

4. Multimodale Interaktionsanalysen im Anschluss an Goffman: Methodologische Überlegungen

Goffmans Arbeiten ermöglichen es, Legitimität aus einer interaktionalen Perspektive zu konzipieren; zur empirischen Erforschung von Interaktion ist das Instrumentarium, das er bereitstellt, jedoch zu wenig umfangreich. Gregory W.H. Smith schlägt vor, Phänomene der sozialen Welt, die von Goffman identifiziert wurden, empirisch mit Instrumenten der Konversationsanalyse zu untersuchen (Smith 2003, 272-273). Dies könnte auch ganz im Sinne Goffmans sein, dem Rod Watson eine »If it works, use it«-Haltung zuschrieb (Watson 1983, 104). Das Vorgehen, die Rekonstruktion von Legitimität im L2-Klassenzimmer auf der Primarstufe im Anschluss an Goffman und unter Verwendung von Instrumenten der Konversationsanalyse (*conversation analysis*, kurz CA) und der Mitgliedschaftskategorisierungsanalyse (*membership categorization analysis*, kurz MCA) zu analysieren, bedarf jedoch einer etwas ausführlicheren Diskussion. Interaktionsanalysen werden zwar schon länger auch in Fremdsprachenklassenzimmern durchgeführt; Analysen, die die soziale Dimension des interaktiven Geschehens im Fremdsprachenklassenzimmer ins Zentrum stellen, sind allerdings relativ jung, und selbstverständlich sind sie nicht. Sie können auch heute noch als »alternative Ansätze« der Erforschung des Fremdsprachenunterrichts bezeichnet werden (Kasper & Wagner 2011). Solche Analysen müssen in Rechnung stellen, dass mehrere Sprachen verwendet werden, von welchen mindestens eine Sprache eine L2 ist. Wird eine Sprache als L2 verwendet, kann sich dies auch auf die Organisation der Interaktion auswirken und damit auch neue Möglichkeiten der Konstruktion von *face*, aber auch der Gefährdung des Selbst bieten. Zudem beinhaltet das Korpus auch Interaktionen, die nur unter Präadoleszenten stattfinden, und diese unterscheiden sich von Interaktionen, an denen Erwachsene beteiligt sind; dies hat die interaktional ausgerichtete Forschung zu Kindern gezeigt (siehe etwa Cromdal 2004b, 34; Goodwin & Kyriatzi 2007; Goodwin 2008; Milroy & Wei 1995). Und letztlich muss auch dargelegt werden, wie eine multimodale Interaktionsanalyse im Anschluss an Goffman unter Verwendung konversationsanalytischer und mitgliedschaftskategorisierungsanalytischer Instrumente hier konzipiert werden kann.

4.1 Analysen von Interaktionen unter L2-Lernenden

Wie Willis Edmondson (2001) in seinem Überblick über »konversationale« Analysen des Fremdsprachenunterrichts schreibt, sind deren Anfänge schon in den 1950er Jahren zu finden.¹ Zahlreicher werden Untersuchungen und Ansätze aber erst ab den 1970er Jahren, was sich mit der kommunikativen Wende im Fremdsprachenunterricht einerseits und der Etablierung soziologischer und linguistischer Ansätze, die sich mit Kommunikation beschäftigen, erklären lässt. Einige der noch heute gebräuchlichen Ansätze sind wenig elaboriert und weisen ein spezifisches Manko auf: Sie kategorisieren Teile des Klassenzimmergesprächs etisch im Rückgriff auf didaktische Modelle. Das hat zur Folge, dass schon im Voraus feststeht, welche Aspekte der Interaktion relevant sind, nämlich diejenigen, welche durch die jeweilige Didaktik oder Lerntheorie festgelegt sind (Edmondson 2001, 1684).

Dies zeigt sich etwa in der Forschung zu »Bedeutungsaushandlungssequenzen« (Pica 1987; Long 1996; Long & Robinson 1998). Bedeutungsaushandlungssequenzen wird deswegen besonders viel Aufmerksamkeit geschenkt, weil »Bedeutungsaushandlung« als für den L2-Erwerb in dreifacher Hinsicht wichtig erachtet wird: Erstens, weil Lernende in Bedeutungsaushandlungssequenzen mit der Aufgabe konfrontiert sind, den L2-Input verstehen zu müssen, zweitens, weil sie modifizierten Output produzieren müssen, drittens, weil in Bedeutungsaushandlungssequenzen der Fokus auch auf Formen der L2 gerichtet wird. Pica (1994) und Long (1996; 1998) stützen sich theoretisch auf Krashens Inputtheorie (Krashen 1981) und greifen auf Swain (1985) zurück, die postuliert, dass es für Lernende wichtig sei, verständlichen Output zu produzieren, das heißt hier, Gesagtes in modifizierter Weise zu wiederholen. Dass durch Bedeutungsaushandlung gleichzeitig auch der Fokus auf die Form der L2 gerichtet wird, ist eine der Kernthesen dieses Ansatzes: »Both comprehension and modified production are tied closely to learners' attention to L2 form« (Pica 1994, 510; siehe auch Long 1996, 451ff.). Relevant für die Analyse ist in diesem Ansatz daher nicht die Klassenzimmerinteraktion *per se*. Relevant sind nur diejenigen kommunikativen Phänomene, die unter der Kategorie »Bedeutungsaushandlung« subsumiert werden können (Pica 1994, 515 et passim; Oliver 2000).

Dieser Ansatz lässt sich aus unterschiedlichen Perspektiven kritisieren. Edmondson (2001, 1686) merkt an, dass es bis heute keine schlüssigen Forschungsergebnisse gibt, die beweisen, dass »Bedeutungsaushandlungssequenzen« für den Erwerb der L2 so wichtig sind, wie dies postuliert wird. Zweitens sind die Analysekategorien, die verwendet werden, vage. Welche interaktionalen Phänomene genau zur Kategorie »Bedeutungsaushandlung« gehören, bleibt unklar. Die Liste von Bedeutungsaushandlungsstrategien, die etwa Oliver (2000, 375) bietet, beinhaltet Bitten um Klärung (»NNS: A little line in the leave. NS: A what?«), bestätigendes Nachfragen (»NNS1: Where does the um, glasses go? NNS2: The glasses?«), Nachfragen zwecks Verstehenssicherung (»NNS: You know what, you know?«), aber auch Selbst- und Fremdwiederholungen. Dies sind heterogene Phänomene, die in unterschiedlichen Interaktionskontexten ganz unterschiedliche Funktionen aufweisen können. Drittens besteht das Problem, dass dieser Ansatz präskriptiv definiert, was »guter« Fremdsprachenunterricht ist.

¹ Im Prinzip könnte man hier zwischen L2, L3 etc. differenzieren. Um die Lektüre nicht unnötig zu verkomplizieren, wird hier »L2« als generischer Terminus verwendet, der auch L3, L4 etc. benennt.

Dieser soll offensichtlich vor allem darin bestehen, Möglichkeiten zur Bedeutungsaushandlung zu schaffen. Wie dies genau geschehen soll, ist unklar. Aus kommunikationstheoretischer Perspektive lässt sich, viertens, monieren, dass der Kommunikationsbegriff unterentwickelt ist. Denn die Vorstellung, dass Bedeutungsaushandlung etwas ist, das nur ab und zu geschieht, kann nur vor der Folie einer Kommunikationstheorie entwickelt worden sein, gemäß welcher Kommunikation als Übermittlung von Information betrachtet wird. Denn nur dann kann die Idee auftreten, dass bei Problemen Bedeutung ausgehandelt werden muss. Kritisch merkt Wagner an:

It allows to rewrite interaction as information transfer, to isolate meaning transfer from meaning construction activities which are restricted to trouble solving, to code and quantify these episodes in the flow of interaction, and to calculate them against the baseline. The information transfer model surfaces again from below the negotiation metaphor. Actually, the negotiation metaphor fits nicely into the information transfer model, which is an exchange model – as is the ›market‹. When exchange processes are interrupted, negotiation and bargaining can begin! (Wagner 1996, 223)

Obwohl der Bedeutungsaushandlungs-Ansatz postuliert, Lernen als soziales, interaktives Ereignis zu verstehen, nimmt er die soziale Dimension der Interaktion nicht ernst genug. Geschieht Lernen in der Interaktion, dann ist Lernen oftmals nur ein Teil dessen, was geschieht. In Interaktionen werden soziale Beziehungen ausgehandelt, Hierarchien aufgebaut und wieder zerstört, auch in »learning related episodes«. Wenn etwa eine Schülerin eine andere fortwährend korrigiert, was aus der Bedeutungsaushandlungsperspektive positiv erscheinen muss, da so ja ständig Bedeutung ausgehandelt wird, dann ist dies unter Umständen für die Korrigierte und ihr Selbst nicht nur positiv. Wenn Lernen einer Sprache ein sozialer Prozess ist, dann ist dieses Lernen in die Teilnahme an der Organisation der Interaktion verwoben, ereignet sich in der Konstruktion der unterschiedlichen Partizipationsstrukturen (»participation frameworks«), ist verknüpft mit der Konfiguration der Diskursaufgaben und eingebettet in die situative und interktionale Definition der Identitäten (Mondada & Pekarek Doehler 2004, 504). Um diese Aspekte zu analysieren, wird heute oft auf die Konversationsanalyse zurückgegriffen (Schwab, Hoffmann & Schön 2017, 8).

4.2 Konversationsanalytisch orientierte Ansätze der Erforschung von L2-Interaktionen

Aufgrund der spezifischen theoretischen und methodischen Grundannahmen der Konversationsanalyse ist es nicht selbstverständlich, dass konversationsanalytisch orientierte Analyse von Interaktionen in L2-Klassenzimmern und anderen Kontexten durchgeführt werden, denn die CA setzte sich lange Zeit ausschließlich mit L1-Sprechenden auseinander. Ziel war die Analyse der Interaktionsordnung und der -methoden kompetenter Mitglieder einer Gemeinschaft (Firth 1996, 252). Als in den 1990er Jahren erste Versuche unternommen werden, Daten von Fremdsprachensprechenden konversationsanalytisch zu untersuchen, stellt etwa Wagner (1996) die Frage, inwiefern Konversationsanalyse dazu überhaupt in der Lage sei, denn CA setze die linguistische Kompetenz der Interagierenden voraus, und das von der CA typi-

schwerweise untersuchte Gespräch sei das monolinguale. Das monolinguale Gespräch unter L1-Sprechenden könne deswegen konversationsanalytisch analysiert werden, weil sie als Angehörige derselben Sprachgemeinschaften über das notwendige Mitgliederwissen verfügten, um die Interaktion überhaupt verstehen zu können. Dies sei bei Analysen von L2-Interaktionen nicht gegeben. Weiter behauptet er, dass sich L2-Interaktionen im kulturellen Niemandsland ereignen würden, wo »native speaker« und »non-native speaker« interagierten. Hier würde zwar oft ein formales, nicht aber ein substantielles Verstehen erreicht. Die Beteiligten würden so tun, als ob sie sich verstanden hätten, ohne sich *de facto* verstanden zu haben. Das könne konversationsanalytisch nicht analysiert werden, weil die CA aus methodologischen Gründen nur dann ein Kommunikationsproblem erkennen könne, wenn die Beteiligten selbst auf das Kommunikationsproblem referierten. Wagner schloss: »For the time being, it is unclear to me how data of this type might be analyzed. The methodological implications have to be carefully considered.« (Wagner 1996, 232)

Diese Kritik an der Brauchbarkeit der CA in Analysen von L2-Interaktionen adressierte etwa Seedhouse (1998) und skizzierte gleichzeitig ein methodologisches Programm für eine CA für L2-Interaktionen. Seedhouse wies Wagners Vorwurf zurück, dass die CA von einem monolingualen Verständnis ausgehe. Er verwies dazu auf die Grundannahme der Konversationsanalyse, dass zwischen Interaktion und Kontext eine reflexive Beziehung bestehe (Seedhouse 1998, 86) und Interagierende einander in der laufenden Interaktion anzeigen würden, an welchem Kontext, an welchen Interaktionsnormen, an welchen sozialen Ordnungen sie sich orientierten (siehe dazu auch Garfinkel & Wieder 1992, 186). Ordnung werde interaktiv konstituiert. Gleichzeitig orientierten sich die Beteiligten an dieser Ordnung und seien so in der Lage, Rechenschaft darüber abzugeben, was sie tun. Die konstituierte Ordnung sei, so würde Garfinkel schreiben, »reflexive and naturally accountable« (Garfinkel & Wieder 1992, 202; Psathas 1999, 55-65). Laut Seedhouse (1998, 98-100) sind auch L2-Interaktion in diesem Sinn reflexiv. Auch hier würden sich Interagierende Orientierungshinweise geben, um anzusehen, auf welche Kontexte, Rahmen, Normen, Ordnungen sie verweisen würden, etwa aufgrund der Sprachwahl. Diese Orientierungshinweise ließen sich deshalb mit Hilfe der CA analysieren, weil die CA keinen kulturellen Rahmen voraussetze, auch keinen monolingualen, sondern analysiere, wie diese Rahmen in der Interaktion konstituiert würden. Für Seedhouse ist klar: Das Interesse der CA besteht darin, die Orientierungen der Interagierenden und die Art und Weise der Organisation der Interaktion zu analysieren. Und selbst wenn ein Großteil der Analysen, welche die Basis des konversationsanalytischen Grundstocks an Wissen bilden, im angloamerikanischen Raum entstanden sei, heiße dies nicht, dass die CA und ihre Methodologie *per se* eine monolinguale und soziokulturelle Voreingenommenheit haben müsse. Denn selbst wenn die CA herausgefunden habe, dass es spezifische strukturelle Organisationen im Gespräch, dass es Gesprächsschritte, Präferenz und Reparatur im Gespräch gebe, heiße dies nicht, dass die Kategorien, welche sie bisher qua Analyse herausgearbeitet habe, ein Codierungsrauster bilden würden, das auf alle sprachlichen Interaktionen angewendet werden könne (Seedhouse 1998, 100).

Seedhouse berücksichtigt in seiner Argumentation die Ergebnisse der Bi- und Multilingualismusforschung, die schon seit den 1980er Jahren mit konversationsanalytischen Instrumenten arbeitet, zwar nicht; diese zeigte aber ähnliche Ergebnisse. In Abgrenzung von der Markiertheitstheorie des Codeswitchings entwickelten Peter

Auer, Li Wei, Joseph Gafaranga und andere einen Zugang zur Sprachalternation, in welchem der Sinns des Codeswitches nicht einfach vor der Folie bestehender sozialer Strukturen interpretiert wird, sondern als kontextuell qua sequentieller Position determiniert analysiert wird. Li Wei (1998, 162) schreibt: »[T]he meaning of codeswitching is conveyed as part of the interactive process and cannot be discussed without referring to the conversational context.« Das heißt nicht, dass der Rekurs auf die soziale Struktur in der Analyse keine Rolle spielen dürfe. Doch müsse dieser unter Verweis auf die konversationale Struktur begründet werden (Li Wei 1998, 162). Die Frage danach, welches die »base language« einer Interaktion oder, in Gafarangas Terminologie, welches das »medium of bilingual conversation« sei, sei damit nicht obsolet geworden. Sie müsse aber empirisch unter Berücksichtigung der Interaktion geklärt werden (Gafaranga 2007, 305; siehe auch Auer 1998; Liebscher & Dailey-O'Cain 2008).

Die in den späten 1990er Jahren etablierte konversationsanalytisch orientierte Forschung zu Interaktionen zwischen L2- und zwischen L2- und L1-Sprechenden deckte eine Reihe von Spezifika auf. Diskurspartikeln werden anders verwendet (Wong 2000b). Verzögerungen bei Turnübergängen treten auf, ohne dass dies durch die Struktur des Turns oder durch die Intonation angezeigt würde (Wong 2004). Reparaturen und Korrekturen in Interaktionen zwischen L1-Sprechenden und L2-Sprechenden werden von L1-Sprechenden zugunsten der Kommunikation vermieden (Wong 2005); Ähnliches zeigt auch Kurhila: Während Finnisch L2-Sprechende Wort- und Formsuchen anzeigen, werden sie von den Finnisch L1-Sprechenden mit Hilfe von »continuers« aufgefordert, fortzufahren. Wenn die L1-Sprechenden reparieren, dann tun sie dies, ohne die Reparatur explizit zu thematisieren, wie dies etwa im Klassenzimmer der Fall wäre. Kurhila hat dafür eine plausible Erklärung:

[...] by only minimally orienting to the form of the NNSs' utterances, the NSs avoid ›doing being second language teachers‹, and maintain the relevance of their role as a secretary or a friend. (Kurhila 2005, 156)

Mit Goffman würde man sagen, dass hier *face* beachtet und spezifische soziale Marker ins Zentrum gestellt werden, um das eigene Selbst und das Selbst des anderen nicht zu bedrohen.

Ein jüngerer Zweig der konversationsanalytischen L2-Forschung setzt sich mit Frage auseinander, wie sich die interktionale Kompetenz von L2-Lernenden über einen kürzeren oder einen längeren Zeitraum hinweg entwickelt (siehe auch Cekaite 2007; Hellermann 2008, 6-14), und in solchen Arbeiten trifft sich die Konversationsanalyse auch mit der *language socialization theory*, die untersucht, wie Sprachen im Kontext erworben werden. Zwar monieren traditionell ausgerichtete Konversationsanalytikerinnen und -analytiker, dass in der Sprachsozialisationsforschung die Interaktionen nicht mehr aus eigenem Recht analysiert würden und die Gefahr bestehe, dass die Perspektive der Beteiligten hinter der Forscherperspektive verschwinde, die sich primär für die Entwicklung der Lernenden interessiere (siehe dazu auch Lerner, Zimmerman & Kidwell 2011). He erwidert darauf: »CA can provide highly detailed insights into how opportunities for language learning and teaching are socially constructed by participants who observably orient to the practices of institutional varieties of talk.« (He 2004, 573) Dies zeigen die Untersuchungen von Gajo und Mondada (2000) und Mondada und Pekarek Doehler (2004). Sie analysieren, wie Lernen als soziales,

interaktionales Handeln in institutionellen Kontexten organisiert ist und zeigen, dass das Durchführen einer gemeinhin nicht als sonderlich kommunikativ verstandenen Grammatikübung eine soziale und interaktive Aufgabe ist, welche ein bestimmtes Format aufweisen muss, damit sie als akzeptabel oder legitim gilt. Ihre Analysen lassen erkennen, welche Kompetenzen ein Schüler, eine Schülerin haben muss, um der Kategorie der »guten Schülerinnen und Schüler« zugerechnet zu werden. Die richtige Antwort zu geben, reicht nämlich nicht. Wichtig sind, was ebenfalls schon Mehan (1979) thematisierte, sozioinstitutionelle Kompetenzen. Mondada und Gajo (2000) und Mondada und Pekarek Doepler (2004) weisen damit konversationsanalytisch nach, wie eng Lernen, Institution und Identität verknüpft sind.

Die konversationsanalytisch orientierte Forschung zu Diskursen im Klassenzimmer bietet eine Fülle von Ergebnissen, die berücksichtigt werden können, um Klassenzimmerinteraktionen auch aus einer Goffman'schen Perspektive zu analysieren und zu interpretieren. Seedhouse (2004) unterscheidet vier verschiedene Kontexte des Klassenzimmers, die durch den pädagogischen Fokus der Lehrperson initiiert und in und durch Interaktion konstituiert werden. Dass sich im Klassenzimmer aber nicht immer alle Interaktionen voll und ganz am pädagogischen Fokus ausrichten, verdeutlicht Markee (2005) in seiner Analyse der Struktur einer »off-task«-Interaktion (Markee 2005, 210-211). Weitere Themen, die untersucht wurden, sind Reparaturen im Klassenzimmer (Kasper 1985; Liebscher & Dailey O'Cain 2005; Hall 2007; Kääntä 2010), die Organisation der Teilnahme am Unterricht (Cekaite 2008; Kääntä 2010; Schwab & Hoffmann 2017), der kommunikative und interaktive Einsatz der teilweise begrenzten Ressourcen (Mori 2004), aber auch Moralität im Unterricht (Mortensen & Hazel 2017). Der auch von Hellerman und Pekarek Doepler (2010) thematisierte Zusammenhang zwischen »task-as-workplan« und »task-as-process« rückt ebenfalls zunehmend in den Fokus der konversationsanalytischen L2-Forschung (Pochon-Berger 2011). Genauso wichtig ist die Erforschung von »learnables«, deren Struktur und Bearbeitung geworden (Majlesi & Broth 2012; Majlesi 2014; Pelchat 2017). Diese Einzelanalysen vergrößern das Wissen um Strukturen sozialen Handelns in L2-Klassenzimmern und liefern damit auch eine wichtige empirische Basis für die Weiterentwicklung pädagogischen und didaktischen Handelns.

Viele der Untersuchungen stammen aus Universitätskontexten und sind in immersiv orientierten Klassenzimmern erhoben, in welchen die L1 eine untergeordnete Rolle spielt (Kasper 2004a; Kääntä 2010; Markee 2005; Seedhouse 2004, 2005, 2006). Wie L2-Interaktionen auf der Primarstufe und in nicht-immersiven Kontexten strukturiert sind, ist bisher noch wenig erforscht. Gleichzeitig besteht innerhalb der L2-Forschung die Tendenz, das Forschungsfeld zu beschränken. So schreiben Wagner und Kasper: »Language learning as a social activity, refers to interactional moments where the participants make learning the focal concern of their interaction.« (Kasper & Wagner 2011, 127) Würde man nur diese interaktionalen Momente analysieren wollen, in welchen die Beteiligten sich gegenseitig zeigen, dass sie am Lernen sind, dann müsste man alle Interaktionen, in welche diese eingebettet sind, aus der Untersuchung ausschließen. Die Ergebnisse würden dadurch sicherlich nicht reichhaltiger, und das Verständnis dafür, wie L2-Interaktionen in institutionellen Kontexten organisiert sind, auch nicht. Ein umfassenderes Wissen darum, wie das L2-Klassenzimmer als soziale Arena strukturiert ist, ist wichtig, um zu verstehen, wie Lernen sozial organisiert ist.

Dies hat die Sprachsozialisierungsforschung hinlänglich bewiesen (siehe dazu umfassend Duff 2012).

4.3 Mitgliedschaftskategorisierungsanalyse und Goffman

Goffman zeigt in vielen seiner Texte, wie verschiedene Facetten des Selbst in der Interaktion konstruiert werden. Er dekonstruiert Sprechende und Hörende in seinen Analysen zum Redestatus und zeigt, wie vielfältig die Teilnahmeformate sind, die in Interaktionen eingenommen werden (Levinson 1988; Zimmerman 1998). Goffman geht von der relativen Autonomie der Interaktionsordnung und einer losen Koppelung zwischen Sozialstruktur und Interaktionsordnung aus. Diese Koppelung gelte es, genau zu analysieren (Goffman 1983, 11), und dies tut Goffman auch. Er zeichnet nach, wie Menschen sozial unerwünschte Attribute in der Interaktion zu verstecken versuchen und Informationen über sich selbst managen, um bestimmten »Identitätsnormen« zu entsprechen (Goffman 1975, 161). Er zeigt, wie sich soziale Marker auf die Interaktion auswirken können. Er illustriert, wie Alter (etwa Goffman 1973, 118ff.), Geschlecht (insbesondere Goffman 2001), Hautfarbe (etwa Goffman 1982, 285, 2005, 145), sexuelle Orientierung (etwa Goffman 1982, 93) in verbaler und nonverbaler Interaktion sichtbar gemacht werden. Er unterstreicht, dass geteiltes Wissen über die Möglichkeiten der Kategorisierung von Menschen und den damit verknüpften Rechten und Pflichten eine der Bedingungen der Möglichkeit der Herstellung von Verstehen ist (Goffman 1983, 12). »Wissen über die soziale Struktur« kann interaktional relevant werden und muss notwendigerweise in die Interaktion »importiert« werden, damit Interaktion funktionieren kann (Goffman 2005, 222-223). Goffman zeigt aber auch, dass sich Menschen den Kategorisierungen und den Rollen, die ihnen situativ zugeteilt werden, entziehen können (Goffman 1973).

Goffman setzte sich unter anderem in *Rahmen-Analyse, Forms of Talk* und seinem späten Aufsatz zu *Glückensbedingungen* mit sozialem Wissen auseinander, das in und für Interaktionen relevant wird. Sein Schüler Sacks setzte sich in seiner Dissertation ebenfalls mit sozialem Wissen und der Frage, wie dies in Interaktionen relevant und damit rekonstruiert wird, auseinander; er betitelte seinen Ansatz als *membership categorization analysis* (siehe etwa Sacks 1965). Goffman lehnte laut Schegloff Sacks Argumentation der Begründung der Kategorien zwar in den 1970er Jahren als zirkulär und das Vorgehen als zu formalistisch ab (Schegloff 1992, xiv, note 18; Smith 2003, 255-256). In *Glückensbedingungen* weist Goffman aber explizit auf Sacks' Mitgliedschaftskategorisierungsanalyse hin und nennt sie eine Methode, mit welcher die Funktion identifikatorischer Ausdrücke in Interaktionen untersucht werden könne (Goffman 2005, 223). Sacks' Methode der Analyse sozialer Kategorien, der Implikationen, die damit verbunden sind, sowie deren Konsequenzen für die Interaktion kann, wie Tilbury (2007, 629) schreibt, die Instrumente liefern, um die Artikulation dieser Kategorien in Interaktionen zu analysieren, um zu zeigen, wie Kategorien zusammenhängen und um aufzudecken, mit welchen Normen sie belegt sind.

Vor seiner Hinwendung zu eher strukturellen Aspekten von Konversationen fokussierte Sacks die Kategorien, welche Menschen im Alltag verwenden, um sich und ihre Mitmenschen zu beschreiben. Dies ist interessant, weil sich ein jeder Mensch mit einer Reihe unterschiedlicher Kategorien beschreiben lässt. Jemand kann ein Mensch, ein Lehrer, ein Hockeyspieler, ein Tierschützer, ein Ehemann und noch vieles mehr

sein. Doch nicht alle Kategorien sind immer relevant, und in den meisten Fällen reicht es vollkommen aus, eine einzige Kategorie zu verwenden. Interessant ist es nun zu fragen, welche Kategorien in Interaktionen relevant gesetzt werden, wie, wann und warum diese Kategorien relevant gesetzt werden. Die Analyse der verwendeten Kategorien, der Beziehungen zwischen diesen Kategorien und der Schlussfolgerungen, die Menschen im Alltag aus diesen Kategorien ableiten, deckt auf, wie alltägliche Sinnkonstitution und alltägliches Denken funktionieren (Butler 2008, 28-29) und wie die aktualisierten Kategorien das Handeln beeinflussen. Sacks schreibt:

[...] a great deal of the knowledge that members of a society have about the society is stored in terms of these categories. And by 'stored in terms of' I mean that much knowledge has some category term from this class as its subject. And the inference-rich character of these categories constitutes another warrant for their occurrence in early parts of first conversations: When you get some category as an answer to a 'which'-type question, you can feel that you know a great deal about the person, and can readily formulate topics of conversation based on the knowledge stored in terms of that category. (Sacks 1992a, 40-41)

Sacks nimmt an, dass Kategorien in sogenannten »membership categorization devices« oder kurz »categorization devices«, von Eberle als »Mitgliedschaftskategorisierungen« übersetzt (Eberle 1997, 251; Schegloff 2007b, 467), strukturiert sind. Dies sind Kollektionen von Kategorien, die als zusammengehörig betrachtet werden (»categories members do use together«) und Anwendungsregeln, die festlegen, wann jemand zu einer bestimmten Kategorie gehört (Sacks 1992a, 238; 1972b, 332). Wenn wir die von Sacks mehrmals analysierte Mini-Kindergeschichte »The baby cried. The mommy picked it up« (Sacks 1992a, 236-266; 1972b) hören, gehen wir davon aus, dass Baby und Mutter zusammengehören, obwohl dies nicht der Fall sein muss, weil wir in der Interpretation auf kategoriales Wissen und Mitgliedschaftskategorisierungen zurückgreifen: Ein Baby schreit, eine Mutter kümmert sich um ihr schreiendes Kind. »Mutter« und »Baby« gehören zur Kategorienkollektion »Familie«. Die Kollektion »Familie« besteht nicht nur aus diesen Kategorien, sondern beinhaltet auch die Kategorien »Vater«, »Großvater«, »Großmutter« etc. Andere wichtige Kategorienkollektionen sind etwa »Geschlecht«, »Konfession« oder »Nationalität«. In unterschiedlichen Kulturen werden Kategorisierungen unterschiedlich strukturiert. Die Kategorisierung »Geschlecht« umfasste in westeuropäischen Kulturen lange Zeit »männlich« und »weiblich«, seit Kurzem in Deutschland auch die Kategorie »divers«. Die Analyse der Art und Weise, wie Interagierende unterschiedliche Kategorien zu Mitgliedschaftskategorisierungen zusammenfassen, eröffnet einen Blick auf die kulturelle Basis, auf welcher Menschen handeln (Sacks 1992a, 237).

Menschen orientieren sich im Alltag an Kategorien und den damit verbundenen Normen, sie ziehen Schlussfolgerungen basierend auf Mitgliedschaft zu einer Kategorie (Butler 2008, 28-29). Laut Sacks ist eine große Menge an Alltagswissen, das auch stereotype Züge trägt, mit den Kategorien verbunden. Dieses wird etwa dann sichtbar, wenn Menschen versuchen, das mit einer Kategorie verbundene stereotype Wissen von sich selbst abzuwenden: »Ich bin 48, aber ich sehe viel jünger aus« (Sacks 1992a, 44). Gleichzeitig ist es laut Sacks aber schwierig, dieses Wissen zu verändern. Denn die Modifikationen, die eingeführt würden, seien prekär. Behauptete jemand: »He is a

Negro, but the things you can say about Negros you can't say about him«, dann bestünde immer die Möglichkeit, dass dieser Ausnahmestatus annulliert würde, und dass am Ende gesagt würde: »Well, now it really comes out ...« (Sacks 1992a, 45). Kategorien sind gemäß Sacks mit Wissen über »kategorienbundene Aktivitäten« verbunden, die üblicherweise von Mitgliedern einer bestimmten Kategorie ausgeübt werden. Die Verbindung zwischen kategorienbundener Aktivität und Mitgliedern einer Kategorie ist nicht eine rein empirische, sondern kann auch deontische Aspekte aufweisen: Wer sich nicht so verhalte, wie dies von einem Mitglied einer bestimmten Kategorie erwartet werde, müsse mindestens mit Verunglimpfung rechnen (Sacks 1992a, 241). Das schreiende Kind auf den Arm zu nehmen ist also nicht einfach etwas, was die Mutter tut, sondern etwas, was sie tun soll.

Mitgliedschaftskategorisierungen funktionieren nach Ökonomie- und Konsistenzregeln. Die Ökonomieregel besagt, dass die Kategorisierung von jemandem als Mitglied einer einzigen Kategorie oftmals völlig ausreichend ist. Stellt eine Lehrperson einer außenstehenden Person ihr Kollegium vor, dann kann es völlig ausreichen, wenn sie auf die Funktionen ihrer Kolleginnen und Kollegen in der Schule verweist: Das ist die Mathelehrerin, dies der Werklehrer, dieses die Französischlehrerin. Weitere Angaben, etwa darüber, welchen Sport die betreffenden Personen betreiben oder was sie gerne essen, ist unerheblich (dazu Sacks 1972b, 333). Die Konsistenzregel besagt, wenn man jemanden als Mitglied einer bestimmten Kategorie kategorisiert, dann wird man diese Kategorie oder eine andere Kategorie derselben Kategorisierungen verwenden, um andere Personen im selben Kontext zu kategorisieren (Sacks 1972b, 333; Sacks 1992a, 313). Sacks leitet daraus eine Maxime ab, die Folgendes besagt: »If two or more categories are used to categorize two or more members of some population, and those categories can be heard as categories from the same collection, then: Hear them that way.« (Sacks 1972b, 333)

Zwischen Kategorisierungen bestehen Relationen: Wer männlich ist, kann zum Ehemann, aber nicht zur Ehefrau werden, es sei denn, er wechselt die Kategorie des Geschlechts (dazu auch Garfinkel 1967). Einige Kategorisierungen können in einem bestimmten Setting »omni-relevant« sein:

An »omni-relevant device« is one that is relevant to a setting via the fact that there are some activities that are known to get done in that setting, that have no special slot in it, i.e., do not follow any given last occurrence, but when they are appropriate, they have priority. (Sacks 1992a, 313-314)

Solche Kategorisierungen können in einem spezifischen Setting jederzeit relevant werden, etwa durch eine spezifische Form der Strukturierung der Interaktion, welche eine Orientierung an dieser Kategorisierung sichtbar macht. Sacks nennt in diesem Zusammenhang explizit auch die Kategorisierung »Lehrperson/Schüler-/in« und die mit den jeweiligen Mitgliedschaften verbundenen Rechte und Pflichten hinsichtlich der Art und Weise, wer die Interaktion für eine Pause unterbrechen kann und darf (Sacks 1992a, 314-319). Kategorisierungen, auch omni-relevante, können Menschengruppen in spezifischer Weise partitionieren, etwa als Lehrkräfte und Schülerinnen und Schüler, als Therapeutinnen und Patienten, als Zuschauer und Schauspielerinnen. Zwischen diesen Partitionierungen bestehen wiederum Relationen. Sacks führt das Beispiel der Therapiesitzung an, die er hinter einem Einwegspiegel mitverfolgt. Hier

werden unterschiedliche Kategorisierungen relevant: »Zuschauerin/Handelnder auf einer Bühne« einerseits, aber auch »Patientin/Beobachter«. Die Szene lässt sich aus unterschiedlichen Perspektiven beschreiben. Relevant für Sacks ist hier, dass die Kategorienzuteilung hier konsistent zwischen Patientin/Beobachter einerseits und Schauspieler/Zuschauerin andererseits verläuft. Sacks nennt dies »partitioning constancy« (Sacks 1992a, 317; Schegloff 2007b, 468).

Sacks nennt eine Reihe weiterer Eigenschaften, welche Kategorisierungen aufweisen können. Sie können etwa PN-adäquat sein, das heißt, die Kategorie N ist in jeder Population so anwendbar, dass jedes Mitglied der Population P mit N kategorisiert werden kann; dies trifft etwa auf die Kategorienkollektionen »Alter« zu (Sacks 1972a, 32-33; Schegloff 2007b). Dabei müssen, wie Butler verdeutlicht, PN-adäquate Kategorisierungen keineswegs uneingeschränkt auf jede Population anwendbar sein, sondern können sich auch nur auf alle anwesenden und relevanten Mitglieder beziehen (Sacks 1972a, 38-39; Butler 2008, 36). Manche Kategorienkollektionen treten paarweise auf, so etwa Ehemann-Ehefrau, Eltern-Kind, Lehrperson-Schülerin oder Schüler. Die Relationen zwischen Mitgliedern dieser Kategorien zeichnen sich durch spezifische Rechte und Pflichten aus (Sacks 1972a).

Manche Kategorisierungen sind wie Teams organisiert und können Restriktionen hinsichtlich ihrer Mitglieder aufweisen. In einer Quidditchmannschaft auf dem Feld gibt es drei Jägerinnen, einen Hüter, zwei Treiberinnen und einen Sucher. Diese Kategorisierungen werden als Einheit verstanden, und sie sind »duplicatively organized« (Sacks 1972b, 334). Bei der Anwendung duplikativ organisierter Kategorien spielt Konsistenz eine wichtige Rolle. Wenn wir eine solche Kategorisierung anwenden, dann betrachten wir Mitglieder, die in eine der Kategorien dieser Kategorisierungen gehören, als Teil der Einheit, die durch die Kategorisierung beschrieben wird. Wenden wir duplikativ organisierte Kategorisierungen an, dann führt dies auch dazu, dass wir Mitglieder einer Kategorie, die Teil einer teamartig organisierten Kategorisierung ist, als Mitglied eines Teams verstehen: Höre ich jemanden über »die Mutter, den Vater und die vier Kinder« sprechen, dann werde ich davon ausgehen, dass alle, die so beschrieben werden, zu einer einzigen Familie gehören. Dies ist eine weitere Sack'schen Maximen (Sacks 1972b, 334).

Weitere Regeln beziehen sich auf die Art und Weise, wie wir Ereignisse wahrnehmen und daraus Mitgliedschaften zu bestimmten Kategorien ableiten, also mögliche korrekte Beschreibungen für andere. Dies sind die sogenannten Zuschauermaximen oder Relevanzregeln (Butler 2008, 32-33). Die erste besagt: »If a Member sees a category-bound activity being done, then, if one can see it being done by a member of a category to which the activity is bound, see it that way. (Sacks 1992a, 259)« Die zweite Maxime bezieht explizit den normativen Aspekt kategorienbundenen Handelns mit ein:

If one sees a pair of actions which can be related via the operation of a norm that provides for the second given the first, where the doers can be seen as members of the categories the norm provides as proper for that pair of actions, then (a) see that the doers are such Members [sic!], and (b) see the second as done in conformity with the norm. (Sacks 1992a, 260)

Sacks' Darstellungen der Kategorien und deren Anwendungsregeln ist nicht immer ganz eindeutig zu entnehmen, welchen ontologischen Status er ihnen zuschreibt

(Hester & Eglin 1997a). Teilweise scheint es der Fall zu sein, als ob er Kategorisierungen als generative Grammatiken konzipiert, die wie kulturgebundene Transformationsgrammatiken funktionieren (Sacks 1992a, 237). Housley und Fitzgerald betonen, dass man eine dekontextualisierte Sichtweise auf Kategorisierungen zugunsten einer solchen aufgeben müsse, die den lokalen Gebrauch der Kategorien durch die Beteiligten *in situ* analysiere. Kategorisierungen und die Verwendung von Kategorien kämen dann als »occasioned phenomenon« in den Blick, als Kategorien »im Kontext« (Housley & Fitzgerald 2002, 68).

Doch wie kann man in der Analyse zeigen, welche Kategorien für Interagierende relevant sind, ohne dass sie selbst auf diese Kategorie verweisen, etwa wenn sie sagen »Ich als Mann sehe das so« (Schegloff 2007b, 477)? Hester und Eglin (1997a), Watson (1997), Housley und Fitzgerald (2002) und Butler (Butler 2008) gehen davon aus, dass Sequenz- und Kategorisierungsanalyse eng miteinander verbunden werden müssen, da sequentielle und kategoriale Aspekte sich in sozialen Interaktionen gegenseitig durchdringen. Denn die Interagierenden orientieren sich in der Interaktion logischerweise nicht nur an der sequentiellen Struktur. Ihre Handlungen richten sich gleichzeitig auch an Mitgliedschaftskategorien aus. Das heißt, die Orientierung an relevanten Kategorien kann die Struktur der Äußerungen beeinflussen, und die Struktur wiederum kann gleichzeitig die Kategorien, die gegenwärtig relevant sind, indizieren. Dies ist sowohl aus einer konversationsanalytischen Perspektive, welche die Wichtigkeit des *recipient designs* betont, wie auch aus einer stärker an Goffman orientierten Perspektive klar. Denn mit Goffman könnte man etwa sagen, dass sich die Partizipationsstruktur und damit die gegenwärtige Rolle, die jemand in der Interaktion einnimmt, in der Struktur der Äußerungen manifestieren kann (siehe etwa Goffman 1981, 140). Nähe und Distanz, über welche Zugehörigkeit ausgehandelt wird, manifestieren sich ebenfalls anhand der sprachlichen Ressourcen, die verwendet werden (siehe etwa Goffman 1967). Gender wird, wie Goffman zeigte, ebenfalls in und durch Interaktionen inszeniert (Goffman 2001). Watson nennt dies mit Cicourel einen »folding back effect« (Watson 1997, 54).

Interagierende kategorisieren sich in der Interaktion oftmals gegenseitig direkt. Doch die Mitgliedschaft in einer Kategorie wird auch auf andere Weise relevant gesetzt, etwa indem die Interagierenden etwas von jemandem prädizieren. Die Prädikate, die verwendet werden, können mit bestimmten Rollen und damit mit Mitgliedschaften in bestimmten Kategorien verknüpft sein (Butler 2008, 99-106).

4.4 Interaktionen unter Präadoleszenten und ihre Analyse

Kinder werden zwar in die Gesellschaft sozialisiert, leben aber in ihren je eigenen Welten mit je eigenen Normen. »Kinder« wurden zwar schon lange, wenigstens peripher, thematisiert, dennoch waren sie lange Zeit zumindest in der Soziologie kein Forschungsgebiet *sui generis*. Kamen sie in den Blick, dann oft nur in Abhängigkeit von ihren Eltern (Schweizer 2007, 72, 168). Das Gleiche gilt grundsätzlich für die Anthropologie. Denn obwohl schon Malinowski (1923) und Mead (1933) dafür plädiert hatten, das soziale Leben der Kinder zu untersuchen, haben Anthropologen erst in den letzten Jahrzehnten begonnen, die Welt der Kinder aus eigenem Recht zu analysieren (Goodwin 2008, 21-24). Das heißt nicht, dass Kinder nicht erforscht wurden. Doch war im zwanzigsten Jahrhundert eine, wie Mackay (1974) sie nennt, normative Pers-

pektive vorherrschend, welche vor allem an der Entwicklung des Kindes interessiert war und welche dem Kind den Erwachsenen als Ziel voraus- und als Norm entgegensezte. Mackay führt diese Perspektive auf Parsons und Bales zurück (1955), die Kinder als unvollständig, das heißt unreif, irrational, inkompetent, asozial und akulturell betrachteten (Mackay 1974, 28). Kinder wurden aus dieser Perspektive vor allem als »Sozialisationsobjekte« konzipiert, die passiv gesellschaftliche Werte und Normen internalisieren (Schweizer 2007, 52). Dass aus dieser Perspektive auch die Sozialwelt der Kinder und die ihr möglicherweise inhärente »Kinderkultur« nicht interessierten, liegt auf der Hand.

Nicht nur die alltäglichen, sondern auch die wissenschaftlichen Vorstellungen davon, was ein »Kind« ist, haben sich in den letzten Jahrzehnten massiv geändert. Zwar geht wohl heute noch die Mehrheit der Pädagoginnen und Pädagogen und Eltern davon aus, dass Kinder zu erziehen seien (Hutchby & Moran-Ellis 1998, 9). Doch die Vorstellung, dass Kinder nur »Sozialisationsobjekte«, unschuldig und inkompetent seien, ist in dieser Deutlichkeit nicht mehr anzutreffen (James & Prout 1997b, xiv).² Die sozialwissenschaftliche Kindheitsforschung begreift Kinder heute vielmehr als kompetente Akteurinnen und Akteure oder Konstrukteurinnen und Konstrukteure, die sich die Welt erschließen, die verhandeln und interagieren, und die unter sich, gemeinsam mit oder auch gegen die Erwachsenen, ihre eigenen sozialen Welten konstruieren. Das »Wechselspiel von Aktivität, Passivität und Rezeptivität [...] rückt in den Mittelpunkt eines sozialkonstruktivistischen Verständnisses von Kindheit« (Schweizer 2007a, 54).

Kindheit wird verstanden als Schauplatz sozialer Aktivitäten und Handlungen, wo Macht, Bedeutung und Beziehungen ausgehandelt werden, und Kindheit wird nicht mehr einfach als transitorische Phase aufgefasst, in welcher Kinder sich zu entwickeln haben, sondern Kinder kommen als kompetent Handelnde in den Blick (Hutchby & Moran-Ellis 1998, 9). James und Prout (1997), die zur Etablierung des »Kompetenzparadigmas« maßgeblich beitrugen, formulieren die Eckpunkte dieses Paradigmas folgendermaßen: Erstens wird Kindheit als soziale Konstruktion aufgefasst, die als interpretativer Rahmen dazu dient, die frühen Jahre des menschlichen Lebens zu kontextualisieren. Zweitens wird Kindheit nur als eine soziale Variable betrachtet, die mit weiteren Variablen wie Gender, Klasse oder Ethnizität interagiert. Drittens werden die sozialen Relationen und die Kulturen der Kinder als Forschungsbereiche definiert, die es aus eigenem Recht zu untersuchen gilt. Viertens müssen Kinder als Konstrukteure ihrer sozialen Welt in den Blick genommen werden. Sie sind nicht bloß passive Subjekte sozialer Strukturen und Prozesse (James & Prout 1997a, 8).

Mit der Hinwendung zur Analyse der Sozialwelt der Kinder aus eigenem Recht ist die methodologische Frage verbunden, wie es denn möglich sein kann, als Forschender an der Welt der Kinder teilzuhaben oder gar die Perspektive der Kinder einzunehmen. Die Antwort, die in der Soziologie der Kindheit vertreten wird, ist, dass dies mit teilnehmender Beobachtung, Interviews und mit Hilfe der Analyse von Zeugnissen, die Kinder selbst über ihr Leben ablegen, möglich sei (James & Prout 1997b). Hutchby führt überzeugend aus, weshalb dieses Vorgehen problematisch ist. Auch diejenigen, die nicht die Rolle unbeteiligter Beobachterinnen und Beobachter einnehmen und so signalisieren, dass zwischen den Kindern und den Forschenden ein sozialer und in-

² Schweizer schreibt, dass die strukturfunktionalistische Sozialisationstheorie »auch heute noch – oft undurchschaubar – mindestens unterschwellig nachwirkt« (Schweizer 2007, 157).

telektueller Graben besteht, sondern versuchen, Mitglieder der Welt der Kinder zu werden, werden die Welt der Kinder perspektiviert wahrnehmen. Als Teilnehmende sind sie zudem selbst daran beteiligt, diese soziale Welt zu konstruieren, gerade auch im Prozess der Datenerhebung. Das heißt nicht, dass solchermaßen erhobene ethnographische Daten wertlos sind. Wie Hutchby schreibt, besteht aber auch hier die Möglichkeit, die Perspektiven der Beteiligten anhand der Analyse ihrer Interaktionen zu rekonstruieren, um so Einblicke in die soziale Organisation der Welt der Kinder zu erhalten, zu sehen, welchen Sinn Kinder mit ihren verbalen und nonverbalen sozialen Handlungen verbinden, welchen Sinn sie den Handlungen anderer, auch den Handlungen Erwachsener, attribuieren, wie sie ihre Welt strukturieren und wie sie in und durch Interaktion zu kompetenten Mitgliedern sozialer Gruppen werden (Hutchby & Moran-Ellis 1998, 10-11; Hutchby 2007, 8-9).

Es gibt heute, grob gesprochen, zwei Ausrichtungen, die sich aus interaktional-emischer Perspektive der sozialen Welt der Kinder nähern, eine eher linguistisch-anthropologische und eine eher entwicklungsorientierte; die Grenzen sind indessen fließend. Marjorie Goodwin (1990) hat aus linguistisch-anthropologischer Perspektive Kulturen der Kinder aus eigenem Recht schon seit den 1980er Jahren erforscht, und zwar auch unter Verwendung von Instrumenten der Konversationsanalyse und auch im Anschluss an Goffman. Sie zeigt, wie afroamerikanische Kinder in einem Vorort von Philadelphia ihre Sozialwelt in Interaktionen organisieren und strukturieren, wie Kinder Spielregeln konstituieren und Teilnahme an Spielen regulieren, sie untersucht, wie Kinder klatschen und miteinander Konflikte aushandeln, und demonstriert, welche komplexen Regeln den kindlichen Sprachspielen, wie sie sie im Anschluss an Wittgenstein nennt, zugrunde liegen (Goodwin 1990). In späteren Untersuchungen fokussiert sie nicht mehr nur die sprachliche Interaktion. Sie untersucht die multimodale Organisation von Himmel-und-Hölle-Spielen unter Kindern der US-amerikanischen Latino-Bevölkerung (Goodwin 2001), analysiert Positionierungen und Ausschlussmechanismen unter Mädchen in einer Grundschule in Los Angeles, Interaktionen zwischen Mädchen und Jungen beim Seilspringen und bei Klatschsequenzen unter Kindern (Goodwin 2008).

Weitere Studien, die mit interaktionsanalytischem Instrumentarium unterschiedliche Aspekte verschiedener Kinderkulturen untersuchen, befassen sich mit der Aushandlung von Konflikten unter Kindern (Cromdal 2004a), der Konstruktion von Hierarchien und Machtstrukturen (Danby & Baker 2000; Griswold 2007), und zwar auch in Rollenspielen und durch Rollenspiele (Kyratzis 2007). Andere fokussieren die soziale Organisation von Spielen (Butler 2008; Björk-Willen & Cromdal 2009; schon Sacks 1966), die moralischen Ordnungen, die Kinder etablieren (Evaldsson 2007) oder, noch einmal, wie Kinder klatschen (Tholander 2003). Andere Untersuchungen, die ebenfalls mit konversationsanalytischen Instrumenten arbeiten, fokussieren, was in Erwachsenen-Kinder-Interaktionen geschieht, und zwar sowohl in familiären (Corrin 2009) wie auch in institutionellen Settings innerhalb des Klassenzimmers (Cekaite 2008), außerhalb des Klassenzimmers, so etwa in Eltern-Kinder-Lehrer-Kontexten (Silverman, Baker & Keogh 1998), in der Kindertagesstätte (Lerner, Zimmerman & Kidwell 2011), aber auch in Fernsehshows, in welchen Kinder zu aktiven Beteiligten werden (Thornborrow 1998), oder in Beratungssitzungen (Hutchby 2007).

Der prominenteste Ansatz, der sich vor allem aus entwicklungsorientierter Perspektive mit Interaktionen unter Kindern befasst, ist der Sprachsozialisations-

ansatz von Schieffelin und Ochs. Dieser Ansatz untersucht, wie Sozialisation im und durch den Sprachgebrauch vonstatten geht, und wie Kinder in den Sprachgebrauch hineinsozialisiert werden: »socialization through the use of language and socialization to use language« (Schieffelin & Ochs 1986, 163). Er bedient sich soziologischer, anthropologischer und psychologischer Ansätze, um die sozialen und die sprachlichen Kompetenzen und deren Entwicklungen innerhalb spezifischer sozialer Gruppen zu untersuchen. *Language socialization* geht davon aus, dass Sprache und Interaktion zentrale Instanzen in der Entwicklung der kulturellen und sozialen Kompetenzen von Kindern sind. Anders als Ansätze, die sich ausschließlich mit der Entwicklung konversationeller Fähigkeiten auseinandersetzen (Filipi 2009; Wootton 1997) und die Gefahr laufen, Interaktionen als Vehikel mentaler Prozesse zu verstehen und dabei auch die Perspektive der Beteiligten aus den Augen verlieren (Lerner, Zimmerman & Kidwell (2011, 45), will die Sprachsozialisierung analysieren, wie Kinder und andere Novizinnen und Novizen die Kontexte, in denen sie handeln, wahrnehmen und aktiv an deren Konstruktion mitarbeiten. Sie schreiben:

In so doing, the study of language socialization research integrates discourse and ethnographic methods to capture the social structuring and cultural interpretations of semiotic forms, practices, and ideologies that inform novices' practical engagements with one another. (Ochs & Schieffelin 2012, 1)

Auch wenn im Paradigma der *language socialization* die Entwicklungsperspektive zentral ist, so geht sie keineswegs davon aus, dass Novizinnen und Novizen, die sich in einer neuen Situation befinden, zuerst bestimmte Kompetenzen erwerben müssten, um überhaupt an der Interaktion teilhaben und teilnehmen zu können. Kinder und Novizinnen und Novizen werden als aktiv Handelnde betrachtet (Ochs & Schieffelin 1994, 504). Dabei wird Kompetenz nicht als kognitive, sondern als interktionale und soziale Kompetenz aufgefasst, die auch Emotionen umfassen kann. Dies zeigt sich paradigmatisch in den Arbeiten von Cekaite (2008, 2012). Cekaite untersucht in ihren sprachsozialisationsorientierten Untersuchungen, wie sich Kinder mit Zuwanderungsgeschichte in ihrer neuen schulischen Umgebung sozialisieren. Sie analysiert diese Sozialisierung gerade nicht als kognitive Entwicklung, sondern rekonstruiert, wie eine Schülerin in und durch situierte Praktiken im Klassenzimmer einen bestimmten Platz erwirbt (Cekaite 2012). Die Ergebnisse dieser Untersuchungen zeigen nicht nur, wie Kinder ihre linguistischen und interaktionalen Ressourcen in spezifischen Kontexten einsetzen, wie sie auf kulturelle Ressourcen der Erwachsenenwelt zurückgreifen und diese für ganz spezifische Zwecke adaptieren, welche Kategorisierungen in ihren Welten eine Rolle spielen, sondern auch, wie sich die Kinder gegenseitig sozialisieren und damit ihre sozialen Welten strukturieren und mit Werten, Normen, Hierarchien versehen und dadurch auch Unterschiede schaffen zwischen denjenigen, die dazugehören, und denjenigen, die nicht dazugehören.

4.5 Die Multimodalität der Interaktion

Goffman verweist immer wieder darauf, dass Interaktionen, auch verbale, in einem physikalischen Umfeld stattfinden und von einer Reihe weiterer nonverbaler Aktivitäten begleitet sind – auch im Klassenzimmer (siehe etwa Goffman 1981, 40). Die Ana-

lyse von L2-Interaktionen im Klassenzimmer muss zwingend multimodal sein. Denn wie die neuere Forschung zeigt, spielen Blick und Gesten etwa in der Redezuteilung im Klassenzimmer eine wichtige Rolle (Kääntä 2010); genauso werden Aufmerksamkeit und damit potentielle Zugänglichkeit mit Hilfe von Blicken und Körperhaltungen angezeigt (vgl. Goffman 1982, 111). Spezifische Partizipationsstrukturen kommen durch Körperhaltung und Bewegungen des Körpers im Raum zustande (Goodwin & Goodwin 2004): Kinder, die während des Unterrichts miteinander tuscheln, verändern nicht nur die Lautstärke des Sprechens, sondern auch die Körperhaltung und die Blickrichtung.

Charles Goodwin beginnt schon Ende der 1970er Jahre im Anschluss an Goffman, den Blick und seine Funktionen in der Organisation der Interaktion in seine Analysen zu integrieren (Goodwin 1979). Er zeigt empirisch, was Goffman immer wieder postuliert hat, nämlich dass auch die Körperhaltung eine wichtige Rolle in der Koordination der Handlungen von Interagierenden spielt (Goodwin 1981). Später erweitert er seinen Fokus und plädiert in Anlehnung an Goffman, für den das physikalische Umfeld, der Raum, in welchem die Interaktion stattfindet, eine zentrale Bedeutung hat, für eine multidimensionale Analyse von Interaktion, die die semiotischen Felder der Interaktion berücksichtigt; ein semiotisches Feld konstituiert sich aus den benutzten Zeichen, die an ein bestimmtes Medium gebunden sind (Goodwin 2000, 1494). In der Interaktion unterscheidet er drei Felder: Erstens den sprachlichen Teil der menschlichen Interaktion und seine unterschiedlichen Aspekte, das heißt lexiko-semantische, syntaktische und prosodische; zweitens den Körper und die verschiedenen Zeichen, die sich an und in ihm verkörpern: Körperhaltung, Gestik, Mimik, Proxemik, Blick. Drittens identifiziert er weitere in der Umgebung der Interagierenden auffindbare semiotische Strukturen, in seinen Beispielen etwa die Bodenbemalung des Himmel- und Hölle-Spiels oder eine Munsell-Farbkarte, welche in der Archäologie benutzt wird. Diese Ensembles semiotischer Felder oder Ressourcen, die in Interaktionen für die Beteiligten relevant werden, nennt Goodwin »contextual configurations«:

It is argued that actions are both assembled and understood through a process in which different kinds of sign phenomena instantiated in diverse media, which I call semiotic fields, are juxtaposed in a way that enables them to mutually elaborate each other. A particular, locally relevant array of semiotic fields that participants demonstrably orient to (not simply a hypothetical set of fields that an analyst might impose to code context) is called a contextual configuration. (Goodwin 2000, 1490)

Goodwin, genauso wie Goffman, elaboriert seinen eigenen Zeichenbegriff nicht, und hinsichtlich des semiotischen Begriffsinstrumentariums lässt sich, wie auch sonst in der interaktionsanalytischen Diskussion, konstatieren, dass »ein elaborierter Begriff sprachlicher Multimodalität immer noch ein Desiderat« ist (Fricke 2012, 41).

Goodwins Medienbegriff ist weit und nicht differenziert, und seine Vorstellung des Verhältnisses zwischen Medium, Modus und Code ist etwas unklar. Goodwin setzt offensichtlich bei den Sinnesmodalitäten an, um verschiedene Modi zu unterscheiden, ein Vorgehen, das etwa auch von Stivers und Sidnell (2005, 2) gewählt wurde. Klar ist, dass Goodwin ein handlungstheoretisches, funktionales Zeichenverständnis hat. Er verweist darauf, dass die Bodenbemalung im Himmel-und-Hölle-Spiel oder die Munsell-Karte funktional analysiert werden müssen, und zwar im Hinblick darauf,

wie sie im Zusammenspiel mit den anderen semiotischen Ressourcen, die vorhanden sind, verwendet werden. Semiotische Strukturen, wie er sie nennt, sind aus seiner Perspektive also Affordanzen, verstanden im Neo-Gibson'schen Sinn als Eigenschaften eines Objektes, welche Handlungen ermöglichen (Zipoli Caiani 2013, 5). Diesen wird erst in der und durch die Interaktion Sinn zuteil (Goodwin 2000, 1516). Zudem ist sein Verständnis von Multimodalität zumindest teilweise anschlussfähig an Frickes (2012, 42-50) semiotisches Verständnis von Multimodalität, welches sie im Anschluss an Posners (1986) Differenzierung unterschiedlicher Medienbegriffe entwickelt.

Posner definiert das Medium als »System von Mitteln für die Produktion, Distribution und Rezeption von Zeichen, das den in ihm erzeugten Zeichenprozessen bestimmte gleichbleibende Beschränkungen auferlegt« (Posner 1986, 293). Relevant im vorliegenden Zusammenhang ist, dass Posner unterschiedliche Medienbegriffe differenziert und dass er unter anderem einen biologischen, einen physikalischen, einen technischen und einen codebezogenen Medienbegriff unterscheidet. Der biologische Zeichenbegriff »charakterisiert Zeichensysteme nach den Körperorganen, welche zur Produktion, Distribution und Rezeption der jeweiligen Zeichensysteme« eingesetzt werden. So lassen sich etwa visuelle, auditive, olfaktorische oder gustatorische und taktile Medien unterscheiden. Der physikalische Medienbegriff hebt auf die Unterscheidung der »physischen Kontaktmaterie« (Fricke 2012, 42) ab, die zur Übertragung der Zeichen notwendig ist. Der technologischen Medienbegriff »charakterisiert die Zeichensysteme nach den technischen Mitteln«, die zwischen Senderin und Rezeptionsorgan des Empfängers eingesetzt werden, und mit Hilfe des codebezogenen Medienbegriffs lassen sich etwa unterschiedliche natürliche Sprachen voneinander unterscheiden. Multimodalität nun setzt einerseits das Vorliegen mindestens zweier Medien voraus, beispielsweise der auditiv verarbeiteten Lautsprache einerseits sowie der visuell verarbeiteten Gesten.

Bezieht man diese Begriffe auf sprachliche Interaktion, so lassen sich die natürlichen Sprachen erstens aus codebezogener Perspektive als Medien bezeichnen, die auf bestimmte physikalische Medien zu ihrer Realisierung angewiesen sind. Werden diese Medienbegriffe nun mit dem biologischen Medienbegriff verschränkt, so lassen sich »verschiedene Medien des Sprachlichen« (Fricke 2012, 45) unterscheiden und begrifflich voneinander abgrenzen: Schriftsprache ist primär monomedial und monomodal. Die Lautsprache ist, wenn am Telefon verwendet, primär monomedial und monomodal. Allerdings ist dies nicht die »kanonische Äußerungssituation«, denn die *face-to-face*-Interaktionssituation ist in der Tat multimodal und, nach der von Fricke gelieferten Definition, ebenfalls multimedial, da hier nicht nur die auditive Sinneswahrnehmung betroffen ist, sondern auch die visuelle und, je nachdem, ebenfalls die taktile, die olfaktorische oder gar die gustatorische. Es sind also mindestens zwei Medien an der Rezeption beteiligt. Um nun aber Multimodalität und Multimedialität, die nun gleichbedeutend sind, voneinander zu unterscheiden, führt Fricke ein weiteres Merkmal ein, nämlich das Merkmal der »strukturellen und funktionalen Integration in ein und denselben Matrixkode«. Die strukturelle und/oder funktionale Integration ist für Fricke notwendige Bedingung des Vorliegens von Multimodalität. Das Fehlen einer strukturellen und/oder funktionalen Integration bei gleichzeitigem Vorliegen der Verwendung unterschiedlicher Medien ist für sie ebenfalls notwendige Bedingung von Multimodalität (Fricke 2012, 46).

Die Vorstellung der Unterscheidung unterschiedlicher Medien und die Idee, dass zum Vorliegen von Multimodalität eine funktionale Integration unterschiedlicher Medien notwendig ist, sind mit der bei Goodwin skizzierten Vorstellung von Multimodalität kompatibel. Allerdings ist die Rede von der Integration in einen Matrixcode unnötig. Denn die Aufgabe der Analyse ist die Deskription der integrierten Verwendung unterschiedlicher semiotischer Ressourcen aus emischer Perspektive. Eine multimodale Analyse stellt die Analysierenden jedoch vor Herausforderungen, die sich beim monomodal verengten Blick auf den verbalsprachlichen Aspekt von Interaktion nicht stellen. Erstens ist es unmöglich, das Gesamt der in der Umgebung der Interagierenden vorhandenen semiotischen Ressourcen zu katalogisieren und in die Analyse zu integrieren. Laut Goodwin kann es deshalb nur darum gehen, diejenigen semiotischen Felder in die Analyse zu integrieren, welche die Interagierenden selbst als relevant und als für ihr Handeln folgenreich anzeigen. Auch hier geht es also darum, die sichtbare Orientierung der Teilnehmenden als »Scheinwerfer« zu verwenden, der den Analysierenden anzeigt, welche Aspekte des Kontexts in die Analyse integriert werden müssen (Goodwin 2000, 1508-1509).

Gleichzeitig zu dieser Fülle an semiotischen Ressourcen, die berücksichtigt werden müssten, steht die multimodale Analyse vor einem anderen Problem. Wie Goffman betont, sind Gespräche oft in unterschiedliche Handlungsstränge integriert (z.B. Goffman 1981, 40). Dies lässt sich auch interaktional analysieren: Mondada (2008) etwa zeigt anhand von multimodalen Analysen von Telefongesprächen in einem institutionellen Setting, dass Telefongespräche in unterschiedlichste handlungspraktische, räumliche, aber auch semiotisch potentiell relevante materielle Kontexte eingebettet sind. Das Klingeln des Telefons muss nicht dazu führen, dass die gegenwärtige Aktivität aufgegeben wird, sondern kann zu einer Beschleunigung dieser Aktivität führen. Während des Telefonierens werden andere Aktivitäten parallel ausgeführt, die mit dem Telefongespräch sequentiell koordiniert werden. Das Telefongespräch wird nicht, wie dies die Audioaufnahme suggeriert, durch die Beendigung der Telefonverbindung beendet, sondern setzt sich in »post-closing«-Sequenzen fort, die verbaler, lautlicher und mimisch-gestischer Art sein können. Mondada verdeutlicht, dass eine Reduktion aufs Sprachliche nur einen Bruchteil der komplexen Organisationsform des in weitere Aktivitäten eingebetteten Telefongesprächs zu rekonstruieren vermag (Mondada 2008).

Doch öffnet die multimodale Erweiterung des Fokus nicht nur den Blick für weitere semiotische Felder und simultan durchgeführte Handlungen, sondern zeigt auch, dass die Rollen der an den Interaktionen Beteiligten komplexer und umfassender sind, als dies auf den lautsprachlichen Code der verbalsprachlichen Interaktionen fokussierende Analysen suggerieren. Die Interagierenden kommen nämlich nicht mehr nur als Sprechende oder Hörende in den Blick, sondern als Beteiligte, die verbal und nonverbal aktiv sind, die nonverbal adressiert werden können oder auch nicht, die beteiligt oder unbeteiligt sind. Die Analyse der Partizipationsstruktur kann ohne die multimodale Analyse kaum funktionieren. Selbst wenn als Kernaktivität eine sprachliche Handlung analysiert wird, so ist in Betracht zu ziehen, dass die Sprechenden sehr genau verfolgen, was die anderen an der Interaktion Beteiligten tun (Goffman 1981, 124ff.). Und was diese anderen tun, kann auch nicht-sprachlicher Art sein. Die interaktive Ordnung ist somit als »Gesamtzusammenhang aller simultan realisierten, se-

quenziell strukturierten und aufeinander bezogenen interaktiven Beteiligungsweisen aller Teilnehmer» zu rekonstruieren (Deppermann & Schmitt 2007, 17).

Für Deppermann und Schmitt bedeutet dies, dass eine multimodale Analyse zumindest *idealiter* die Berücksichtigung aller Beteiligter anstrebt und nach der Rolle der Einzelnen in diesen Interaktionsensembles fragt (Deppermann & Schmitt 2007, 20-21). Es kann dabei nicht darum gehen, als desinteressierter Beobachter Interaktionsensembles zu beschreiben. Dies ist schon aus technischen Gründen nicht gut möglich. Denn die Kamera-Einstellungen bedingen eine Vorauswahl darüber, was überhaupt analysiert werden kann, da sie festlegen, welche Ereignisse wie fokussiert werden (Mondada 2008, 50). Doch eine analytische Orientierung an der Orientierung der Beteiligten lässt auch multimodale Daten zu, das rekonstruktive Prinzip lässt sich auch auf multimodale Daten anwenden. Allerdings sind dabei die Restriktionen zu lockern. Nicht mehr die Sequenzorganisation steht im Zentrum, sondern »a broader set that concerns ordering and positioning of any kind of moves and utterances« (Arminen, Koskela & Palukka 2014, 48). Der Ansatz, der dann entsteht, ist nicht mehr konversationsanalytisch *per se* – dafür aber reichhaltiger.

4.6 Multimodale Interaktionsanalysen im L2-Unterricht in der Nachfolge Goffmans

Die Beziehung zwischen ethnomethodologischer Konversationsanalyse und Mitgliedschaftskategorisierung einerseits und Goffman andererseits sind komplex und waren lange Zeit ziemlich vergiftet (siehe etwa Goffman 1981; Schegloff 1988). Das Urteil darüber, ob sich die beiden Ansätze verbinden lassen, hängt von der Perspektive und der eigenen Position ab. Für den orthodoxen Konversationsanalytiker Schegloff sind sie inkommensurabel (Schegloff 1988), für Anne Rawls (1989) und Gregory Smith (2003) dagegen komplementär, denn Goffman, Garfinkel und Sacks seien darum bemüht, zu zeigen, wie Ordnung lokal produziert werde. Goffman gehe von der Prämisse aus, dass Ordnung dadurch in der Interaktion konstituiert wird, dass das Selbst in der Interaktion geschützt und bestätigt werde; Garfinkel und Sacks gingen davon aus, dass die Konstitution von Ordnung im Gespräch darauf beruhe, dass in der Interaktion Intelligenz hergestellt werden müsse. Hier werden Goffmans Vorstellungen von Interaktion, Selbst und Normalität verwendet, um zu analysieren und interpretieren, wie Legitimität hergestellt werden. Da Goffman nicht in der Lage war, die strukturellen und sprachlichen Aspekte der Interaktion so detailliert zu analysieren, wie dies der ethnomethodologischen Konversationsanalyse gelungen ist (Rawls 1989, 152), wird das Instrumentarium der CA und der MCA verwendet, um genau diese Aspekte der Interaktion zu analysieren. Dies ist kein methodologisches Novum. Goodwin analysierte schon früh sprachliche Interaktionen multimodal (Goodwin 1981). Zimmerman (1998) untersuchte das *footing* und Lerner (1996) fand *face* in den Präferenzstrukturen der Interaktion.

Die methodologische Restriktion des orthodoxen konversationsanalytischen Blicks auf Turn-Konstruktionen, Sequenzen und Präferenzen sowie Reparaturen werden zugunsten eines umfassenderen Blicks auf die Interaktion aufgegeben (siehe auch Goodwin 2000; Cekaite 2012; Goodwin & Cekaite 2013). Goffmans Insistenz auf der Notwendigkeit des Einbezugs auch des physischen Kontexts, aber auch des weiteren Kontexts der Interaktion für die Analyse soll berücksichtigt werden (Goffman

2005, 260-264). Sequenzanalysen, oder, wie mit Goffman gesagt werden könnte, eine Sequenz von Reaktionszügen, die nicht notwendigerweise sprachlicher Natur sind (Goffman 2005, 124), werden als in größere Zusammenhänge eingebettet betrachtet: In einen physikalischen Kontext, in einen kulturellen Kontext und damit einen Wissenszusammenhang, welchen die Interagierenden teilen oder auch nicht, aber auch in einen Gesprächskontext, der nicht nur aus der vorangehenden Äußerung oder dem vorangehenden Turn besteht (Goffman 2005, 262). Es muss und kann auch nach anderen Begründungen gesucht werden, denn, wie Goffman etwas spöttisch schreibt:

Jedenfalls lässt jede Erklärung der zweiten Äußerung aus ihrer Abhängigkeit von der Ersten außer Betracht, wie es denn überhaupt zur ersten kommen konnte. Woraus lässt sich denn deren Gestalt erkennen? Konversationen könnten nie beginnen. Oder, hätten sie einmal begonnen, wären sie nur eine Äußerung von ihrem Ende entfernt. Die Schwänze wüssten, wie sie wedeln, doch es gäbe keinen Hund. (Goffman 2005, 263)

Turns und Sequenzen werden daher als in »extended joint projects«, hier kommunikative Projekte genannt, eingebettet betrachtet (Clark 1996, 206; Levinson 2013, 120-122), in welchen unterschiedliche semiotische »contextual configurations« (Goodwin 2000) eine Rolle spielen und die Interagierenden auch mit verkörperten »participation frameworks« (Goodwin 2000) anzeigen, wie sie sich zur Interaktion verhalten. Gleichzeitig wird in Betracht gezogen, dass die Analysen der Züge der Interagierenden *de facto* auf Interpretationen beruhen, die zwar plausibilisiert werden müssen, aber eben Interpretationen bleiben. Zentral für die gesamte Untersuchung ist dabei, dass es auch in Interaktionen im L2-Klassenzimmer darum geht, das eigene Selbst sowie das Selbst der anderen zu wahren, dass hier Normalitäten rekonstruiert und damit Ordnungen legitimiert werden, und dass dies unter der Rekurrenz auf Wissensbestände geschieht, die von handelnden Subjekten eingesetzt werden. Konstruktionen des Selbst und *face* bilden dabei theoretische Basis und methodische Vorgabe: *face* wird in und durch Interaktion konstruiert (Goffman 1967, 44) und zeigt sich auch in Design und Struktur von *turn* und *turn-taking* (Lerner 1996). Der Ansatz, der dabei entsteht, lässt sich als multimodale Interaktionsanalyse bezeichnen (Cekaite 2012, 649) – das »multimodal« hier in programmatischer, wenn auch pleonastischer Weise verwendet, denn Interaktionsanalysen müssen immer multimodal sein. Goffman folgend, der immer auch am Einzelfall interessiert war, werden zwar vergleichende Fallanalysen durchgeführt; gleichzeitig soll aber auch das Besondere im Alltag des L2-Klassenzimmers gezeigt werden.

Multimodale Interaktionsanalysen im hier entwickelten Sinn rekonstruieren, wie Beteiligte ihre Interaktionen organisieren, im Hinblick auf übergeordnete kommunikative Projekte, im Hinblick auf die sequentielle Organisation, im Hinblick auf die Turnkonstruktion. Sie berücksichtigen nicht nur die lautsprachlichen Ressourcen, die eingesetzt werden, sondern fokussieren auch die anderen semiotischen Ressourcen, die für die Beteiligten in der Organisation der Interaktionen relevant werden. Der Fokus des Interesses liegt darauf, wie Beteiligte institutionell vorgegebene Ziele in und durch die kommunikativen Projekte, an welchen sie beteiligt sind, zu realisieren versuchen oder wie sie sich diesen Zielen entziehen, auf welche sprachlichen und nicht-sprachlichen Ressourcen sie dabei zurückgreifen, wie sie dabei *face* bewahren oder zerstören und welche lokal ausgehandelten Identitäten dabei relevant werden.

Die Analysen zeigen, wie im Handeln der Beteiligten und dessen Organisation spezifische institutionelle Orientierungen, spezifisches institutionelles Wissen, spezifische institutionelle Abläufe und institutionelle Normen rekonstruiert und damit legitimiert werden. Dabei wird berücksichtigt, dass an diesen Interaktionen L2-Sprechende beteiligt sind, und dass die meisten der Beteiligten Präadolescente sind, die einerseits in den Gebrauch der L2 sozialisiert werden, gleichzeitig aber längst als Mitglieder bestimmter »Communities of Practice« (Wenger 2006) agieren.