

4 Methodologische Einordnung

Eine praxeologisch ausgerichtete Arbeit scheint zunächst auf das Problem zu stoßen, dass eine an Praktiken ausgerichtete Methodologie sich noch nicht fest etabliert hat und die Diskussion dazu erst in den letzten Jahren beginnt (vgl. Schäfer/Daniel 2015: 40). Wie eine besondere Herausforderung kann zudem die Verbindung einer praxistheoretischen Grundierung mit den sprachwissenschaftlichen Erkenntnisinteressen dieser Arbeit wirken, die methodologische Implikationen einer ganz anderen Tradition mit sich bringen. Doch bei genauerem Hinsehen zeigt sich gerade die Verbindung von Praxeologie als theoretischer Ausgangslage mit einem sprachwissenschaftlich-empirischen Zugang als besonders fruchtbar, wenn für die methodologische Situierung als Bindeglied die Ethnomethodologie eingesetzt wird (vgl. auch Hector 2022). Diese ist nicht nur eine wichtige Ressource praxeologischer Theoriebildung (vgl. Hillebrandt 2014: 43), sondern hat schon früh Thesen aufgestellt, die bei der Entwicklung einer Methodologie zur Analyse sprachlicher als Bestandteil sozialer Praktiken hilfreich sein werden. Außerdem, und das ist am wichtigsten, liefert sie „bewährte Analyseverfahren“ (Schäfer/Daniel 2015: 42), die bereits in den 60er- und 70er-Jahren in der Sprachwissenschaft aufgegriffen wurden und sich etablieren konnten, was nahelegte, deren Mentalität zu folgen, u. a. mit der Sequenzanalyse als zentralem Verfahren der Datenauswertung.

Das Verhältnis von Praxeologie und Ethnomethodologie sowie der Weg Letzterer in die Linguistik hinein soll im Folgenden kurz vorgestellt werden, um den Zusammenhang zwischen einer soziologischen Forschungsrichtung und sprachlichen Analysen deutlich zu machen und die Entscheidung, diesen methodologischen Prinzipien zu folgen und sie in den Dienst einer praxeologischen Untersuchung zu stellen, zu legitimieren (Kap. 4.1). Es werden anschließend auch die darauf aufbauenden methodologischen Ausgangspunkte vorgestellt. Damit ist ein Bündel von Strömungen gemeint, das Imo (2013) als „Sprache-in-Interaktion“ zusammenfasst – die Konversations- bzw. Gesprächsanalyse, die Interaktionale Linguistik sowie die Gesprochene-Sprache-Forschung (Kap. 4.2). Darauf aufbauend wird Ethnografie als Forschungsstil in der Gesprächsanalyse diskutiert, der die vorliegende Arbeit maßgeblich geprägt hat und Perspektiven über die Gesprächsanalyse hinaus bietet (Kap. 4.3). Abschließend wird die Rolle von Multimodalität in der sprachlichen Praxis diskutiert (Kap. 4.4), um so die methodologische Positionsbestimmung zu komplettieren. Im Zwischenfazit (Kap. 4.5) werden die einzelnen Teile noch einmal mit Blick auf die vorliegende Arbeit reflektiert.

4.1 Ethnomethodologie als praxeologische Methodologie

Zu betonen ist zunächst noch einmal die immer wieder herausgestellte Nähe der Arbeiten des Begründers der Ethnomethodologie, Harold Garfinkel, zu den späteren sprachphilosophischen Arbeiten Ludwig Wittgensteins (vgl. Bergmann 2007b: 54; 2011: 229; Auer 2013: 132), die beide dadurch charakterisiert sind, dass sie den alltäglichen *Gebrauch* sprachlicher oder nicht-sprachlicher Ressourcen zur Durchführung sozialer Handlungen in den Mittelpunkt stellen und die Betrachtung auf neuropsychologischer bzw. kognitiver Ebene ablehnen. Die sozialen und mithin sprachlichen Praktiken können nicht unabhängig vom alltäglichen *Vollzug* dieser betrachtet werden: Experimentelle Untersuchungsdesigns können darüber – diese Überzeugung teilt der späte Wittgenstein mit Garfinkel – keinen Aufschluss geben. Harold Garfinkel verweist also, ganz ähnlich wie Wittgenstein, auf den *Gebrauch* der Zeichen und misst der Sprache eine entsprechend hohe Bedeutung bei. Die Ähnlichkeit der praxeologischen und ethnomethodologischen Wurzeln, die sich in den Arbeiten von Wittgenstein und Garfinkel zeigen, ist bemerkenswert, denn „obwohl sie wissenschaftsgeschichtlich nichts miteinander zu tun haben“ (Auer 2013: 132), liefern sie beide die Grundlage für bis heute produktive Paradigmen der Sozialtheorie, die immer wieder aktualisiert und neu zugerichtet wurden. Auch Meyer (2015: 91) stellt heraus, dass die Ethnomethodologie als „praxistheoretischer Ansatz“ gesehen werden muss: „Praktiken werden in der Ethnomethodologie sowohl sozialen Strukturen als auch sozialen Handlungen bzw. Akteuren (und ihren Intentionalitäten und Identitäten) vorgeordnet“ – damit folgt sie ganz dem Ziel, „Praxis allen anderen [...] Erklärungsgrößen vorzuordnen“, das Schüttpelz/Meyer (2017: 158) für eine praxistheoretische „Konstitutionsanalyse“ in ihrem bereits erwähnten Glossar zur Praxistheorie formulieren. Die Ethnomethodologie eignet sich damit bestens als Verbindungsstück zwischen den dargelegten praxissoziologischen Grundlagen und der sprachwissenschaftlichen Beschreibungsebene, auf der sprachliche *als* soziale Praktiken herausgearbeitet werden sollen. Die Verschränkung von Praxeologie und Ethnomethodologie wird dabei auch in den folgenden Ausführungen immer wieder durchscheinen.

Die Ideen Harold Garfinkels und die v. a. durch sein 1967 erschienenes Hauptwerk *Studies in Ethnomethodology* begründete Forschungsrichtung grenzte sich anfänglich v. a. gegen den Strukturfunctionalismus Talcott Parsons ab, bei dem Garfinkel selbst studiert hatte. In der Linie der phänomenologischen Soziologie nach Alfred Schütz war er jedoch mit diesem der Auffassung, dass in den Arbeiten Parsons zur sozialen Ordnung die subjektiven Perspektiven der einzelnen (und immer einen konkreten situativen Zusammenhang eingebetteten) Individuen zu kurz kommen und durch ein Verständnis von Wissenschaftsrationalität

überschrieben werden.¹²¹ Garfinkel führt stattdessen eine alltägliche Rationalität ein, die von der wissenschaftlichen Rationalität strikt zu trennen ist. In der Untersuchung der alltäglichen Rationalität werden die konkreten sozialen und alltäglichen Situationen ernst genommen, in die gesellschaftliche Ordnungsprinzipien hinein diffundieren, in denen sie aber auch erst wirkmächtig und real werden und eine Weiterentwicklung und ein Zurückdiffundieren in die Ordnungsprinzipien erfahren (vgl. Bergmann 2007b: 53). Nur diese an der Oberfläche beobachtbaren sozialen Handlungen können herangezogen werden, um die Praktiken zu beschreiben, mit denen soziale Ordnung immer wieder situativ neu hergestellt und ausgehandelt wird.

Damit ist die Grundidee der Ethnomethodologie nach Garfinkel (1967) bereits angerissen: Es sollen die Praktiken und Vorgehensweisen, d. h. die *Methoden*, beschrieben werden, die die Mitglieder einer Gesellschaft (*Ethnos*) nutzen, um ihr eigenes Handeln erkenn- und verstehbar zu machen und dabei gleichsam die Wirklichkeit der Situation, in der sie sich befinden, zu strukturieren (vgl. Bergmann 2007b: 51). Diese Praktiken und Vorgehensweisen können sich übersituativ stabilisieren und somit gesellschaftliche „Rollen“ bilden, die wiederum ebenfalls zu stabileren Gefügen konglomerieren können und somit „Institutionen“ ausbilden (vgl. Patzelt 1987: 10). Damit folgt auch die Ethnomethodologie einer flachen Ontologie im Sinne Schatzkis (2016), weil auch hier zwischen den ‚kleineren‘ und den ‚größeren‘ Einheiten ein konstitutiver Zusammenhang besteht, nicht aber eine Kausalitäts- oder Abhängigkeitsbeziehung zwischen beiden. Die Unterscheidung zwischen Mikro- und Makroebenen bleibt insofern aus Sicht der überwiegenden Zahl der Ethnomethodolog*innen irrelevant (vgl. Bergmann 2011: 231). Die Praktiken bzw. *Ethno-Methoden* sind die entscheidenden Einheiten, die ‚größere‘ Einheiten konstituieren, während die größeren reflexiv auf die Ethno-Methoden zurückwirken. Die Wirklichkeit und ihre Geordnetheit wird ethnomethodologisch insofern als „fortwährendes Resultat situativer, aber methodischer Alltagspraktiken“ (Bergmann 2011: 229) untersucht. Mit dem Konzept der „accountability“ (Garfinkel 1967: 34) wird dieser zentrale ethnomethodologische Gedanke auf den Punkt gebracht: die These, dass diese methodisch eingesetzten Praktiken, die einer bestimmten sozialen Wirklichkeit folgen und in einer spezifischen sozialen Situation angemessen sind, gleichsam diejenigen sind, die diese Situation erst als solche erkenn- und darstellbar machen (vgl. Bergmann 2007a: 125–126). Die Ethnomethodologie geht also davon aus, dass „die Tätigkeiten, mit denen Gesell-

¹²¹ Das Verhältnis von Harold Garfinkel zu seinem Lehrer Talcott Parsons, von dem er sich zugleich mit Vehemenz abgrenzt, ist vielschichtig und kann hier nicht zufriedenstellend dargestellt werden. Zur Beziehung der beiden siehe Heritage (1984b); zur wissenschaftlichen Auseinandersetzung zwischen Alfred Schütz und Parsons siehe Schütz/Parsons (1977).

schaftsmitglieder die Szenerien organisierter Alltagsangelegenheiten schaffen und bewältigen, identisch mit den Vorgehensweisen sind, mit denen Mitglieder diese Szenerien zurechenbar (*account-able*) machen“ (Garfinkel 2020[1967]: 35),¹²² und dass dies als eine „endlose fortlaufende, kontingente Leistung“ (Garfinkel 2020[1967]: 35) anzusehen ist. So ist das Verhältnis von Subjekt bzw. subjektivem Handlungsplan, Umständen und sozialen Regeln als prozessual zu begreifende Eigenrationalität des Alltäglichen charakterisiert (vgl. Bergmann 2007a: 127).

Da nach Auffassung der Ethnomethodologie in allen gesellschaftlichen Lebensbereichen Ethno-Methoden zum Einsatz kommen und diese durch ihre inhaltliche „Ladung“ (siehe Patzelt 1987: 10) spezifisch für den Wirklichkeitsbereich sind, in dem sie sich vollziehen, ist kein Phänomengebiet wichtiger als ein anderes. Überall müssen die situativen Gegebenheiten, der soziale Kontext und die Hintergründe der handelnden Personen ernst genommen werden – Garfinkel/Sacks (1979: 138) bezeichnen dies als „ethnomethodologische Indifferenz“. Wie Patzelt (1987: 11–13) ausführt, ist es aufgrund der gesellschaftlichen Pluralität, auf die Garfinkel und Sacks mit diesem Konzept eine Antwort finden, unmöglich, Zusammenhänge allgemein und unspezifisch zu erfassen. Stattdessen hat sich ausgehend von der „Allgemeinen Ethnomethodologie“ eine Reihe ethnomethodologischer Subdisziplinen gebildet, die in der Soziologie sehr produktiv waren und sind.¹²³ Diese arbeiten alle strikt empirisch mit ‚ungestellten‘ bzw. ‚natürlichen‘ Daten,¹²⁴ häufig mit Video- oder Audioaufnahmen als Fixierungen des Prozesses, und nehmen in jedem Fall den realen zeitlichen Verlauf in den Blick. Ethnografische Methoden wie die teilnehmende Beobachtung sind ebenso verbreitet, auch wenn die technische Fixierung der Abläufe als Ton- und Videoaufnahmen und die damit verbundene Nachvollziehbarkeit der realen zeitlichen Abläufe vorzuziehen ist (vgl. Bergmann 2001: 922). Das Datenmaterial wird induktiv bearbeitet und die Analysekategorien ohne Vorannahmen erst aus dem Material heraus entwickelt.

Trotz dieser methodischen Gemeinsamkeiten ethnomethodologischer Untersuchungen ist, auch als Reaktion auf eine Methodenkritik, der Formulierung eines klaren Regelwerks von Garfinkel stets eine Absage erteilt worden, mit dem Verweis auf ein „unique adequacy requirement“ (Garfinkel 2002: 175–176): Die Methoden müssen immer den gegebenen Umständen angepasst werden. Damit

¹²² Hier wird auf die von Schüttpelz/Rawls/Thielmann (2020) herausgegebene Übersetzung ins Deutsche verwiesen.

¹²³ Für Beispiele aus sehr unterschiedlichen Bereichen wie ethnomethodologische Studien in der Kriminologie, einer „Ethnomethodologie des Lesens“ oder einer ethnomethodologischen Wissenschaftsanalyse siehe Patzelt (1987: 12).

¹²⁴ Zum Natürlichkeitspostulat siehe Kap. 5.3.

nahm Garfinkel auch Stellung zur dahingehend geäußerten Kritik, dass die methodische Ausleuchtung von Ethno-Methoden ja selbst eine Methode ist, die es ja gerade zu hinterfragen galt, aber in sich schon eingeschriebene Vorannahmen enthalte (siehe dazu Bergmann 2007b: 57). Eine methodologische Offenheit wird auch in der aufkeimenden Methodologie-Diskussion in der Praxissoziologie eingefordert, die insofern auch eine Rückbesinnung auf ethnomethodologische Prinzipien darzustellen scheint: Eine unflexible Methodentreue wird ebenso abgelehnt, demgegenüber eine Entwicklung von Methoden entlang empirischer Untersuchungen vorgeschlagen (vgl. Schäfer/Daniel 2015: 52), was genau dem Vorschlag eines „unique adequacy requirement“ entspricht, für das ein „möglichst breites Spektrum an methodischen Herangehensweisen“ (Schäfer/Daniel 2015: 52) zur Verfügung stehen sollte. Der theorie- und methodenbildende Anspruch wurde auch in der Ethnomethodologie äußerst gering gehalten – besonders mit dem Anspruch, keine Vorprägungen durch vorgefertigte Konzepte zuzulassen. Dieses Prinzip zieht sich durch die Praxeologie ebenso wie durch die ethnomethodologischen Subdisziplinen, wenngleich es dort gelegentlich bei genauerer Betrachtung auch zugunsten einer stärkeren Methodenlehre verletzt wurde.

4.2 Sprache-in-Interaktion

Eine der Strömungen, die sich als ethnomethodologische Subdisziplin entfaltete, ist die ethnomethodologische Konversationsanalyse.¹²⁵ Von dieser inspiriert und von einer sprachwissenschaftlichen Weiterentwicklung der ursprünglich soziologischen Erkenntnisinteressen geprägt, entwickelte sich zudem eine Reihe weiterer Forschungsansätze, die Imo (2013: 59) zusammen mit der Konversationsanalyse unter dem Oberbegriff „Sprache-in-Interaktion“ zusammenfasst. Diese Traditionen greifen auch Forschungslinien und Sprachbeschreibungstheorien auf, die sich unabhängig davon bereits in der Sprachwissenschaft entwickelt hatten und teilweise eine Wiederbelebung durch neue Verbindungen erfuhren. Die Ansätze lassen sich nicht trennscharf voneinander abgrenzen – die nachfolgende Gliederung orientiert sich an der terminologischen Handhabung bei Imo (2013: 59), auch wenn klar ist, dass zwischen den Teilbereichen nicht nur wissenschaftshistorische Entwicklungen verbindend sind, sondern auch aktuell Übergänge im Hinblick auf die Bearbeitung gemeinsamer Erkenntnisinteressen bestehen. Für die vorliegende Arbeit sind Prin-

¹²⁵ Ihre ethnomethodologischen Wurzeln sind, je nach Ausprägung, mehr oder weniger präsent. Der Zusatz „ethnomethodologisch“ ist insofern nicht in jedem Fall gerechtfertigt und auch von den Autor*innen selbst nicht verwendet, weshalb er auch im Folgenden getilgt wird.

zipien der ethnomethodologischen Konversationsanalyse und der damit eng verbundenen Gesprächsanalyse sowie das Forschungsprogramm der Interaktionalen Linguistik entscheidend. Zurückgegriffen wird auch auf Teile der älteren Gesprochene-Sprache-Forschung. Diese Strömungen und Programme sollen als Teil der methodologischen Einordnung kurz vorgestellt werden. Es werden deren Entwicklungslinien nachgezeichnet und ihre Relevanz für die vorliegende Arbeit begründet. Weitere, mehr oder minder scharf abgegrenzte Strömungen und Schattierungen im Umfeld von Sprache-in-Interaktion, wie z. B. die Dialoganalyse (Fritz/Hundsnurscher 1994b),¹²⁶ die Gesprächslinguistik (Brinker/Sager 2010; Mroczynski 2014) oder die Konstruktionsgrammatik (Ziem/Lasch 2013), teilen zwar die analytische Arbeit an Audio- bzw. Videoaufnahmen von gesprochensprachlichen Daten, einen empirischen Anspruch und eine Engführung auf sprachwissenschaftliche Interessen. Sie unterscheiden sich aber auch deutlich, zum einen in der Ausprägung ihres Anwendungsbezugs und mit Blick auf die sprachanalytische Ebene der Untersuchungen (vgl. Imo 2013: 71), zum anderen grenzen sie sich von der Konversations- bzw. Gesprächsanalyse und der Interaktionalen Linguistik insbesondere dadurch ab, dass sie nicht die sprachliche Praxis bzw. den sprachlichen Zuschnitt auf die Interaktion in den Mittelpunkt rücken, sondern auf andere Aspekte fokussieren. Sie sind insofern für die Bearbeitung der Fragestellungen nicht geeignet, auch wenn sie vereinzelt Impulse liefern können; auf sie wird daher in der nachfolgenden Darstellung nur peripher eingegangen.

4.2.1 Konversationsanalyse

Wie Patzelt (1987: 11) anmerkt, handelt es sich bei der Konversationsanalyse zwar einerseits um eine der „speziellen Ethnomethodologie[n]“, indem sie Sprachverwendung ins Zentrum rückt und die Rolle sprachlicher Zeichen bei der Herstellung sozialer Wirklichkeit untersucht, damit aber andererseits nahezu alle Lebensbereiche tangiert und deshalb zwar eine „spezielle“ Ethnomethodologie ist, gleichzeitig aber nicht in einem thematisch gebundenen Feld Halt macht (vgl. Patzelt 1987: 12). Die ethnomethodologische Konversationsanalyse ist auch insofern eine etwas anders gelagerte Teildisziplin, als sie nach ihrer Gründung v. a. von Harvey Sacks, Gail Jefferson und Emanuel Schegloff als eigenständiges Teilgebiet

¹²⁶ Es ist bezeichnend für die Unschärfe der verschiedenen Ansätze, dass das erste Kapitel dieses Sammelbandes eine Einführung in die ethnomethodologische Konversationsanalyse von Jörg Bergmann ist. In den weiteren Kapiteln werden, u. a. von Johannes Schwitalla, Heinrich Löffler und Jacques Moeschier, weitere Ausdifferenzierungen dieser Ansätze eingeführt, denen hier nicht im Detail nachgegangen werden soll.

äußerst produktiv war und bis heute ist und sich zunehmend als eigene Forschungsrichtung verstand. Sie grenzte sich mit ihrem Fokus auf die inhärente Struktur von Gesprächen zunehmend von der Ethnomethodologie ab. Von der Konversationsanalyse haben sich weitere Forschungszweige wie die Interaktionale Linguistik (Selting/Couper-Kuhlen 2001a) inspirieren lassen. Sie wurde zunehmend methodisch wie theoretisch autonom von der Ethnomethodologie Garfinkels, was auch zu methodologischen Auseinandersetzungen und Spannungen zwischen Ethnomethodologie und Konversationsanalyse führte (vgl. Bergmann 2001: 921; 2007a: 134). Die Konversationsanalyse brach so auch mit einigen ethnomethodologischen Grundsätzen – so sollten in der konversationsanalytischen Datenauswertung gerade keine Hintergrundinformationen einbezogen werden, sondern nur das für die Analyse relevant gemacht werden, was auch die Sprechenden selbst in der Konversation relevant setzen (Bergmann 2001: 922). Während die Ethnomethodologie im Sinne Garfinkels die Einmaligkeit und Situiertheit der Ereignisse betont und Zuschreibungen übergeordneter Kategorien ablehnte, versuchte die Konversationsanalyse abstraktere Mechanismen der Interaktionsordnung herauszuarbeiten (vgl. Pollner 1991: 373). Für diesen Versuch, allgemeinere Verfahren zu finden, wurde sie seitens der Ethnomethodologie kritisiert; im Gegenzug bemängelten Konversationsanalytiker*innen die im engeren Sinne ethnomethodologischen Ansätze als zu partikularistisch und argumentierten mit der geringen Aussagekraft der Ergebnisse, die eine Überbetonung der Einzigartigkeit von Situationen impliziere.

Die Entwicklung der ethnomethodologischen Konversationsanalyse war einer der wesentlichen Faktoren und leistete Pionierarbeit für die weitere Etablierung sequenzanalytischer Verfahren in Soziologie und Sprachwissenschaft (vgl. Ploder/McElvenny 2022: 292). Mit ihrer Abstraktion von der konkreten Situation und mit ihrer Engführung auf das *Sprachliche* zur Herstellung von Sinn und Ordnung in der sozialen Interaktion war die Konversationsanalyse durchaus in der Lage, wesentliche Prinzipien sprachlicher Interaktion zu entdecken und auf unterschiedliche Kontexte anzuwenden. So wurden etwa Verfahren des Turn-Takings, d. h. des Sprecherwechsels, systematisiert (vgl. Sacks/Schegloff/Jefferson 1974), selbst- und fremdinitiierte Reparaturleistungen klassifiziert (vgl. Schegloff/Jefferson/Sacks 1977; Schegloff 1997a) sowie insbesondere von Schegloff (1968; 2007) die Bedeutung des sequenziellen Ablaufs von Gesprächen und Adjazenzpaaren (z. B. Frage – Antwort) herausgestellt (siehe auch Imo 2013: 74–75).¹²⁷ Letztere war als Strukturmerkmal nicht nur für kleinere Einheiten wie Begrüßungen und Verab-

¹²⁷ Auch Birkner (2020a: 16–17) nennt diese drei Themen als Gegenstandsbereiche der frühen Konversationsanalyse.

schiedungen (vgl. Schegloff/Sacks 1973) zentral, sondern auch für größere Einheiten wie Erzählungen (vgl. Quasthoff 1980a; 2001), Beschwerden (vgl. Günthner 1999a) oder Witze (vgl. Kotthoff 2006a; Norrick 2001). Diese Arbeiten waren möglich aufgrund der von Sacks (1984: 21) zusammengeführten Erkenntnisse, dass soziale Aktivitäten formell beschrieben werden können, dass sie außerdem methodisch und rekurrent auftreten und dass sie insofern generalisierbare Eigenschaften aufweisen, die es herauszuarbeiten gelte. Dazu sei es nötig, die Daten anzuschauen, ohne Vorannahmen oder konkrete Fragestellungen an sie heranzutragen, idealerweise nicht einmal über den Kontext der Daten genau Bescheid zu wissen, um dann aus den Daten heraus die Aspekte herauszufiltern, die für die Teilnehmer*innen selbst relevant sind und die sie sich gegenseitig aufzeigen und sie so gleichzeitig den Analytiker*innen zugänglich machen (vgl. Sacks/Schegloff/Jefferson 1974: 728). Dieses Konzept ist als *display*-These ein Eckpfeiler der konversationsanalytischen Forschungsarbeiten und bezieht sich nicht nur auf die reflexive Herstellung von Kontexten, sondern auch für das gegenseitige Aufzeigen von Sinn und Ordnung sowie das Verstehen dieser Ordnung. Das *display*-Konzept ist unmittelbar mit dem Konzept des *doing* verwoben, mit dem die Regelmäßigkeit und Konventionalisierung bestimmter kontextualisierender Verfahren betont wird, die ‚auf dem *display*‘ angezeigt werden (vgl. Birkner et al. 2020: 18). Dies kann von kleineren, aber wiederkehrenden sozialen Ereignissen bis hin zu gesellschaftlichen Rollenbildern – etwa *doing gender* (vgl. West/Zimmermann 1987) – reichen (siehe dazu auch Schegloff 1997a).¹²⁸ Ein anderer wesentlicher Eckpfeiler ist das Konzept des Adressatenzuschnitts, auch Recipient-Design, ebenfalls eingeführt in dem bereits erwähnten Aufsatz von Sacks/Schegloff/Jefferson (1974: 727). Dieses Konzept umfasst eine ganze Reihe von gesprächsorganisatorischen Erscheinungen, die dazu beitragen, einen Turn, d. h. einen Redezug, immer auf die Gesprächspartner*innen zuzuschneiden – bzw. auf die Bilder, die sich die Sprecher*innen von den Empfänger*innen machen.¹²⁹ Diese Anforderung an die Sprecher*innen verlangt von diesen, den sozialen Kontext, in dem sich diese befinden, zu erkennen und gleichzeitig, fortlaufend mit zu produzieren und damit auch die Rolle, die das Gegenüber in diesem Kontext hat. Dieses auch als „Adressatenzu-

¹²⁸ Zu Kritik und Umgang mit der *display*-These siehe Kap. 4.3.

¹²⁹ Damit ist, wie Hitzler (2013: 112) ausführt, ein sprecher*innenspezifisches Konzept umrissen, nicht eine „analytische Maxime“ – auch wenn es in den meisten Arbeiten so behandelt wird und damit auch gegen den bereits angerissenen ethnomethodologischen Anspruch der Freiheit von abstrakten methodischen Prinzipien verstößt.

schnitt“ bekannte Konzept ist Ausgangslage für viele konversationsanalytische Untersuchungen und wird auch in den vorliegenden Analysen eine Rolle spielen.

Die ethnomethodologische Fokussierung auf die sprachlichen Ressourcen, die bei der Herstellung sozialer Interaktion genutzt wurden, war anfänglich nicht nur dem Interesse an einem fundamentalen Aspekt des Sozialen geschuldet, sondern auch eine Form von analytischer Bequemlichkeit (vgl. Day/Wagner 2008: 44): Tonbandaufnahmen waren leicht zu beschaffen und gut zu analysieren und die Transkriptionen gaben viel Analysematerial her. Ton- und Videoaufnahmen dokumentieren das Gesprochene unvoreingenommen im Sinne einer „registrierende[n] Konservierung“ (Bergmann 1985: 305) und ermöglichen damit eine detaillierte Rekonstruktion der Aufzeigehandlungen von Gesprächsteilnehmer*innen. Von dort aus konnten konversationsanalytische Untersuchungsansätze, die sich trotz des ethnomethodologischen Anspruchs der Reduktion generalisierbarer Konzepte entwickelt hatten, in benachbarte Disziplinen diffundieren. Sie wurden so etwa in der empirischen Linguistik aufgegriffen und führten zur Entwicklung von noch stärker auf Sprache fokussierten Ansätzen, z. B. die Interaktionale Linguistik (siehe übernächster Abschnitt).

4.2.2 Gesprächsanalyse

Der Begriff *Gesprächsanalyse* wird erstens nahezu synonym für „Konversationsanalyse“ verwendet – als Übersetzung des Ausdrucks „conversation analysis“ – und ist mit diesen Konzepten eng verwandt (vgl. Imo 2013: 71–76). Auch in gesprächsanalytischen Arbeiten ist ein starker Begriff von Empirie gesetzt und bildet die Sequenzanalyse das grundlegende Verfahren zur Datenauswertung. Allerdings ist mit der Übertragung in den deutschsprachigen Raum bei Arbeiten aus dem Bereich der Gesprächsanalyse in den meisten Verwendungsweisen eine methodische Öffnung einhergegangen, etwa wenn ethnografisches Wissen in die Analysen einbezogen wird (siehe Deppermann 2000), kognitive Ansätze ergänzend verfolgt werden oder andere qualitative Forschungsmethoden komplementär hinzugezogen werden (vgl. Deppermann 2008: 10). Er steht also dafür, die bereits erwähnte „analytische Mentalität“ (Schenkein 1978: 1) der Konversationsanalyse beizubehalten, behält sich als Forschungsparadigma aber gleichzeitig Abweichungen vor und kann damit auch als eine Rückbesinnung zum „unique adequacy requirement“ (Garfinkel 2002: 175–176) gesehen werden, das in der Ethnomethodologie von Garfinkel postuliert worden war. Genau auf diese Form der Analyse stützt sich auch die vorliegende Arbeit, die in der strikten Befolgung konversationsanalytischer Maximen keinen Mehrwert für die Bearbeitung der eingangs gestellten Leitfragen sieht. Während die Maximen sich für die Erforschung konversationeller Universa-

lien etabliert hatten (im Überblick Schegloff 2006), steht in der vorliegenden Arbeit gerade nicht die Erarbeitung sprachlicher Universalien, sondern die ihre Anwendung unter spezifischen Kooperationsbedingungen im Fokus. Insofern scheint es geradezu geboten, Hintergrundwissen über die Sprecher*innen, ihre Vorherfahrungen und ihre Beziehung zueinander, über die lokalen, materiellen Gegebenheiten der Haushalte und über die Funktionsweise der VUIs wo erforderlich hinzu-zuziehen und die Analyse davon zu informieren.¹³⁰

Mit *Gesprächsanalyse* wird darüber hinaus zweitens ein weiteres, spezifisches Forschungsprogramm verstanden, das die ethnomethodologische Konversationsanalyse adaptiert und sie um Fragen der Sprechakttheorie und der Textlinguistik erweitert: „Eine umfassende Gesprächsanalyse hat [...] beide Ereignisse des sozialen Ereignisses ‚Gespräch‘ zu berücksichtigen und zu beschreiben: das Handlungsergebnis und den Handlungsvollzug“ (Brinker/Sager 2010: 20). Diese Weiterentwicklung der Konversationsanalyse war insbesondere in den 70er- und 80er-Jahren produktiv und versuchte zwischen der ethnomethodologischen Konversationsanalyse und der philologischen Perspektive zu vermitteln. Sie kann insofern auch als soziologisch interessierter Arm der Gesprochene-Sprache-Forschung betrachtet werden, der für das Forschungsprogramm der Ethnomethodologie empfänglich war und sich etwas später entwickelt hat, gleichzeitig jedoch nicht so sehr in der sozialwissenschaftlichen Tradition verwurzelt war, dass er den Anschluss etwa an die Arbeiten von Dell Hymes oder John Gumperz (s. o.) gefunden hätte. Dieses Programm ist für die vorliegende Arbeit nur eingeschränkt relevant.

4.2.3 Interaktionale Linguistik

Als sprachwissenschaftliche Wendung des konversationsanalytischen Ansatzes ist die *Interaktionale Linguistik* zu verstehen, die sich um die Jahrtausendwende etablierte und insbesondere methodische Erkenntnisse aus der Linguistik der späten 90er-Jahre bündelte (Selting/Couper-Kuhlen 2000; 2001a; 2001b). Die Interaktionale Linguistik nahm Grundgedanken der ethnomethodologischen Konversationsanalyse auf, verschob aber das Erkenntnisinteresse der Konversationsanalyse: Im Forschungsprogramm, wie es von Selting/Couper-Kuhlen (2001a) umrissen wurde,

130 Ferner wird der Terminus Gesprächsanalyse auch als Sammelbegriff für verschiedene Ansätze zur Untersuchung von Gesprächen verwendet und schließt damit die bis hierher vorgestellten Forschungslinien ein (vgl. Imo/Lanwer 2019: 18). Als Sammelbegriff für all diese Forschungslinien soll jedoch der bereits eingeführte, von Imo (2013) geprägte Begriff „Sprache-in-Interaktion“ oder im Folgenden auch der Begriff der ‚Gesprächsforschung‘ verwendet werden, um methodologische Implikationen anderer Begriffe zu vermeiden.

geht es zwar auch, aber nicht primär um die Herstellung von Sinn und Ordnung in der sozialen Interaktion und das Sprachliche als Mittel zum Zweck, sondern vielmehr um die Frage, welche konkreten sprachlichen Muster sich im Sprachgebrauch herausbilden, empirisch nachweisbar sind und von den Sprecher*innen auf die Interaktion zugeschnitten eingesetzt werden. Wie diese sprachlichen Muster durch ihren Interaktionszuschnitt entstehen und geformt werden und gleichzeitig wie die Interaktionen durch entstandene Muster geformt werden, sind die Kernfragen der Interaktionalen Linguistik (vgl. Barth-Weingarten 2008: 78), die insofern einen deutlich fokussierteren Blick auf die sprachlichen Details einnimmt, sich aber, wie Barth-Weingarten (2008: 78) ausführt, auf allen sprachwissenschaftlichen Beschreibungsebenen bewegen kann.

Innerhalb des sprachwissenschaftlichen Forschungsdiskurses setzte die Interaktionale Linguistik besonders dem „written language bias“ (Linell 2005) ein konsequent auf Empirie und Mündlichkeit eingestelltes Programm entgegen. Die Interaktionale Linguistik verfolgt somit, mit Couper-Kuhlen/Selting (2017: 3), „an interest in spoken language and a desire to see it studied in its natural habitat: in social interaction“. Durch die starke Verbindung zwischen sprachlicher Form und sozialer Interaktion, auf die in diesem Sinne alle sprachlichen Äußerungen zugeschnitten sind, bleibt auch eine starke Verbindung zwischen Konversationsanalyse und Interaktionaler Linguistik bestehen; die Interaktionale Linguistik ist „heavily indebted in conversation analysis“ (Couper-Kuhlen/Selting 2017: 3). So sehen die Begründerinnen dieses Forschungsprogramms die Interaktionale Linguistik auch als „Interface“ zwischen Sprachwissenschaft und Interaktionsanalysen, wobei sie sich disziplinar klar der Linguistik verschreiben (Selting/Couper-Kuhlen 2001a: 260). Zugleich betonen sie die Interdisziplinarität und schlagen vor, sich nicht um eine detaillierte Abgrenzung zu bemühen, sondern sich auf das Zusammenwirken der Disziplinen zu konzentrieren: „Rather than trying to define borders, the field of Interactional Linguistics should be looked upon as an interdisciplinary space [...]“ (Couper-Kuhlen/Selting 2017: 18).

Sprache wird als zugeschnitten auf ihren Verwendungszweck betrachtet, ihr ‚natürlicher Lebensraum‘ sind kontinuierlich neu entstehende kommunikative Probleme, zu deren Lösung kompetenten Sprecher*innen ein bestimmtes Repertoire an sedimentierten Verfestigungen zur Verfügung steht (vgl. Couper-Kuhlen/Selting 2017: 3). Die Lösungen wiederum werden durch die Sprechenden an die Individualität und Singularität jeder kommunikativen Konstellation, der sozialen Hintergründe und der materialen Gegebenheiten angepasst und haben somit zugleich auch immer Züge von Emergenz (vgl. Barth-Weingarten 2008: 82; siehe auch Günthner 2018b). Die Interaktionale Linguistik nimmt mit Auer (2000) insofern die Prozessualität der gesprochenen Sprache ernst (und das nicht nur, wie Auer in den Mittelpunkt stellt, im Bereich der Syntax). Dies hat zur Konsequenz, dass etwa ein

strukturalistischer Satzbegriff als Segmentierungskategorie zu kurz greift und andere Regelwerke, wie sie normativ in Grammatiken niedergeschrieben sind, auf der empirischen Ebene nicht nur nicht ‚eingehalten‘ werden, sondern systematisch Regeln, die erst durch den Gebrauch entstehen, übersehen. Grammatik ist insofern ein „ständig sich veränderndes Inventar an Orientierungsmustern“ (Imo 2013: 79), das sich im Spannungsfeld von situativer Emergenz auf der einen Seite und grammatischer, musterhafter Verfestigung auf der anderen Seite bewegt (siehe auch Günthner/Bücker 2009).

Als Wegbereiter und Grundlage der Interaktionalen Linguistik können andere soziologische Arbeiten gesehen werden, die einen Fokus auf die sprachliche Beschreibungsebene legten (vgl. Imo/Lanwer 2019: 59). Dazu zählen die Kontextualisierungstheorie, die Arbeiten von John Gumperz und Dell Hymes zur Interaktionalen Soziolinguistik und zur Ethnographie der Kommunikation (Gumperz/Hymes 1972; Hymes 2004; Gumperz 1982; 1999), die ebenfalls einen gebrauchsbasierten Ansatz verfolgten, sich gleichermaßen gegen die abstrakten, kontextlos gelösten Sprachstrukturen richteten und linguistische Formen in ihrem lokalen, aber auch ihrem sozialen und kulturellen Diskurs berücksichtigen wollten (vgl. Günthner 2008: 53). Besonders prominent und fruchtbar erwies sich das Konzept der Kontextualisierungshinweise (*contextualization cues*), die von Gumperz (1982) und für das Deutsche von Auer (1986) in den Blick genommen wurden und zeigen, dass auf allen sprachlichen Ebenen, von der Prosodie bis zur Stilistik, indexikalische Zeichen auftreten können, die als relationale, an einen sozialkulturellen Kontext gebundene, inferentielle Äußerungen zu verstehen sind und gleichzeitig diesen Kontext auch immer wieder konstituieren. Diese Grundannahmen sind nicht nur für die Entwicklung des „Forschungsprogramms Interaktionale Linguistik“ (Selting/Couper-Kuhlen 2001a) wesentlich gewesen (siehe auch Couper-Kuhlen/Selting 2017: 8–10), sondern sie sollen auch in den Analysen der vorliegenden Arbeit fruchtbar gemacht werden. Sie sind Teil einer anthropologischen Linguistik, die Sprache als Teil kultureller Praktiken begreift und deren Rolle bei der Entstehung und Entwicklung von sozialen und kulturellen Identitäten, in der Kulturverständigung und in der alltäglichen Situierung empirisch untersucht (vgl. Günthner 2008: 53; Foley 1997: 3).

Die Interaktionale Linguistik hat, so kann konstatiert werden, viele, weit verzweigte Wurzeln. In den Analysen wird nicht streng allen methodischen Geboten der Interaktionalen Linguistik gefolgt – dazu bietet sie sich auch mit der praxeologischen Fundierung dieser Arbeit nicht optimal an. Gleichwohl sollen wesentliche Prinzipien, insbesondere ein sequenzanalytisches Vorgehen und ein Interesse an Zusammenhängen zwischen sprachlicher Form und kommunikativer Funktion, übernommen werden. Dies kann aber nicht isoliert von anderen methodologischen Paradigmen geschehen, auf die nachfolgend noch einzugehen ist.

4.2.4 Gesprochene-Sprache-Forschung

Von der philologischen Linie der Forschung her beeinflusst waren unterschiedliche Arbeiten, die als *Gesprochene-Sprache-Forschung* zusammengefasst werden können (vgl. Imo/Lanwer 2019: 19–20). Auch auf diese wird in Arbeiten der Interaktionalen Linguistik regelmäßig zurückgegriffen, teilweise können sie auch zur Interaktionalen Linguistik hinzugerechnet werden. Sie wurzeln in einem Aufsatz des Germanisten Otto Behagel (1927), der in einer Festschrift die Domänen der Mündlichkeit und der Schriftlichkeit voneinander abgrenzte und auf wesentliche Unterschiede aufmerksam machte. Auch wenn dies in der philologischen Forschung zunächst keine große Rolle spielte, wurden diese Gedanken wieder aufgegriffen und Versuche einer Einordnung der gesprochenen Sprache in die Sprachgesamtheit (vgl. Moser 1969), Typisierungen (vgl. Schröder/Steger 1981) und syntaktische Beschreibungen (vgl. Eggers 1962) unternommen. In dieser linguistischen Auseinandersetzung, unbeeinflusst von den zeitlich parallel stattfindenden Prozessen, die zur Ausbildung der Ethnomethodologie, der Konversationsanalyse oder der Interaktionalen Soziolinguistik führten, rang man um die Kriterien, mit denen die Mündlichkeit in Abgrenzung von der Schriftlichkeit überhaupt bestimmt werden sollte (etwa Koch/Oesterreicher 1985). Zur Gesprochene-Sprache-Forschung in dieser Tradition können in jüngerer Zeit etwa die Arbeiten von Schwitalla (2012; 2018) oder der Band von Ágel/Hennig (2006) gezählt werden, deren Untersuchungen nicht so sehr von den soziologischen Forschungstraditionen beeinflusst waren, sondern stärker in der Germanistik verwurzelt sind. Die Empirie ist hier häufig nicht leitend für die Analyse, sondern Fragestellungen der Typisierung und Charakterisierung gesprochen-sprachlicher Funktionen.

An gesprochener Sprache interessiert und gebrauchsbasiert arbeitet auch die *Funktionale Pragmatik*, die Gesprächsforschung mit Kontexten und gesellschaftlichen Zusammenhängen über die einzelne Situation hinaus verbunden hat (vgl. Meer/Pick 2019: 31). Im Mittelpunkt der Untersuchungen der u. a. von Konrad Ehlich und Jochen Rehbein begründeten Funktionalen Pragmatik stehen der Zusammenhang von Sprache und Handeln bzw. Handlungen und (gesellschaftlichen) Zwecken des in Gespräche eingebundenen Handelns (vgl. Ehlich 2019: 304). Damit konnten insbesondere über Gespräche in gesellschaftlichen Institutionen (u. a. Schule und Hochschule, Gericht, Sozialberatung, aber auch Familie) und somit über den Zusammenhang von gesellschaftlichen Zwecken und damit verbundenen Handlungsmustern Erkenntnisse gewonnen werden (vgl. Meer/Pick 2019: 32). Dafür wurde ein spezifisches Set an Begriffen und Methoden entwickelt (einführend Ehlich 2007), das teilweise große Ähnlichkeit mit anderen Ansätzen aus der Gesprächsforschung aufweist und diesen insbesondere durch die starke Handlungs- und Anwendungsorientierung ähnelt. Forschungsergebnisse aus den

Forschungslinien der Gesprochene-Sprache-Forschung und der Funktionalen Pragmatik werden zwar in den Analysen dieser Arbeit teilweise herangezogen, deren Erkenntnisinteressen und methodologische Vorgehensweisen aber nicht geteilt. Im Mittelpunkt der vorliegenden Arbeit steht eine induktive Vorgehensweise und die sequenzanalytische Datenauswertung, was (mit guten Gründen) in den genannten Strömungen nicht der Fall war.

4.3 Sprache und Ethnografie

Gesprächsanalytische Arbeiten im weiter oben vorgestellten, ersten Sinne (als Fortsetzung der Konversationsanalyse) zeigten sich etwas offener für methodologische Erweiterungen aus dem Bereich der Ethnografie, ganz besonders durch den wegweisenden Aufsatz von Deppermann (2000). Damit sind sie etwas anders nuanciert als Arbeiten der ‚strengen‘ Konversationsanalyse am anderen Ende eines ausdifferenzierten Spektrums von Arbeiten, die nicht immer eindeutig klassifiziert werden können und müssen. Dennoch war eine solche Öffnung nicht selbstverständlich: Der Anspruch der Konversationsanalyse ist es, nur die Informationen für die Analyse zu verwenden, die die Sprecher*innen auch im Gespräch selbst für die Organisation der Konversation relevant machen – Schlussfolgerungen sollen also ausschließlich anhand der Gesprächsaufzeichnungen gezogen werden. Keine zusätzlichen Informationen sollen in die Analysen einfließen. Diese Regeln werden auch in der vorliegenden Arbeit ganz bewusst nicht befolgt. Denn schon aus den theoretischen Verortungen wurde deutlich, dass dies einer praxeologisch ausgerichteten Arbeit mit einem Interesse an einer ganzheitlichen Betrachtungsweise nicht gerecht werden kann: Es wird in dieser Arbeit explizit nicht nur nach den Prinzipien gefragt, mit denen eine Interaktion bzw. ein VUI-Dialog organisiert werden, sondern auch danach, wie interaktionale und soziotechnische Arrangements entstehen, wie sprachliche Praktiken Ausdruck einer sozialen Vollzugswirklichkeit an der Oberfläche sind und wie sich das kommunikative Repertoire der Interagierenden entwickelt. Ohne Hintergrundinformationen über die Sprecher*innen und deren Beziehungsgeflecht lassen sich diese Praktiken und Prinzipien nicht erfassen. Es wird mit Deppermann (2000: 99) davon ausgegangen, „dass das Erfassen von Sinn und Ordnung im Gespräch keine Frage eines einfachen Ablesens oder Hörens ist. Das Verständnis von Sinn und Ordnung bedarf vielmehr der Interpretation“. Interpretationen, so konstatiert Deppermann, werden in der strengen Konversationsanalyse für einen Teil der untersuchten Fragestellungen nicht hinreichend berücksichtigt. Zwar wird, z. B. von Schegloff (1997b), die Relevanz von Kontextwissen durchaus einbezogen, dies soll jedoch lediglich aus den Daten heraus rekonstruiert werden. Depper-

mann (2000: 100) bezweifelt demgegenüber, dass dies möglich ist, aus drei Gründen: Erstens, weil die Aufzeigeleistungen nicht immer für Gesprächsteilnehmer*innen wie auch für Forscher*innen eindeutig sind, zweitens, weil nicht trennscharf unterschieden werden kann zwischen Kontextaspekten, die im Gespräch relevant sind, und solchen, die dies nicht sind, und drittens, weil Aufzeigeleistungen nicht ohne Vorwissen *erkannt* werden können.

Deppermann steht mit dieser Kritik nicht allein, schon Gumperz (1982; 1999) und Gumperz/Hymes (1972) argumentieren so bei der Begründung der Interaktionalen Soziolinguistik und der Ethnographie des Sprechens (siehe auch Hymes 2004). Sie argumentieren, dass das Erkennen von Kontextualisierungshinweisen nur möglich ist, wenn auch Hintergründe der Teilnehmendenkonstellation sowie kulturelles Vorwissen einfließen.¹³¹ Auch die Forschungen an der Schnittstelle von Linguistik und Anthropologie, z. B. von Duranti (1997; 2008; 2009) oder Foley (1997), die sich der Konstruktion von sozialen und kulturellen Identitäten über sprachliche Praktiken widmen und die Interaktionale Linguistik wesentlich beeinflusst haben, arbeiten mit einem hohen Maß an ethnografischer Anreicherung der linguistischen Daten.

Der überwiegende Teil der *medienlinguistischen* Arbeiten mit gesprächslinguistischen Anteilen aus den letzten Jahren berücksichtigt Hintergrundinformationen, auch wenn der ‚Grad der Ethnografisierung‘ der Arbeit unterschiedlich ausgeprägt ist: Er reicht von knappen Vorabinformationen zu Beziehung und Interaktionsgeschichte, wie sie exemplarisch bei König (2019: 41–45) zu finden sind, bis hin zu umfangreichen Ethnografien, wie etwa bei den Arbeiten von Dang-Anh (2017; 2019) oder Huynh (2020), in denen Interviews, Situationsbeobachtungen oder Feldnotizen eine monatelange Annäherung und ein tiefgehendes Verständnis der kommunikativen Praktiken ermöglichten. Diese verschiedenen Umgänge mit dem Einbezug der Ethnografie sollten nicht gegeneinander ausgespielt werden, sondern sie sind im Sinne des „unique adequacy requirement“ (Garfinkel 2002: 175–176) Möglichkeiten, auf die Fragestellung und den zu untersuchenden Gegenstandsbereich angepasst zu reagieren.

Ethnografie etabliert sich in der empirischen, angewandten Linguistik also zunehmend als Bestandteil der Forschung, wie auch neuere Lehrbücher zur angewandten Sprachwissenschaft (etwa Meer/Pick 2019) oder zur Konversationsanalyse (etwa Birkner et al. 2020) zeigen. In der Ethnologie, Anthropologie und Soziologie hat ethnografisches Arbeiten demgegenüber bereits eine lange Tradition

¹³¹ In diesem Stil arbeiten u. a. auch Kallmeyer/Keim (1994) und Kallmeyer (1995), die eine Verbindung von ethnografischen Beobachtungen und sprachlichen Analysen durchführen, um städtische Sozialkonfigurationen zu untersuchen.

zur Erforschung sozialer Praktiken. Sie ist hier methodologisch schon länger etabliert und hat ihre Wurzeln in der älteren ethnologischen Kulturanalyse, der Chicagoer Schule sowie der phänomenologischen Alltagssoziologie nach Schütz (vgl. Breidenstein et al. 2013: 13; Ayaß 2016: 335–336). Diese Ansätze eint eine methodische Vielfalt, denn „Ethnografie ist keine Methode“ (Breidenstein et al. 2013: 34), sondern vereint vielmehr ein flexibles Methodenspektrum zu einem Forschungsstil, der sich durch einen „feldspezifischen Methodenopportunismus“ (Breidenstein et al. 2013: 34) auszeichnet und insofern weder auf ein bestimmtes Vorgehen noch einen bestimmten Datentyp festgelegt ist. Ethnografisch zu arbeiten, bedeutet, sich den Bedingungen im Feld anzupassen. Es wird versucht, „die Komplexität des untersuchten Felds zu erfassen“ (Ayaß 2016: 337). Das setzt eine zweite Bedingung ethnografischer Forschung voraus: Ethnografische Arbeit zeichnet sich dadurch aus, dass Forscher*innen ins ‚Feld‘ gehen und dort unmittelbare und vergleichsweise lange Erfahrungen im direkten Kontakt mit den Beforschten machen (vgl. Breidenstein et al. 2013; siehe dazu auch Charmaz 2006). Es entsteht, bei allen Arten qualitativer Sozialforschung, aber bei ethnografischen Verfahren im Besonderen, eine andere Beziehung und ein längerer Kontakt zwischen den Forscher*innen und den Beforschten bzw. dem beforschten Feld (vgl. Scholl 2016: 25). Teilnehmende Beobachtung und Protokolle darüber sind üblicherweise Kern einer ethnografischen Vorgehensweise, aber auch Interviews, Aufnahmen, gesammelte Dokumente und Artefakte und auch die eigene Erprobung und Erfahrungen von Praktiken des Feldes können Bestandteile einer Ethnografie sein (vgl. Ayaß 2016: 337). Einwirkungen der Forscher*innen auf das Feld, also deren „Reaktivität“, werden durch reflexive Verfahren, mit denen die Forscher*innen ihr Agieren im Feld in die Analyse einfließen lassen, bearbeitet. Es entsteht so eine „*intellektuelle Distanz* [Herv. i. O.]“ (Breidenstein et al. 2013: 109) zu den Erlebnissen und Erfahrungen; diese Verfahren gehören zur Anatomie der ethnografischen Arbeit (siehe auch Reichertz 1992).

Ethnografisches Arbeiten betont außerdem eine Art Transkription des sozialen Geschehens. Das Geschehen selbst kann, aufgrund seiner Situiertheit, seiner Zeitlichkeit und Flüchtigkeit, seiner Materialität und seiner Singularität, nicht erfasst und untersucht werden, es kann nur greifbar gemacht werden, wenn es darstellbar gemacht wird, und erst durch die Darstellbarkeit wird das Offensichtliche überhaupt sichtbar. Allein dieser Darstellungsprozess ist wesentlicher Bestandteil der reflexiven Leistungen der Ethnograf*innen. Ethnografische Darstellung, die klassischerweise schriftlich erfolgt, aber darauf nicht beschränkt ist (vgl. Dwelling/Prus 2012: 194–196), ist von Einfachheit im Stil geprägt und soll v. a. eine leichte Zugänglichkeit zu den gemachten Beobachtungen ermöglichen. Ethnografisches Arbeiten ist als „offene Forschung“ verstanden explorativ (vgl. Dwelling/Prus 2012: 53–54), d. h. hypothesengenerierend, und geht – ganz ähnlich wie auch

die vorgestellten methodologischen Konzepte der Konversationsanalyse – ohne vorab gebildete Vorannahmen an die Daten heran. Diese würden den Blick verstellen und mit *a priori* angenommenen Kategorien die Analysen zu sehr verzerren. Ethnografie bzw. ethnografische Methoden eignen sich also als Forschungsstil bestens zur Analyse sozialer Praktiken, wie auch die von Schäfer/Daniel/Hillebrandt (2015) zusammengestellten Arbeiten von Lengersdorf (2015), Both (2015) und Porsché (2015) deutlich zeigen. Der Stil erlaubt es, methodisch flexibel auf das beobachtete Geschehen reagieren zu können; er kann unterschiedliche Schwerpunkte setzen und eine ganzheitliche Betrachtung von *in situ* verfertigtem Geschehen ermöglichen, und so praxeologisch die Lebenswirklichkeit einfangen.

In Verbindung mit dem sprachwissenschaftlichen Erkenntnisinteresse bedeutet das: Ethnografie ist erstens ein *Forschungsstil*, dessen methodologische ‚Denkweise‘ auf die Linguistik übertragen werden soll. Er muss sich den linguistischen Zielen nicht unterordnen, sondern soll, an den Methodenzwang des Gegenstands angepasst (vgl. Amann/Hirschauer 1997) – also ausgerichtet auf die Beschreibung des *Sprachlichen* –, eine Haltung sein, mit der sprachliche als soziale Praktiken erfasst werden können. Zweitens ist Ethnografie eine *Notwendigkeit*, um überhaupt sprachliche Praktiken als Bestandteil sozialer Praktiken einfangen zu können; dazu wurde bereits auf den deutlichen Appell von Deppermann (2000) und dessen breite Rezeption in der Sprachwissenschaft, auch über die Gesprächsforschung hinaus, hingewiesen. Drittens ist in ethnografischen Arbeiten eine Vielzahl von Methoden angewendet worden, mit denen unterschiedliche Erfahrungen gemacht werden konnten und die für verschiedene Zwecke mehr oder minder geeignet waren. Auf diese Erfahrungen soll bei der Auswahl der Methoden ebenfalls zurückgegriffen werden, auch wenn sie sich zugleich an den interaktional-linguistischen Notwendigkeiten orientieren muss. Das bedeutet auch, dass bei Sprache und Gesprächen als Forschungsgegenstand Abwägungen zwischen einem intensiven, partizipativen Feldkontakt und z. B. einer in solchen Settings unvermeidbaren Beeinflussung auch der sprachlichen Ausdrücke getroffen werden müssen. Dieses als Beobachterparadoxon (vgl. Labov 1972) formulierte Phänomen wird, auch bei einer Relativierung des etwa von Potter/Wetherell (1987) erhobenen Anspruchs „natürlicher Daten“,¹³² einige ethnografische Methoden ausschließen. Im Gegenzug kommen für eine umfassende Beschreibung sprachlicher Praktiken, die die soziale und materiale Situiertheit ernst nimmt, besonders Methoden in Betracht, die das Geschehen aufzeichnen und es so einer feingranularen Untersuchung zugänglich machen, die durch bloße Beobachtung

132 Für eine kritische Betrachtung dazu siehe etwa ten Have (1999) und sehr differenziert Gerwinski/Linz (2018). Siehe auch die Ausführungen in Kap. 5.3.

nicht ermöglicht wird. Nur so können parallel stattfindende Bestandteile einer sprachlichen Praktik (etwa Verbalisierung und kinesische Bestandteile), die körperlich-räumliche Organisation und Konfiguration von Interaktionen und der Einbezug der materialen Umgebung hinreichend erfasst werden.

4.4 Multimodalität

Goodwin (2012) zeigt, dass sprachliche Äußerungen eingebunden sind „in eine soziale Ökologie der Praktiken zur Bedeutungsherstellung“ (Goodwin 2012: 418), die alle Mitglieder einer Gesellschaft umgibt und an der wir uns alle ebenso orientieren wie wir sie gleichermaßen mitgestalten. Diese Ökologie umfasst zwar sprachliche Ausdrücke und auch wenn diese unbestritten einen hohen bedeutungsstragenden Anteil haben, so argumentiert Goodwin (2000: 1490), dass Sprache nicht isoliert betrachtet werden kann, sondern einen Analyseansatz benötigt, „that takes into account the simultaneous use of multiple semiotic resources by participants“; dazu gehören verschiedene Zeichensysteme, die situative und sozialräumliche Umgebung sowie die sequenzielle Organisation selbst. Goodwin wendet sich damit gegen einen Dualismus von Text und Kontext und plädiert für eine ganzheitliche Betrachtungsweise der Interaktion und ihrer Beteiligten.¹³³ Goodwin postuliert, dass die Kategorien Text und Kontext nicht von Analytiker*innen *a priori* gesetzt werden können, sondern dass diese Unterscheidung Gegenstand einer gemeinsamen Hervorbringung und interaktiven Aushandlung ist (vgl. Goodwin 2000: 1519; 1981; 1984).

Der Mainstream der Sprachwissenschaft hat diese Untersuchungsbereiche, auf die Goodwin bereits in seinen frühen Arbeiten eingeht, erst um die Jahrtausendwende für sich entdeckt: Aspekte von Gestik, Mimik, Körperlichkeit und Raum sind erst spät in den Blick der Linguistik gerückt (siehe etwa Mondada 2007). Die technischen Entwicklungen im Bereich der Datenerhebung und -aufbereitung machten es möglich, auch Videoaufnahmen von Interaktionen zu erheben und den Aufwand

¹³³ Schüttpelz (2020: 139) betont die Zusammengehörigkeit des (nach Goodwin 2018) *akkumulativen* Charakters der sprachlichen Interaktion mit deren *altruistischen* Eigenschaften: „Vielleicht [...] kann man sagen, dass ‚Sprache‘ nichts anderes ist als menschlicher ‚Altruismus‘, oder im Gegenzug, dass sich der ‚Altruismus‘ des Menschen zuerst und zuletzt in seiner ‚Sprachlichkeit‘ verwirklicht“. Es ist dieser Altruismus, die Möglichkeit einer Assistenz zum Gelingen einer erfolgreichen Interaktion, der sich ganz besonders aus der sozialen Ökologie speist: Sprachliche Erscheinungen sind ein Teil davon, ebenso die Interaktionsgeschichte, die gemeinsamen Beziehungen und die räumliche Umgebung, die über Gestik und Mimik in die Interaktion einbezogen werden kann (vgl. Schüttpelz 2020: 138–142).

für ihre Verarbeitung erheblich zu reduzieren (vgl. Deppermann/Schmitt 2007: 15), und führten so zu einer Ausdehnung der Analysebereiche auch in den Bereich der Gestik, Mimik und des interaktionalen Raums. Die Mechanismen der Organisation sprachlicher Interaktion wurden in der Tradition der frühen konversationsanalytischen Arbeiten zu großen Teilen anhand von Telefongesprächen erschlossen (vgl. Sacks/Schegloff/Jefferson 1974; Schegloff/Jefferson/Sacks 1977; Jefferson 1978). Es konnten auf diese Weise grundlegende Erkenntnisse über die konversationelle, sequenzielle Organisation der Interaktion gewonnen werden, die Rolle anderer Ausdrucksressourcen konnte dabei allerdings aufgrund der Charakteristika von Telefongesprächen kaum bearbeitet werden. Dabei soll nicht übersehen werden, dass Gestenforschung bereits weitaus früher begann, etwa mit den wegweisenden Arbeiten von Kendon (1972; 1980; 2004), Kendon/Sebeok/Umiker-Sebeok (1981) oder auch Ekman/Friesen (1969); diese nehmen jedoch keine an der Praxis orientierte Perspektive ein, gehen nicht sequenzanalytisch vor und sind insofern weniger an der interaktiven Hervorbringung sozialer (und insofern kommunikativer und sprachlicher) Praktiken interessiert.¹³⁴ Einschlägigere Vorarbeiten im Bereich der Multimodalität leistete insbesondere Charles Goodwin; sein Zusammendenken von Sprache und materialer und räumlicher Umgebung (Goodwin 1994; 1981; 1996; 2000; 2003; 2007; 2012) war eine wesentliche Vorbereitung, bevor der *multimodal turn* insbesondere den empirischen Zweig der Linguistik erfasste.¹³⁵ Dieser wirkte sprachwissenschaftlich prägend: Die Bedeutung der Multimodalität in der Interaktion wird in heutigen Arbeiten kaum noch unterschätzt. Gleichwohl gibt es eine fortlaufende Debatte darüber, was genau unter Multimodalität zu verstehen ist. Ein einheitliches Verständnis dieses Terminus ist bislang nicht gefunden, wie auch Deppermann (2013: 2) betont und eine Verwendung des Begriffs angepasst auf den jeweiligen Kontext (hier: Konversationsanalyse, Gleiches gilt aber auch für die Adaption an andere Subdisziplinen) anmahnt:

[...] „[M]ultimodality“ is a label which is already worn out and has become most fuzzy by its use in various strands of semiotics, discourse and media analysis. So, there may be the danger that under the heading of ‚multimodal interaction‘ a lot of things will gather which do not have to do much with CA. (Deppermann 2013: 2)

¹³⁴ Theoretische Grundlagen für deren Erforschung im linguistischen Kontext lieferte bereits Bühler (1965[1934]) in seiner Sprachtheorie.

¹³⁵ Wie Selting (2013: 590) betont, nahmen gelegentlich auch andere Konversationsanalytiker*innen zusätzliche Aspekte abseits des verbal-vokalen Ausdrucks in den Blick, so etwa Schegloff (1984) oder Lerner (2003), doch wurden diese hier eher als Mittel zum Zweck einbezogen und nicht systematisch in ihrer Relevanz für die Interaktionsorganisation untersucht.

Der vorliegenden Arbeit liegt ein Verständnis von Multimodalität zugrunde, das sich an der Pluralität und Gleichzeitigkeit der verkörperten Ausdrucksressourcen orientiert. Mondada (2021: 44) konstatiert, dass deren Ausgangspunkt in der Erkenntnis ethnomethodologischer Arbeiten über die Körperlichkeit des Sprechens liegt, die die „artificiality of separating language and the body in action“ hervorheben. Die Organisation der Interaktion beruht nicht nur auf sprachlichen Verfahren (in denen Lexik, Syntax, Prosodie usw. kombiniert werden), sondern auch auf Gestik, Blick, Körper(haltung), räumlichen Bewegungen und der physischen Umgebung und ist somit inhärent *multimodal* (vgl. Mondada 2021: 45).¹³⁶ Entscheidend ist, und so soll Multimodalität verstanden und in den Analysen bearbeitet werden, woran die Gesprächsbeteiligten *in der Praxis* die Sinnkonstitution festmachen und welche Ressourcen sie dafür einsetzen. Dabei kann in einem praxeologischen Sinne sowohl der eigene Körper wie auch die materiale und imaginäre Umgebung relevant werden und auch die Vergangenheiten und imaginierten Zukünfte dieser (vgl. Goodwin 2018), auf die ebenfalls mit diesen Ressourcen verwiesen werden kann.¹³⁷ Keine der Modalitäten kann dabei *a priori* als dominant gegenüber einer anderen Modalität aufgefasst werden (vgl. Stivers/Sidnell 2005: 2) – dies kann nur empirisch gezeigt werden:

When people interact within embodied social frameworks that are structured and changed through their shifting co-presence, analysts should attend to what the participants themselves are treating as important. (Streeck/Goodwin/LeBaron 2011: 12)

In einem methodologisch ganzheitlichen Ansatz, wie er bisher skizziert wurde, sollen die verschiedenen Ausdrucksressourcen also gleichermaßen Berücksichtigung finden – je nachdem, in welchem Maße die Beteiligten selbst Sinn an diese koppeln. Ausgehend von der bisherigen Multimodalitätsforschung und den dabei berücksichtigten Modalitäten ist insbesondere auf folgende Aspekte zu achten, die schon in den bisherigen Ausführungen immer wieder genannt wurden: Körperlichkeit, Räumlichkeit und Simultaneität sowie das Wechselspiel mit Sprache (siehe dazu etwa Fricke 2013; 2015).¹³⁸ Auf diese drei Aspekte wird nachfolgend kurz eingegangen.

136 Für dieses in Studien zur sozialen Interaktion überwiegende Verständnis von Multimodalität als Kombination verschiedener Ausdrucksressourcen in der Praxis der Interaktionsorganisation siehe Mondada (2014b). In Abgrenzung zu disziplinär anders gelagerten Verständnissen, etwa in der Computerwissenschaft oder der Semiotik, siehe Mondada (2021: 45).

137 Siehe dazu auch das Konzept der „Deixis am Phantasma“ (vgl. Bühler 1965[1934]) und neuere Adaptionen, z. B. von Stukenbrock (2015) zur Deixis in der Face-to-Face-Interaktion; siehe auch Heller (2018) sowie Selting (2013: 590).

138 Nach Fricke (2012: 47–49), die ein stärker semiotisch orientiertes Multimodalitätsverständnis vertritt, ist Multimodalität im engeren Sinn erst dann gegeben, wenn sich *ein* Zeichensystem auf mehreren wahrnehmbaren Kanälen darstellt, z. B. bei redebegleitenden Gesten. Um dadurch ent-

Körperlichkeit: Interaktion ist nach Hausendorf (2015: 52) durch die „Inanspruchnahme humanspezifischer Körperlichkeit“ materialisiert. Damit sind zum einen die sprachlichen Äußerungen selbst gemeint, die zu ihrer Produktion wie Rezeption auf einen menschlichen Körper angewiesen sind – in ihrer einfachsten Form sind dafür keine weiteren Hilfsmittel notwendig. Darüber hinaus sind aber andererseits nicht nur die verbalsprachlichen Äußerungen, sondern auch ganze Körper mit ihrer Materialität an der Sinnkonstitution beteiligt. Dazu kann zunächst die Gestik gezählt werden, mit deren Bedeutung in der Interaktion und praktischer Sinnkonstitution sich schon früh Charles Goodwin auseinandersetzt und die auch Jürgen Streeck eingehend untersucht hat.¹³⁹ Streeck (2016: 77) sieht in sprachbegleitenden Gesten Praktiken des Beschreibens, d. h. „habitualisierte[] Methoden [Herv. i. O.]“, die wiederkehrend zur Lösung bestimmter kommunikativer Aufgaben, d. h. in der kommunikativen Praxis, eingesetzt werden, die aber kontextlosgelöst nicht als konventionalisierte Formen beschrieben werden können. Zeigegesten können in hohem Maße in eine Lokaldeixis unterschiedlicher Art eingebunden sein (vgl. Uhlmann 2013; Stukenbrock 2015; 2016). Streeck (2013: 679) lenkt in seiner „Praxeologie der Gestik“ die Aufmerksamkeit weg von der Beziehung, die verbale, visuelle, kognitive oder andere Teile untereinander haben; eine Praxeologie der Gestik ist stattdessen „interested in the ways in which gestures participate in the bringing about of shared practical understandings“ (Streeck 2013: 679) – Gestik wird als Praxis, Gesten werden als Praktiken betrachtet. Insbesondere bei bildhaften Gesten wird, so Streeck (2013: 681), der Vorteil einer praxeologischen Beschreibung deutlich: Es gibt allenfalls in Ansätzen ein kodifiziertes, kanonisiertes Repertoire von bildhaften Gesten, in der Mehrzahl der Fälle „when a specific gestural depiction is needed, it is improvised. The depiction may or may not include established depictive forms“ (Streeck 2013: 679). Diese Betrachtung

stehende Abgrenzungsproblematiken zu vermeiden, soll den Abschattungen von Multimodalität (engerer/weiterer Sinn), wie sie auch Imo/Lanwer (2019: 248) einführen, an dieser Stelle nicht gefolgt werden, sondern all diese Fälle als multimodale Interaktion begriffen werden. Sprachliche Interaktion ist mithin *immer* multimodal – auch sprachliche Äußerungen kombinieren verschiedene Ausdrucksressourcen (etwa Lexik und Prosodie, um nur die auffälligsten zu nennen).

¹³⁹ Erneut sollen die Arbeiten von Adam Kendon, auf die bereits einführend verwiesen wurde, nicht übersehen werden, auch weil Kendon (2004) eine durchaus differenzierte Klassifizierung verschiedener Gestentypen vorlegt. Doch werden diese in Ermangelung einer interaktional oder praxeologisch orientierten Untermauerung in den Analysen nicht fruchtbar gemacht und auf eine genauere Darlegung insofern an dieser Stelle verzichtet; für eine fruchtbare Anwendung von Kendons Arbeiten in der Linguistik siehe etwa Huynh (2020), die eben diese Unterfütterung selbst als Gegenstand der Arbeit im Kontext der Darstellung und interaktional-praktischen Aushandlung von Emotionen leistet.

tungsweise soll helfen, die *in situ* vollzogenen Gesten zu erfassen und den praxeologisch-theoretischen Anspruch der Arbeit auch methodologisch einzulösen.

Körperlichkeit bleibt allerdings nicht auf Verbalsprache und Gestik beschränkt, auch wenn diese besondere Aufmerksamkeit erfahren haben. Mit einem kompatiblen theoretischen Hintergrund wurde auch zu anderen körperlichen Ressourcen gearbeitet, z. B. wurde bereits früh die Rolle von Blickverhalten in der sozialen Interaktion erkannt (vgl. Kendon 1967) und in seiner Rolle für die Turn-Koordination (vgl. Goodwin 1980; Mondada 2007) oder für den Spracherwerb (vgl. Tomasello 1988) untersucht und weitere Arbeiten darauf aufgebaut, die auch das Zusammenspiel dieser beiden Ressourcen in den Blick nehmen (vgl. Sidnell 2006; Mondada 2007; Streeck 2014). Methodologisch wurde die Einführung auf den Blick v. a. durch den Einsatz von Eye-Tracking-Technologien bereichert, Brône (2018) legen in ihrem Band eine Reihe aktueller Forschungen vor, die z. B. auch die Rolle von Blick beim Turn-Taking (vgl. Auer 2018) und bei Mensch-Computer-Dialogen berücksichtigt, wie auch die Arbeiten von Gehle et al. (2015) oder Pitsch et al. (2009b; 2017) zeigen. Darüber hinaus ist, wie Streeck/Goodwin/LeBaron (2011) argumentieren, Körperlichkeit in der Interaktion auch über Gestik und Blickverhalten hinaus eine entscheidende Voraussetzung für eine gemeinsame Ausrichtung aufeinander. Diese bildet, zeitlich entkoppelt von den unmittelbaren Äußerungen und Bewegungen der Gesprächsteilnehmer*innen, als „embodied participation framework“ (Streeck/Goodwin/LeBaron 2011: 2; siehe auch Goodwin 2002) eine Grundlage für die Gesprächsorganisation und Aushandlung der Teilnahme am Gespräch, auch für – insbesondere von Stukenbrock (2015; 2018 u. a.) untersuchte – multimodale deiktische Praktiken, die erst dadurch verständlich werden, dass Körperhaltung, Blick und räumliche Umgebung einbezogen werden; „embodied interaction“ unterscheidet sich somit in seiner Gesamtheit von spezifischen Gesten und Blick und nimmt die übergreifende Struktur der Körperlichkeit mit in den Blick:

One phenomenon that quickly emerges from records that preserve not only the talk but also the bodies of actors, is that action is built through the mutual elaboration of diverse semiotic resources with quite different properties, each of which, including language, can make only a partial, incomplete contribution to the action in progress. (Streeck/Goodwin/LeBaron 2011: 3)

Ein weiterer Hinweis auf die Relevanz dieser körperlichen Praktiken ist in der Forschungstradition der Ethnografie zu finden, die sich bereits früh mit Körpern beschäftigt und deren Rolle bei dem Vollzug sozialer Praktiken ernst genommen hat. Im Rahmen verschiedener Studien haben ethnografische Arbeiten herausgestellt, dass körperlicher Ausdruck als „culturally shaped and socially acquired“ (Meyer 2013: 235) betrachtet werden muss.

Räumlichkeit: Eng verbunden mit den körperlichen Ressourcen, der Körperhaltung und dem Blickverhalten ist die räumliche Umgebung, teilweise ist auch von „visuospatial modalities“ die Rede (Stivers/Sidnell 2005: 2), was die Verbindung von Sichtbarkeit und Räumlichkeit betont. Dabei soll Stukenbrock (2013a: 402) gefolgt werden, wenn sie die These aufstellt, dass eine Gemeinsamkeit der Ansätze zur Betrachtung von Raum und Interaktion die (überwiegend) geteilte Annahme ist, „that space is not a given entity, but an interactive achievement of the participants“ und insofern nicht vom „interactional space“, sondern vom „interactive achievement of space“ die Rede sein sollte (siehe auch Mondada 2013; Hausendorf 2013). Darüber hinaus kann der Raum nicht nur ein gemeinsames „achievement“ sein, sondern auch eine interaktional genutzte Ressource zur Herstellung von „co-orientation, co-ordination and co-operation“ (Hausendorf 2013: 278). Diese wird von Sprecher*innen zwar systematisch eingesetzt, deren Bedeutung wurde aber im Vergleich zu visuellen und verbalen Ressourcen lange übersehen (vgl. Hausendorf 2013: 278). Um den Raum als Ressource nutzen zu können, d. h. ihn als „räumliche Praktik“ zu funktionalisieren, muss er wechselseitig von den Interagierenden als solcher verfertigt werden (siehe dazu auch Stukenbrock 2012). Daher ist die räumliche Umgebung sowohl interaktionales „achievement“ wie auch Ressource, interaktionaler Raum ist ein „shared ,here‘ for perception, movement and action“ (Hausendorf 2013: 277) und – so Hausendorf weiter – gebunden an das „emplacement“ der Gesprächsteilnehmenden, d. h. leiblich in den Vollzug der Praxis eingebunden. Erst in diesem Raum können Körperlichkeit und Körperhaltungen zu Körperformationen und somit zu interaktionalen Arrangements werden.

Simultaneität: Körperlichkeit und Räumlichkeit sind nur insofern mögliche, konstituierende Bestandteile von Multimodalität, als sie *parallel* zu anderen Modalitäten auftreten. Das bedeutet nicht, dass die Modalitäten durchgängig zueinander synchron sein müssen. Zwar sind sie es teilweise, z. B. bei der hohen Synchronität von prosodischen Eigenschaften wie Akzent und gestischen Eigenschaften wie Signifikanzpunkten, die rhythmischen „Taktstockgesten“ eigen sind (vgl. Müller 1998: 103; Loehr 2007). Doch, wie Levinson (2006: 53) ausführt, ist Temporalität allein nicht ausreichend:

[...] synchrony alone will not do the trick of hooking up the bits in the different signal streams: gestures, facial expressions, nods, and the like can come earlier or later than the words they go with. If temporal binding is not sufficient, what will do the trick?

Nach Levinson ist es nicht eine Frage des Ablesens, der genauen zeitlichen Parallelität oder Sequenzialität von bestimmten Ausdrücken, sondern auch hier eine Frage der Sinnkonstitution. In der Praxis können visuelle, geräuschhafte, räumliche oder körperliche Praktiken entsprechend nicht nur untermalend gebraucht

werden, sondern dafür eingesetzt werden, zwei kommunikative Probleme gleichzeitig zu lösen – z. B. metakommunikative Aufgaben wie Turn-Taking parallel zu bearbeiten (vgl. etwa Mondada 2007). Somit laufen nicht nur intrapersonelle, sondern auch die interpersonelle Koordination in hohem Maße multimodal ab. Im Sinne einer „opportunistischen und gefräßigen Praxis“ (Goodwin 2018: 445) können alle in einer bestimmten situativen Umgebung zur Verfügung stehenden Ressourcen – verbal-vokale ebenso wie visuelle, körperliche und räumliche – funktionalisiert werden. Es ist insofern auch methodologisch zentral, nicht nur eine sich temporal-sequenziell entfaltende Praxis im „Nacheinander“ anzunehmen, sondern auch ein „Miteinander“, so wie auch Charles Goodwin in seiner Methodologie ausgeht von „social events as dynamic, emergent, intertwined totalities, not as aggregations of individual acts that are sequentially performed“ (Schüttpelz/Meyer 2018: 174).

Ergänzend zu diesen drei für die vorliegende Arbeit wesentlichen methodologischen Dimensionen von Multimodalität soll noch einmal auf das Konzept der verbalen, vokalen visuellen Praktiken und Ressourcen von Selting (2013) verwiesen werden,¹⁴⁰ das insbesondere hilfreich ist, um zu unterscheiden zwischen den rein visuellen Ausdrucksmöglichkeiten (Blick, Gesichtsausdruck, Gestik, Haltung, Einbezug von Objekten usw.) bzw. verbalen (d. h. auf allen linguistischen Beschreibungsebenen sprachlichen) und vokalen (d. h. insbesondere stimmlich-qualitativen und prosodischen) Ressourcen auf der einen und ihrer Anwendung in einem bestimmten Kontext „as they are deployed for the signalling of interactional meaning“ (Selting 2013: 591) auf der anderen Seite. Dieser Begriff der Praktik reduziert die Rolle der sozialkulturellen Hintergründe (ohne sie auszublenden) und stellt stattdessen auf die Interaktionsorganisation auf der mikrostrukturellen Ebene scharf, was für die Beschreibung der sprachlichen Interaktion ebenfalls eine notwendige Voraussetzung ist.

4.5 Zwischenfazit

Es soll also in diesem Sinne mit einer aus der ethnomethodologischen Konversationsanalyse entsprungenen und von der Interaktionalen Linguistik informierten Gesprächsanalyse gearbeitet werden, d. h. primär mit Sequenzanalysen von Beispielen aus dem erhobenen Korpus, die – wo von den Daten und den Unterfragestellungen her passend – um Aspekte der Ethnografie sowie qualitativ-interpretative Videoanalysen (siehe Schmitt 2013; Tuma 2018) erweitert wird. Es

¹⁴⁰ Siehe auch Kap. 2.1.3.

wurde deutlich, dass für die Analysen die „Prägnanz und die Wucht der visuellen Anteile an der Interaktionskonstitution“ (Schmitt/Hausendorf 2016: 9) berücksichtigt werden muss, soweit dies im Rahmen der erhobenen Daten möglich ist. Eine multimodale Herangehensweise ist – wie Schmitt/Hausendorf (2016: 9) betonen – Voraussetzung für die Untersuchung von Praktiken in einem visuell-räumlichen, interaktionalen Setting, bei dem u. a. auch die Konfiguration soziotechnischer Arrangements im Raum in einer laufenden Interaktion untersucht werden soll, um der „sichtbaren Verankerung verbaler Strukturen in räumlich-situativ strukturierten Handlungszusammenhängen“ (Schmitt/Hausendorf 2016: 10) Rechnung zu tragen.

Tuma (2018: 426) unterscheidet bei qualitativ-interpretativen Videoanalysen, die der ethnomethodologischen Konversationsanalyse in ihren Grundannahmen folgen, die „linguistische Multimodalitätsforschung“ sowie die „Video-Interaktionsanalyse“, wobei sich die vorliegende Arbeit nicht zum Ziel gesetzt hat, eine durchgehend multimodale Beschreibung zu liefern, sondern diese einzubeziehen, wo es im Hinblick auf die Erhebungsmodalitäten und die Fragestellungen sinnvoll und möglich ist. Entsprechend werden multimodal ausgerichtete Analysen vornehmlich zu Situationen angefertigt, in denen die Ersteinrichtung des Smart Speakers im Fokus steht – hierzu liegt Videomaterial vor.¹⁴¹ Videoaufnahmen, die auch die kontinuierliche Nutzung von VUIs dokumentieren, wären wünschenswert, konnten jedoch im Rahmen dieser Arbeit nicht erhoben und ausgewertet werden.¹⁴² Um sprachliche Praktiken umfassend einzufangen, bietet sich gerade eine Kombination der sprachwissenschaftlichen Einführung mit ethnografisch integrierten Bestandteilen besonders an (vgl. Deppermann 2000). So soll ein ganzheitliches Bild der Praxis gewonnen werden, durch das in den einzelnen Untersuchungen möglichst viele der für die sprachlichen Praktiken relevanten Aspekte aufgedeckt werden können. Eine solche Vorgehensweise erweist sich gerade für medienlinguistische Fragestellungen als tauglich, wie bereits in exemplarischen praxeologisch grundierten Analysen von Videoaufnahmen der Smart Speaker-Ersteinrichtung gezeigt werden konnte (vgl. Hector 2022; Habscheid et al. 2021; Habscheid/Hector/Hrncal 2023).

Die Methodologie dieser Arbeit ist insgesamt qualitativ-explorativ und somit hypothesengenerierend ausgerichtet. Ziel ist es, Phänomene zu entdecken, die für die untersuchten Teilnehmer*innen von Bedeutung sind, mit denen sie Sinn und Ordnung in VUI-Dialogen und Gesprächen unter Einbezug eines VUIs herstellen. Dazu stehen sequenzanalytische Detailbetrachtungen einzelner Ausschnitte im Vordergrund, die ein verstehendes Betrachten von VUI-Dialogen in ihrem Kontext

¹⁴¹ Siehe Kap. 5.2.

¹⁴² Zu sich daraus ergebenden Limitationen und Folgefragen siehe Kap. 7.

ermöglichen. Ein Anspruch auf Repräsentativität besteht insofern nicht und war auch nie das Ziel der Untersuchung. Im Sinne qualitativer Forschung ist aber sehr wohl die „Relevanz der gefundenen Strukturen“ (Scholl 2016: 28) sicherzustellen und durch eine Diversität im Hinblick auf (*a posteriori* gebildete) Strukturen und Materialien sowie durch Abstraktion und Reflexion der Forschungspraxis zu festigen. Gütekriterien für die Arbeit sollen im Rahmen des Untersuchungsdesigns festgelegt werden – sie können für qualitative Arbeiten nicht auf der abstrakten Ebene vordefiniert, sondern nur „unter Berücksichtigung der jeweiligen Fragestellung, Methode, der Spezifik des Forschungsfelds und des Untersuchungsgegenstands“ (Steinke 2007: 323) diskutiert werden. Auch Deppermann (2008: 105) stellt fest: „Die Güte einer Untersuchung ist nicht an punktuellen Merkmalen festzumachen“. Orientierung können bei dieser im Folgekapitel zu führenden Diskussion die von Steinke (2007) formulierten Kriterien geben, ebenso wie drei von Deppermann (2008: 105–110) genannte Aspekte: „Qualität der Datenmaterialien“, „Durchführung und Präsentation der Gesprächsanalyse“ sowie die „Frage der Generalisierbarkeit“. Wie das bisher zur methodologischen Einordnung Gesagte, heruntergebrochen auf ein konkretes Untersuchungsdesign und die Auswahl bzw. Zurichtung von bestimmten Methoden, in die konkrete Forschungspraxis und die Analysen einfließt, soll im nachfolgenden Kapitel bearbeitet werden.