von Sprachpartner\*innen in Tandemtreffen einen interessanten Rahmen für Forschungsprojekte darstellt. Tianyi Bai bereichert den Forschungsdiskurs durch die Anwendbarkeit seiner Analyse auf den Spracherwerb.

**Julia Schrittenlacher (Universität Duisburg-Essen)  
Anrede in kollaborativen Schreib-Interaktionen (Datensitzung)**

Julia Schrittenlacher arbeitet als wissenschaftliche Mitarbeiterin im Fachbereich Linguistik an der Universität Duisburg-Essen. Sie forscht zu kollaborativen Schreiberninteraktionen in Lernsituationen.

Bei den von Julia Schrittenlacher mitgebrachten Daten handelte es sich um eine mehrperspektivische Videoaufnahme von kollaborativen Sprachlerntandems der 8. Klasse. Nacheinander wurden jeweils zwei Schüler\*innen von der restlichen Klasse getrennt und beauftragt, in einem eigenen Klassenraum eine vorgegebene Aufgabe zu bearbeiten. Die Lernenden sollten in der Rolle als Polizist\*in einen Brief an ihren Chef schreiben, um den Tathergang eines fiktiven Kriminalfalls zu erklären. Dafür standen Ihnen als Material die gedruckte Aufgabe sowie ein Laptop zur Verfügung, mit dem sie den Brief erstellen sollten. Jedes Tandem wurden dabei

schreibleistungsheterogen zusammengestellt und alle erhielten die gleiche Aufgabenstellung. Für die Untersuchung wurden die Eingaben der Schüler\*innen sowie die Konversationen aufgezeichnet. Als Material lagen anonymisierte Videoausschnitte sowie Transkripte für die Datensitzung vor. Der Betrachtungsschwerpunkt bei den beiden vorbereiteten Beispielen sollte auf der Verschränkung von Mündlichkeit und Schriftlichkeit und der Nutzung multimodaler Ressourcen liegen.

Im ersten Fallbeispiel konnte man zwei Lernende in den ersten Minuten der Aufgabenbearbeitung beobachten. Die Schüler\*innen lasen die Aufgabenstellung und diskutierten kurz darüber, wer das Schreiben übernehmen soll. Eine Person bot sich an und begann am Laptop *Hallo* zu tippen, woraufhin beide über eine geeignete Anrede für den Brief diskutierten. Die Verantwortung über den Schreibprozess wechselte in diesem Moment zur/zum zweiten Lernenden. Die Bearbeitung lief daraufhin ohne eine erneute Thematisierung der Verantwortlichkeit weiter. Die Teilnehmenden der Datensitzung sprachen sehr lange über diese Situation und wiesen auf die Vielschichtigkeit der Daten hin, die die Analyse erschwerte. Als erster möglicher Ansatz wurde die Interpretation des Blickkontakts besprochen. Die Lernenden schauten abwechselnd auf die Aufgabenstellung, den Bildschirm und einander an, was einen interessanten Ansatz darstellen könnte. Auch das erste getippte Wort *Hallo* war Thema der Gespräche, welches als Exploration der Verschriftungstätigkeit gewertet wurde. Andererseits könnte der/die Lernende dies auch als vorgeschlagene Anrede für den Brief gemeint haben. Neben weiteren Momenten wurde auch das haptische Verrücken des Laptops nach dem Aufgabenwechsel besprochen. Dieses Einrichten des Arbeitsplatzes, bei dem auch die Aufgabenstellung zurechtgelegt wurde, wirkte als eine finale Klärung der Rolle der verschriftenden Person, welche jedoch im Prozess weiterhin dynamisch blieb.

Die zweite Sequenz fand im gleichen Setting wie die erste statt. Im Unterschied zur ersten Szene fiel hier das planvolle und kooperative Vorgehen der Lernenden auf. Die Aufgaben zwischen den beiden waren klar verteilt. Eine Person schrieb am Laptop, die andere plante den Inhalt des Textes und beobachtete den Schreibprozess. Trotz der Verteilung hob die Gruppe die kooperative und zielführende Arbeitsweise des Tandems hervor. Die Schüler\*innen gingen wertschätzend miteinander um, schlugen sich gegenseitig Formulierungen vor und fanden gemeinsam Lösungen. Auch hier wurde die Vielschichtigkeit der Daten thematisiert. Julia Schrittenlacher fragte daraufhin die Teilnehmenden, ob der selbst entwickelte Transkriptionsstil die multimodalen Aspekte der Situation gut darstelle, was mit wenigen Hinweisen bestätigt wurde. Die Gespräche zeigten insgesamt eindrucksvoll, wie vielperspektivisch und reichhaltig die Gespräche in dem kollaborativen Setting sind.

**Jean-Marie Nau (Universität Luxemburg)**  
**Bahá'í consultation in various settings (Datensitzung)**

Jean-Marie Nau ist Doktorand der Bildungswissenschaften an der Universität Luxemburg. Er untersucht Bahá'í Konsultationen aus konversationsanalytischer Perspektive.

Bahá'í Konsultationen kommen aus der Anwendung der religiösen Gemeinschaft des Bahá'ítums. Das Bahá'ítum (Stockman 2021) ist eine weltweite Religion, die im 19. Jahrhundert gegründet wurde und deren zentrale Prinzipien die Einheit der Menschheit, die Gleichberechtigung der Geschlechter und der Einklang

von Wissenschaft und Religion sind. Bei den Konsultationen handelt es sich um einen strukturierten und prinzipiengeleiteten Austausch, der darauf abzielt, Probleme zu lösen und Entscheidungen zu treffen, indem man kollektives Wissen und unterschiedliche Perspektiven einbezieht. Dabei soll letztendlich versucht werden, sich von persönlichen Ansichten zu lösen und nicht auf der eigenen Meinung zu beharren. Die Prinzipien der Bahá'í-Konsultation umfassen unter anderem die Suche nach Wahrheit, das Zuhören ohne Unterbrechung, die Vermeidung von Konflikten und das Streben nach Einheit. Dadurch entsteht eine spezifische Gesprächsform, die weder vollständig institutionell noch rein informell ist.

Das Interessante an diesen Konsultationen ist, dass sie Merkmale institutioneller Settings (z.B. Prinzipien und Abläufe) aufweisen, gleichzeitig aber in familiären oder anderen informellen Kontexten stattfinden. Es gibt keine festen Rollen und keine Leitung, was sie von klassischen hierarchischen Beratungen unterscheidet.

Die englischsprachigen Daten enthalten ein Gespräch zwischen Mutter und Tochter sowie dem Ehemann der Tochter, der hinter der Kamera sitzt und sich nur geringfügig am Gespräch beteiligt. Die beiden Frauen reflektieren die Möglichkeiten einer medizinischen Behandlung, die der Mutter nach ärztlicher Konsultation vorgeschlagen wurde. Der Verlauf des Gesprächs zeigt, wie die Teilnehmerinnen sich den Möglichkeiten der Behandlungsform annähern, ihre Wissensstände und Expertise aushandeln, ihre Meinungen in Bezug auf Vor- und Nachteile der jeweiligen Entscheidung einbringen und gemeinsam abwägen.

In der Diskussion der Daten wurde zunächst im Zuge des *unmotivated looking* beleuchtet, welche Meinungen die Gesprächsteilnehmenden mitbringen und wie sie sich in Bezug auf Wissensstände und Expertise positionieren. Dabei wurde über die Herstellung und Aushandlung von *common ground* gesprochen und welche Rolle körperliche und verbale Ko-Konstruktionen in der Interaktion einnehmen. Beispielsweise wurde der rekurrente Verweis auf die Einzelheiten und Unterschiede der Behandlungsformen durch multimodale Äußerungen, also das Zusammenspiel aus Sprache und Geste, eingegangen, wobei einzelne Gesten repräsentativ für die Behandlungsmethode festgelegt wurden. Der Port, der für die Dauer der Behandlung in der Brust der Patientin verbleiben würde, wurde mehrmals durch Berührung an der Brust dargestellt. Gleiches gilt für die Nadel, die in der anderen Behandlungsform über Stunden im Arm verbleiben würde, aber zwischen den Sitzungen entfernt werden würde.

Abschließend wurde die Sequenz mit Hinblick auf *upgrades* und *downgrades* der Möglichkeiten durch die Gesprächsteilnehmenden analysiert. Insgesamt argumentiert die Tochter für die Option des Ports und hebt diese durch diverse *upgrades* hervor, während die andere Behandlungsform *downgrades* durch die Tochter erfährt. Beide Meinungsäußerungen unterstützt sie durch Argumente, die ihre Sorge zum Ausdruck bringen. Die Mutter hingegen argumentiert für die von der Tochter abgelehnte Behandlungsform mittels intravenöser Nadel, wobei sie ihre Argumente hauptsächlich durch *downgrades* des Ports ausdrückt.

Die Analyse des Videoausschnitts zeigt, wie die Gesprächsteilnehmenden zwischen *upgrades* und *downgrades* alternieren und wie sie multimodale Ko-Konstruktionen in ihrer Argumentation einsetzen. Dabei wird auch deutlich, dass es den Gesprächspartnerinnen schwer fällt, das Ideal der Bahá'í-Konsultationen, sich von persönlichen Ansichten zu lösen, zu erreichen. Jean-Marie Nau reflektiert, dass die Diskussion seiner englischsprachigen Daten auf Deutsch ihm neue Einblicke in die

Interaktion gegeben hat. Die Diskussion hat vorherige Erkenntnisse und Annahmen bestätigt.

**Vivien Lohmer (Universität Bielefeld)**

**Gestik und Multimodalität in Erklärungen (technischer) Artefakte<sup>10</sup> (Datensitzung)**

Vivien Lohmer ist an der Universität Bielefeld im interdisziplinären Forschungsbereich SFB TRR318 *Constructing Explainability* tätig. Im Projekt A04 *Integration des technischen Modells in das Partnermodell beim Erklären von digitalen Artefakten* arbeitet sie gemeinsam mit ihren Kolleg\*innen an Spielerklärungen. Unter Rückbezug auf die initiale Annahme des Projekts, dass technische Artefakte eine inhärent duale Natur besitzen, forscht sie zu diskursiven Strukturen, Multimodalität und Gestik in Alltagserklärungen mit dem Fokus auf die gestische Herstellung eben dieser Artefakte (Kendon 2004; McNeill 1992, 2005; Quasthoff/Heller/Morek 2017). Unter dualer Natur technischer Artefakte ist zu verstehen, dass diese Artefakte auf der einen Seite eine physische und/oder digitale Form bzw. Struktur (*Architektur*) aufweisen. Auf der anderen Seite sind Artefakte zur Erfüllung bzw. Verwendung einer inhärenten oder individuellen Funktion (*Relevanz*) entsprechend gestaltet (Kroes 1998, 2009; Schulte/Budde 2018; Terfloth et al. 2023).

Für ihre Dissertation zu diskursiven Strukturen und Multimodalität in Erklärungen von (technischen) Artefakten schaut sich Vivien Lohmer Spielerklärungen des 2-Personen-Brettspiels *Quarto!* an. Eine der Personen (*Explainer*, kurz: EX) wurde gebeten, *Quarto!* einer anderen Person (*Explainee*, kurz: EE) zu erklären, ohne dass das Spiel während der Erklärung vorhanden war. In ihrer Datensitzung zeigte sie zwei Beispiele aus diesem Korpus ihres Dissertationsprojekts. Die Grundlage ihrer Arbeit bilden 22 videografierte dyadische Erklärsequenzen zum Brettspiel *Quarto!*. Mit dem Ziel sowohl diskursive Strukturen (ein Schwerpunkt des Dissertationsprojekts) als auch die strukturelle Komponente der dualen Natur des Spiels *Quarto!* als Beispiel eines (technischen) Artefakts zu explorieren, betrachtet sie die Daten nicht nur aus gesprächsanalytischer Perspektive heraus, sondern folgt methodisch ebenfalls der qualitativen Inhaltsanalyse. Daher wurde der gesamte Korpus anhand verschiedener Kodierschemata aufbereitet. So entwickelte Vivien Lohmer ein Schema zur Kodierung hinsichtlich der diskursiven Strukturen in Anlehnung an das GLOBE-Modell (Morek 2012; Quasthoff/Heller/Morek 2017). Auf inhaltlich-semantischer Ebene wurde entsprechend der Theorie zur dualen Natur technischer Artefakte nach *Architektur* und *Relevanz* kodiert (Kroes 1998, 2009; Terfloth et al. 2023).

Im Rahmen ihrer Datensitzung standen besonders die multimodalen und gestischen Aushandlungen der Interaktant\*innen in Bezug auf die Beschreibung des inhaltlich-semantisch vergleichbaren Elements der Spielzüge (*Architektur*) und die daraus resultierenden strategischen Implikationen (*Relevanz*) im Fokus. Gestaltet als offene Datensitzung gingen die Teilnehmer\*innen der Datensitzung ohne Vorwissen über die theoretischen Vorannahmen in die Betrachtung und Diskussion der Daten. In der Diskussion der Daten wurden sowohl die verbalen als auch multimodalen Unterschiede zwischen der normativen Beschreibung von Spielelementen

---

<sup>10</sup> Der Bericht zur sechsten Datensitzung wurde von Vivien Lohmer selbst verfasst, da den Autorinnen ein Protokoll zur Sitzung nicht vorlag. Wir danken Vivien Lohmer für ihre Arbeit und interessanten Einblicke in ein spannendes Forschungsfeld.

(bspw. Spielfeld oder Figuren) und bewertenden Formulierungen im Bereich strategischer Implikationen herausgearbeitet. Letztere resultieren aus Sicht der Teilnehmer\*innen der Datensitzung direkt aus den normativen Beschreibungen der Spielelemente. So zeigte sich in der Diskussion, dass der Wechsel zwischen normativer Beschreibung hin zu daraus resultierenden strategischen Implikationen, Praktiken der Bewertung oder der *stance*-Formulierung mit einherzugehen scheint. In beiden Sequenzen haben die Teilnehmer\*innen der Datensitzung diese Praktiken in unmittelbarer interaktiver Umgebung herausarbeiten können. Beispielsweise wurde besonders die interaktive Funktion des Adverbs *also* als eine Art Markierung des Übergangs stark diskutiert. Auch mit Fokus auf Gestik und Multimodalität wurden Unterschiede bezüglich normativer Beschreibung und strategischen Implikationen erörtert. Die Gestik der Interaktant\*innen verläuft während der Deskription von Spielelementen zeitlich synchron, während sie bei der Darstellung strategischer Implikationen eher vage, unbestimmt und zeitlich leicht versetzt bleibt.

**Regina Geisler-Knünz (Universität Wien)**  
**Patient\*innenseitiger Widerstand in Ärzt\*innen-Patient\*innen-Interaktionen**  
**(Datensitzung)**

Regina Geisler-Knünz arbeitet als DaF/DaZ-Trainerin in Dornbirn (Österreich). Innerhalb ihres Promotionsprojektes an der Universität Wien erforscht sie Entscheidungsfindungsprozesse in Ärzt\*innen-Patient\*innen-Interaktion. Dem zugrunde liegt ein Korpus von Gesprächen über Behandlungs- und Therapieentscheidungen im niedergelassenen Bereich (bspw. ob ein spezifisches Medikament weiter verschrieben wird). Während im wissenschaftlichen Diskurs spezifische sprachliche Muster z.B. bei Begrüßung oder Anamnese in Ärzt\*innen-Patient\*innen-Interaktionen bereits beschrieben wurden, richtet Regina Geisler-Knünz den Fokus auf die Entscheidungsfindungsprozesse der Ärzt\*innen bzw. die Aushandlungsprozesse mit Patient\*innen.

Für die Datensitzung wurde ein spezifischer Analysefokus auf Dissens und Widerstand bei Behandlungs- und Therapieentscheidungen durch Patient\*innen gelegt. Als Diskussionsgegenstand lag die Tonaufnahme einer Gesprächssequenz zwischen einem Hausarzt und einer Patientin und deren Transkription nach HIAT-Konventionen vor. Die Patientin war ohne Termin in die Praxis gekommen, um eine Hautirritation untersuchen zu lassen. Während deren Untersuchung erfragt der Arzt den Stand zur Einnahme eines Nervenmedikaments, das er ihr verschrieben hatte. Die Patientin legte begründet dar, warum sie das Medikament nicht weiter genommen hatte, während der Arzt argumentierte, warum eine weitere Einnahme sinnvoll sei und die empfohlene Dosis beschließt.

Wie üblich für institutionelle Gespräche bedingt die Zeitknappheit häufige Unterbrechungen des Arztes. Ebenso kennzeichnet sich die Sequenz durch häufiges simultanes Sprechen, was beidseitigen Widerstand gegen die Argumentation der anderen Person vermuten lässt. Der anfänglich zaghaft gezeigte Widerstand der Patientin (inhaltlich gekennzeichnet durch den Hinweis auf Mangel eines positiven Effekts) wird im weiteren Verlauf des Gespräches durch minimale compliance marker wie "hmmm" abgelöst. Zwar spricht der Arzt eine eindeutige Empfehlung aus, verwendet allerdings "Wir"-Formulierungen - mutmaßlich, um einen kollaborativen Entscheidungsprozess zu signalisieren. Ebenfalls scheint er durch häufige Wiederholung seiner Argumente eine Zustimmung hervorzurufen.

Insgesamt fällt auf, dass Patientin und Arzt in unterschiedlichen zeitlichen Perspektiven argumentieren. Während die Patientin über ihre bisherigen Erfahrungen berichtet, geht der Arzt nur bedingt auf diesen vergangenheitsbezogenen Bericht ein, sondern setzt die zukünftige Einnahme in den inhaltlichen Fokus des Gesprächs. Im Kontext eines funktional-pragmatischen Verständnisses der Interaktion könnte man den Aktant\*innen unterschiedliche Ziele in der Gesprächsgestaltung unterstellen: Planung der Behandlung als Ziel des Arztes und Teilen der bisherigen Erfahrungen als Ziel der Patientin.

### **Abschluss und informeller Austausch mit Fokus auf Transkriptions- und Analysesoftware in der Gesprächsforschung**

Um dem Wunsch nach informeller Vernetzung auch im digitalen Format entsprechen zu können, ging die Veranstaltung in den finalen Agendapunkt des Treffens über: *Networking* und *Socializing*. Hierzu wurden drei Breakout-Räume mit thematischen Schwerpunkten eröffnet. Neben Themen wie der persönlichen Entwicklung und Karriere im universitären Kontext gab es auch ein großes Interesse daran, sich über das softwaregestützte Vorgehen in der Gesprächsanalyse auszutauschen. Eigene Erfahrungen mit verschiedenen technologischen Werkzeugen und Methoden zur Transkription und Aufbereitung von Daten wurden geteilt.

Die Debatte konzentrierte sich auf die Herausforderungen und Lösungen bei der Integration von Transkriptionssoftware in verschiedene Arbeitsprozesse. Erfahrungen mit Software-Tools wie EXMARaLDA, FOLKER, MAXQDA, Transana, ELAN oder f4 wurden besprochen. Dabei wurde deutlich, dass die Wahl des jeweils präferierten Tools häufig stärker von entsprechenden Vorkenntnissen der Forschenden abhängt, als dass das geeignetste Tool für das jeweilige Forschungsdesign ausgesucht würde.

Dennoch ließen sich eindeutige Vor- und Nachteile der jeweiligen Software-Lösungen vergleichen. Während EXMARaLDA (bzw. ebenso wie dessen Adaption FOLKER) die Transkription nach GAT2-Konventionen besonders gut unterstützt, bietet die Software keine Lösung zur weiteren Analyse. Wiederum eignen sich MAXQDA und Transana besonders effektiv für die Markierung, Kodierung und Kategorisierung von Daten. Die MAXQDA-Transkripte lassen sich zwar leicht als Word exportieren, die Übertragung der gespeicherten Kategorien lassen sich jedoch nicht als Tabelle in andere Tools wie z.B. Excel übertragen. Dieser Umstand führt dazu, dass viele Anwender\*innen ihre Daten direkt in Excel aufbereiten, wobei sie Timestamps als Zeilenangaben verwenden, um Daten effizient organisieren zu können. Ein wertvoller Hinweis für die Praxis war die Nutzung von Makros, um in Word Texte zu nummerieren und besser für die Analyse vorzubereiten.

Im Laufe der Diskussion wurde deutlich, dass es an ganzheitlichen End-to-end-Lösungen für Transkriptions- und Analyseprozesse fehlt. Viele kleine Tools und Lösungen sind zwar auf Plattformen wie GitHub verfügbar, jedoch fehlt es ihnen oft an vollständiger Integration oder deren Anwendung bedürfte vertiefte IT-Kenntnisse. Das verzögert den Transkriptions- und Analyseprozess ungemein, insbesondere, wenn nach einem umständlichen Daten-Übertragungs- und Formatierungsprozess im finalen Textverarbeitungsprogramm ein Änderungsbedarf (bspw. Hinzufügen einer Annotation) entdeckt wird, der manuell in allen Tools korrigiert werden muss, um eine einheitliche Arbeitsgrundlage schaffen zu können.

Es bleibt zu resümieren, dass insbesondere für die spezifischen Anforderungen und Zwecke der Gesprächsforschung bisher keine Software entwickelt wurde, die unter Usability-Aspekten für den gesamten Transkriptionsprozess und die Datenaufbereitung vollständig überzeugen kann. Wissenschaftler\*innen in der Gesprächsforschung bleiben oftmals gezwungen, ihre Daten umständlich manuell zu bearbeiten und Workarounds zu finden oder sich ein vertieftes IT-Wissen anzueignen (bzw. das jener IT-Expert\*innen ihres jeweiligen Netzwerkes zu nutzen), um Bearbeitungsschritte (teil-)automatisiert ausführen zu lassen. Die Teilnehmenden verbleiben in der stillen Hoffnung, dass zukünftig bestehende Softwareanbieter ihre Transkriptions- und Analysetools stärker an die Bedürfnisse der Gesprächsforschung anpassen oder vorhandene Varianten anwendungsfreundlicher gestaltet werden. Auch in den anderen beiden Breakout-Räume mit privaten Themen fand ein sehr guter und produktiver Austausch statt.

## **Fazit und Ausblick**

Im Abschluss der Veranstaltung wurde gemeinsam dessen Erfolg rekapituliert und der Wunsch nach weiteren Netzwerktreffen als sehr förderliche Austauschmöglichkeit für Nachwuchsforscher\*innen geteilt. Einheitlicher Konsens war, dass das Medium der Online-Tagung über Zoom zwar sehr zugänglich und praktikabel ist, eine Präsenz-Veranstaltung unter dem Aspekt der informellen Vernetzung dennoch wünschenswert bleibt. Dementsprechend findet das 11. Treffen des Netzwerkes für Doktorand\*innen der Gesprächsforschung (DokGF) am 23. und 24. Mai 2025 in Osnabrück statt. Zur Organisation haben sich freundlicherweise Henning Vahlenkamp (Bremen) und Marius Müller (Osnabrück) bereiterklärt.

## **Teilnehmer\*innen des 10. DokGF-Netzwerktreffens:**

*Tianyi Bai* (Universität Mannheim; Prof. Dr. Florence Oloff): Kooperation und Koordination in herausfordernden Situationen: Multiaktivität und Progressivität von Alltagshandlungen in der Tandem-Interaktion (Arbeitstitel).

*Wifek Bouaziz* (Universität Freiburg; Prof. Dr. Uta Reinöhl): Conversational Practices in Kera'a (Arbeitstitel).

*Helena Konstanze Budde* (Universität Mannheim/Leibniz-Institut für Deutsche Sprache Mannheim; Prof. Dr. Florence Oloff): Multimodale Praktiken des Verstehens älterer Erwachsener im Umgang mit neuen Technologien (Arbeitstitel).

*Rebekka Gallusser* (geb. Heil) (Universität Duisburg-Essen, Prof. Dr. Heike Roll): Pflegekräfte aus Mexiko - eine diskursanalytische Studie zu Pflege-Patient\*in-Gesprächen in mehrsprachiger Sprachenkonstellation.

*Regina Geisler-Knünz* (Universität Wien, Prof. Helmut Gruber): Entscheidungsfindungsprozesse in Ärzt:innen-Patient:innen-Interaktion (Arbeitstitel).

*Myriam Goll* (RPTU Kaiserslautern-Landau, Campus Landau; Prof. Dr. Jan Georg Schneider): Sprecherwechselorganisation in (Polit-)Talkshows – eine medienlinguistische und multimodal interaktionale Analyse.

*Undine Kuhlmann* (Friedrich-Schiller-Universität Jena): Im Projekt von Prof. Dr. Pia Bergmann "Prosodie von initialen Diskursmarkern".

*Constanze Lechler* (Universität Potsdam; Prof. Dr. Martin Pfeiffer): Kindliche Fremdwiederholungen in der Erwachsenen-Kleinkind-Interaktion (Arbeitstitel).

*Vivien Lohmer* (Universität Bielefeld; Prof. Dr. Friederike Kern): Diskursive Strukturen und Multimodalität in Erklärungen (technischer) Artefakte (Arbeitstitel).

*Liliana Lovallo* (Universität Heidelberg; Prof. Dr. Anja Stukenbrock): Verbale und visuelle Praktiken der Raumkonsitution in Virtual-Reality-Videospielen.

*Marius Müller* (Universität Osnabrück; Prof. Dr. Oliver Ehmer): Themensuche im Bereich Morphologie im Spanischen aus interaktional-linguistischer Sicht.

*Veronika Sahlbach* (Technische Universität Dresden; Prof. Dr. Alexander Lasch): Wirklichkeitskonstruktionen im Gespräch. Transkription und Analyse von deutschen Podcasts (Arbeitstitel).

*Hannah Sawall* (Europa-Universität Viadrina, Frankfurt (Oder); Dr. Miriam Lind): Promotion im Projekt "Posthumanistische Linguistik. Kommunikative Praktiken zwischen Menschen, Tieren und Maschinen".

*Laura Schleicher* (Universität Freiburg; Prof. Dr. Stefan Pfänder): Multimodal Grounding. The interactional management of common ground with co-speech gestures (Arbeitstitel).

*Miriam Schöps* (Martin-Luther-Universität Halle-Wittenberg; Prof. Dr. Maxi Kupetz, Prof. Dr. Anne-Kathrin Lindau): Die sprachlich-interaktionale Dimension fachlichen Lernens im systemorientierten Geographieunterricht.

*Julia Schrittenlacher* (Universität Duisburg-Essen, Prof. Miriam Morek): Kollaboratives Schreiben im inklusiven Deutschunterricht.

*Sina Spiekemeier Gimenes* (Universität Bremen; Prof. Dr. Andrea Daase): Zweitsprachaneignung in der betrieblichen Berufsausbildung.

*Tabea Stolte* (Friedrich-Schiller-Universität Jena; Prof. Dr. Pia Bergmann, Dr. Kathrin Weber): Multimodale Ressourcen in der Aushandlung von Dissensäußerungen im Sprachkontakt (Arbeitstitel).

*Simon Arvo Titze* (Universität Heidelberg; Prof. Dr. Elwys De Stefani): Prosodie von Bewertungen in französischen Varietäten (Arbeitstitel).

*Celina Tschiedel* (Universität Bielefeld; Prof. Dr. Friederike Kern): Partizipation autistischer Schüler\*innen in inklusiven Unterrichtsettings (Arbeitstitel).

*Henning Vahlenkamp* (Universität Bremen; Prof. Dr. C. Harsch): Autonomie und Kooperation in der Lernberatung zum selbständigen Sprachenlernen (Arbeitstitel).

*Regina Warmers* (Universität Hildesheim/Hochschule Magdeburg Stendal; PD Dr. Sylvia Jaki, Prof. Dr. Christiane Zehrer): Haltung in der Gesprächsführung von Mediationsverfahren (Arbeitstitel).

*Anne Weiland* (Martin-Luther-Universität Halle-Wittenberg; Graduiertenkolleg INTERFACH, Prof. Dr. Georg Breidenstein): Sprachunterstützende Praktiken in fachlichen Peer-Gesprächen im jahrgangsgemischtem Unterricht

## Literaturverzeichnis

- Birkner, Karin / Auer, Peter / Bauer, Angelika / Kotthoff, Helga (2020): Einführung in die Konversationsanalyse. Berlin; Boston: De Gruyter.
- Dröge, Kai (2025): noScribe. AI-powered Audio Transcription (Version 0.5). In: <https://github.com/kaixxx/noScribe> (letzter Zugriff 31.03.2025).
- Ehlich, Konrad / Rehbein, Jochen (1976): Halbinterpretative Arbeitstranskriptionen (HIAT). In: *Linguistische Berichte* 45, 21-41.
- Haberl, Armin / Fleiß, Jürgen / Kowald, Dominik / Thalmann, Stefan (2024): Take the aTrain. Introducing an interface for the Accessible Transcription of Interviews. In: *Journal of Behavioral and Experimental Finance* 41 (100891). <https://doi.org/10.1016/j.jbef.2024.100891>.
- Kendon, Adam (2004): *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511807572>.
- Kroes, Peter (1998): Technological Explanations: The Relation between Structure and Function of Technological Objects. In: *Society for Philosophy and Technology Quarterly Electronic Journal* 3 (3), 124-134. <https://doi.org/10.5840/techne19983325>.
- Kroes, Peter (2009): Engineering and the dual nature of technical artefacts. In: *Cambridge Journal of Economics* 34 (1), 51-62. <https://doi.org/10.1093/cje/bep019>.
- Kupetz, Maxi (2015): *Empathie im Gespräch: eine interaktionslinguistische Perspektive*. Tübingen: Stauffenburg Verlag.
- Malik, Mishaim / Malik, Muhammad Kamran / Mehmood, Khawar / Makhdoom, Imran (2021): Automatic speech recognition: a survey. In: *Multimedia Tools and Applications* 80 (6), 9411-9457. <https://doi.org/10.1007/s11042-020-10073-7>
- McNeill, David (1992): Hand and Mind: What Gestures Reveal about Thought. In: *Leonardo* 27 (4), 105-133. <https://doi.org/10.2307/1576015>.
- McNeill, David (2005): *Gesture and Thought*. University of Chicago Press.
- Morek, Miriam (2012): *Kinder erklären: Interaktionen in Familie und Unterricht im Vergleich*. Tübingen: Stauffenburg.
- Nolda, Andreas (2024): Die Dulko-Tools des Exmaralda-Partitur-Editors. Von einer externen Toolsammlung zum integrierten Bestandteil. In: 4 (2), 224-235. <https://doi.org/10.48694/KORDAF.4139>.
- Poerner, Nina / Schiel, Florian (2016): An automatic chunk segmentation tool for long transcribed speech recordings. *Proceedings of the Phonetics & Phonology Conference (P&P)*, München.
- Quasthoff, Uta / Heller, Vivien / Morek, Miriam (2017): On the sequential organization and genre-orientation of discourse units in interaction: An analytic framework. In: *Discourse Studies* 19 (1), 84-110. <https://doi.org/10.1177/1461445616683596>.
- Radford, Alec / Kim, Jong Wook / Xu, Tao / Brockman, Greg / McLeavey, Christine / Sutskever, Ilya (2022): *Robust Speech Recognition via Large-Scale Weak Supervision*. <https://dx.doi.org/10.48550/arXiv.2212.04356>.

- Schegloff, Emanuel A. (2000): Overlapping talk and the organization of turn-taking for conversation. In: *Language in Society* 29 (1), 1-63.  
<https://doi.org/10.1017/S0047404500001019>.
- Schulte, Carsten/Budde, Lea (2018): A Framework for Computing Education: Hybrid Interaction System: The need for a bigger picture in computing education. In: *Koli Calling '18: Proceedings of the 18th Koli Calling International Conference on Computing Education Research*, 1-10.  
<https://doi.org/10.1145/3279720.3279733>.
- Selting, Margret et al. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung* 10, 252-402.
- Stockman, Robert H. (2021): *The World of the Bahá'í Faith*. London: Routledge.
- Terfloth, Lutz / Schaffer, Michael / Buhl, Heike M. / Schulte, Carsten (2023): Adding Why to What? Analyses of an Everyday Explanation. In: Longo, Luca (Hg.): *Explainable Artificial Intelligence*. Cham: Springer Nature Switzerland. 256-279. [https://doi.org/10.1007/978-3-031-44070-0\\_13](https://doi.org/10.1007/978-3-031-44070-0_13).

Veronika Sahlbach  
Technische Universität Dresden  
Wiener Straße 48  
01219 Dresden

[Veronika.sahlbach@tu-dresden.de](mailto:Veronika.sahlbach@tu-dresden.de)

Laura Schleicher  
Romanisches Seminar der Universität Freiburg  
Platz der Universität 3  
79085 Freiburg

[laura.schleicher@romanistik.uni-freiburg.de](mailto:laura.schleicher@romanistik.uni-freiburg.de)

Regina Warmers  
Hochschule Magdeburg-Stendal  
Breitscheidstraße 2  
39114 Magdeburg

[regina.warmers@h2.de](mailto:regina.warmers@h2.de)

Veröffentlicht am 5.5.2025