

Oliver Schnoor

Die Pädagogik der Stimme

Auditive Ethnographie frühpädagogischer Ordnungsbildung am Beispiel sprechmusikalischer Adressierungspraktiken

The educative voice

An auditive-ethnographical approach to the constitution of pedagogical orders (in early education) as speech-musical practices of addressing children

Zusammenfassung:

Der Artikel trägt zur qualitativen Empirierung des Auditiven (insbesondere von Stimmlichkeit) in sozialen Praktiken bei und demonstriert zugleich einen Ansatz zur Rekonstruktion pädagogischer Ordnungsbildung in frühkindlichen Betreuungseinrichtungen. Nach einer vorangestellten ethnographischen Vignette, in der sich Gegenstand und Ansatz des Beitrags schon verdichten, folgen zunächst eine Sondierung relevanter Theorie- und Methodologiebezüge, dann ein Umriss einer auditiven Ethnographie der Pädagogik (die das Konzept der „Adressierung“ nicht-metaphorisch wendet) sowie Angaben zu Forschungskontext und methodischem Vorgehen. Der erste empirische Abschnitt entwickelt den Begriff von Praktiken sprechmusikalischer Adressierung anhand der Analyse ethnographischer Feldnotizen aus einer luxemburgischen Kindertagesstätte. Vokale Adressierungsmuster, insbesondere Variationen der so genannten „Rufterz“, spielen eine zentrale Rolle bei der Entwicklung pädagogischer Routinen und der Festigung der Differenz zwischen Kindern und Erwachsenen. Der zweite empirische Abschnitt zeigt Relationen dieser Muster zu im engeren Sinne musikalischen Ereignissen im Einrichtungsalltag auf und verweist auf die wechselseitige soziale und sozialisatorische Konstitution von Sprache und Musik.

Abstract:

The article contributes to qualitative research on the auditive dimension (particularly on vocality) of social practices. At the same time it demonstrates an empirical approach to reconstructing the constitution of pedagogical orders in child care centers (crèches). An introductory ethnographic note already conveys the subject and approach of this contribution in a nutshell. Then, a discussion of relevant theoretical and methodological positions is given, followed by a sketch of an auditive ethnographical approach to education (which turns the concept of “addressing” non-metaphorically) and some information on the overall research context and the methods of the study. The first empirical section develops the notion of speech-musical addressing practices by analyzing ethnographical fieldnotes from a luxembourgian crèche. Vocal patterns of addressing, in particular varying uses of the “call contour”, play a crucial role in developing pedagogical routines and in strengthening the practical differentiation of children and adults. The second empirical part relates these patterns to musical events in the narrower sense and points to the mutual social and socialising constitution of speech and music.

Schlagwörter: Frühpädagogik, Stimme, Multimodalität, Ethnographie, Generationale Ordnung

Keywords: early childhood education and care, voice, multimodality, ethnography, generational order

1. Präludium

„*Maison Relais* Butzegaertchen, d’Martine, Moiën?“ – Im Hintergrund Geräusche und Kinderstimmen, ich bin also am Telefon des Gruppenraums gelandet. Eigentlich wollte ich Nathalie sprechen, die Leiterin, um meinen nächsten Aufenthalt für teilnehmende Beobachtung anzukündigen, aber wenn im Büro niemand abhebt, wird der Anruf weitergestellt.

„Moi – ën! Hier ist der Oliver.“ – Die luxemburgische Begrüßungsformel singe ich immer noch, im Bemühen, den herzlichen Tonfall, den ich während der ersten Wochen nach dem Start (der „Geburt“) der *Maison Relais pour Enfants* so häufig gehört habe, einigermaßen zu treffen. Das „Moiën“, mit drei spitz artikulierten Vokalen, aber in einer fließenden Bewegung gesprochen, um förmlich auf mich zuzuschnellen, wenn ich zum Beispiel gerade die Haustür der Einrichtung aufgeschlossen hatte und ein freundlich winkender Mensch vom anderen Ende des langen Flurs her grüßte, oder das langgezogene, aus vielen Richtungen kommende „Äddi“, bevor ich die Tür zum Gruppenraum hinter mir zumachte, solche mit Hingabe intonierten Grußrituale schallten in der ersten Zeit fast ständig durch die Einrichtung: Man begrüßte sich, als ob man sich lange nicht mehr gesehen hätte, und verabschiedete sich, als träfe man sich so schnell nicht wieder. In einer Zeit, da ein gerade zusammengestelltes pädagogisches Team in einem noch karg eingerichteten ehemaligen Vorschulgebäude miteinander zu arbeiten begonnen hatte und nicht nur mit seinen ersten sieben Kindern, sondern auch mit deren Eltern, die im Zuge der „Eingewöhnungsphase“ häufig anwesend waren, sowie nicht zuletzt mit uns Ethnologen, die ebenso vom ersten Tag an dabei waren, zurecht kommen musste, hatte intensives Grüßen offensichtlich auch den Zweck, Unsicherheiten abzubauen, Misstrauen vorzubeugen und sich – in einer allgemeinen großen Phase der „Eingewöhnung“ – gegenseitig des Willkommenseins zu versichern. Meine Eindrücke, die ich noch häufig während der acht Zugstunden zwischen neuem und altem Wohnort notiert hatte, waren stark auf non-verbale Aspekte von Interaktionen fokussiert, denn das Luxemburgische erwies sich als zu fern von meinem Deutsch, um es schnell zu verstehen, und mein Französisch als weniger konversationstauglich als erwartet; also hielt ich mich zunächst an das, was ich konnte: das „Wie“ der Blicke, der Bewegungen und besonders der Stimmen beobachten, dem Musikalischen im Alltagsgeschehen nachgehen.

Ein trockenes, eher monotones „Hallo“ kommt durch den Hörer zurück, ein höflicher Wechsel in meine Sprache. „Ich würde gern Nathalie sprechen, wenn sie im Haus ist“, trage ich mein Anliegen vor, da gellt ihre Stimme auf einmal hoch und laut in mein Ohr. Noch bevor ich verstanden habe, was sie ruft, weiß ich aber, dass ich nicht gemeint sein kann. Ein helles, leicht nach oben schleifendes „i“ erweist sich als Bestandteil eines Namens: „Ni – no!“ Nach einer kurzen Pause wendet Martine sich mir wieder zu: „Sie ... telefoniert gerade. Sie kann dich zurückrufen.“ – „Okay. Das wäre gut.“ – „Ja. Ich sag ihr Bescheid“,

gibt sie im nüchternen, kurz angebandenen Sprechtonfall zurück. Ich habe das Gefühl, Martine möchte mir gerade zeigen, dass sie bei der Arbeit ist.

Nach dem Auflegen notiere ich das Gespräch und die ungefähren Tonverläufe von markanten Sätzen. Der höchste Ton am Ende der begrüßenden Frageintonation („...d’Martine, Moiën?“) ist gerade der niedrigste des lauten Ermahnens eines Kindes gewesen. Es ist nicht irgendeine Arbeit, sondern pädagogische Arbeit, und das kann man hören.

2. Empirisierung des Auditiven: Bausteine einer qualitativen Methodologie

2.1. Sozialität als Vokalität: Stand der Diskussionen

Die Stimme ist im Hinblick auf die „Materialität der Kommunikation“ (Gumbrecht/Pfeiffer 1988) besonders eng mit dem Körper verbunden, da sie nicht nur körperliches Trägermedium von sprachlichen Signifikanten ist, sondern als leiblicher Ausdruck (vgl. Schützeichel 2011) zugleich auf den Körper zurück verweist sowie mit anderen Körpern – wenn man es phänomenologisch fassen will – im Sinne einer „Zwischenleiblichkeit“ in direkter Verbindung steht (vgl. Kolesch 2009). Mit diesen Eigenschaften spielt sie aber auch eine herausragende Rolle in der physischen Konstitution und Veränderung sozialer Räume (etwa im „spacing“ genannten interaktiven Positionierungsgeschehen, vgl. Löw 2001, S. 158) – eine Funktion, die in der dominant medienwissenschaftlichen Thematisierung der Stimme in den letzten Jahren eher unterbelichtet blieb (z.B. Kittler/Macho/Weigel 2002). Im Verbund mit der sprachlichen Oraltät ist die stimmliche Kommunikation somit *per se* „multimodal“, wie es sich sonst allenfalls für Formen der „Schriftbildlichkeit“ (Krämer 2003) zeigen lässt. Überdies ist die Stimme in zwei Arten von höheren Zeichensystemen eingebunden, die sprachlichen und die musikalischen, welche freilich verschiedene Überschneidungsflächen aufweisen bzw. in spezifischen kommunikativen Praktiken (um welche es im Folgenden auch gehen soll) gar nicht voneinander zu trennen sind.

Mit Blick auf die grundlegende Frage nach der Möglichkeit bzw. Herstellung von sozialer Ordnung hat das Stimmliche aber bislang verhältnismäßig wenig Aufmerksamkeit erhalten. Mikrosoziale, insbesondere ethnomethodologische Betrachtungsweisen („ordered at all points“) haben vornehmlich das sprachliche (d.h. verschriftsprachliche) Geschehen in den Blick genommen. Gelegentliche (aber nicht gesonderte) Hinweise finden sich jedoch bei Erving Goffman, etwa im Hinblick auf Verletzungen von Territorien als „Lauträumen“ (1982, S. 77) oder zu den sozialen Bedeutungen von scheinbar nur individuell-spontanen „response cries“ (1981). Am Ehesten wurde die Stimme im Bereich der gesprächsanalytischen und (sozio-)linguistischen Forschung berücksichtigt (z.B. Gumperz 1982), allerdings vornehmlich als Träger prosodischer Zeichen, welche ansonsten, wie auch die wortsprachlichen Transkriptionen, weitgehend vom komplexen physischen Geschehen gelöst oder nur im Kontext typischer Gesprächssituationen untersucht wurden.

Lässt sich bei diesen qualitativen Methodologien und Studien bereits anknüpfen, so sollte das Thema der Stimme bei solchen aktuell viel diskutierten soziologischen Theorieströmungen auf einen breiteren Resonanzboden treffen, die das Soziale in seinen komplexen „Zusammensetzungen“ in den Blick nehmen und dabei dem Körper, der Materialität und der Räumlichkeit konstitutive Bedeutungen zuschreiben: Einerseits wird in einer Reihe von Praxistheorien, welche über das bewusste und unbewusste Handeln von Akteuren hinausgehen, das Soziale von vornherein in Praktiken von prozessual agierenden Ensembles aus Menschen und Dingen verortet (vgl. Reckwitz 2008), andererseits wird, von der Semiotik her kommend, das Zusammenspiel vielfältiger nicht-sprachlicher Modi der Kommunikation und Sinnkonstruktion analysiert (vgl. Kress 2009; für klangliche Zeichen van Leeuwen 1999). Derart sollte auch Stimmlichkeit nicht mehr lediglich als Quelle expressiver Besonderheiten zu betrachten sein, sondern als konstitutiv in den Prozessen sozialer Ordnungsbildung mitwirkend.

Zentral ist die Stimme in den letzten Jahren im Zeichen des „performative turn“ und hier vor allem in den Perspektiven der Kulturgeschichte, der Medientheorie und der historischen Anthropologie behandelt worden (vgl. u.a. Kolesch 2009; Kolesch/Krämer 2006; Schulze/Wulf 2007). Die in diesem Kontext entwickelten Reflexionen zu Leiblichkeit und Räumlichkeit, Ereignishaftigkeit und Appellcharakter des Stimmlichen sind anschlussfähig für eine Forschung, die diesem auch in den praktischen Vollzügen sozialer Ordnungsbildung nachgehen will.

Aus einer philosophischen Richtung kommend wurde nicht nur im Performativen, sondern auch in den vorherigen „turns“ (linguistic, pictorial) ein bereits vollzogener, in den jeweiligen Paradigmen enthaltener „acoustic turn“ entdeckt (Meyer 2008, S. 14). Es solle lediglich „eine neue Phase markiert werden, in der eine erste Bewegung der Wiederbesinnung auf das Hörbare in den 70er und 80er Jahren überwunden wird“, indem dieses nicht dem Sichtbaren gegenüber privilegiert oder der Schrift gegenüber als das „Unmittelbare“ verklärt würde (ebd., S. 18f.). Auch in der ethnographischen Repräsentationsdebatte, welche die Sinnesmodalitäten in ganz ähnlicher Weise gegeneinander in Stellung gebracht hatte, ist gemahnt worden, eine Erforschung von „Hearing Cultures“ sei auf einem „long way toward allowing the ear ‘an unromanticized place alongside the eye’“ (Erlmann 2004, S. 5).

Allerdings mag mit dem gleichzeitigen Aus- und Rückruf eines „acoustic turn“ die Chance einer systematischen, insbesondere sozialwissenschaftlichen Empirisierung zu schnell vergeben werden. Der gleichnamige Band bündelt im Großen und Ganzen bereits vorhandene Ansätze zur klanglichen Dimension in verschiedenen Disziplinen, wobei deutliche Schwerpunkte bei künstlerischen, musikalischen und medialen Themen liegen. Die Entwicklung methodologisch reflektierter Zugänge im Dienst sozial- und erziehungswissenschaftlicher Forschung, wie sie in der qualitativen Forschung etwa auf dem Terrain der Bildlichkeit passiert ist, steht noch aus.

Vorbehalten gegenüber einer anthropologisch, medientheoretisch oder wie auch immer begründeten Festlegung auf bestimmte Sinn- und Erkenntnisdimensionen und einer dann naheliegenden methodischen Aussonderung des Klanglichen ist dabei gleichwohl zuzustimmen. Überzogene Erwartungen an die Eigenständigkeit nicht-sprachlicher Ausdrucks- und Darstellungsmodi, auch die Unterschätzung der Übersetzungs- bzw. Transformationsprozesse von Sinn in praktischen Vollzügen, die zum Teil die Diskussion über Methoden der Bildin-

terpretation geprägt haben (vgl. auch Schnoor 2007), müssen bei der sozial- und erziehungswissenschaftlichen Einholung des Klanglichen nicht erneuert werden. „Methodisch interessant sind die porösen Grenzen zwischen Bild, Ton und Sprache, an denen Gemeinsamkeiten wie Differenzen unterschiedlicher symbolischer Ordnungen hervortreten.“ (Rustemeyer 2003, S. 182) In diesen Zusammenhang gehört auch der Stellenwert anthropologischer Aussagen über die Stimme, über den Blick usw., die m.E. heuristisch wertvoll sein können, aber nicht verabsolutiert werden sollten: „Die herausgehobene Stellung der Stimme“, so Rainer Schützeichel jüngst in einem begründenden Aufsatz für eine „Soziologie der Stimme“, „beruht auf der Anthropologie unserer Sinne. Das Auge ist ein Distanzorgan. Um etwas zu sehen, benötigt man Distanz zu diesem Objekt. Die Stimme hingegen umgibt den Hörer.“ (Schützeichel 2011, S. 99; ähnlich van Leeuwen 1999, S. 195ff.) Wie aber die spezifischen leiblich-physischen Qualitäten der Sinne und des Ausdrucks soziologisch relevant werden, in welche sozialen Interaktionen und Praktiken sie eingebunden und mit welchen sozialen Bedeutungen sie versehen werden, ist erst empirisch zu klären. So erweist sich die schnelle und direkte Herstellbarkeit physischer Kontakte mit der Stimme im Folgenden als eine kommunikative Ressource, die sozial gerade auch zur Distanzierung verwendet wird.

2.2. Auditive „Ethnographie der Pädagogik“

Dem Alltagsbewusstsein kommen sofort Assoziationen, wenn es der Bedeutung der Stimme im Zusammenhang mit Erziehung nachgeht. Die dann in den Sinn kommenden Phänomene und Szenen mögen aber kaum einen angemessenen Platz in pädagogischen Theorien, Didaktiken und Ansätzen finden. Die Suche nach der spezifisch „pädagogischen Ordnung“ in Bezug auf Stimmlichkeit soll auch nicht mit pädagogischen Begriffen, sondern mit den Instrumentarien sozialwissenschaftlicher Theorien und Methodologien beginnen. Pädagogische Ordnung soll als soziale Ordnung untersucht und auf diesem Weg in ihrer Spezifik aufgeklärt werden (vgl. Neumann 2011). Um zu beantworten, wie genau die tatsächlichen praktischen Gebrauchsweisen der Stimme in die Erzeugung des Pädagogischen eingebunden sind, scheint zunächst einmal eine „Ethnographie der Pädagogik“ und eine diese kennzeichnende „Dezentrierung des pädagogischen Blicks“ (Hünersdorf/Maeder/Müller 2008) am vielversprechendsten. Sie hätte auch die Aufgabe, das pädagogisch-anthropologische Wissen über Vokalität (vgl. die Beiträge in Bilstein [Hrsg.] 2011) stärker im Kontext der Beobachtung sozialer Praktiken zu empirisieren.¹

Dass sich eine ethnographische Beobachtung des Pädagogischen nicht an pädagogischen Programmen und Konzepten orientiert, es also nicht von deren korrekter Umsetzung in der Praxis oder gar den Erfolgen abhängig macht, ob sie es mit Pädagogischem zu tun hat oder nicht, bedeutet jedoch nicht, dass diese Beobachtung auf die Sensibilisierung durch bereits erreichtes theoretisches und empirisches Wissen über die Praktiken, in denen Soziales als Pädagogisches beobachtbar (gemacht) wird, verzichten muss. So kann als unumstritten gelten, dass zum gesuchten Gegenstand immer die Konzipierung und Behandlung einer bestimmten Personengruppe als Adressaten einer Praxis gehört, welche im Horizont einer verbessernden Personenveränderung steht. Ebenfalls ist

bereits bekannt, dass Pädagogik sich in solchen praktischen Vollzügen konstituiert, in denen sie eine verbessernde Personenveränderung für die direkt Beteiligten wie für Außenstehende wahrnehmbar macht, ausweist und dokumentiert, und zwar als ihre eigene Wirksamkeit (vgl. Bollig 2004; Cloos/Schulz 2011; Neumann 2012). Schließlich ist auch bekannt, dass Pädagogik in ihrem basalen Muster der Adressatenkonstruktion auf der Unterscheidung von Kindern und Erwachsenen aufsetzt, sie in eine „pädagogische Differenz“ verwandelt und „damit zugleich die sozialen und kognitiven Bedingungen der Möglichkeit pädagogischer Intervention [schafft]“ (Honig 2009, S. 46), etwa indem sie Konzeptionen vom Kind als einem in besonderer Weise verbesserungsbedürftigen Wesen (bildsam, lern-, erziehungsbedürftig usf.) zur Grundlage besonderer Praxen macht.

In einem auditiven Modus stößt ethnographische Beobachtung im Alltagsgeschehen einer Tagesstätte zunächst darauf, dass sich die mündliche Kommunikation der Erwachsenen mit Kindern von derjenigen zwischen Erwachsenen schon klanglich deutlich unterscheidet. Von diesem Hintergrund, so muss angenommen werden, wird sich die „Pädagogik der Stimme“ hinsichtlich des operativen Vollzugs und vermutlich auch der Ausdrucksgestalten noch einmal abheben. Die praktische Unterscheidung der Generationen im Aspekt des Vokalen zu beobachten führt im Folgenden zu einem bestimmten Bereich sozialer Situationen, die im engeren Sinne als *Adressierungen* bezeichnet werden können (darauf hat die einleitende Vignette schon hingedeutet). Aktuell wird das Pädagogische auch allgemeiner, z.B. als Lernkultur im Unterricht, als „Adressierungsgeschehen“ analysiert (vgl. Fritzsche/Idel/Rabenstein 2011; Dinkelaker/Idel/Rabenstein in diesem Band). Im Folgenden wird „Adressierung“ jedoch weniger metaphorisch und generalisiert verstanden, sondern konkreter als „Ansprechen“ zum Zweck der Herstellung oder Aktualisierung eines Kontakts (auch zur Einleitung einer Kontaktbeendigung), das etwa die Form einer Anrede, eines Rufs, eines Grußes oder auch eines Appells annehmen kann. Es handelt sich häufig, in der linguistischen Terminologie, um *Vokative* (vgl. Zwicky 1974).

An solchen Adressierungsakten – als Schaltstellen sozialen Geschehens – lassen sich zugleich Positionen und Positionierungen, soziale Ordnungen und Praktiken des Ordners beobachten. Der Mehrwert eines Einbezugs des Stimmlichen erweist sich dann darin, dass solche Adressierungen auch als material-räumliche und/oder metaphorisch-räumliche Herstellungen von Distanzen analysierbar werden, die sich wiederum als Elemente der praktischen sozialen Differenzierung erkennen lassen. Damit eröffnet sich möglicherweise auch die Frage, in welcher Weise formal-relationalistische Ordnungsbegriffe mit „konkretistischen“ Alltagsvorstellungen von materieller und körperlicher Ordnung und Ordentlichkeit vermittelt werden können. In diesem Sinne wird die hier präsentierte Studie etwa an die Ethnographie von Kai Schmidt (2004) zum „geordneten Raum“ in Kindertagesstätten anschließen, welche Löws Raumbegriff als „Ordnung des Nahen und Fernen“ auslegt (ebd., S. 166) und die fortlaufende Raumkonstruktion am Beispiel des Status von Objekten analysiert, der vor allem im wiederholten *Aufräumen* hergestellt wird (S. 173ff.; vgl. Abschnitt 4).

2.3. Forschungskontext, methodische Strategien und Datenerzeugung

Die für diesen Text verwendeten Beobachtungsprotokolle stammen mehrheitlich aus einem frühen Zeitraum der teilnehmenden Beobachtung in einer Luxemburger Kindertagesstätte für null- bis vierjährige Kinder, einer so genannten *Maison Relais Crèche*.² Die *Maison Relais pour Enfants* (MRE) sind ein im Jahr 2005 in Luxemburg gesetzlich etablierter Typus von Betreuungseinrichtungen, der durch längere Öffnungszeiten und flexible Einschreibe- und Aufenthaltszeiten gekennzeichnet ist. Er dient zum einen arbeitsmarkt- und gleichstellungspolitischen Zwecken („Vereinbarkeit von Familie und Beruf“), zum anderen soll er aber auch eine höhere pädagogische Qualität bieten. Das Begleitforschungsprojekt fragt übergreifend, wie pädagogische Qualität unter deregulierten organisationalen Bedingungen in den alltäglichen Praktiken erreicht wird.

Die hier herangezogenen Feldnotizen erlauben es zum einen, die internen Vergleichs- und Kontrastierungsmöglichkeiten auszunutzen, die sich aus der seltenen Gelegenheit ergeben haben, einen Prozess der Gründung und der raschen Entwicklung und Veränderung einer Kindertageseinrichtung begleitend verfolgen zu können. Zum anderen haben die Feldnotizen anfänglich von einer sprachlichen „Fremdheit“ (die später nicht mehr bzw. weniger gegeben war), insofern profitiert, als sie die Fokussierungen und Akzentuierungen, auf denen dieser Beitrag basiert, möglich gemacht oder zumindest begünstigt haben.

Solange wenig über die Rolle des Akustischen und Stimmlichen in sozialen, insbesondere pädagogischen Praktiken bekannt ist, ist es sinnvoll, das Hörbare schon auf einer phänomenalen Ebene in den Mittelpunkt zu stellen und nach (unerwarteten) Gemeinsamkeiten und Unterschieden zu befragen. Der Autor konnte für dieses Unterfangen – komplementär zum sprachlichen Kompetenzrückstand – sein musikwissenschaftliches Studium und seine musikalische Gehörbildung nutzen, womit es ihm auch ohne durchgängige technische Aufnahmen möglich war, hinreichend genaue Notizen und Transkriptionen anzufertigen. In der Praxis des ethnographischen Protokollierens wird aber selbstverständlich nicht ausschließlich die audielle Dimension erfasst, sondern ebenso andere Dimensionen, wie die zeitliche, räumliche, körperliche, die im engeren Sinne sprachliche etc. Die für Ethnographie entscheidende Be- oder Verfremdung liegt hier jedoch auch in einer besonderen Art der Interpretation und Verknüpfung des Alltagsgeschehens, die nicht unbedingt den Zäsuren und Markierungen der institutionellen Praktiken folgt, auch nicht den Vorgaben von pädagogischen und didaktischen Konzepten (etwa musikalische Aktivitäten, Sprachförderung, Erziehung zur Befolgung von Regeln und Werten etc.), sondern Muster und Zusammenhänge über die Orte, Zeiten und Settings hinweg entdeckt und diesen immer auch schon in der Erzeugung von Daten nachgeht. Die so entstehenden Beobachtungsprotokolle erlauben es, spezifische kommunikative Akte, Interaktionsmuster und deren Variationen – hier vor allem: Adressierungen – zu vergleichen und unter dem Gesichtspunkt sozialer Ordnungsbildung zu interpretieren. Dabei werden die Beobachtungen jedoch auch während des Forschungsprozesses theoretisch informiert, um sie wiederum mit sensibilisiertem Blick respektive Ohr im Feld verfolgen und differenzieren zu können. Spuren dieses Prozesses werden auch in der folgenden Präsentation empirischer Ergeb-

nisse sichtbar, so wechseln konkretere Beschreibungen mit (notwendigen) theoretisch-begrifflichen Klärungen ab.

Alles in allem hat dieses Vorgehen (welches *einen* Zugang innerhalb des genannten Forschungsprojekts darstellt) eher die „Gunst der Stunde“ genutzt, als einen genauen Erhebungsplan verfolgt, hat eher den Wahrnehmungs- und Aufzeichnungskapazitäten des teilnehmenden Beobachters vertraut, als einer in vieler Hinsicht schwerfälligen Aufnahmetechnik, hat schließlich eher von abduktiven Schlüssen profitiert als von methodisch geleiteten, für dieses Forschungsfeld auch noch gar nicht entwickelten Auswertungsschritten.³

3. Register der Distanz: Das vokale Aufrufen der Erwachsenen-Kind-Differenz

Die eingangs mit einer narrativen Vergewärtigung zweier Zeitpunkte der ethnographischen Feldforschung angeklungenen Kontraste vokaler Muster haben in verdichteter Form gezeigt, wie eine verhältnismäßig schwach betonte Differenz zwischen Kindern und Erwachsenen im Verlauf der Entwicklung pädagogischer Routinen in einer Betreuungseinrichtung zu einer gefestigten Differenz wird. Sie kann auf Basis geteilten Wissens schnell reaktualisiert, etwa in einem kurzen Telefonat krass exponiert werden, ohne dass dies als bedeutsamer Bruch erlebt werden muss.

In der Zusammenschau vergleichbarer Beobachtungsprotokolle innerhalb dieses Zeitraums stellt sich der Kontrast als stärker kontinuierlich und differenzierbar dar. Dabei lässt sich über die Differenzen im stimmlichen Ausdruck hinweg zeigen, dass in den variierenden Sprech- und Anspracheweisen auch *gemeinsame* sprechmusikalische Mittel verwendet werden, deren genauere Analyse zu den zugrunde liegenden Mustern der Adressierung und Positionierung von Kindern als Kindern (und von Erwachsenen als Erwachsenen) führt.

Zu Beginn der teilnehmenden Beobachtung, als in der Einrichtung oft die Atmosphäre eines „Familientreffens“ vorherrschte (viele Erwachsene, insbesondere auch Eltern, wenige Kinder), konnten häufig sogar wechselnde Blicke ausreichen, um sich gegenseitig der Anwesenheit und der aufeinander gerichteten Aufmerksamkeit zu versichern, leise Gespräche oder gar Stille waren keine Seltenheit. Es bildeten sich oft räumliche Konstellationen, die durch eine Zentrierung von Kindern durch außen sitzende erwachsene Beobachter gekennzeichnet waren. Als „Aufsicht“ auf Kinder intendiert (zur Beobachtung ihres Verhaltens, ihrer Entwicklung usw.), bemerkte ich, wie diese Situationen gelegentlich von Kindern transformiert wurden, indem sie plötzlich ostentativ die Blicke aufnahmen und Blickverbindungen stabilisierten, womit die momentane räumliche Distanz „überbrückt“ und zugleich erst realisierbar wurde.

Die auffällig betonten Grußrituale des Personals (intern und zum Teil mit anderen Erwachsenen) erscheinen im Vergleich mit anderen interaktiven Szenen des gleichen Beobachtungszeitraums (also der ersten Wochen nach Inbetriebnahme der Einrichtung) als Varianten von Interaktionen mit Kindern, und zwar solchen, die für gewöhnlich der Eltern-Kind-Interaktion zugerechnet und im Hinblick auf ihre Entwicklungsförderlichkeit interpretiert werden: Inszenie-

rungen von Verschwinden und Wiederkommen, von „Abschied“ und „Wiedersehen“, wobei das erneute Erscheinen immer erst von einer hörbaren „Begrüßung“ zu einem Wieder-Sehen gemacht wird (vgl. Bruner/Sherwood 1978).

Diese Struktur ließ sich in dem sozialen Arrangement, das von der „Eingewöhnung“ der ersten Kinder bestimmt war, sehr häufig und in unterschiedlichen Varianten beobachten. Es waren nicht nur Eltern, die mit ihren eigenen oder auch fremden Kindern solche Interaktionen initiierten, sondern vor allem Erzieherinnen (weniger Erzieher), die Spiele des Wiedersehens und Begrüßens inszenierten und dabei insbesondere auch vielfältige räumlich-materielle Gelegenheiten ihres Hauses einbanden. So wurde an Kinder, die sich in ein kleines, mit transparentem Stoff bezogenes „Haus“ verkrochen hatten oder durch im Bettenmobiliar ausgesägte sternförmige Kucklöcher hindurch schauten, aus einigem Abstand ein typisches „Moi-ën-!“ oder auch „Cou-cou-!“ gerichtet, oder wurden solche älteren Kinder, die den Weg durch den Tunnel einer Podestlandschaft zurückgelegt hatten und am anderen Ende durch die Luke wieder auftauchten, mit einem förmlicheren „Bon-jour-, Madame!“, „Bon-jour-, Monsieur!“ in Empfang genommen.⁴

Die *Stimmqualität*⁵ (vgl. Laver 1980) bei diesen freundlichen Hinwendungen ist in der Regel entspannt, oft aspiratorisch, seitens der Frauen häufig im Falsett; die Lautstärke ist nicht übermäßig – was für die Rolle solcher Anspruchsmuster bei der Herstellung von Nähe, Intimität und, in institutionellen Begriffen, von Familialität oder Familienähnlichkeit spricht. Auch in entwicklungspsychologischen Beobachtungen wurde der Ausgangspunkt von „Kuckuck“-Spielen, die oft repetitiv und zugleich variierend gespielt werden, in der „Annäherung“ der Mutter an ihr Kind verortet (Bruner/Sherwood 1978, S. 165). In diesen Gebrauchsweisen der Stimme im frühkindlichen Kontext könnte man zunächst also eine Bestätigung und Nutzung der anthropologisch-phänomenologischen Bestimmungen erkennen, die dem Hörsinn *Distanzlosigkeit* bzw. mangelnde Distanzierbarkeit zuordnen und der Stimme die „Vereinigung“ und „Solidarisierung“ zusprechen (Schmitz 2008, S. 86; van Leeuwen 1999, S. 196).

Um das Allgemeine und die spezifischen Varianzen dieser Stimmäußerungen empirisch weiter aufklären zu können, ist ein Exkurs zur prosodischen Form der besagten Äußerungen hilfreich. Diese ist stets in etwa dieselbe: Die Vokale zweier Silben werden in annähernd gleichbleibender Tonhöhe in einem absteigenden Intervall gleichsam gesungen: „Hal lo-!“ (bei einsilbigen Wörtern wird der Vokal verlängert: „Tschü üss-!“ Bei mehrsilbigen Wörtern wird häufig eine optionale tiefere Tonhöhe vorangesetzt: „Wie der seh'n-!“ Dieser prosodischen Kontur wird in der Phonetik eine Nähe zum Musikalischen zugesprochen (Ladd 1978, S. 525), zumal das Intervall zwischen den beiden langen Tönen häufig annähernd eine kleine Terz trifft und in dieser Form auch vielfältig in der Musik wiederkehrt.⁶ In Bezug auf das Englische wird diese Figur auch *spoken chant* oder *vocative chant* genannt. Der für das Deutsche gebräuchliche Ausdruck „Rufertz“ weist schon auf ein allgemeines Funktionsmerkmal der Sprechakte hin, die in dieser Form geäußert werden, sie sind auf die eine oder andere zu klärende Weise immer ein „Rufen“. Auch im Englischen wird mit dem Ausdruck „call contour“ (Abe 1962) spezifiziert, um welche Art von Vokativ es sich handelt.

Rufen verweist nun aber nicht auf Nähe, sondern im Gegenteil auf einen großen Abstand. In der Grundsituation des Ausrufens, z.B. eines Namens, ist die gerufene Person außer Reichweite und/oder außer Sichtweite, gar an einem un-

bekanntem Ort. Sprachwissenschaftlich ist auch erkannt worden, dass sich diese räumliche Konstellation in einer metaphorisierten Weise auch noch in denjenigen Äußerungen mit der rufenden Intonationskontur wiederfindet, die an eine Person in unmittelbarer Nähe gerichtet ist: „Distance between the person calling and the person being called is, no matter whether this distance is a real thing or an imagined one, a vital factor for prescribing [the call contour]“ (Abe 1962, S. 520). Demnach hat die prosodische Form, um die es geht, immer den Sinn einer Distanzüberbrückung. Dieser „stylized fall“ (Ladd 1978) intoniert meist nicht tatsächlich einen Sprechakt des Rufens, sondern repräsentiert ein Rufen, d.h. er sagt: „Ich spreche dich an, als ob ich dich rufe.“ Wenn es aber so ist, dass die Distanz keine physische, reale Distanz sein muss, sondern auch eine „vorgestellte“ sein kann, dann wird interessant, in welchen sozialen Situationen nach welchen Kriterien etwas durch den Gebrauch der „Rufterz“ als Distanz interpretiert und definiert wird.

Exempel in phonetischen Abhandlungen geben schon entscheidende Hinweise. So finden sich im Kapitel „Calls“ bei Gibbon (1976, S. 274ff.) neben Grüßen und verwandten Formeln (die im Grunde genommen auf einen zeitlich zurückliegenden oder bevorstehenden räumlichen Abstand verweisen) eine eigene Kategorie „Talking to babes-in-arms“, d.h. also Sprechen zu Personen, die sich in denkbarer körperlicher Nähe befinden. Beispielsätze lauten: „Was macht_ er_ denn-? [...] Wo ist_ er_ denn-?“ (ebd., S. 277) Die singende Ruf-Intonation hat hier die Funktion, Aufmerksamkeit herzustellen und interaktiven Austausch zu stiften, in diesem Sinne eine „vorgestellte Distanz“ immer wieder neu zu überwinden.⁷

Ähnlich wie mit Säuglingen kam das spielerische Rufen in den ersten Tagen der Luxemburger Tagesstätte auch mit den älteren *toddlers* und Kleinkindern zum Einsatz, und zwar typischerweise nicht einmalig, sondern repetitiv oder perennierend. Damit wird die Distanzüberbrückung auf Dauer gestellt. Insofern deuten die Beobachtungen auch auf ein praktisches „Konzept“ vom Kind (Säugling) hin, das Kinder als *a priori* fremd entwirft, „in einer anderen Welt“ verortet, aus welcher sie erst in „diese Welt“ hineinzubringen sind (auch die zitierten Beispielsätze bei Gibbon illustrieren dies). Damit ist dieses Kommunizieren mit Kindern nicht schon pädagogisch; es naturalisiert aber bereits jene Differenz, die sich frühpädagogische Ansätze und Praktiken zum Ausgangspunkt nehmen: „Kinder sind anders“ (Liegle 2006, S. 11ff.).

Das pädagogische Potenzial dieser Sprechweisen mit Kindern kann in der Analyse des folgenden Beobachtungsprotokolls, das einige Wochen nach dem Betriebsstart der *Maison Relais* während des Frühstücks entstanden ist, genauer zum Vorschein kommen.

Christine fragt Lisa (die neben ihr sitzt), was sie essen möchte. Diese schaut in die richtige Richtung zu den Frühstücksaufstrichen und -belägen, sagt aber nichts. Bonnie (Lisa gegenüber sitzend) schaut sie fest an und sagt laut: „Li_ sa_!“, im Tonfall wie zu jemandem, der nicht hört bzw. den man wiederholt ansprechen muss, um dessen Aufmerksamkeit zu erreichen. Auf Französisch (die Sprache, die Lisa am besten versteht) sagt Bonnie zu Lisa: „Wenn du nicht sagst, was du willst, können wir dir nichts geben.“ Christine, zu Bonnie und mir gesprochen: „Das kommt von zuhause. Da bekommt sie einfach alles und wird bedient. Sie muss lernen, zu sagen, was sie will.“

In der beschriebenen Szene wird ein bereits sprachfähiges Kind mit einer Ruf-Intonation angesprochen. Offensichtlich gibt es aber weder eine tatsächliche Distanz, noch muss die Aufmerksamkeit des Kindes eigentlich erst erreicht

werden. Die Intervention der Erzieherin scheint eher zu bezwecken, Lisa die spezifische Interaktionsordnung der Kinderkrippe „in Erinnerung zu rufen“ und sie in eine bestimmte Routine der Kommunikation einzugliedern: Erwachsene fragen Kinder nach ihren Präferenzen (in einer vorgegebenen Auswahl von Dingen oder auch Aktivitäten), Kinder antworten, indem sie ihre Auswahl benennen oder darauf zeigen. Allgemeiner: Solche Vokative sind Techniken, die angemessene Äußerungen von Kindern *e-vozieren* sollen, d.h. „Stimmen“ der Kinder, die in diesem institutionellen pädagogischen Setting als kindgemäß, z.B. als „autonome“ Entscheidung, gelten können.

Die akustischen Aspekte der gerufenen Adressierung können jedoch zunächst die spezifische Ordnung konkretisieren helfen, die im *Akt der Adressierung selbst* hergestellt wird. Die hohe Lautstärke (die dem Notierenden aufgefallen ist) heißt im Grunde, dass gerufen wird, als sei eine Rufdistanz tatsächlich vorhanden. Der Vokativ wird mit stimmlichen Mitteln *nicht* als ein „als-ob“-Rufen gerahmt. Die derart verstärkte Erzeugung von Aufmerksamkeit führt zugleich die Annahme mit, dass die Aufmerksamkeit unter besonders schwierigen Bedingungen zu erlangen sei, und diese würden, da es offensichtlich keine räumliche Distanz gibt, beim adressierten Kind liegen.

In obigem Beispiel soll jedoch mehr als eine bloße Aufmerksamkeit für die Interaktionspartner und verhandelten Dinge erzeugt werden, vielmehr soll die Aufmerksamkeit auf eine bestimmte, gültige Verhaltenserwartung gelenkt werden. Diese scheint umso stärker artikuliert zu sein, als das *Tonhöhenintervall* im geschilderten Beispiel nicht mehr näherungsweise eine Kleinterz ist, sondern deutlich größer. Es bewegt sich – wie in vergleichbaren Situationen in der Tagesstätte häufig, insbesondere aber bei wiederholtem Rufen – im Raum um die Quinte herum (beispielsweise ungefähr ein Tritonus oder eine kleine Sexte). Der Dissonanz dieser Intervalle, wie auch der angespannten Stimmqualität, scheint mir eine stärkere Diskrepanz zwischen erwartetem und tatsächlichem Verhalten des Kindes zu entsprechen.⁸

Werden Erwachsene schon mit zurückgenommenen Stimmqualitäten nur in Ausnahmefällen, in der Regel im öffentlichen Raum, mit der *call contour* adressiert, so zeigt sich nun als Spezifikum des Umgangs mit Kindern, dass sie in den geschlossenen Räumen der Betreuungsinstitution häufig und mit nahezu uneingeschränkter Stimmkraft tatsächlich gerufen werden. Eine konversationsanalytische Studie hat das Adressieren und Rufen als „attention-getting device“ detailliert im Zusammenhang mit *Eröffnungen* von Gesprächen untersucht (Schegloff 1968). Gegen diese Studie gestellt machen die Beobachtungen in der Kinderkrippe deutlich, dass die dort anwesenden Kinder permanent mit konversationseröffnenden Sprechakten adressiert werden.

Die sprechmusikalische Praktik hat weitere Implikationen. Die Verwendung der „Rufterz“, vor allem in der lauten Variante, die die reale Nähe zwischen den Personen übergeht, bringt gleichsam erst die Distanz hervor, die sie „überbrückt“. Zugleich wird diese Distanz nicht auf einer „reinen“ Bedeutungsebene kommuniziert, sondern – einhergehend mit der Überschreitung der „Territorien des Selbst“ und der Erzeugung von physiologischem und sozialem Stress – gleichsam *verkörperlicht*. Zudem bewirkt die Vergrößerung des akustischen Raums und die Signalwirkung an die allgemeine Umgebung auch eine *Veröffentlichung* des sozialen Geschehens: Die Interaktionssituation erhält im Gruppenraum der Krippe potenziell viele „bystanders“ (Goffman 1981), die – nicht unwillkommenermaßen – auf die ablaufende Handlung aufmerksam werden.

Da das Geschehen in der *Maison Relais* mit der Veränderung des Gesamtsettings, nicht zuletzt durch die steigende Zahl eingeschriebener Kinder, bald weniger übersichtlich wurde, war auch zunehmend weniger gesichert, dass alle „da sind“, d.h. den erwarteten Verhaltensnormen entsprechen. Im Verlauf der teilnehmenden Beobachtung hat sich der Eindruck eingestellt, dass laute, feste Ermahnungen von der Ausnahme zum allenthalben hörbaren Alltagsbestandteil wurden. Sie konnten bald auch ganz unvermittelt in den Raum hereinbrechen, ohne dass dies sichtlich als unterbrechendes Ereignis erlebt worden wäre. Insgesamt verlief, grob gesagt, eine Entwicklungslinie vom Blick und der „intimen“ Stimme hin zum Rufen und der „öffentlichen“ Stimme.

In Bezug auf die Positionierung der Kinder, die in den sprechmusikalischen Praktiken erkennbar wird, stellten sich die zwar regelmäßigen, aber anlassbezogenen ermahnenenden Rufe jetzt stärker als Interventionen mit Bezug auf Personenveränderung dar, sie schienen Kinder auf einen bereits erreichten Stand „zurückzurufen“ bzw. die stets prekären Fortschritte zu sichern.

In Bezug auf die Aufgabe der Vereindeutigung von sozialen Praktiken als pädagogischen lässt sich hinzufügen, dass die gelegentlich erschallenden Ermahnungsrufe auch als Aktualisierungen und Bestätigungen der gemeinsamen Zwecke der Einrichtung gehört werden konnten. Sie haben dann – auf der Ebene des „flüchtigen“, den Alltag durchziehenden Auditiven – eine ähnliche Funktion wie die handfesteren Dokumentationsbögen und Formulare unterschiedlichster Art, deren Produktion ebenso stetig und oft beiläufig vorstättengeht (vgl. Bollig 2011).

Der Begriff des Sprechmusikalischen ist bisher im künstlerischen Performancebereich für verschiedene Gattungen und Gattungsüberschreitungen verwendet worden, hier habe ich ihn auf den Bereich alltäglicher sozialer Kommunikation übertragen. Jedoch ist diese Begriffsübertragung kein Verfahren, das Aspekte des sozialen Lebens mit einer Metapher aus dem Kunstbereich erhellen soll, analog etwa der Theater-Metaphorik von Goffman. Vielmehr deutet der Begriff auf Momente der mündlichen Kommunikation hin, die sich nicht eindeutig der Sprache oder Musik zurechnen lassen. Die hier empirisch fokussierte Grundform der Rufferz mit ihren verschiedenen, mit stimmlichen Mitteln erzeugten Variationen in Akten der direkten Adressierung hat sich dabei als Mittel erwiesen, das eine Absetzung und Hervorhebung des Sprechens von anderem Sprechen erlaubt. Insofern dieses Mittel, wie gezeigt, zur differenziellen und distinktiven Ansprache von Erwachsenen und Kindern gebraucht wird, spielt es eine zentrale Rolle in der praktischen Herstellung von (Früh-)Pädagogik.

4. Frühpädagogisches Kontinuum von Sprache und Musik

Mit dieser Perspektive auf die Multimodalität der Stimme lassen sich in einem erweiternden empirischen Schritt Anschlüsse von sprechmusikalischen Momenten der Kommunikation an deutlicher musikalische Formen sichtbar machen. Dabei weist diese Perspektive zugleich auf die generelle Musikalität des Sprechens und das Sprechende der Musik hin und damit auf den Sachverhalt, dass

die Trennung und zugleich Relationierung von Sprache und Musik in sozialen Praktiken vollzogen wird.

Über den Einrichtungsalltag hinweg lässt sich ein melodisches Kontinuum ausmachen, das von den ermahnenden Ansprachen bis zu gesungenen oder abgespielten Musikstücken reicht. So taucht die „Rufertz“ zentral in einem Lied auf, das in der Einrichtung über lange Zeit ein regelrechter Hit war (und vermutlich weiterhin bleiben wird).⁹ Dort heißt die stets gleichlautende, von einzelnen Kindern gesungene erste Zeile der Strophen: „Kuck, wat ech kann!“, die Hörer werden also angesprochen, um bei einer Könnensvorführung zuzuschauen. In der folgenden Zeile wird nun jeweils bezeichnet, was das gerade singende Kind in gekonnter Weise tut: „Ech klopfen, ech klopfen!“, bzw. in den Variationen der Strophen unter anderem: „Ech hopsen“, „Ech sprangen“ (springen), „Ech dréinen“ (drehen). Eben diese Kundgabe wird im Kleinterzmotiv gerufen: „Ech klopfen, ech klopfen!“ Dann fordert die Kinderstimme auf: „Fänkt all mat mir un!“ (Fangt alle mit mir an), woraufhin ein Kinderchor einfällt: „Mir klopfen, mir klopfen. Den nächsten kënnt drun“ (Der nächste kommt dran).

Das Lied wird in der *Maison Relais* gern zu denjenigen Zeiten abgespielt, die von Warten geprägt sind, beispielsweise in der Viertelstunde vor dem Mittagessen, in der die Kinder, wenn sie pflegetechnisch bereits versorgt sind, am Esstisch ausharren. Dann werden die Textzeilen von Erziehern und Kindern mitgesungen und dazu die wechselnden Tätigkeiten mit entsprechenden Hand- oder Beinbewegungen angedeutet. Zu einem Zeitpunkt, an dem mir das Lied bereits hinreichend vertraut war, wurde ich darauf aufmerksam, wie einzelne Erzieherinnen situativ neue Strophen erfanden und in den übrigen „Alltag“ einbanden. So wurde etwa die Anweisung an die Kinder, mit dem Aufräumen der Spielsachen zu beginnen (generell ist diese Anweisung oft ein lautes Rufen in verschiedenen Betonungen), auch in Liedform gegeben, wobei dieses Singen oftmals auch das eigentliche gemeinsame Aufräumen noch begleitete: „Mir rau men, mir rau men, fänkt all mat mir un ...“.

Zu einem dieser Zeitpunkte, an denen eine Spielphase beendet wird und das allgemeine Aufräumen die Essenszeit vorbereiten soll, konnte ich beobachten, wie die Erzieherin Anouk besonders langgedehnte, unüberhörbare Ausrufe zum Besten gab. Sie stellte sich zunächst in die Raummitte und sang bzw. rief (das ist nicht entscheidbar) mehrere Male hintereinander, während sie ihren Kopf langsam drehte, die bekannte Formel in folgender Weise:

„Mir_ rau _____ men——“

Anouk intonierte lediglich einen Durdreiklang, eher eine Tritonusspanne, dies jedoch auffällig gleichbleibend in den Wiederholungen. Die erste Silbe hielt sie ungefähr vier Sekunden lang, die zweite Silbe ungefähr zwei Sekunden. Der repetierte Ruf wirkte auf mich wie eine „Sirene“, vergleichbar mit einem Feueralarm, einem Wecken in der Kaserne oder auch einem Gong nach der Opernpause. Die Länge der Silbe „rau“ sprengt die Grenzen von Sprechakten, die an Individuen gerichtet sein können, die „Sirene“ ist ein offizielles, administratives Signal. Dabei fallen in frappierender Weise die Semantik – *Raumen* – und die akustische Transformation des Raums zusammen, um eine soziale Ordnung der andauernden kollektiven Adressierung zu erzeugen, die in das materielle Transformieren des Raums als Ordnen der Spielsachen mündet.

Bei derart differenzierten Ausformungen von professionellen „Stimmen der Distanz“ (Zollna 2003) mag eine auditive Ethnographie der Pädagogik beinahe schon an musikethnologische Studien anknüpfen können, wie sie etwa zu den „arbeitsfunktionalen musikalischen Signalisierpraktiken“ von Viehhirten in Mittelgebirgslandschaften der DDR durchgeführt wurden (Kaden 1977). Allerdings ist in pädagogischen Praktiken stärker mit nicht-intendierten mimetischen Wirkungen der Signalkommunikation zu rechnen:

Nachdem Anouk aufgehört hat, steht Inka ebenfalls noch mitten im Raum, singt das sirenenartige „Mir raumen“ nach und schaut dabei langsam um sich. Sie scheint sich eher zur gemeinsamen Performance aufgefordert zu fühlen als zum Ordnung schaffen. Anouk wendet sich bald in ihre Richtung und schnauzt sie an: „Inka, dann *hëllef* raumen!“ (dann hilf aufräumen)

Rufsignale ertönen allenthalben auch in technisierter Form. So wird in den Gruppenräumen ein absteigender Durdreiklang abgespielt, wenn beispielsweise die Mutter oder der Vater eines Kindes die Klingel an der meist verschlossenen Haustür betätigt; hin und wieder imitieren auch Kinder die Melodie der Klingel.

Die Reihe der aufeinander beziehbaren Klangphänomene in der Einrichtung kann – zum Schluss – durchaus auch aus einer Sozialisationsperspektive betrachtet werden, die dann die Frage aufwirft, wie sich ästhetische und nicht-ästhetische Wahrnehmungsweisen in früher Kindheit ausdifferenzieren, zunächst aber auch, wie sie sich gegenseitig konstituieren. Was eine Konsonanz, was eine Dissonanz ist, scheint in einem kontinuierlichen Klanggeschehen zwischen Musik und Sprache unmittelbar *sozial* bestimmt zu sein und in diesen sozialen Kontexten gelernt zu werden. Die Pädagogik der Stimme (und des Klangs) im frühpädagogischen Feld erweist sich somit auch darin, dass musikalische Formen und ein stark musikalisches Sprechgeschehen nicht nur zeitlich und räumlich eng beieinander ablaufen, sondern auch in einem dichten praktischen und kommunikativen Verweisungszusammenhang stehen.

Anmerkungen

- 1 In diesem erziehungswissenschaftlichen Forschungsbereich gibt es bisher lediglich eine Studie zur Sprechgestik von Jugendlichen im Kontext der Schule, die an der „ethnography of speaking“ im Sinne von Hymes orientiert ist (vgl. Dietrich 2010).
- 2 Das Begleitforschungsprojekt „Betreuungswirklichkeit und Bildungswirklichkeit. Die Pädagogik der Maison Relais pour Enfants“ wurde im August 2009 von Michael-Sebastian Honig und Sascha Neumann an der Universität Luxemburg (Forschungsachse „Early Childhood: Education and Care“) ins Leben gerufen und hat eine Laufzeit bis Ende 2012. Dem Projektteam sowie meiner Kollegin Sabine Bollig und nicht zuletzt den Gutachtern und Herausgebern danke ich für wertvolle Hinweise zur ersten Fassung dieses Textes.
- 3 Eine stärkere Methodisierung des hier in Angriff genommenen Forschungsbereichs ist nicht ausgeschlossen. So kann etwa konkretisiert werden, wie entsprechende audielle bzw. audiovisuelle Aufnahmen gemacht werden sollten, welche Transkriptionstechniken sinnvoll sind und welche methodischen Analyseschritte erkenntnisversprechend und generalisierbar sind. Eine Differenzierung der Thesen und Ergebnisse auf Basis später gemachter Audioaufnahmen ist geplant. All dies ist jedoch notwendigerweise im Stadium der Exploration.
- 4 Die Frage der Transkription klanglicher und prosodischer Elemente der Kommunikation kann an dieser Stelle nicht grundsätzlich behandelt werden. Pragmatisch wird

sie hier mit der Verwendung eines effektiven und für die Analyse der hier fokussierten Phänomene völlig ausreichenden Zeichens gelöst, welches linguistischen und gesprächsanalytischen Kontexten entlehnt ist: Lange Striche zeigen an, dass die vorherige Silbe mit einer gleichbleibenden Tonhöhe, außerdem länger als für reguläres Sprechen gewöhnlich gesprochen wird; die Position dieser Striche in der Zeile (oben, in der Mitte, unten) zeigt einen relativen Tonhöhenunterschied an.

- 5 Die Linguistik hat die Stimmqualität über lange Zeit wenig beachtet und auch nur bedingt ihrem Gegenstandsbereich, der Sprache, zuordnen können, Stimmqualitäten sind enger mit Körperlichkeit verbunden und weniger arbiträr (vgl. Laver 1980). Ähnlich gilt dies für die Lautstärke; diese modifiziert aber vor allem auch die Räumlichkeit des Stimmklangs.
- 6 In der Musik, zumal in der Musikpädagogik, wird die kleine Terz häufig „Kuckucks-terz“ genannt. Mit dem entsprechenden Volkslied („Kuckuck ... ruft's aus dem Wald“) kann man sich leicht „in Erinnerung rufen“, wie das Intervall im tonalen Zusammenhang klingt.
- 7 Die melodische Form des Rufens oder Adressierens hat sicherlich kultur- und sprachspezifisch variante Ausprägungen, jedoch gibt es über verschiedene Sprachräume hinweg auch starke Gemeinsamkeiten (gerade auch in asiatischen Sprachen, etwa Japanisch, Thai), wie man aus einem Compendium von Abhandlungen über die Intonationssysteme in zwanzig Sprachen ersehen kann (Hirst/di Christo 1998). Die Intonationskontur des Rufens wird in der Regel mit dem von Ladd (1978) geprägten Begriff der „stylized intonation“ bezeichnet und beinhaltet stets auf einer Tonhöhe gehaltene, „gesungene“ Silben. (Ich folge nicht der weithin übernommenen Interpretation von Ladd, die besagt, der stereotypisierte Stimmgebrauch bedeute, dass die Botschaft *vorhersehbar* sei.) Bemerkenswert ist ebenfalls, dass die illustrierenden Exempel in verschiedenen Sprachen häufig im Zusammenhang mit der Adressierung von Kindern durch Erwachsene oder auch durch Kinder stehen. Ob sich auch für andere Sprachkulturen ein ähnlicher Zusammenhang zwischen Intervallstrukturen und der jeweiligen Musikkultur zeigen lässt, wie hier im Folgenden nahegelegt, wäre eine spannende Frage für Anschlussforschung.
- 8 Dabei mag der sich vergrößernde Tonabstand auch damit zu tun haben, dass der Sprechenden/rufenden Person hierdurch ein physiologischer Spannungsabfall, eine „Entladung“, ermöglicht wird. Möglicherweise ist eine willkommene, kultivierte Begleiterscheinung, dass hierdurch oft Intervalle produziert werden, die im konventionell-musikalischen Sinne gerade als sehr spannungsreich gelten (und zumindest von diesen Sprechenden/Rufenden aufgrund ihrer musikalischen Sozialisation auch so erlebt werden dürften). Es gibt meines Wissens keine Forschung zu diesem Zusammenhang, lediglich eine Kontroverse in der Linguistik darüber, ob die genauen Tonhöhenverhältnisse im Sprechen eine analytische Relevanz haben oder nicht.
- 9 Es handelt sich um ein Lied vom Album „Metti Spaghetti“, das moderne Kinderlieder in luxemburgischer Sprache enthält und in Luxemburg allseits bekannt ist. Als ich die Vermutung geäußert habe, dass die Musik dem Personal „doch sicher schon zu den Ohren herauskommen“ würde, hat mir die Einrichtungsleiterin entschieden widersprochen und ergänzt, auch sie kenne „Metti Spaghetti“ schon aus ihrer Kindheit, das sei ein absolutes „Muss“.

Literatur

- Abe, I. (1962): Call contours. In: Proceedings of the 4th International Congress of Phonetic Sciences. The Hague, pp. 519–523.
- Bilstein, J. (Hrsg.) (2011): Anthropologie und Pädagogik der Sinne. Opladen.
- Bollig, S. (2004): Zeigepraktiken. How to Do Quality with Things. In: Honig, M.-S./Joos, M./Schreiber, N. (Hrsg.): Was ist ein guter Kindergarten? Weinheim, S. 193–225.
- Bollig, S. (2011): Notizen machen, Bögen ausfüllen, Geschichten schreiben. Praxisanalytische Perspektiven auf die Materialität der bildungsbezogenen Beobachtung im Ele-

- mentarbereich. In: Cloos, P./Schulz, M. (Hrsg.): Kindliches Tun beobachten und dokumentieren. Perspektiven auf die Bildungsbegleitung in Kindertageseinrichtungen. Weinheim, S. 33–48.
- Bruner, J. S./Sherwood, V. (1978): Das Erlernen von Regelstrukturen in den frühesten Spielen von Mutter und Kind (Guckuck-da). In: Flitner, A. (Hrsg.): Das Kinderspiel. München, S. 158–167.
- Cloos, P./Schulz, M. (Hrsg.) (2011): Kindliches Tun beobachten und dokumentieren. Perspektiven auf die Bildungsbegleitung in Kindertageseinrichtungen. Weinheim.
- Dietrich, C. (2010): Zur Sprache kommen. Sprechgestik in jugendlichen Bildungsprozessen in und außerhalb der Schule. Weinheim.
- Erlmann, V. (2004): But what of the ethnographic ear? Anthropology, Sound, and the Senses. In: Erlmann, V. (ed.): Hearing Cultures. Essays on Sound, Listening, and Modernity. Oxford/New York, pp. 1–20.
- Fritzsche, B./Idel, T.-S./Rabenstein, K. (2011): Ordnungsbildung in pädagogischen Praktiken. Praxistheoretische Überlegungen zur Konstitution und Beobachtung von Lernkulturen. In: Zeitschrift für Soziologie der Erziehung und Sozialisation 31(1), S. 28–44.
- Gibbon, D. (1976): Perspectives of Intonation Analysis. Frankfurt a.M.
- Goffman, E. (1981): Forms of Talk. Pennsylvania.
- Goffman, E. (1982): Das Individuum im öffentlichen Austausch. Mikrostudien zur öffentlichen Ordnung. Frankfurt a.M.
- Gumperz, J. J. (1982): Discourse strategies. Cambridge.
- Hirst, D./Christo, A. d. (ed.) (1998): Intonation Systems. A Survey of Twenty Languages. Cambridge.
- Honig, M.-S. (2009): Das Kind der Kindheitsforschung. Gegenstandskonstitution in den *childhood studies*. In: Honig, M.-S. (Hrsg.): Ordnungen der Kindheit. Problemstellungen und Perspektiven der Kindheitsforschung. Weinheim, S. 25–51.
- Hünersdorf, B./Maeder, Ch./Müller, B. (2008): Ethnographie der Pädagogik: Eine Einführung. In: Hünersdorf, B./Maeder, Ch./Müller, B. (Hrsg.): Ethnographie und Erziehungswissenschaft. Methodologische Reflexionen und empirische Annäherungen. Weinheim, S. 11–28.
- Kaden, Ch. (1977): Hirtensignale – Musikalische Syntax und kommunikative Praxis. Leipzig.
- Kittler, F./Macho, Th./Weigel, S. (Hrsg.) (2002): Zwischen Rauschen und Offenbarung. Zur Kultur- und Mediengeschichte der Stimme. Berlin.
- Kolesch, D./Krämer, S. (Hrsg.) (2006): Stimme. Annäherung an ein Phänomen. Frankfurt a.M.
- Kolesch, D. (2009): Zwischenzonen. Leiblichkeit – Räumlichkeit – Aisthesis. In: Kolesch, D./Pinto, V./Schrödl, J. (Hrsg.): Stimm-Welten. Philosophische, medientheoretische und ästhetische Perspektiven. Bielefeld, S. 13–22.
- Krämer, S. (2003): „Schriftbildlichkeit“ oder: Über eine (fast) vergessene Dimension der Schrift. In: Krämer, S./Bredekamp, H. (Hrsg.): Bild, Schrift, Zahl. München, S. 157–176.
- Kress, G. (2009): Multimodality: A Social Semiotic Approach to Contemporary Communication. London.
- Ladd, R. D. (1978): Stylized Intonation. In: Language 54(3), pp. 517–540.
- Laver, J. (1980): The phonetic description of voice quality. Cambridge.
- Leeuwen, T. van (1999): Speech, Music, Sound. Houndmills.
- Liegle, L. (2006): Bildung und Erziehung in früher Kindheit. Stuttgart.
- Löw, M. (2001): Raumsoziologie. Frankfurt a.M.
- Meyer, P. (Hrsg.) (2008): acoustic turn. München.
- Neumann, S. (2011): Ordnungen der Erziehungswirklichkeit. Einführung in den Themenschwerpunkt. In: Zeitschrift für Soziologie der Erziehung und Sozialisation 31(1), S. 4–11.
- Neumann, S. (2012): Beobachtungsverhältnisse. Feldtheoretische Erkundungen zu einer Empirie des Pädagogischen. In: Bernhard, St./Schmidt-Wellenburg, Ch. (Hrsg.): Feld-

- analyse als Forschungsprogramm. Band 1: Der programmatische Kern. Wiesbaden (im Erscheinen).
- Reckwitz, A. (2008): Grundelemente einer Theorie sozialer Praktiken. In: Reckwitz, A. (Hrsg.): *Unschärfe Grenzen. Perspektiven der Kulturosoziologie*. Bielefeld, S. 97–130.
- Rustemeyer, D. (2003): Medialität des Sinns. In: Rustemeyer, D. (Hrsg.): *Bildlichkeit. Aspekte einer Theorie der Darstellung*. Würzburg.
- Schegloff, E. A. (1968): Sequencing in Conversational Openings. In: *American Anthropologist* 70(6), pp. 1075–1095.
- Schmidt, K. (2004): Das Freispiel und der geordnete Raum. Die Praxis eines Programms. In: Honig, M.-S./Joos, M./Schreiber, N. (Hrsg.): *Was ist ein guter Kindergarten?* Weinheim, S. 157–192.
- Schmitz, H. (2008): Leibliche Kommunikation im Medium des Schalls. In: Meyer, P. (Hrsg.): *acoustic turn*. München, S. 75–88.
- Schnoor, O. (2007): Sammelbesprechung: Bildinterpretation im erziehungswissenschaftlichen Kanon? In: *Zeitschrift für Erziehungswissenschaft* 10(4), S. 581–592.
- Schulze, H./Wulf, Ch. (Hrsg.) (2007): *Klanganthropologie. Performativität – Imagination – Narration*. Paragrana 16(2).
- Schützeichel, R. (2010): Soziologie der Stimme. Über den Körper in der Kommunikation. In: Keller, R./Meuser, M. (Hrsg.): *Körperwissen*. Wiesbaden, S. 85–104.
- Zollna, I. (2003): *Stimmen der Distanz. Professionelle monologische Sprechstile. Eine vergleichende Untersuchung zu Wiederholung und Expressivität. Prosodische Gestaltung in spanischen, französischen, englischen und deutschen Gebeten, Durchsagen und Verkaufsrufen*. Tübingen.
- Zwicky, A.M. (1974): Hey, Whatsyourname! In: *Chicago Linguistic Society* 10, pp. 787–801.