

Lehr-/Lernprozesse in sprachlich heterogenen Lerngruppen
von ‚Deutsch als Zweitsprache‘-Kursen der
Erwachsenenbildung und der Schule

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von
Moritz Stephan Sahlender
aus Thannhausen

Tübingen
Dezember 2022

Tag der mündlichen Prüfung:

26.03.2024

Dekan:

Prof. Dr. Ansgar Thiel

1. Gutachter:

Prof. Dr. Josef Schrader

2. Gutachter:

Prof. Dr. Hannes Schröter

Danksagung

Ich möchte mich an erster Stelle ausdrücklich bei meinem Doktorvater Prof. Dr. Josef Schrader und meinem Zweitbetreuer Prof. Dr. Hannes Schröter für die Ermöglichung der vorliegenden Arbeit und die vertrauensvolle Zusammenarbeit bedanken. Vielen Dank für die konstruktive Unterstützung und die stets klugen Ratschläge, die entschieden zum Gelingen dieser Arbeit beigetragen haben.

Mein zweiter Dank gilt den Doktorandinnen aus dem COLD-Projekt Inga ten Hagen, Zarah Weiss, Anastasia Knaus und Stefanie Helbert für die intensive Zusammenarbeit im Projekt und in den unzähligen Videokonferenzen, die uns institutsübergreifend und pandemiebegleitend einen fruchtbaren wissenschaftlichen Austausch ermöglicht haben. Ich danke auch Zeena Abdulghani, die durch ihren unermüdlichen Einsatz als studentische Hilfskraft über die gesamte Projektlaufzeit hinweg einen wertvollen Beitrag zur Datenaufbereitung der vorliegenden Studien geleistet hat. Auch möchte ich mich bei den Lehrkräften und den Lernenden aus den Sprachkursen bedanken, die durch ihre Bereitschaft zur Studienteilnahme die Einblicke in den realen Unterrichtsalltag erst ermöglicht haben.

Ich möchte mich bei meinen Arbeitskolleginnen und -kollegen am Deutschen Institut für Erwachsenenbildung bedanken, die mir in vielen fachlichen und überfachlichen Diskussionen hilfreiche Anregungen gegeben haben; insbesondere bei Dr. Martin Merkt und Magdalena Schmitz für die sehr wertvolle Unterstützung in der Schlussphase dieser Arbeit.

Mein letzter Dank gilt meinen beiden lieben Eltern Claudia und Michael, die mich zeitlebens ermutigt haben, den Dingen, die mich begeistern auf den Grund zu gehen, meinen Freunden, die mir den Alltag so erholsam machen, und vor allem meiner Partnerin Natalie und meiner Tochter Marie, dafür, dass sie mir auf ihre liebevolle Weise so viel Kraft gegeben haben, dass diese Arbeit möglich wurde. Danke!

Zusammenfassung

Das Beherrschen der Sprache ist ein zentraler Schlüssel für die Bewohnerinnen und Bewohner eines Landes beim Zugang zu Bildung, beruflichem Erfolg und gesellschaftlicher Teilhabe. Um diese gesamtgesellschaftliche Aufgabe zu erfüllen, wird der Erwerb der deutschen Verkehrs- und Bildungssprache sowohl für zugewanderte Menschen als auch für Menschen, die in Deutschland aufgewachsen sind, in Form von Sprachunterricht angeboten. Damit Lehrkräfte Spracherwerbsprozesse erfolgreich gestalten können, ist mehr Wissen über die Gelingensbedingungen der im Sprachunterricht stattfindenden Lehr-/Lernprozesse notwendig. Dies besonders, da die Herausforderungen für die Lehrkräfte beim Sprachunterricht durch eine zunehmende sprachliche Heterogenität der Lernenden eher steigen. Der erste Beitrag der vorliegenden kumulativen Dissertationsschrift bietet einen breit angelegten Literaturüberblick, um die Frage zu beantworten, welche Bedingungen die aktuelle Forschung für einen erfolgreichen Spracherwerb von Zugewanderten und gering literalisierten Menschen bereits identifiziert hat. Es fehlt an empirisch fundierten Informationen sowohl über erfolgreiche Gelingensbedingungen des Spracherwerbs als auch über die Frage, wie Lehrkräfte Spracherwerb erfolgreich unterstützen können. Zwar gibt es eine hohe Anzahl an Forschungsaktivitäten aus verschiedenen Disziplinen, jedoch ohne wechselseitige Bezugnahme der Befunde aufeinander.

Im einleitenden Rahmentext wird zunächst auf die Bedeutung von Sprache und Spracherwerb eingegangen. Anschließend, nachdem die grundsätzlichen Dimensionen von Sprachunterricht vorgestellt wurden, werden drei zentrale Sprachunterrichtsformen (Alphabetisierungskurse, Vorbereitungsklassen und Integrationskurse) dargestellt. Um die in Beitrag #1 identifizierten Desiderate einzuordnen, werden zunächst die Lehr-/Lernprozesse im Unterricht in das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts eingeordnet, um in Orientierung daran auf die Rolle der Lehrkraft und die besondere Doppelfunktion von Sprache als Inhalt und Medium im Sprachunterricht einzugehen. Auch gelingt es, anhand des Modells zu zeigen, wie unterschiedlich und vielschichtig die sprachliche Heterogenität im Sprachunterricht wirken kann, und es wird aufgeführt, welche Möglichkeiten und Herausforderungen die Lehrkraft hat, diese Heterogenität zu adressieren. Empirische Befunde, die sich an diesem Modell orientieren, ermöglichen es, unter Zuhilfenahme verschiedener soziokultureller und kommunikationstheoretischer Konzepte, Konsequenzen für den Umgang mit sprachlich heterogenen Lerngruppen im Sprachunterricht abzuleiten. Eine Möglichkeit ist die Adaptivität des Lehrkräfthandelns. Dies wird

abschließend anhand Lehrkraftgesten, als ein Teil der nonverbalen Kommunikation von Lehrkräften, erörtert. Das dritte Kapitel der Einführung behandelt die methodische Untersuchung von Lehr-/Lernprozessen durch videogestützte Unterrichtsbeobachtung in der empirischen Bildungsforschung, die Erstellung von gesprochenen Lernerkorpora in der (Computer-)Linguistik sowie schlussendlich die Beschreibung der Datenerhebung für die Datengrundlage des zweiten und dritten Beitrags. In Beitrag #2 werden Forschungsdesiderate aufgegriffen, die im ersten Beitrag identifiziert wurden. So wird am Thema Deutsch als Zweitsprache (DaZ) untersucht, wie videogestützte Unterrichtsaufnahmen so aufbereitet werden können, dass das Handeln der Lehrkraft in Lehr-/Lernsituationen unter besonderer Berücksichtigung der sprachlichen Heterogenität analysiert werden kann. Durch interdisziplinäre Zusammenarbeit zwischen der Erwachsenenbildungsforschung, der Computerlinguistik, der Psychologie und der DaZ-Didaktik wurde dabei ein Transkriptions- und Annotationsleitfaden entwickelt, der es ermöglicht, Lehr-/Lernprozesse in sprachlich heterogenen Lerngruppen in der Erwachsenenbildung in Integrationskursen sowie in der Schule in Vorbereitungsklassen der Schule zu untersuchen. Im Rahmen des dritten Beitrags findet das Instrument Anwendung in der konkreten Analyse von Lehrkraft Handeln in Lehr-/Lernprozessen. Am Beispiel von zehn aufbereiteten Unterrichtsvideos aus Integrationskursen und Vorbereitungsklassen kann gezeigt werden, dass Lehrkräfte ihre nonverbale Kommunikation in Teilen an die sprachlichen Fähigkeiten der Lernenden anpassen und je nach sprachlicher Leistungsfähigkeit der einzelnen Lernenden unterschiedlich komplexe Gesten verwenden.

Abschließend werden in einer Gesamtdiskussion die Befunde der drei Beiträge diskutiert und in den Forschungszusammenhang eingeordnet. Dabei erfolgt neben der kritischen Betrachtung der Befunde auch die Auflistung möglicher Anknüpfungspunkte an die in dieser Arbeit aufgeführten Untersuchungen zu Lehr-/Lernprozessen in sprachlich heterogenen Lerngruppen von DaZ-Kursen der Erwachsenenbildung und der Schule.

Inhaltsverzeichnis

1. Einleitung	1
2. Sprache und Spracherwerb	4
2.1. Inhalte und Gegenstandsbereiche des Sprachunterrichts	4
2.2. Alphabetisierungskurse für Menschen ohne Zuwanderungsgeschichte	6
2.3. Vorbereitungsklassen für Menschen mit Zuwanderungsgeschichte	8
2.4. Integrationskurse für Menschen mit Zuwanderungsgeschichte	9
3. Lehr-/Lernprozesse im Unterricht	12
3.1. Lehr-/Lernprozesse und Unterrichtsqualität	13
3.2. Die Rolle der Lehrkraft beim Spracherwerb	15
3.3. (Sprachliche) Heterogenität im Unterricht	16
3.4. (Sprachliche) Heterogenität als Herausforderung für die Lehrkraft	19
3.5. Lehrkraftgesten und adaptive Lehr-/Lernprozesse	21
3.6. Lehrkraftgesten und sprachliche Heterogenität	26
4. Videogestützte Erfassung von Lehr-/Lernprozessen im Unterricht	29
4.1. Videogestützte Unterrichtsforschung in der empirischen Bildungsforschung und der Erwachsenenbildung	30
4.2. Interdisziplinäre Ansätze zur Analyse von Lehr-/Lernprozessen	33
4.3. Kompetenzen von Lehrkräften in Schule und Erwachsenenbildung im Bereich Deutsch als Zweitsprache beim Unterrichten von sprachlich heterogenen Lerngruppen	35
5. Fragestellungen der empirischen Studien	37
6. Beitrag #1 Spracherwerb bei Zugewanderten und funktionalen Analphabeten	40
7. Beitrag #2 Investigating spoken classroom interactions	68
8. Beitrag #3 Do teachers adapt their gestures?	133
9. Gesamtdiskussion – Fragestellungen, Befunde und Perspektiven	170
9.1. Zusammenfassung und Einordnung zentraler Ergebnisse	178
9.2. Kritische Betrachtung der Studien und Desiderate für zukünftige Forschung	183
9.3. Fazit	189
10. Literaturverzeichnis	191

1. Einleitung

Der erfolgreiche Spracherwerb der jeweiligen Landessprache ist für die Bewohnerinnen und Bewohner eines Landes ein relevanter Faktor für eine positive Entwicklung der Bildung, des beruflichen Erfolgs und der gesellschaftlichen Partizipation. Die Gewährleistung, dass die Landessprache beherrscht wird, ist deshalb eine wichtige gesamtgesellschaftliche Aufgabe. Ziel der vorliegenden Dissertationsschrift ist es, die *Lehr-/Lernprozesse in sprachlich heterogenen Lerngruppen von ‚Deutsch als Zweitsprache (DaZ)‘-Kursen in der Erwachsenenbildung und der Schule* zu untersuchen und dadurch ein besseres Verständnis über Lehr- und Lernprozesse in Spracherwerbssituationen zu gewinnen. Die Arbeit ist im Rahmen einer Anstellung im Projekt *Competencies of School Teachers and Adult Educators in Teaching German as a Second Language in Linguistically Diverse Classrooms (COLD)* entstanden. Das Verbundprojekt wurde im Rahmen des Förderprogramms *Kooperative Exzellenz* mit Mitteln des Leibniz-Wettbewerbs von 2019–2022 finanziert (Fördernummer: K113/2018) und vom *Deutschen Institut für Erwachsenenbildung – Leibniz Zentrum für Lebenslanges Lernen e.V.* gemeinsam mit dem *Mercator Institut für Sprachförderung und Deutsch als Zweitsprache* initiiert. Beteiligt waren zudem das DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, das IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, sowie die Universitäten Dortmund, Köln, Tübingen und Zürich. In COLD wurden die professionellen Kompetenzen von Lehrkräften in Schule und Erwachsenenbildung beim Unterrichten von Deutsch als Zweitsprache in sprachlich heterogenen Lerngruppen unter anderem videogestützt erfasst. Das Projekt adressiert dabei die speziellen Anforderungen an Lehre und Didaktik, die bildungsbereichsübergreifend durch neu zugewanderte Kinder, Jugendliche und Erwachsene entstanden sind. Die Untersuchungen fanden dabei in authentischen Unterrichtsettings mit Lehrkräften in Vorbereitungsklassen und Integrationskursen statt. Die im Rahmen des COLD-Projekt entstandenen Daten erlauben eine intensive empirische Analyse von Lehr-/Lernprozessen in Spracherwerbssituationen. Die Innovation der Datenaufbereitung liegt dabei in der Erfassung der Information über die Empfänger der im Unterricht getätigten verbalen Aussagen und nonverbalen Handlungen. Dadurch kann differenziertes Lehrkrafthandeln sicht- und messbar gemacht werden. Beitrag #2 und Beitrag #3 der vorliegenden Arbeit sind im Rahmen des Projekts COLD entstanden.

Der erste Beitrag bildet in Form eines Reviews eine Analyse des Forschungsstandes und liefert durch die darin durchgeführte Untersuchung Beschreibungswissen über ein Desiderat für die Lehr-/Lernforschung: fehlende empirisch gesicherte Ergebnisse über Gelingensbedingungen erfolgreichen Spracherwerbs bei gleichzeitig hohem interdisziplinärem Potenzial aus verschiedenen Nachbardisziplinen der Erwachsenenbildung. Das Fazit des Beitrags ist die Empfehlung, mehr Erklärungswissen (Bromme et al., 2017) über die Gelingensbedingungen erfolgreichen Spracherwerbs bereitzustellen.

Im zweiten Beitrag werden die in Beitrag #1 aufgeführten Desiderate aufgegriffen und über einen interdisziplinären Zugang ein Instrument (in Form eines Leitfadens zur Analyse von Unterrichtsvideos) entwickelt, mit dem die wachsenden Herausforderungen bearbeitet werden können, denen sich Lehrkräfte beim Unterrichten aufgrund einer zunehmenden sprachlichen Heterogenität gegenübersehen. Dadurch wird die empirische Analyse von Lehrkraftshandeln in sprachlich heterogenen Lerngruppen ermöglicht. Ferner wird das Instrument auf wissenschaftliche Gütekriterien geprüft, damit es sowohl in der Bildungsforschung als auch in der (Computer-)Linguistik Weiterverwendung finden kann.

Der dritte Beitrag beschreibt schließlich die exemplarische Anwendung des Instruments an einer konkreten Forschungsfrage, die auf die nonverbale Kommunikation von Lehrkräften gerichtet ist und die auf Erklärungswissen über den adaptiven Einsatz von Gesten bei Lehrkräften im Sprachunterricht zielt. Und zwar durch die Beantwortung der Frage, wie und ob Lehrkräfte die sprachliche Heterogenität ihrer Lernenden in ihrem Unterrichtshandeln berücksichtigen. Es wird gezeigt, dass Lehrkräfte ihre nonverbale Kommunikation (am Beispiel von Gesten) in Teilen an die sprachlichen Fähigkeiten der Lernenden individuell anpassen.

Im Vorfeld der Beiträge werden die Phänomene Sprache und Spracherwerb beleuchtet und im Kontext von Unterricht als organisierter Bildungsprozess verortet. Dadurch kann die Bedeutung von sprachlicher Heterogenität als Herausforderung für Lehrkräfte wissenschaftlich eingeordnet und bearbeitbar gemacht werden. Ebenfalls wird aufgezeigt, wie Lehrkräfte durch adaptives Unterrichtshandeln verbal und nonverbal die sprachliche Heterogenität adressieren können. Nachdem im anschließenden Teil Möglichkeiten aufgezeigt werden, wie diese Lehr- und Lernprozesse mit videogestützter Unterrichtsbeobachtung in der Bildungsforschung und aus interdisziplinärer Perspektive mit Hilfe von gesprochenen Lerner-Korpora der (Computer-)Linguistik empirisch messbar gemacht werden können, wird abschließend das

Verbundprojekt COLD vorgestellt, in dessen Rahmen auch der zweite und dritte Beitrag der Dissertationsschrift erstellt worden sind.

2. Sprache und Spracherwerb

International sind Migrationsbewegungen in wachsendem Ausmaß beobachtbar, ausgelöst durch Krisen wie etwa veränderte klimatische Bedingungen und damit einhergehende Hungersnöte, Kriege um die Verteilung knapper Ressourcen, politische Instabilitäten. Seit dem Jahr 2010 verzeichnet das Statistische Bundesamt über 16,7 Millionen Zuzüge¹ von Zuwanderinnen und Zuwanderern nach Deutschland (Statistisches Bundesamt [Destatis], 2022).

Ein zentraler Schlüssel, um Menschen erfolgreich in eine Gesellschaft zu integrieren, ist die Sprache. Denn wenn Menschen die Verkehrs- bzw. Bildungssprache des Landes, in dem sie leben, beherrschen, verbessern sie ihre Chancen auf dem Arbeitsmarkt und erfahren gesellschaftliche Anerkennung und Partizipationsmöglichkeiten (z. B. Bleakley & Chin, 2004; Guven & Islam, 2015). Darüber hinaus fördert ein erfolgreicher Spracherwerb die Identifikation mit der Gesellschaft und die soziale Teilhabe an der Gesellschaft (Esser, 2006; Hochman & Davidov, 2014; Martinovic et al., 2009). Ebenfalls sind Sprachkenntnisse mit Erfolg im Bereich Bildung verknüpft (für einen Überblick siehe Esser, 2006). Das Beherrschen der jeweiligen Landessprache ist somit eine zentrale Fähigkeit, die durch die steigende Zahl an Zuwanderinnen und Zuwanderern in den letzten Jahren zunehmend Bedeutung erlangt. Dies gilt sowohl für gesprochene als auch für schriftsprachliche Fähigkeiten und hier zeigt sich für Deutschland, dass neben den migrationsbedingten Erfordernissen für erfolgreiche Spracherwerbsbiographien auch weitere Handlungsnotwendigkeiten bestehen. So zeigt die jüngste LEO-Studie, dass es in Deutschland mehr als 6,2 Millionen Menschen gibt, die Schriftsprache nicht ausreichend sinnentnehmend verstehen können. Das entspricht 12,1 Prozent der erwerbsfähigen Bevölkerung. Knapp die Hälfte dieser Personen, 2,9 Millionen Menschen (47,4 %), sind Zugewanderte, die nicht in deutschsprachenden familiären Umfeldern aufgewachsen sind (Grotlüschen & Buddeberg, 2020).

2.1. Inhalte und Gegenstandsbereiche des Sprachunterrichts

Bei der Beschreibung des Lehrens und Lernens von Sprachen wie auch bei der Beurteilung von Sprachkompetenzen wird in der Regel auf den Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER) verwiesen. Neben einer

¹Im Jahr 2021 waren es etwas mehr als 1,3 Millionen Zuzüge

länderübergreifenden Vergleichbarkeit von sprachlichen Qualifikationen (A1 und A2 = elementare, B1 und B2 = selbstständige, C1 und C2 = kompetente Sprachverwendung) beschreibt dieser auch die Inhalte und Bereiche des Spracherwerbs bei Lernenden (Trim et al., 2009). Zusätzlich zu der sechsstufigen Globalskala ermöglicht der GER die verschiedenen Gegenstandsbereiche des Sprachunterrichts dezidiert zu beschreiben. So finden sich für alle Niveaustufen detaillierte Beschreibungen über die Fähigkeiten der Lernenden beim Lese- und Hörverstehen, beim Sprechen und beim Schreiben. Weiterhin werden im Bereich des mündlichen Sprachgebrauchs explizit qualitative Aspekte hinsichtlich des Spektrums, der Korrektheit, der Redeflüssigkeit, der Interaktion und der Kohärenz definiert (Trim et al., 2009).

Das grundlegende Ziel des DaZ-Unterrichts ist die Aneignung der Kommunikationsfähigkeit in der Zielsprache Deutsch (Kniffka & Siebert-Ott, 2012). Die Aneignung der Kommunikationsfähigkeit lässt sich dabei durch die verschiedenen Niveaustufen des GER einschätzen. Die Fachdidaktik des DaZ-Unterrichts berücksichtigt bei der Curriculums-Entwicklung und in der Gestaltung von Unterrichtsstunden die Förderung der vier sprachlichen Dimensionen *Hören, Sprechen, Lesen* und *Schreiben*, weshalb sich DaZ-bezogene Unterrichtsinhalte auf den Erwerb dieser Fähigkeiten oder den Erwerb von Strategien zu diesen Fähigkeiten einteilen lassen (Jeuk, 2015). Diese lassen sich weiterhin in produktive, rezeptive, interaktive und sprachmittelnde Aktivitäten unterteilen, wodurch alle sprachbezogenen Aktivitäten des Unterrichts erfasst werden können (Trim et al., 2009). Sprechen und Schreiben werden dabei als *produktive Aktivitäten* zusammengefasst und beziehen sich auf alle Arten der mündlichen und schriftlichen Produktion von Sprache (z.B. zusammenhängendes monologisches Sprechen und Schreiben, argumentieren, öffentliche Ankündigungen machen und vor Publikum sprechen sowie das Verfassen von Berichten und Aufsätzen). Hören und Lesen werden als *rezeptive Aktivitäten* zusammengefasst und beziehen sich auf alle Arten des Hör- und Leseverstehens (z.B. Verstehen von Gesprächen zwischen Muttersprachlern, Ankündigungen, Durchsagen und Anweisungen, Radiosendungen und Tonaufnahmen sowie das Verstehen von schriftlichen Korrespondenzen, Informationen und Anweisungen). Mündliche und schriftliche Interaktionen werden darüber hinaus als *interaktive Aktivitäten* zusammengefasst und verknüpfen produktive und rezeptive Aktivitäten (z.B. formelle und informelle Konversationen und Diskussionen, Informationsaustausch und Interviewgespräche oder die Interaktionen via Notizen, Mitteilungen und Formulare). Formen der Vermittlung zwischen Gesprächspartnern verschiedener Sprachen werden als *sprachmittelnde Aktivitäten* zusammengefasst (z.B.

das formelle und informelle Dolmetschen in sozialen Situationen oder die Übersetzung von Verträgen und das Paraphrasieren von Texten) (Trim et al., 2009).

In Deutschland ist die Gewährleistung des Spracherwerbs für Deutsch als Erstsprache, Deutsch als Zweitsprache sowie Deutsch als Bildungssprache (Gogolin & Lange, 2011) eine zentrale gemeinsame, bildungsbereichsübergreifende Aufgabe, die je nach Bildungsbereich unterschiedlich ausgestaltet ist. In der Schule wird DaZ in sogenannten Vorbereitungsklassen unterrichtet. Erwachsene Zugewanderte können im Rahmen der Integrationskurse die deutsche Sprache erlernen. Erwachsene ohne Zuwanderungsgeschichte können in Alphabetisierungskursen Deutsch lernen bzw. ihre Deutschkenntnisse verbessern.

Im Folgenden werden die verschiedenen Sprachlernformate für Deutsch als Erstsprache für Erwachsene (Alphabetisierungskurse), für DaZ für zugewanderte Schülerinnen und Schüler in Vorbereitungsklassen sowie für DaZ für zugewanderte Erwachsene in Integrationskursen vorgestellt. Diese drei Formen organisierter Spracherwerbsprozesse bilden die inhaltliche Basis des ersten Beitrags der Dissertationsschrift. In Beitrag #2 und Beitrag #3 wird der thematische Fokus auf den DaZ-Bereich ausgerichtet. Vorbereitungsklassen und Integrationskurse rücken dabei in den Vordergrund.

2.2. Alphabetisierungskurse für Menschen ohne Zuwanderungsgeschichte

Wie eingangs erwähnt gibt es in Deutschland eine große Zahl an Personen, die als gering literalisiert bezeichnet werden. Diese Menschen verfügen nur über unzureichende schriftsprachliche Kompetenzen und haben häufig auch Defizite in der sinnentnehmenden Verarbeitung von Schriftsprache. Geringe Literalisierung alleine ist noch kein Kriterium für soziale Ausgrenzung, jedoch gibt es, je nach Schwere der Defizite, unterschiedlich starke Ausschlussmechanismen, deren Funktionsweise noch genauer untersucht werden muss (Grotlüschen et al., 2020).

Da es keine Gesamtzahlen über Belegungszahlen von Alphabetisierungskursen gibt, hilft ein Blick in die VHS-Statistik des Jahres 2020 des Deutschen Instituts für Erwachsenenbildung, da diese eine umfassende Datensammlung zu Teilnahmen an Alphabetisierungskursen liefert (Echarti et al., 2022). Die Statistik zeigt für den dokumentierten Bereich (848 von 895 Volkshochschulen) die Anzahl von Teilnahmen in Alphabetisierungskursen in den beiden Programmbereichen Sprachen sowie Grundbildung. Addiert man die Anzahl der Teilnehmenden dieser beiden Bereiche aus dem

Jahr 2020 und berücksichtigt die Teilnehmenden aus den Integrationskursen mit Alphabetisierung nicht, da diese qua Definition zugewandert sind, erhält man eine Zahl von 15.697 Teilnehmenden aus Alphabetisierungskursen aus dem Grundbildungsbe- reich und 6.749 Teilnehmenden aus Alphabetisierungskursen im Programmbereich Sprachen. Weitere 3.102 Menschen gaben an, im Sprachbereich einen ‚Deutsch als Muttersprache‘-Kurs zu besuchen. Insgesamt ergibt das 25.548 Teilnehmende in 3.294 durchgeführten Alphabetisierungskursen (Echarti et al., 2022). Vergleicht man diese Zahl mit den 3,2 Millionen als gering literarisiert identifizierten Menschen, die in „deutsch sprechenden familiären Umfeldern aufgewachsen sind“ (Grotlüschen et al., 2020, S. 23), bedeutet dies, dass nur 0,79 % aller gering literarisierten Menschen einen Deutschsprachkurs besuchen. Nun gibt es neben den Volkshochschulen noch weitere Träger, die Deutschkurse für Menschen ohne Migrationshintergrund anbieten, die Volkshochschulen sind jedoch in diesem Bereich der größte Anbieter. Daher ist davon auszugehen, dass die wahre Zahl an Teilnehmenden an Alphabetisierungskur- sen nicht sehr viel größer ist als der Wert von 0,77 %, der in der LEO2018 Studie identifiziert wurde. Zwar gehen die Autoren der LEO2018 Studie in einer Teilfrage ex- plizit auf die überraschend geringe Zahl ein, kommen allerdings zu einem wenig be- lastbaren Fazit, das ebenfalls nur eine Vermutung darstellt: „Insgesamt findet sich re- lativ wenig *Alphabetisierung*, demgegenüber jedoch entschieden höhere Belegungs- zahlen in der *Grundbildung*. Letztere zielt allerdings nicht unbedingt auf Literalität, sie kann insofern ein Einflussfaktor auf die verbesserten Werte [im Vergleich zur vorheri- gen LEO-Studie; Anmerkung des Autors] sein, muss es aber nicht.“ (Grotlüschen et al., 2020, S. 50) Eine weitere Möglichkeit, die Teilnahmezahlen aus der Perspektive der Kursanbieter einzuschätzen, erlaubt die DIE-Verbundstatistik (Horn et al., 2021). Sie liefert Daten zur gemeinsamen Erfassung von Alphabetisierungs- und Grundbil- dungskursen und erfasst 53.595 Teilnahmen in 5.938 Kursen im Bereich Grundbil- dung und Alphabetisierung im Jahr 2019. Berücksichtigt man, dass es 6,2 Millionen gering literarisierte Menschen in Deutschland gibt, wovon 3,2 Millionen in einem Deutsch sprechenden Umfeld aufgewachsen sind, so beläuft sich die Quote von Per- sonen, die an einer Alphabetisierungsveranstaltung teilgenommen haben, auf 0,86 %. Die Grundgesamtheit der Studie bilden 1.696 Einrichtungen.

Es gibt intensive bildungspolitische Bemühungen, die Zahl der gering literali- sierten Menschen mit und ohne Migrationshintergrund zu verringern. So wurde mit dem Einberufen der ‚AlphaDekade‘ ein Instrument entwickelt, um gezielt Projekte zur Erforschung und Entwicklung von Lösungsansätzen für eine Verbesserung der

Situation gering literalisierter Menschen zu ermöglichen. Im Zeitraum von zehn Jahren wird von 2016 bis 2026 eine Vielzahl an Projekten gefördert. Die Bundesregierung hat für den gesamten Förderzeitraum ein Finanzvolumen von 180 Millionen Euro veranschlagt (Deutscher Bundestag, 2019). Allein im Jahr 2020 wurden insgesamt 1.155 Angebote für Lernende, Multiplikatorinnen und Multiplikatoren und Lehrende durchgeführt. Erreicht wurden dabei nach Angaben der Koordinierungsstelle der AlphaDekade über 15.000 Personen (Koordinierungsstelle der AlphaDekade, 2022).

Zusammengefasst lässt sich festhalten, dass mit Blick auf die Spracherwerbssituationen von Zugewanderten und bereits in Deutschland lebenden Menschen ohne ausreichende Kenntnisse in der deutschen Sprache deutlich wird, dass es eine Vielzahl an Herausforderungen gibt, um Spracherwerb erfolgreicher zu fördern. Neben der großen Zahl an Zugewanderten, die die Lehrkräfte durch die reine Menge an Lernenden vor große Herausforderungen stellt, ist auch die große sprachliche Heterogenität eine lehr-/lerntheoretisch einflussreiche Größe bei der Bedeutung des Unterrichtsgeschehens.

2.3. Vorbereitungsklassen für Menschen mit Zuwanderungsgeschichte

Neu zugewanderten Schulpflichtigen, die noch nicht über ausreichende sprachliche Deutschfähigkeiten verfügen lernen DaZ in sogenannten Vorbereitungsklassen (El-Mafaalani et al., 2020). Diese Klassen gibt es deutschlandweit in verschiedenen Bezeichnungen: Vorbereitungsklassen, Willkommensklassen, Sprachfördergruppen, Förderkurse (Dewitz & Bredthauer, 2020). Die jeweiligen gesetzlichen Vorgaben zur Ausgestaltung sind dem zuständigen Kultusministerium überlassen und somit in jedem Bundesland unterschiedlich ausgestaltet. Das Spektrum reicht von getrenntem Unterricht, bei dem die Kinder nur in ihrem Vorbereitungs-Klassenverband verbleiben, oder integrativen Modellen, bei denen die Lernenden in Teilen am Regelunterricht partizipieren, bis zu Modellen ohne separate Klassen, bei denen die Kinder vollständig im Regelunterricht lernen und zusätzliche Förderung erhalten (Karakayali & Heller, 2020). Bundesweite Zahlen über die Anzahl von Schülerinnen und Schülern in Vorbereitungsklassen sind aufgrund der landesspezifischen Regelungen und einer fehlenden bundesweiten Schulstatistik mit Informationen zum sprachlichen Hintergrund der Lernenden kaum vorhanden (Massumi & Dewitz, 2015).

2.4. Integrationskurse für Menschen mit Zuwanderungsgeschichte

Zugewanderte, die nicht mehr schulpflichtig sind, können in Deutschland einen Integrationskurs besuchen, um die deutsche Sprache zu erlernen. Seit dem Inkrafttreten des Zuwanderungsgesetzes 2005 haben ca. 3,4 Millionen Menschen in Deutschland eine Teilnahmeberechtigung für einen Integrationskurs erhalten. Im Jahr 2020 waren es 168.279 teilnahmeberechtigte Personen (Bundesamt für Migration und Flüchtlinge, 2021). Die größten Anbieter für Integrationskurse sind Volkshochschulen und Sprach-/Fachschulen. Im Jahr 2021 haben 36,2 % aller Lernenden Kurse an Volkshochschulen bzw. 21,8 % aller Lernenden Kurse an Sprach-/ Fachschulen besucht (Bundesamt für Migration und Flüchtlinge, 2022).

Ein allgemeiner Integrationskurs besteht aus insgesamt 700 Unterrichtsstunden à 45 Minuten und teilt sich in einen Basis- (300 Unterrichtsstunden) und in einen Aufbaukurs (300 Unterrichtsstunden) auf. Anschließend daran folgt ein hundertständiger Orientierungskurs, in dem Kenntnisse über die grundlegenden Werte der Gesellschaft, der Kultur, der Geschichte sowie der politischen Institutionen Deutschlands vermittelt werden. Der Integrationskurs wird mit dem Deutschtest für Zuwanderer (DTZ) abgeschlossen, wobei in einem schriftlichen und mündlichen Teil die zentralen Bereiche des GER (Hören, Schreiben, Lesen und Sprechen) abgedeckt werden.

Neben allgemeinen Integrationskursen, die in Beitrag #2 und Beitrag #3 im Fokus stehen, gibt es auf spezifische Zielgruppen ausgerichtete Kursprogramme, die nachfolgend genannt werden, aber nicht zentral für die Fragestellung der Dissertation sind. So gibt es Integrationskurse mit einem vorgeschalteten, 300 Unterrichtsstunden umfassenden Alphabetisierungsmodul für Zugewanderte, die bisher keine Schule besucht haben und die über keine oder nur sehr geringe Lese- und Schreibkenntnisse in ihrer Herkunftssprache verfügen. Für Menschen, die zwar alphabetisiert sind, deren Herkunftssprache aber auf einem anderen Schriftsystem beruht (d.h. nicht lateinisch), gibt es Kurse für Zweitschriftlernende (Zweitschriftlernerkurse) mit einem Umfang von 900 Unterrichtseinheiten. Des Weiteren gibt es Kurse mit einem ebenfalls erweiterten Stundenvolumen von 900 Unterrichtseinheiten für Jugendliche bis 27 Jahre (Jugendintegrationskurse), Frauen und Eltern (Integrationskurse für Frauen und Eltern) und Personen, die schon länger in Deutschland leben und durch den ungesteuerten Spracherwerb häufig Schwächen hinsichtlich der Schriftsprache und der Grammatik aufweisen (Förderkurse). Schließlich gibt es sogenannte Intensivkurse, die im Umfang von 400 Unterrichtsstunden explizit schnell lernende Zuwanderinnen und Zuwanderer

adressieren, mit dem Ziel, in bereits kurzer Zeit die erforderlichen Sprachkenntnisse zu erwerben (Tissot et al., 2019).

Im Zuge der großen Migrationsbewegung seit Jahr 2015 und der Folgejahre wurde die Wirksamkeit des Systems der bundesweiten Integrationskurse kritisch diskutiert und die Herausforderungen, denen Lehrkräfte beim Unterrichten in Integrationskursen begegnen, thematisiert. Herausfordernd waren laut Tissot et al. (2019) die zunehmende Heterogenität der Lernenden aus vielen verschiedenen Herkunftsländern sowie eine Zunahme an Personen aus Ländern ohne lateinisches Schriftsystem. Darüber hinaus sind die Rahmenbedingungen der Zugewanderten (Qualität der Unterbringung und unsichere Aufenthaltstitel) mögliche Erklärungsfaktoren für den Rückgang der erfolgreichen Sprachkursabschlüsse (Tissot et al., 2019). Als Maß für die Wirksamkeit der Integrationskurse wurde dabei häufig der Deutschtest für Zuwanderer (DTZ) benannt (Tissot et al., 2019). In der aktuellen Integrationskursgeschäftsstatistik zeigt sich, dass die Prüfungsergebnisse im DTZ in den letzten Jahren verbessert werden konnten². Während im Jahr 2019 nur 50,6 % das Niveau B1 erreicht haben, sind es im Jahr 2021 bereits 60,2 % (Bundesamt für Migration und Flüchtlinge, 2022). Weitere 27,7 % der Kursteilnehmenden haben im Jahr 2021 Deutsch auf Niveau A2 beherrscht. Der Anteil der Lernenden, die das Niveau A2 nicht erreichten, konnte von 17,9 % (im Jahr 2019) auf 12,2 % (im Jahr 2021) gesenkt werden (Bundesamt für Migration und Flüchtlinge, 2022). Im Rahmen der am Institut für Deutsche Sprache durchgeführten *IDS-Goethe-Studie* konnte gezeigt werden, dass es mit tendenziell älteren, bildungsfernen Lernenden ohne Fremdsprachenerfahrung systematisch bestimmte Gruppen sind, die das Sprachniveau B1 nicht erreichen (Cindark et al., 2019).

Um in allgemeinen Integrationskursen unterrichten zu dürfen, müssen gesetzlich vorgeschriebene Zulassungskriterien erfüllen. Dies ist erreicht, wenn die Lehrkräfte entweder eine Direktzulassung erhalten, indem sie auf einen Hochschulabschluss im Bereich Deutsch als Fremdsprache/Deutsch als Zweitsprache verweisen können, einschlägige anerkannte (Hochschul-)Zertifikate im Bereich DaZ erworben haben oder einer Lehrbefähigung für das Fach Deutsch aus dem Schulbereich besitzen. Für Lehrkräfte, die diese Qualifikationen nicht erworben haben, gibt es die Möglichkeit durch eine Zusatzqualifizierung im Rahmen von 140 Unterrichtseinheiten die Lehrbefähigung in Integrationskursen zu erhalten (Bundesamt für Migration und

² Zum ersten Quartal 2018 wurde die Berechnungsmethode geändert und Teilnehmende werden nur noch einmal in der Statistik erfasst, wenn sie das Niveau B1 erreicht haben, gleichgültig wie oft sie am Test teilgenommen haben (Bundesamt für Migration und Flüchtlinge, 2022).

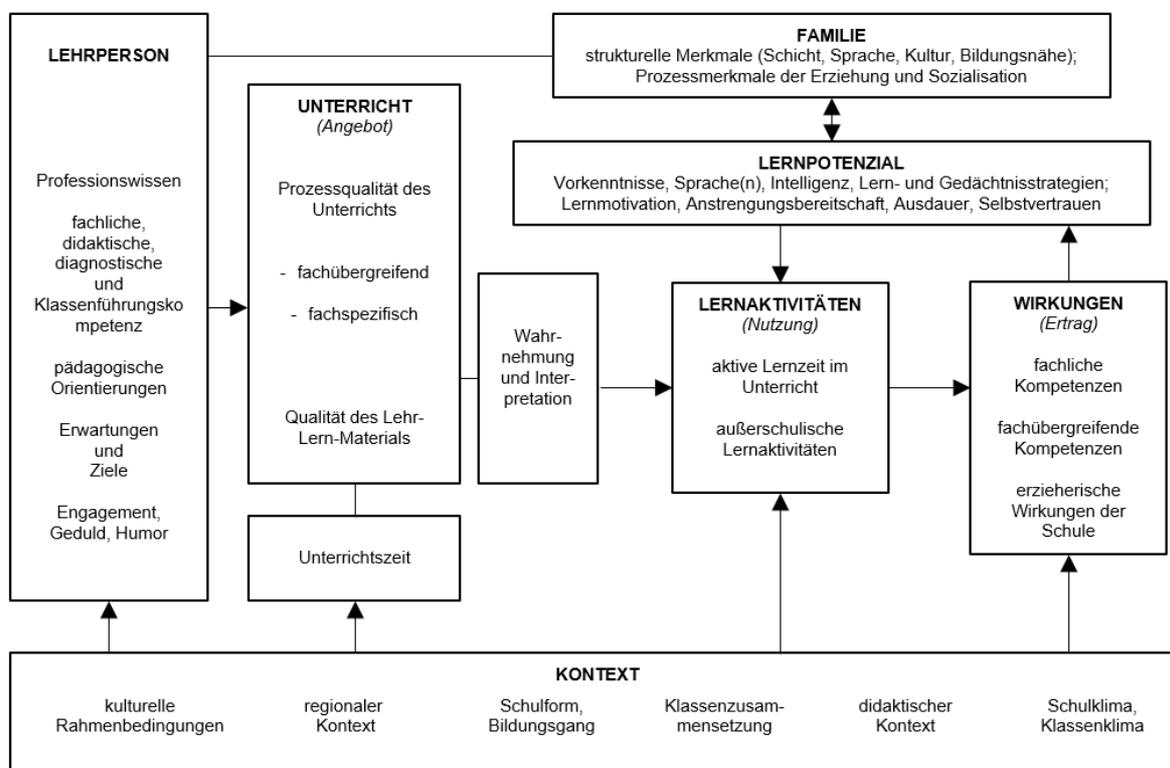
Flüchtlinge, 2020). Inwieweit die Qualifizierung auf den erfolgreichen Umgang mit sprachlicher Heterogenität vorbereitet, ist zurzeit noch nicht empirisch untersucht. Die Vermittlung von Sprachkompetenzen im Bereich DaZ ist eine aktuelle und zentrale, bildungsbereichsübergreifende Herausforderung, die sowohl in der Schule als auch in der Erwachsenen- und Weiterbildung die Lehrkräfte vor verschiedene Herausforderungen stellt (Faltis & Valdés, 2016; Kniffka & Siebert-Ott, 2012). Zum einen unterrichten Lehrkräfte in Klassenzimmern und Weiterbildungskursen häufig Lerngruppen, die eine hohe Heterogenität hinsichtlich vorhandener Sprachkompetenzen in Erst-, Zweit- und Drittsprachen sowie unterschiedliche Bildungsbiographien aufweisen (Rjosk et al., 2017; Stanat et al., 2017), zum anderen sind die Lehr-/Lernprozesse für einen erfolgreichen DaZ-Erwerb bisher noch unzureichend erforscht (Becker-Mrotzek & Roth, 2017).

3. Lehr-/Lernprozesse im Unterricht

Da die oben aufgeführten Formen des Spracherwerbs in Integrationskursen, Vorbereitungsklassen und Alphabetisierungskursen alle in geordneten formalen Bildungsprozessen in klassischem Unterricht stattfinden, ist es hilfreich, die Wirkungsweise von Unterricht genauer zu betrachten. Dies hat zum einen den Vorteil, die bisher bekannten relevanten Einflussfaktoren gelingender Lehr-/Lernprozesse im Unterricht in einem theoretischen Rahmen zu verorten. Darüber hinaus erlaubt es auch zu verdeutlichen, inwieweit die Ausprägung sprachlicher Fähigkeiten auf das System Unterricht wirken kann: als kontextuelle Einflussvariable bei der Zusammensetzung der Lerngruppe, als Kommunikationsmedium zwischen Lehrperson und Lernenden sowie als unterschiedlich verteiltes Lernpotenzial bei den Lernenden selbst.

Abbildung 1

Ein Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts (Helmke, 2017, S. 71)



Das Angebots-Nutzungs-Modell der Wirkungsweise von Unterricht von Helmke (2017) in Abbildung 1 zeigt dabei aus einer konstruktivistischen Perspektive die

Wirkungsbeziehung im Unterricht. Dabei beeinflusst das Angebot der Lehrkraft (der Unterricht), das durch Professionswissen und diverse Kompetenzen geprägt ist, den Unterricht nur zu einem Teil. Die Lehrperson verantwortet, steuert und strukturiert zwar den Unterricht; ob dieser jedoch zur gewünschten Wirkung (dem Ertrag) führt, hängt auch von den Lernaktivitäten (der Nutzung) der Lernenden ab. Die Lernaktivitäten der Lernenden sind einerseits von den Wahrnehmungen und Interpretationen durch die Lernenden geprägt und werden andererseits gleichzeitig durch das individuelle Lernpotenzial der einzelnen Lernenden (Vorkenntnisse, Sprachkenntnisse, motivationale Aspekte etc.) beeinflusst. Schlussendlich verdeutlicht die Abbildung auch, dass der jeweilige Kontext (kulturelle Rahmenbedingungen, Region, Klassenzusammensetzung etc.) nicht nur die Lehrperson, sondern über die Dauer des gemeinsamen Unterrichts (Unterrichtszeit) auch den Unterricht, die Lernaktivitäten und unmittelbar den Ertrag beeinflusst (Helmke, 2017).

3.1. Lehr-/Lernprozesse und Unterrichtsqualität

Die Frage, wie und unter welchen Bedingungen Lehr-/Lernprozesse gelingen können, ist verknüpft mit der Frage nach Qualitätsmerkmalen guten Unterrichts, die in den letzten Jahren intensiv diskutiert wurde (Baumert & Kunter, 2011; Helmke, 2017). Eine weithin geteilte Annahme in diesem Forschungsbereich ist die grundlegende Einteilung von Lehr-/Lernprozessen in Sicht- und Tiefenstrukturen (Oser & Baeriswyl, 2001). Übergeordnete Organisationsmerkmale werden dabei als Sichtstruktur bezeichnet und beschreiben die Rahmenbedingungen, die Muster der Unterrichtsinszenierung sowie methodische Unterrichtselemente inklusive der Sozialform. Als Tiefenstruktur des Unterrichts bezeichnet man hingegen direkte Merkmale des Lehr-/Lernprozesses, wie den Umgang mit den Lerninhalten, die Interaktionen der Lehrkräfte mit den Lernenden sowie zwischen den Lernenden (Kunter & Voss, 2011). Während die Elemente der Sichtstruktur den Rahmen für die Gestaltung der Lehr-/Lernprozesse festlegen, entscheidet die Ausgestaltung der Tiefenstrukturen maßgeblich über die Effektivität, also den Lernerfolg der Teilnehmenden und damit über die Qualität des Unterrichts, wie Kunter und Voss (2011) mit Verweis auf verschiedene Studien konstatieren (Hattie, 2009; Seidel & Shavelson, 2007; Wang et al., 1993). Die Tiefenstruktur, die oft auch als „Dimensionen der Unterrichtsqualität“ (Klieme et al., 2009, S. 139) bezeichnet wird, lässt sich nach Kunter und Voss (2011) in drei Unterkategorien einteilen: die Effizienz der Klassenführung („classroom-management“), das Potenzial zur kognitiven Aktivierung sowie die konstruktive Unterstützung, die manchmal auch als

soziale Unterstützung (z.B. bei Klieme et al., 2009) oder als unterstützendes Unterrichtsklima (Denn et al., 2019) bezeichnet wird.

Um Lehr-/Lernprozesse zu verstehen bzw. die Qualität von Lehr-/Lernprozessen zu bewerten, ist es nicht ausreichend, allein die Sichtstruktur zu betrachten, da diese für sich wenig Erklärungsmacht für die Vorhersage von Lernerfolg bei den Teilnehmenden hat (Hattie, 2009; Seidel & Shavelson, 2007; Wang et al., 1993). Durch die Verknüpfung der Sichtstruktur mit Elementen der Tiefenstruktur (Interaktionsverhalten, Kompetenzniveau der Lernenden) lassen sich Lehr-/Lernprozesse besser verstehen. So konnten Denn et al. (2019) zeigen, dass eine erhöhte Meldebereitschaft der Lernenden positiv mit einem unterstützenden Unterrichtsklima und einer effektiven Klassenführung einhergeht. Weiterhin fanden sie allerdings keinen Zusammenhang zwischen der Dimension der kognitiven Aktivierung und der Schülerbeteiligung. Da die kognitive Aktivierung von Lernenden jedoch stark vom jeweiligen Fachinhalt abhängt, legen bisherige Untersuchungen nahe, hierfür mehrere Messzeitpunkte zu berücksichtigen, um eine valide und generalisierbare Erfassung zu gewährleisten (Praetorius et al., 2014). Anders verhält es sich mit den Dimensionen der Unterrichtsqualität, die durch häufig wiederkehrende Merkmale innerhalb einer Unterrichtsstunde erfasst werden können (z.B. Klassenmanagement, Frage-Antwort-Interaktion). Diese sind mit nur wenigen Aufzeichnungen messbar (Praetorius et al., 2014; Seidel & Prenzel, 2006). Für den Sprachunterricht konnte zudem gezeigt werden, dass die Qualität der Lehr-/Lernprozesse auch von den jeweiligen sprachlichen Fähigkeiten der Lehrkraft beeinflusst wird (für das Unterrichten von Englisch als Fremdsprache siehe König et al., 2016), da die Instruktion durch die Lehrkraft im Sprachunterricht selbst ein zentraler Faktor des Sprachunterrichts ist und damit einen Großteil der Unterrichtszeit ausmacht (Borg, 2006). Der Befund von König et al. (2016) knüpft damit an die Debatte um den Zusammenhang von Lehrkraftkompetenzen und erfolgreichem Unterrichten an (Baumert & Kunter, 2011). Hierbei wird das fachliche Wissen, das im Sprachunterricht das Wissen über die zu erlernende Sprache darstellt, als ein zentraler Wissensbereich ausgewiesen. Da das fachliche Wissen im Sprachunterricht zusätzlich die Interaktion mit den Lernenden beeinflusst, weil die Lehrkraft je nach Fachwissensstand der jeweiligen Sprache, die Sprache unterschiedlich kompetent anwenden kann, bekommen die Kompetenzen der Lehrkraft in sprachlichen Lehr-/Lernprozessen eine besondere Bedeutung.

3.2. Die Rolle der Lehrkraft beim Spracherwerb

Die Gestaltung von Lehr-/Lernprozessen, also das Unterrichten, gilt als das wichtigste Aufgabenfeld von Lehrkräften und wird als die Schaffung von Gelegenheitsstrukturen für erfolgreiches Lernen verstanden. Die zentrale Aufgabe der Lehrenden ist dabei die Steuerung komplexer sozialer Situationen (Kunter & Voss, 2011). Die folgenden Ausführungen gehen zu großen Teilen auf empirische Befunde und Überlegungen der allgemeinen Bildungsforschung zurück, welche sich zumeist mit dem Unterrichtsort Schule beschäftigt. Dezierte Befunde zu Lehr-/Lernprozessen in sprachlich heterogenen Gruppen von erwachsenen Sprachlernenden sind selten. Daran anschließend werden die Übertragbarkeit der Erkenntnisse aus der allgemeinen Bildungsforschung auf die Lehr-/Lernprozesse sprachlich heterogener Gruppen und die dabei zu treffenden Annahmen diskutiert.

Aus sozialkonstruktivistischer Perspektive gelingt Lernen maßgeblich durch das zentrale Element der Interaktion – und zwar zwischen den Lernenden und fachkundigeren, kenntnisreicheren Personen (im Englischen: *more knowledgeable others*) wie den Eltern, der Peer-Gruppe oder den Lehrkräften (Vygotsky & Cole, 1978). Die fachkundigeren Personen ermöglichen den Lernprozess durch linguistische und instruktionale Unterstützung. Für das Lernen von Sprache scheint dies in doppelter Weise zu gelten, da hier die sprachliche Interaktion zugleich Lerngegenstand als auch Lernmedium ist. Das aktive Sprechen der Lernenden fördert die Sprachentwicklung und je mehr die Lernenden mit fähigeren Anderen interagieren, desto eher entwickeln sich ihre sprachlichen Fähigkeiten.

Diese sogenannte Interaktionshypothese wurde in der Bildungsforschung bereits ausgiebig empirisch untersucht. So konnte Seidel (2003) nachweisen, dass mit einer steigenden Schülerbeteiligung auch eine erhöhte Qualität der Lernmotivation sowie erhöhte kognitive Lernaktivitäten einhergehen. Weiterhin wurde gezeigt, dass eine aktive Teilnahme der Lernenden im Unterrichtsgeschehen zu einer Verbesserung der fachlichen Leistung führt (Guo et al., 2015; Wang & Fredricks, 2014). Die zentrale Aufgabe der Lehrkraft ist es, die Interaktionsprozesse zu steuern, an die jeweiligen individuellen Voraussetzungen der Lernenden anzupassen und dadurch erfolgreiche Lernprozesse zu ermöglichen.

Beim Zweitspracherwerb spielen Lehr-/Lerninteraktionen und Partizipationsprozesse eine besonders wichtige Rolle für den Lernerfolg der Lernenden. Die Gestaltung des Lehr-/Lerngeschehens und die Qualität und Quantität der Rückmeldung von

Lehrkräften an Lernende bestimmen das Ausmaß aktiver Lernmöglichkeiten, die von zentraler Bedeutung für den Zweitspracherwerb sind (Loewen & Sato, 2018; Ninio & Bruner, 1978; Thoms, 2012). Interaktive Lehr-/Lernprozesse und aktive Lernmöglichkeiten unterstützen auch bildungssprachliche Kompetenzen (Bailey, 2007) und verbessern die Entwicklung der Lesefähigkeit (Nystrand et al., 2003). Gleichzeitig werden Lehrende bildungsbereichsübergreifend durch die zunehmende sprachliche Heterogenität der Lernenden vor neue Herausforderungen gestellt (Rjosk et al., 2017). So müssen Lehrkräfte die sprachliche Vielfalt zusätzlich zu dem ohnehin bereits komplexen Unterrichtshandeln berücksichtigen (siehe hierzu Becker-Mrotzek & Roth, 2017; Gogolin, 2008; Kniffka & Siebert-Ott, 2012).

3.3. (Sprachliche) Heterogenität im Unterricht

Heterogenität beschreibt aus erziehungswissenschaftlicher Perspektive grundsätzlich die Unterschiedlichkeit zwischen Personen aufgrund verschiedener Merkmale. Als Kategorien werden dabei oft Geschlecht, Ethnizität, Milieu oder Behinderung verstanden und im Schulkontext häufig mit Bildungsungleichheiten in Verbindung gebracht (Budde, 2015). Heterogenität im Bereich Schule (und folglich auch Erwachsenen-/Weiterbildung) thematisiert dabei weiter auch die lern- und leistungsbezogenen Unterschiede der Lernenden (Budde, 2015). Beim Spracherwerb spielen für die Frage der Heterogenität alle spracherwerbsrelevanten Faktoren eine potenzielle Rolle.

Kärner et al. (2021) schlagen vor, dass diese Unterschiede jedoch nicht nur als interindividuelle Variabilität zwischen den Lernenden verstanden werden können (wie z.B. unterschiedliche Niveaus im Fachwissen, das im Sprachunterricht auch das Sprachniveau sein kann). Die Unterschiede können auch intraindividuelle Variabilität innerhalb der Lernenden im Zeitverlauf bedeuten (z.B. unterschiedliche Fachwissensstände zu verschiedenen Zeitpunkten in einem Schuljahr und/oder dem Kursverlauf). Als Drittes identifizieren sie auch die Variabilität in Personen-Kontext-Interdependenzen (z.B. bei Interaktionen zwischen individuellen Fachwissensständen und fachbezogeneren Aufgabenschwierigkeiten).

Zum Umgang mit der zunehmenden Heterogenität empfiehlt die Bildungs- und Schulforschung eine Anpassung der Unterrichtspraxis hin zu offenen Unterrichtskonzepten. Mit selbstgesteuertem, kooperativem, individualisiertem und binnendifferenziertem Lernen sollen Lernende in ihrer Unterschiedlichkeit angemessen gefördert werden können (Bohl & Kucharz, 2010). Mit Blick auf die Situation in Integrations- und Spracherwerbskursen lässt sich feststellen, dass im Integrationskursbereich versucht

wird, der interindividuellen Variabilität durch eine äußere Differenzierung in verschiedene Kursarten gerecht zu werden. So werden mit Hilfe eines Eingangstests zur Feststellung des Sprachniveaus verschiedene, oben bereits ausgeführte niveauangepasste Kurse angeboten (Alphabetisierungskurse, Zweitschriftlernerkurse, Intensivkurse etc.). In der praktischen Umsetzung ist diese Differenzierung jedoch nicht überall möglich. Gerade in ländlichen Regionen, in der nicht überall vor Ort alle Kursarten angeboten werden, kommt es oft zur Abwägung zwischen Dauer der Anfahrtswege zum nächstgelegenen Kursort und idealer Passung. Durch diese ausbleibende Differenzierung kann es dazu kommen, dass Lernende mit sehr unterschiedlichen sprachlichen Fähigkeiten in einem gemeinsamen allgemeinen Kurs unterrichtet werden, der dann sprachlich sehr heterogen sein kann. Doch auch wenn die äußere Differenzierung gelingt, bleibt ein hoher Bedarf an Binnendifferenzierung während des Lehr-/Lerngeschehens.

Auch die Integrationskurslehrkräfte bezeichnen die Angebote oft als nur ungenügend auf die unterschiedlichen Voraussetzungen und Fähigkeiten der Kursteilnehmerinnen und Kursteilnehmer ausgerichtet. Es überwiegen allgemeine Integrationskurse, eine Kursdifferenzierung stelle die Ausnahme dar und es fehle an ausreichend Alphabetisierungskursen. Eine angemessene Leistungsdifferenzierung der angebotenen Kurse sei daher nicht möglich und Über- bzw. Unterforderung der Lernenden häufig (Ohliger et al., 2017).

In Kombination mit der Variabilität in Personen-Kontext-Interdependenzen (Kärner et al., 2021) kann das im DaZ-Unterricht zu zusätzlichen Herausforderungen führen. Die Interaktion zwischen individuellen Fachwissensständen, im DaZ-Unterricht sind das Teilfertigkeiten wie Sprechen, Schreiben, Lesen, Hören in verschiedensten Fähigkeitsausprägungen und unterschiedlichen Aufgabenschwierigkeiten erfordern eine sehr gute Diagnosekompetenz auf Seiten der Lehrenden, um der vorhandenen Heterogenität angemessen gerecht zu werden. Denn die Lehrkraft muss nicht nur das Anforderungsprofil der zu bearbeitenden Aufgabe hinsichtlich benötigter sprachlicher Fähigkeiten in den einzelnen Teilbereichen einschätzen, sondern auch gleichzeitig die individuellen Fähigkeiten bezüglich der einzelnen Teilfertigkeiten korrekt einschätzen, um eine gelungene Passung zwischen Aufgabe und Fähigkeit herzustellen.

Durch den Anstieg der Migration in der jüngeren Vergangenheit ist kulturell und sprachlich heterogener Unterricht in beiden Bildungsbereichen zur Regel geworden. In den deutschen Grundschulen beispielsweise hatte 2016 im Durchschnitt etwa ein Drittel der Kinder in der vierten Klasse einen Migrationshintergrund (Stanat et al.,

2017). Und auch im Integrationskursbereich wurden in einem Projekt des Instituts für Deutsche Sprache (IDS) in einer repräsentativen Stichprobe in 42 allgemeinen Integrationskursen über 151 verschiedene Herkunftssprachen/Sprachkombinationen bei den Teilnehmenden identifiziert (Hünlich et al., 2018).

Diese Entwicklungen stellen neue Anforderungen an die beruflichen Fähigkeiten sowohl von Lehrkräften in der Schule als auch in der Erwachsenenbildung. Dazu gehören die Fähigkeit, sprachsensiblen Unterricht umzusetzen, Unterrichtsinhalte im Hinblick auf sprachliche Anforderungen zu analysieren, mit Mehrsprachigkeit und kulturellen Unterschieden im Unterricht umzugehen und Kenntnisse über den Erst- und Zweitspracherwerb zu besitzen (Faltis & Valdés, 2016; Gogolin, 2008; Kniffka & Siebert-Ott, 2012; Lucas et al., 2008; Reich & Krumm, 2013). Dies gilt insbesondere für Lehrkräfte in Schulvorbereitungsklassen für Kinder und Jugendliche sowie für Lehrkräfte in Integrationskursen für Erwachsene. Diese Lehrkräfte unterstützen und begleiten sprachlich sehr heterogene Lerngruppen beim Erwerb grundlegender Deutschkenntnisse, die einen wichtigen Baustein für ihre weitere schulische, akademische und berufliche Laufbahn darstellen.

In den letzten Jahren hat sich die Bildungspolitik mit dem wichtigen Handlungsfeld der Qualifizierung von Lehrkräften im Bereich DaZ beschäftigt. Auch wenn es keine einheitlichen Regelungen und Standards gibt, haben einige Bundesländer in Deutschland verpflichtende modulare DaZ-Kurse in der Lehrerbildung für alle angehenden Lehrkräfte unabhängig von ihrer Fachrichtung eingeführt. Bei Integrationskursen sollen Vorschriften bezüglich der formalen Qualifikation der Lehrkräfte und die erforderlichen pädagogischen Fähigkeiten eine hohe Qualität des Unterrichts gewährleisten (Bundesamt für Migration und Flüchtlinge, 2020; Bundesamtes für Migration und Flüchtlinge, 2007).

Mit Blick auf den Schulbereich und die Erwachsenenbildung ist weiter zu beachten, dass sich die beiden Bildungsbereiche in mehreren Aspekten sowohl hinsichtlich der Zusammensetzung der Lernenden als auch im Professionalisierungsgrad der Lehrkräfte unterscheiden. Die Zusammensetzung der Lernenden ist bereits im Schulbereich durch eine zunehmende Heterogenität bestimmt (Rjosk et al., 2017; Vock & Gronostaj, 2017); mit fortschreitendem Lebensverlauf der Lernenden durch unterschiedliche Biografien sowie unterschiedlich erworbene Qualifikationen nimmt diese stetig weiter zu. DaZ-Kurse in der Erwachsenenbildung zeichnen sich somit durch ein noch größeres Spektrum an Unterschiedlichkeit als jüngere Lernende im Schulbereich hinsichtlich der Ausgangsvoraussetzungen aus (Demmig, 2008).

Darüber hinaus unterscheidet sich auch der Grad der Professionalisierung der Lehrkräfte in den beiden Bildungsbereichen, da die Ausübung einer Lehrtätigkeit in einem DaZ-Integrationskurs in der Erwachsenenbildung mit einer Zusatzqualifizierung und ohne fachdidaktisches mehrjähriges Studium möglich ist. Demgegenüber ist der Zugang zu einer Lehrtätigkeit im Schulbereich durch ein einschlägiges pädagogisches Studium beschränkt. Ein dritter Unterschied ist der Faktor der gemeinsam verbrachten Zeit. Während Lehrkräfte und Lernende im Schulbereich oft mehrere Schuljahre miteinander lehren und lernen, sind Integrationskurse mit 700 Unterrichtsstunden von kürzerer gemeinsamer Dauer. Deshalb ist davon auszugehen, dass die Beziehung zwischen Lehrenden und Lernenden und damit auch die Kenntnisse, z.B. über lernrelevante Vorlieben und Präferenzen, die die Lehrenden über ihre Lernenden haben in Integrationskursen weitaus weniger gering sind als im Schulbereich.

3.4. (Sprachliche) Heterogenität als Herausforderung für die Lehrkraft

Bezogen auf das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts (Abbildung 1) ist die sprachliche Heterogenität ein Merkmal, das gleich zweimal wirksam wird. Zum einen über den Kontext der Klassenzusammensetzung und zum anderen über die Lernpotenziale der Lernenden, die maßgeblich die Lernaktivitäten beeinflussen und von deren Ausprägung die Wahrnehmung und Interpretation des Unterrichts abhängt.

Helmke (2017) verweist in seiner Übersicht zur Unterrichtsqualität auf Beiträge von Weinert (1997) zum Umgang mit Heterogenität. Dieser beschreibt bereits gegen Ende des letzten Jahrtausends vier grundlegend denkbare Möglichkeiten (von ihm Reaktionsformen genannt), mit Heterogenität im Unterricht umzugehen, wobei die erste Variante, die Lehrkraft könne die Heterogenität ignorieren, direkt als illusorisch beschrieben wird. Die zweite Variante (die substituierende Reaktion auf Heterogenität) sieht vor, die Lernenden an die Anforderungen des jeweiligen Unterrichts anzupassen. Zum einen (Variante 1) in Form einer äußeren Differenzierung, wie das bei Integrationskursen in allgemeine, intensive und Alphabetisierungskurse etc. der Fall wäre. Zum anderen (Variante 2) in Form einer systematischen Intelligenzförderung oder Trainings zu metakognitiven Strategien (z.B. des Lernen-Lernens), um durch leistungswirksame Verbesserungen eine Homogenisierung zu erreichen. Die dritte (aktive Reaktionsform) und vierte (proaktive Reaktionsform) Variante sind ähnlich dem, was wir heute unter adaptivem Unterrichten (vgl. Hardy et al., 2019) verstehen, und beschreibt die unterschiedliche Anpassung des Lehrhandelns (Variante 3) an die

Lernvoraussetzungen der Lernenden, bzw. die noch adaptivere Gestaltung individueller Lernziele (Variante 4) je nach Fähigkeitsniveau der Lernenden (Weinert, 1997). Neben der Anpassung der Lernziele an die Voraussetzungen der Lernenden verweist Leutner (1992) auch auf die individuelle Anpassung der Lehrmethode und der jeweiligen Unterrichtszeit, die die Lehrkräfte verschiedenen Lernenden widmen können, um angemessen auf die Heterogenität zu reagieren. Die Anpassung der Lehrmethode beinhaltet dabei auch explizit die Orientierung der Lehr-/Lernformen an die Lernenden.

Dass Lernende davon profitieren, wenn Lehrkräfte ihren Unterricht adaptiv gestalten, also den Input an den jeweiligen Anforderungsniveaus der einzelnen Lernenden ausrichten, konnte mittlerweile in nahezu allen Schulfächern (Sprachen, Mathematik, natur- und sozialwissenschaftliche Fächer) gezeigt werden (Hardy et al., 2019; Parsons et al., 2018). In sprachlichen Schulfächern wird der Umgang mit Adaptivität häufig unter dem Begriff des individuellen Scaffolding verhandelt (z.B. bei Hammond & Gibbons, 2005), ohne expliziten Verweis auf den Begriff der Adaption. Wie sich Anpassungen im jeweiligen Fachunterricht konkret gestalten und ob Lehrkräfte dies auch in sprachlich sehr heterogenen Klassen wie Integrationskursen und Vorbereitungsklassen überhaupt machen oder leisten können, ist noch nicht untersucht. Hierzu fehlen auch Datengrundlagen, die solche Analysen ermöglichen. Für den DaZ-Unterricht kommt aufgrund der Interaktionshypothese noch eine weitere Anpassungsmöglichkeit in Betracht: Um die Heterogenität angemessen zu berücksichtigen, könnten Lehrkräfte auch ihre Unterrichtssprache (im Sinn des Lerninhalts) an den Lernenden orientieren. Die doppelte Bedeutung der Sprache als Lerngegenstand und Medium unter der Prämisse, den Unterricht an den sprachlichen Fähigkeiten der Lernenden auszurichten, lässt diesen Schluss zu.

Das zentrale Element bei der Gestaltung von Lehr-/Lernprozessen im DaZ-Unterricht, das durch die zunehmende Heterogenität von Lerngruppen an Bedeutung gewinnt, ist die „Zone of Proximal Development“ (ZPD). Im Deutschen lässt sich der Begriff als Zone der nächsten Entwicklung umschreiben. Dieser Bereich beschreibt einen von Vygotsky und Cole (1978) anzunehmenden Anforderungsbereich für Lernende, in denen ihr potenziell erreichbarer Lernfortschritt am größten ist. Dies ist der Fall, wenn der (sprachliche) Inhalt die Lernenden herausfordert, indem er leicht über ihrem derzeitigen Sprachniveau liegt, aber immer noch auf einem erreichbaren Niveau bleibt und sie nicht überfordert (Vygotsky & Cole, 1978). Ein vergleichbares Konzept ist die von Krashen (1989) aufgestellte Input-Hypothese ($i + 1$), die ebenfalls besagt, dass Lernerfolg dann erreicht werden kann, wenn der bereits verständliche (sprachliche)

Input (i) um eine Steigerung (+1) ergänzt wird. Bei der verbalen Kommunikation erfolgen diese sprachlichen Anpassungen in der Regel durch die Regulierung der Länge von Äußerungen und die Anpassung von Lexik und Syntax, um den sprachlichen Input für die Lernenden verständlicher zu machen. Weiterhin gibt es auch die Möglichkeit, durch unterrichtsdidaktische Hilfen wie z.B. Scaffolding³ die sprachlichen Hürden für die Lernenden gezielt zu senken (Gibbons, 2002). Die Herausforderung für die Lehrkraft ist dabei, dass sie durch eine angemessene und zutreffende Diagnostik der sprachlichen Fähigkeiten den idealen Anforderungsbereich, die Zone der nächsten Entwicklung, bei den Lernenden identifizieren muss. Mit zunehmender sprachlicher Heterogenität steigt die Anzahl an individuellen Anforderungsbereichen im gleichen Maße an, wie es Lernende in einer Lerngruppe gibt. Wie und ob Lehrkräfte diese sprachliche Heterogenität im Unterricht adressieren bzw. ob sie in der Lage sind, ihre Äußerungen an die sprachlichen Fähigkeiten der Lernenden anzupassen, ist noch nicht untersucht. Ein Projekt, in dem unter anderem diese Fragestellungen untersucht werden und welches auch die Datengrundlage für Beitrag #2 und Beitrag #3 bildet, ist das Projekt COLD.

3.5. Lehrkraftgesten und adaptive Lehr-/Lernprozesse

Die oben aufgeführten Adaptionstrategien von Lehrkräften hinsichtlich sprachlicher Anpassungen lassen sich aus dem Angebots-Nutzungs-Modell und der Interaktionshypothese ableiten. Blickt man aus einer kommunikationstheoretischen Perspektive auf diese Anpassungsstrategien von Lehrkräften, ist zu erwarten, dass es noch weitere Ausprägungen gibt. Denn Kommunikation ist eine multimodale Handlung (McNeill, 2004) und besteht neben der Sprache auch aus Elementen der Körpersprache, wie z.B. Gesten⁴. Die Bedeutung nonverbaler Elemente soll am Beispiel von Lehrkraftgesten im folgenden Abschnitt verdeutlicht werden. Die Verwendung von Gesten hilft Sprechenden dabei, bestimmte sprachliche Aspekte zu verstärken und Informationen und Absichten hervorzuheben (McNeill, 2004). Dabei werden bis zu 90 % aller Gesten von Sprache begleitet (Tomasello, 2010). Es gibt eine Vielzahl an empirischen Befunden über den lernförderlichen Zusammenhang zwischen dem

³ Dabei werden von der Lehrkraft schrittweise sprachliche Hilfen bereitgestellt, um Lernenden die eigenständige Bewältigung von (sprachbezogenen) Anforderungen zu ermöglichen (Gibbons, 2002).

⁴ Gesten sind im Rahmen des Zeigens als Grundoperation des Pädagogischen, wie es von Klaus Prange intensiv erörtert wurde, ein zentrales Element der Gestaltung pädagogischer Prozesse (ausführlicher hierzu Fuhr, 1999).

Gesteneinsatz von Lehrkräften und dem Verständnis gesprochener Sprache bei Lernenden. So konnten Valenzeno, Alibali und Klatzky (2003) nachweisen, dass Lernende deutlich bessere Wissensleistungen erzielten, wenn Lehrkräfte zur Erklärung im Unterricht Gesten verwenden, als eine Vergleichsgruppe, in der keine Gesten verwendet wurden. Empirisch quantifiziert Hostetter (2011) die Frage, ob Zuhörende eine gesprochene Nachricht besser verstehen, wenn sie mit Gesten unterstützt ist, mit einer mittleren Effektgröße ($d = .61$) und verweist darauf, dass 83 % der 63 Studien einer von ihr durchgeführten Metaanalyse hierzu einen signifikanten positiven Effekt aufweisen. Die genauen Mechanismen, aufgrund derer der Einsatz von Gesten die gesprochene Sprache besser verständlich macht, und ob dies für alle Lernenden gleichermaßen der Fall ist, ist nicht geklärt. Es gibt jedoch Hinweise darauf, dass Gesten eine aufmerksamkeitslenkende und eine komplexitätsreduzierende Funktion haben können (Lazaraton, 2004; Valenzeno et al., 2003).

In der aktuellen Forschung zu Gesten und Sprache werden Gesten in der Regel entlang des sogenannten Kendons-Kontinuums angeordnet, das die Gesten nach ihrer Beziehung zur Sprache sortiert: Gestikulationen, sprachverknüpfte Gesten, Embleme, Pantomime, Zeichensprachen (Stam, 2013). Am Anfang des Kontinuums stehen die Gestikulationen⁵, die in der Regel spontane Bewegungen der Hände und Arme sind, die zusammen mit der Sprache auftreten und die sprachlichen Äußerungen ergänzen oder unterstützen. Sie können ähnlich wie Sprache analysiert werden, da sie dieselben pragmatischen Funktionen wie die Sprache haben und hinsichtlich ihrer semiotischen Eigenschaften unterschieden werden können (Stam, 2013). Es ist erwähnenswert, dass auch Sprache und Gesten nicht immer eins zu eins dieselbe Bedeutung haben müssen und in unterschiedlichen Modalitäten ausgedrückt werden können (Lim, 2021). Sprachverknüpfte Gesten sind in ihrer Form ähnlich, unterscheiden sich aber von Gestikulation durch ihre zeitliche Verschiebung zur Sprache. Embleme bilden ein weiteres Element innerhalb des Kontinuums, da sie mit Sprache auftreten können, aber konventionalisiert und übersetzbar sind. Das macht sie auch ohne Sprache verständlich. Das Ende des Kontinuums bilden Pantomime und Zeichen, die in der Regel ganz ohne Sprache auftreten (wie z.B. die Gebärdensprache) und ihre eigenen linguistischen Eigenschaften haben (McNeill, 2008; Stam, 2013).

In der internationalen Bildungsforschung sind Untersuchungen zur Verwendung von Gesten in Lehr-/Lernprozessen in den letzten Jahren verstärkt in den Fokus

⁵ In Beitrag #3 liegt der Fokus auf unterschiedlich komplexen Gestikulationen von Lehrkräften, die im Folgenden als Lehrkraftgesten bezeichnet werden.

der Aufmerksamkeit gerückt (Dargue et al., 2019; McNeill, 2004; Mercer, 2000; Sato, 2020; Stam & Tellier, 2022). Auch im Bereich der Zweitspracherwerbsforschung wurden Untersuchungen zu Art und Häufigkeit verwendeter Gesten von Lehrkräften durchgeführt (Kita, 2004; Sato, 2020; Singer & Goldin-Meadow, 2005). Ein Review (McCafferty & Stam, 2008) verdeutlicht den positiven Effekt der Gesten von Lehrkräften im Unterricht. So unterstützt der Einsatz von Lehrkraftgesten das Lernen neuer Vokabeln und deren Bedeutung (Lazaraton, 2004; Lazaraton & Ishihara, 2005). Insgesamt bewerten McCafferty und Stam (2008) die Verschränkung von Spracherwerbsprozessen mit der Verwendung von Gesten als substantiell und verweisen auf das große Forschungspotenzial in diesem Bereich. Einen ausführlichen Überblick bieten auch die beiden Metaanalysen von Hostetter (2011) und Dargue et al. (2019). Übereinstimmend attestieren beide Analysen dem Gesteneinsatz einen lernförderlichen Effekt, wobei jüngere Lernende noch etwas mehr als ältere Lernende vom Gesteneinsatz der Lehrkräfte profitieren. Erwähnt sei an dieser Stelle auch, dass die Lernförderlichkeit bestimmter Gestentypen vom jeweiligen Sprachverständnis der Lernenden abhängig ist (Dargue et al., 2019).

Der Einsatz von Lehrkraftgesten wird im Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts nicht explizit verortet. Helmke (2017) weist aber in seinen Dimensionen der Unterrichtsqualität unter dem Bereich Klarheit und Strukturiertheit auf die Bedeutung von Mimik und Gestik (neben Klarheit des Ausdrucks und Aspekten der akustischen Verständlichkeit) als relevantes Merkmal hin. Eine gesonderte Rolle kommen unterschiedlichen Gestentypen dabei bisher noch nicht zu.

Im Zusammenhang der drei Basisdimensionen von Unterrichtsqualität (Klassenführung/Klassenmanagement, kognitive Aktivierung, konstruktive Unterstützung) werden Gesten im Modell von Voss et al. (2011) nicht erwähnt. Ophardt und Thiel (2017) schlagen vor, Gesten im Rahmen des Klassenmanagements bei der Steuerung von Unterrichtsprozessen zu verorten. Dabei werden Gesten in ihrer Funktion als aufmerksamkeitslenkend beschrieben, die der Lehrkraft dazu dienen, Hinweise wie Cues (Signale) und Prompts (Aufforderungen) zu geben. Die oben aufgeführten Einzelbefunde, die nahelegen, dass Gesten neben einer steuernden Funktion auch eine lernförderliche Funktion haben können, werden in den Basisdimensionen bisher nicht weiter berücksichtigt oder erwähnt.

Aufgrund der hohen sprachförderlichen Bedeutung, die Lehrkraftgesten besonders beim Spracherwerb haben können, erscheint es sinnvoll, der Frage nachzugehen, welche Rolle Gesten im Sprachunterricht zukommt. Um Gesten in Lehr-

/Lernprozessen erfassen zu können, gibt es mittlerweile eine Vielzahl an Klassifikationsansätzen, die jedoch in zentralen Punkten übereinstimmen (siehe Überblicksarbeit von Kendon, 2010; McNeill, 1992). Tabelle 2 zeigt eine Möglichkeit der Operationalisierung verschiedener Arten von Gesten nach dem aktuellen und empirisch erprobten Modell von Sato (2020), welches auf den Arbeiten von McNeill (1992, 2008) basiert. Dieses Modell, wie es in Tabelle 2 inklusive Operationalisierungsvorschlag abgebildet ist, scheint deshalb für das Promotionsvorhaben passend, da Sato (2020) es ebenfalls für die Untersuchung des Gesteneinsatzes von Lehrkräften im Zweitsprachunterricht (hier im Fach Englisch) verwendet hat.

Tabelle 1*Gestentypen nach Sato (2020), ergänzt um eigene Beispiele*

Geste der Lehrkraft	Deutsche Übersetzung	Bedeutung und Operationalisierungsvorschlag	Beispiel im DaZ-Unterricht
Deictic	Zeigende Geste	Die Person zeigt mit dem Finger oder der Handfläche auf etwas oder jemanden, wobei sie sowohl auf konkrete als auch auf abstrakte Objekte/Sachverhalte hinweisen kann.	Die Lehrkraft zeigt auf einen Kugelschreiber und sagt dazu: „Das ist ein Kugelschreiber.“
Iconic	Bildliche / Ikonische Geste	Die Person benutzt ihre Hand und/oder Arme, um Bilder der tatsächlichen Form von Objekten und/oder Handlungen zu zeigen, die in enger Beziehung zum semantischen Inhalt stehen.	Die Lehrkraft sagt: „Bildet drei Sätze“, und hebt dabei die Hand, bei der drei Finger ausgestreckt sind.
Metaphoric	Bildhafte Umschreibung	Die Person benutzt ihre Hand und/oder Arme, um Bilder von abstrakten Konzepten und/oder Ideen zu zeigen, und nicht die tatsächliche Form von Objekten und/oder Handlungen, die durch Ikonen dargestellt werden.	Die Lehrkraft sagt: „Wiederholt das“, und kreist ihre Hände vor ihrem Körper in einer fließenden Bewegung übereinander.
Emblems	Universelles Zeichen	Die Person zeigt nonverbale Handlungen, die von allen Mitgliedern der gleichen kulturellen Gruppe verstanden werden.	Die Lehrkraft zeigt den nach oben gerichteten Daumen, was in westlichen Kulturen sehr gut/gut bedeutet.
Beats	Rhythmische Geste	Die Person bewegt ihre Hand mit einem rhythmischen Puls. Typisch ist eine Schlaggeste eine einfache Handbewegung oder eine Bewegung der Finger auf und ab oder vor und zurück, die der Betonung des Sprechens folgt.	Die Lehrkraft betont die Silben des Wortes (z.B.: „Nach-mit-tag“) und die Betonung der Silben geht einher mit der Bewegung der Hand.
Affect display	Emotionaler Ausdruck	Die Person zeigt Emotionen wie Glück, Angst, Traurigkeit, Wut, Ablenkung und Interesse, insbesondere durch den Gesichtsausdruck.	Die Lehrkraft runzelt die Stirn, nachdem ein/e Lernende/r seinen/ihren Beitrag beendet hat.
Head movements	Kopfbewegungen	Die Person bewegt beim Sprechen den Kopf, um den Gesprächspartnerinnen und Gesprächspartnern ihr Unverständnis, ihre Zustimmung oder ihre Ablehnung zu signalisieren.	Die Lehrkraft nickt mit dem Kopf, während sie mit den Lernenden spricht.

Ohne in dieser Einleitung auf die einzelnen Gesten im Detail einzugehen (eine detaillierte Beschreibung findet sich in Beitrag #3), kann festgestellt werden, dass manche Gesten, unter Berücksichtigung ihrer Beziehung zum Sprachverständnis, als unterschiedlich komplex eingeschätzt werden können. So gibt es Gesten, die ein

Sprachverstehen benötigen, und Gesten, die auch ohne Sprachverständnis verständlich sind. Zeigende und ikonische Gesten sind dabei häufig sehr nahe an dem semantischen Inhalt, der zeitgleich durch die Sprache ausgedrückt wird, und können deswegen ohne ausgeprägte sprachliche Fähigkeiten verstanden werden (Dargue et al., 2019). Demgegenüber besitzen metaphorische Gesten eine hohe Komplexität, da sie oft sprachunterstützend eingesetzt werden, das heißt, sie begleiten sprachliche Äußerungen, ersetzen sie aber nicht. Deshalb ist hier häufig ein höheres Maß an sprachlichen Fähigkeiten notwendig, um die Bedeutung der Kombination aus Sprache und Geste zu verstehen, da die sprachliche Bedeutung nicht alleine über die Betrachtung der Geste verstanden werden kann (Tellier et al., 2021).

Bei rhythmischen Gesten, Gesichtsausdrücken und Kopfbewegungen gibt es keine eindeutigen Forschungsbefunde. Die Verwendung von Emblemen, die eigentlich ohne Sprachverständnis zu verstehen sind, wird dann herausfordernd, wenn die Zuhörenden in verschiedenen Kulturkreisen sozialisiert worden sind, da dann keine einheitliche Verständigung anhand von Emblemen gewährleistet ist. Ein Beispiel hierfür ist die Verwendung des nach oben gerichteten Daumens als Zeichen der Zustimmung oder dass etwas als „okay“ bewertet wird: In arabisch-sprachigen Ländern kann der nach oben gestreckte Daumen eine Beleidigung ausdrücken (Müller et al., 2014).

3.6. Lehrkraftgesten und sprachliche Heterogenität

Dass Lehrkräfte im DaZ-Unterricht ihr nonverbales Verhalten an den sprachlichen Fähigkeiten der Lernenden ausrichten, ist zumindest aus Sicht der Kommunikations-Akkommodations-Theorie (CAT: Communication-Accommodation-Theory) (Giles, 2016) naheliegend. Dieser kommunikationswissenschaftliche Zugang geht davon aus, dass Sprecherinnen und Sprecher und Empfängerinnen und Empfänger, die beiden beteiligten Parteien einer Konversation, bzw. das Sprechen und das Zuhören eine verschränkte Aktivität darstellen. Die Annahme ist, dass der Sprecher oder die Sprecherin permanent den/die Zuhörende/n im Blick behält und dabei auf die Signale des Gegenübers achtet. Dabei prüft der Sprechende, ob die Signale Feedback oder Informationen enthalten, die ihn erahnen lassen, wie, ob und auf welche Art das Gesagte bei dem Empfänger oder der Empfängerin ankommt. Dabei passt die sprechende Person konstant ihre Sprache an die wahrgenommenen Reaktionen der Zuhörenden an. Dies konnten Clark und Krych (2004) experimentell belegen, indem sie zeigten, dass sprechende Personen, die einen Zuhörenden ohne Sichtkontakt instruieren,

systematisch mehr Zeit benötigen und die Umsetzung des Erklärten durch die Zuhörenden weniger präzise ist als wenn Sichtkontakt besteht.

Die CAT kann erklären und vorhersagen, weshalb und warum Personen ihr sprachliches/kommunikatives Verhalten an Zuhörenden ausrichten und wie sie verbale und nonverbale Anpassungen vornehmen und sich dabei während der Unterhaltung einander angleichen (Konvergenz) oder voneinander distanzieren (Divergenz). Die Kommunikation zwischen Lehrenden und Lernenden im Unterricht ist jedoch ein Sonderfall, auf den sich diese beiden Muster nicht übertragen lassen. Im Unterricht ist die Beziehung zwischen Lehrkräften und Lernenden hierarchisch, weshalb die CAT hier eine besondere Beziehung zwischen der sendenden Person (Lehrkraft) und der empfangenden Person (Lernende/r) aufweist. Hierbei wird erwartet, dass Lehrkräfte, basierend auf den von ihnen wahrgenommenen Bedürfnissen der Lernenden, ihre Sprache und ihre Körpersprache ausrichten, so dass sie bestmöglich ihrer Rolle als Lehrkraft ausüben. Das bedeutet, dass sie ihre (Körper-)Sprache so anpassen, dass sie von den Lernenden bestmöglich verstanden werden können. Im DaZ-Unterricht würde das bedeuten, dass Lehrkräfte z.B. komplexe Gesten nur an Personen adressieren, von denen sie annehmen, dass sie über die notwendigen DaZ-Fähigkeiten verfügen, um die Gesten zu verstehen.

Quasiexperimentelle Studien belegen, dass Sprechende ihren Gesteneinsatz an den Fähigkeitsniveaus der Zuhörenden orientieren – zumindest wenn man Gruppenvergleiche anstellt und die Sprechenden bittet, Geschichten oder etwa Vokabeln Personen mit der gleichen bzw. anderer Herkunftssprache zu erklären (Adams, 1998; Tellier et al., 2021). Deutlich wurde sowohl bei Adams (1998) als auch bei Tellier et al. (2021), dass die Sprechenden gegenüber Personen anderer Herkunftssprache, also mit geringeren sprachlichen Fähigkeiten als die Sprechenden, mehr zeigende Gesten verwendet haben, als wenn sie mit Personen der gleichen Herkunftssprache gesprochen haben. Untersucht wurde in beiden Studien auch, ob die Personen, wenn sie mit Personen anderer Herkunftssprache sprechen, weniger komplexe metaphorische Gesten verwenden. Diese Hypothese konnte jedoch in beiden Untersuchungen nicht bestätigt werden. Die Autoren vermuten, dass dies an der Unerfahrenheit der Sprechenden im Umgang mit Gesten liegen könnte. Bei Tellier et al. (2021) bestand die Stichprobe aus angehenden Lehrkräften, die noch kaum Lehrerfahrung und damit einhergehend auch wenig Erfahrung im erfolgreichen Umgang mit Gesten haben.

Um dieses Unterrichtshandeln zu verstehen und die oben aufgeführten Interaktionen zwischen Lehrenden und Lernenden sichtbar zu machen, bedarf es der

empirischen Analyse ebendieser. Die Untersuchung des Zusammenspiels der non-verbalen Kommunikation der Lehrkraft mit sprachlich heterogenen Lernenden im Unterricht ist sowohl für die Schulforschung als auch für die Erwachsenenbildung ein Desiderat. Dazu bietet die mittlerweile umfangreiche videobasierte Unterrichtsforschung in der Bildungsforschung viele Anknüpfungspunkte und methodische Grundlagen, die im Folgenden aufgeführt werden.

4. Videogestützte Erfassung von Lehr-/Lernprozessen im Unterricht

Für die wissenschaftlich begründete Beobachtung von Unterricht hat sich in den letzten Jahren aus mehreren Gründen die Videographie als Methode der Wahl durchgesetzt. Die Überblicksbeiträge von Dalehefte und Kobarg (2012) und Kade et al. (2014) tragen dabei die zentralen Argumente zusammen: Seidel und Prenzel (2010) verweisen darauf, dass Videoaufnahmen beliebig oft anschaulich sind. Die permanente Dokumentation des Geschehenen erlaubt die intensive Analyse, ohne dabei wichtige Elemente zu übersehen. Ebenfalls kann bereits aufgezeichnetes Material immer wieder unter weiteren und vielleicht erst im Laufe des Analyseprozess erarbeiteten Gesichtspunkten analysiert werden. Darüber hinaus erlaubt die Unterrichtsvideographie durch den Einsatz mehrerer Kameras, das Unterrichtsgeschehen aus dem Blick verschiedener Akteure festzuhalten und somit beispielsweise das Lehr-/Lerngeschehen aus Lernenden- und aus Lehrenden-Perspektive gleichzeitig zu dokumentieren. Ein einzelner Beobachter wäre dazu nicht im Stande. Drittens lässt sich anmerken, dass die Videodokumentation unabhängiger von personellen Wahrnehmungen eines menschlichen Beobachters ist und somit die Objektivität der Beobachtung erhöht. Videoaufnahmen wird das Potenzial zugeschrieben, dass sie die Komplexität des Unterrichtsgeschehens weitestgehend abbilden können und somit eine detaillierte Beschreibung von Lehr-/Lernsituationen ermöglichen (Evertson & Green, 1986; Petko et al., 2003). Besonders dann, wenn die beobachteten Unterrichtsmerkmale mit weiteren Datenquellen kombiniert werden, können Lehr-/Lernprozesse genauer verstanden und analysiert werden (Ulewicz & Beatty, 2001).

Gleichwohl gibt es auch Einschränkungen des Verfahrens. So sind Videoaufnahmen immer nur eine Momentaufnahme und ein Ausschnitt, der nachträglich nicht erweitert werden kann, weshalb es wichtig ist, die Fragestellung der jeweiligen Untersuchung bei der Planung des Forschungsdesigns im Blick zu haben (Seidel & Prenzel, 2010). Ein weiterer Nachteil von Videoaufnahmen, der öfters aufgeführt wird, ist die Erzeugung von Reaktivität durch den Einsatz sichtbarer Kameras und der damit verbundene Eingriff in das Unterrichtsgeschehen. So wurde erwartet, dass die Anwesenheit von Kameras im Klassen-/Kursraum einen Einfluss auf das Unterrichtsgeschehen hat und die Lehrkräfte ablenkt. Dieser Effekt ist, zumindest im Schulbereich, schon unmittelbar nach Aufzeichnungsbeginn (1 min 20 s) nicht mehr nachweisbar (Praetorius et al., 2017). Allerdings konnte gezeigt werden, dass die Anwesenheit von

Kameras einen Einfluss auf das emotionale Empfinden und die Motivation bei Lehrenden und Lernenden hat. Tendenziell wird die Stimmung bei den Lernenden etwas negativer als in Gruppen ohne Videoaufnahmen⁶ und die Unterrichtsbeteiligung ist eher geringer, sie wird folglich eher unterschätzt. Bei den Lehrkräften gibt es keine eindeutigen Unterschiede emotionalen Befindens und motivationaler Aspekte. Erwähnenswert ist dabei auch, dass kein Zusammenhang zwischen den Videoaufnahmen (Unterricht mit vs. ohne Kameras) und dem Lehrverhalten sowie der Unterrichtsqualität festgestellt wurde (Praetorius et al., 2017).

4.1. Videogestützte Unterrichtsforschung in der empirischen Bildungsforschung und der Erwachsenenbildung

Die Durchführung von videogestützter Unterrichtsbeobachtung als vielversprechender methodischer Zugang für die Analyse von Lehr-/Lernprozessen ist mittlerweile fest in der Bildungsforschung etabliert (Pauli & Reusser, 2006). Dies wird besonders deutlich, wenn man die bisherigen durchgeführten Videostudien in den Blick nimmt. Mehrere internationale und nationale Studien im schulischen Bildungsbereich setzten und setzen Videographie ein, um Dispositionen, Kompetenzen und/oder Performanzen der Lernenden und der Lehrkräfte zu untersuchen. Exemplarisch sei an dieser Stelle auf die größeren Studien hingewiesen wie etwa die international vergleichende *Trends in International Mathematics and Science Study (TIMSS)*, die 1995 und 1997 mit einer ergänzenden Videostudie durchgeführt wurde (Jacobs, Hollingsworth & Givvin, 2007), die *Videostudie "Lehr-Lern-Prozesse im Physikunterricht"* des IPN Kiel (Seidel et al., 2003), sowie die *Pythagoras-Studie (Klieme et al., 2009)*, und die *OECD-Studie TALIS-Video* (Grünkorn et al., 2020). Alle diese Videostudien ermöglichten und ermöglichen immer noch wertvolle Erkenntnisse zu Lehr-/Lernprozessen im Mathematik- und Physikunterricht. Darüber hinaus gibt es auch Videostudien in sprachlichen Schulfächern wie zum Beispiel in Deutsch und Englisch die Studie *Deutsch-Englisch-Schülerleistungen-International (DESI-Studie)* (Helmke et al., 2007; T. Helmke et al., 2008), die *VERA-Studie (VERA = Vergleichsarbeiten in der Schule)* (Moller et al., 2007) sowie die Studie *Persönlichkeits- und Lernentwicklung von Grundschulkindern (PERLE)* (Lotz et al., 2013). Der Bereich DaZ ist mit den bisherigen Videostudien noch nicht erforscht und die Analyse von Lehr-/Lernprozessen unter dem Aspekt der sprachlichen Heterogenität der Lernenden stellte ein Desiderat dar, das

⁶ Die Bewertung erfolgte in Form einer Befragung nach der (nicht) gefilmten Unterrichtsstunde.

erst durch das Projekt COLD 2019–2022 (ausführlich hierzu im Kapitel 4.3) systematisch bearbeitet wurde. Zu allen oben aufgeführten Videostudien gibt es umfangreiche Handbücher, in denen die Datenaufbereitung in technischen Berichten dokumentiert ist. Dadurch wird versucht, das Beobachtungsverfahren an höchsten wissenschaftlichen Gütekriterien auszurichten und somit eine transparente Datennutzung zu gewährleisten. Damit wird auch – die rechtlichen Datennachnutzungsaspekte außen vor gelassen – eine Datennachnutzung möglich. In der Regel werden die Videodaten durch an die Projektfragestellungen angepasste Instrumente (Codierhandbücher, Transkriptionsleitfäden etc.) aufbereitet, so dass die gewünschten Lehr-/Lernprozesse quantifizierbar gemacht werden können (ausführlich hierzu Praetorius, 2014; Seidel & Prenzel, 2010).

Wie Lehrkräfte die Herausforderungen im DaZ-Unterricht bewältigen können, ist ein Forschungsdesiderat, und die bisherigen Erkenntnisse der allgemeinen Lehr-/Lernforschung sind häufig domänenspezifisch. Sie beziehen sich mehrheitlich auf Untersuchungen der Schulforschung in naturwissenschaftlichen Unterrichtsfächern. Übertragungen der Befunde in den DaZ-Bereich ermöglichen die bisherigen Studien kaum. Dies allein mag schon Grund genug sein, die Lehr-/Lernprozesse im DaZ-Bereich im Detail zu analysieren. Hinzu kommt, dass sich mit zunehmender Heterogenität der Lerngruppen auch die Struktur und Organisation von Lehr-/Lernprozessen grundsätzlich zu ändern scheint (siehe hierzu Vock & Gronostaj, 2017). So werden offenbar bestimmte Sozialformen, wie das offene Unterrichtsgespräch, welche lange Jahre einen großen zeitlichen Anteil im Unterrichtsgeschehen eingenommen haben, zunehmend von anderen Sozialformen (Einzel-, Partner- und Gruppenarbeit) verdrängt (Götz et al., 2005). Als gesichert geltende Erkenntnisse über Lehr-/Lernprozesse beziehen sich jedoch häufig nur auf das Geschehen im offenen Unterrichtsgespräch (Ackermann, 2011; Frank Lipowsky et al., 2007). Andere Sozialformen wie Einzel-, Partner-, und Gruppenarbeit wurden bisher wenig berücksichtigt. Für das Feld der Erwachsenen- und Weiterbildungsforschung sind die Analyse von Lehr-/Lernprozessen in Spracherwerbskursen und daraus resultierende Erkenntnisse über Wirkungen und Effekte noch gänzlich Forschungsdesiderate (Schrader, 2015). Die aufgeführten Videostudien der Schulforschung sind primär auf die Analyse didaktischer und pädagogisch-psychologischer Prozesse fokussiert und die Daten sind hinsichtlich der sprachlichen Äußerungen der Lehrkräfte und Lernenden nicht so umfangreich, dass eine Analyse des Umgangs mit sprachlich heterogenen Lerngruppen durch die Lehrkraft möglich wäre.

Im Bereich der Erwachsenenbildungsforschung gibt es zwar durchaus Studien, die sich auf Videoaufnahmen in Erwachsenenbildungskursen beziehen (ausführlich siehe hierzu Kade, Nolda, Dinkelaker & Herrle, 2014), jedoch ohne die Lehr-/Lernprozesse explizit in den Blick zu nehmen. Im Fokus stehen hingegen eher erwachsenendidaktische Aspekte der Raumgestaltung, die Beschreibung der Lernumgebungen und Kursgruppen sowie eine generelle Analyse des Medieneinsatzes. Videogestützte Analysen, die die Partizipation und Mitarbeit in der Lehr-/Lernsituation adressieren, gibt es bisher im Bereich der Erwachsenenbildung nicht.

Möchte man jedoch insbesondere die Interaktionsprozesse und die Adaptivität von Lehrkräften im Unterricht in den Blick nehmen, fehlen in den Datensätzen häufig relevante Daten. Zumeist fehlen auch Informationen zur Einschätzung der Reliabilität von Transkriptionen der Unterrichtsgespräche zwischen Lehrkräften und Lernenden. Somit erlauben die bisherigen Datensätze keine nachträgliche Analyse des Unterrichtsgesprächs unter der Berücksichtigung, ob die Lehrkräfte unterschiedlich mit verschiedenen Lernenden sprechen, da selten Metadaten zur verbalen Interaktion erfasst wurden. Zum Beispiel ist die Information, an welche/n Lernende/n eine verbale oder auch nonverbale Aussage adressiert wurde, in den seltensten Transkripten dokumentiert. Um Aussagen über die Anpassung von Lehrkräften an Lernende mit unterschiedlichen sprachlichen Fähigkeiten zu treffen, sind diese Informationen aber zwingend notwendig. Eine weitere Hürde, die die Nachnutzung bestehender Videodaten erschwert, ist die Notwendigkeit, über ausführliche Meta-Informationen zu den in der gefilmten Unterrichtsstunde beteiligten Akteuren zu verfügen. Für Aussagen über den Umgang von Lehrkräften mit sprachlicher Heterogenität braucht es, neben Informationen über die jeweiligen deutschsprachlichen Fähigkeiten der Lernenden zum Zeitpunkt der gefilmten Videostunde, im Idealfall auch Informationen über die von der Lehrkraft wahrgenommenen sprachlichen Fähigkeiten jedes Lernenden, da davon auszugehen ist, dass Letzteres relevanter ist für das Handeln der Lehrkräfte. Möchte man darüber hinaus untersuchen, inwieweit die Lehrkräfte aufgrund bestimmter Hintergrundinformationen, wie z.B. berufsbiographische Qualifikationen, sprachliche Heterogenität mehr oder weniger adaptiv adressieren, sind weitere Meta-Informationen über die Fähigkeiten der Lehrkräfte selbst notwendig. Schlussendlich scheitert eine mögliche Nachnutzung aber maßgeblich an dem Fakt, dass es in der Bildungsforschung für den DaZ-Bereich keine Videoaufnahmen in Vorbereitungsklassen und Integrationskursen gibt.

In den Ergebnissen des Reviews (Beitrag 1) wurde deutlich, dass Spracherwerbsforschung oft multidisziplinär untersucht wird und Ergebnisse dabei aber häufig ohne wechselseitige Bezugnahme produziert und nicht weiter rezipiert werden. Empirische Untersuchungen aus anderen Disziplinen, wie z.B. aus der Linguistik, zeigen durch ihre Anschlussfähigkeit an erwachsenenpädagogische und psychologische Fragestellungen, dass hier eine interdisziplinäre Zusammenarbeit sinnvoll ist und Synergieeffekte vielversprechend sind. Besonders da die Bedeutung der verschiedenen Faktoren für die Gelingensbedingungen von Spracherwerb aus der bisherigen Forschung noch nicht ableitbar ist und es noch an generalisierenden, theorieprüfenden Studien mangelt (Sahlender & Schrader, 2017).

4.2. Interdisziplinäre Ansätze zur Analyse von Lehr-/Lernprozessen

Neben der erziehungswissenschaftlichen und der pädagogisch-psychologischen empirischen Bildungsforschung beschäftigt sich auch die Linguistik aus spracherwerbstheoretischen Gesichtspunkten mit sprachlicher Interaktion in Klassenzimmern und Kursräumen. Die dabei entstehenden Datenansammlungen von sprachlicher Interaktion werden gesprochene Lernerkorpora genannt und beinhalten neben sprachlichen Äußerungen auch Annotationen und Meta-Informationen über die jeweilige Zielgruppe und deren Sprachgebrauch in aufbereiteter Form, so dass sie (für die Computerlinguistik computerbasiert und automatisch) analysiert werden können (Wisniewski, 2022). Im Vergleich zu schriftlichen Korpora, deren Daten sich leicht(er) erheben lassen (z.B. aus digitalen Zeitschriftenarchiven, Filmuntertiteln oder Twitter-Beiträgen), ist die Erstellung von gesprochenen Lernerkorpora ungleich aufwändiger. Der damit einhergehende Ressourcenaufwand bedingt sich dabei durch den Umfang und die Erschließungstiefe der Inhalte. Je nach Fragestellung bedarf es unterschiedlich detaillierter Transkriptionen und Annotationen (wie z.B. die Erfassung phonetischer Besonderheiten und nonverbaler Kommunikationsmerkmale), die die Aufbereitungszeit von Audio- und Videodaten erhöhen (Wisniewski, 2022).

Die Bedeutung und das Potenzial von gesprochenen Lernerkorpora für die Analyse von DaZ-Unterricht ist als besonders hoch einzuschätzen, da gesprochene Sprache es im Vergleich zu geschriebener Sprache besser ermöglicht, Zusammenhänge zwischen sprachlichen Eigenschaften und sozialen, kognitiven und psycholinguistischen Faktoren der Lernenden zu überprüfen (Hall, 2022). Auch im Bereich der Zweitspracherwerbsforschung kann das dazu beitragen, die Gelingensbedingungen erfolgreicher Lehr-/Lernprozesse besser zu verstehen. Aber auch zur Analyse von

sprachlichen Fähigkeiten, die im Rahmen des Gemeinsamen Europäischen Referenzrahmens ein ausgewiesenes Ziel der Sprachkompetenz ist (Trim et al., 2009), sind gesprochene Lernerkorpora geeignet, da sie wichtige Informationen darüber enthalten, wie und in welcher Komplexität Lernende gesprochene Sprache im Unterrichtskontext anwenden (Wisniewski, 2022).

Betrachtet man die existierenden Korpora, wird deutlich, dass die Zahl der Korpora, welche sprachliche Interaktion in authentischen DaZ-Erwerbssituationen beinhalten, sehr gering ist. Das belegt ein systematisches Review, in dem Paquot und Plonsky (2017) darauf verweisen, dass nur ein Viertel von 378 untersuchten Studien gesprochene Lernerkorpora (die Transkriptionen mündlicher Sprache enthalten) untersuchen. Die übrigen beruhen auf schriftsprachlichen Korpora. Von diesem Viertel verfügen 2 % (7 Studien) der gesprochenen Lernerkorpora über interaktive Daten von mehreren Lehrenden und Lernenden. Der Großteil der anderen gesprochenen Lernerkorpora beinhalten überwiegend Monologe und Dialoge unterschiedlicher thematischer Inhalte. Gründe für die geringe Anzahl an gesprochenen Lernerkorpora liegen in der ressourcenintensiven Erstellung, da besonders interaktive Daten, wie sie in Lehr-/Lernsituationen vorkommen, eine zeitintensive Transkription erfordern (Wisniewski, 2022).

Ein Kritikpunkt an vielen gesprochenen Lernerkorpora ist, dass sie zu wenige oder keine deskriptiven statistischen Angaben über die quantitativen Daten beinhalten beziehungsweise die Parameter zur Datenaufbereitung (wie Annotations- und Transkriptionsreliabilität) fehlen oder wenig transparent sind (Wisniewski, 2022). Weiterhin ist zu bemerken, dass eine Vielzahl an gesprochenen Lernerkorpora keine authentischen, sondern (quasi-)experimentelle Daten enthält (Clercq & Housen, 2019; Lintunen & Mäkilä, 2014; Pallotti, 2019). Letzteres hat den Vorteil, dass eine Vielzahl an Faktoren, die den Spracherwerb beeinflussen können, methodisch kontrolliert werden kann – allerdings auch zu Lasten der externen Validität (Eid et al., 2013) der Untersuchungen. Um Aussagen über den Ist-Zustand von Spracherwerbssituationen treffen zu können, sind gesprochene Lernerkorpora aus authentischen Lehr-/Lernsituationen sehr hilfreich. Vorhandene gesprochene Lernerkorpora sind wenig repräsentativ bezüglich der in den Korpora enthaltenen Stichprobe. Oft sind es Sammlungen von Lernenden aus dem Bereich der universitären Lehre mit gehobener Bildungssprache, wohingegen gesprochene Lernerkorpora für die Gruppe der Zugewanderten fehlen (Granger et al., 2015; Wisniewski, 2022).

Es gibt, bis auf das Projekt COLD, noch kein Transkriptionssystem für die Beantwortung von Fragestellungen, die die Adaptivität der Sprache von Lehrkräften an unterschiedliche Lernende im Sprachbereich untersuchen und dabei gleichzeitig reichhaltige Meta-Informationen über Lehrkräfte und Lernende bereithalten.

4.3. Kompetenzen von Lehrkräften in Schule und Erwachsenenbildung im Bereich Deutsch als Zweitsprache beim Unterrichten von sprachlich heterogenen Lerngruppen

Die beiden dieser Dissertationsschrift zugrunde liegenden Beiträge #2 und #3 sind im Rahmen des Verbundprojekts *Competencies of School Teachers and Adult Educators in Teaching German as a Second Language in Linguistically Diverse Classrooms (COLD)* entstanden. Es wurde im Rahmen des Förderprogramms *Kooperative Exzellenz* mit Mitteln des Leibniz-Wettbewerbs von 2019–2022 finanziert.

In dem Projektteam waren neben dem Deutschen Institut für Erwachsenenbildung – Leibniz Zentrum für Lebenslanges Lernen e.V. auch Wissenschaftlerinnen und Wissenschaftler des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache, der Eberhard Karls Universität Tübingen, der Universität zu Köln, der Technischen Universität Dortmund, der Universität Zürich sowie der beiden Leibniz-Institute für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel und für Bildungsforschung und Bildungsinformation (DIPF) in Frankfurt beteiligt. Dadurch konnte das Projektteam interdisziplinär mit Expertise aus den Fachbereichen Erwachsenenbildung/Weiterbildung, der Fachdidaktik Deutsch/Deutsch als Zweitsprache, der empirischen Bildungsforschung, der Linguistik, der Computerlinguistik und der Psychologie arbeiten.

Im Rahmen des Projekts COLD wurden die professionellen Kompetenzen von Lehrenden in Erwachsenenbildung und Schule beim Unterrichten von Deutsch als Zweitsprache in sprachlich heterogenen Lerngruppen untersucht. Das Projekt wurde mit dem Ziel durchgeführt, die Fragen zu untersuchen, über welche professionellen Kompetenzen Sprachlehrkräfte in Vorbereitungsklassen und Integrationskursen verfügen, um schlussendlich, basierend auf den Befunden, Empfehlungen zur Professionalisierung und Unterstützung von Sprachlehrkräften abzuleiten.

Für Beitrag #2 und Beitrag #3 der vorliegenden Arbeit wurden Daten aus der COLD-Pilotstudie verwendet. Die Stichprobe dafür haben je fünf Lehrkräfte aus Vorbereitungsklassen und fünf Lehrkräfte aus Integrationskursen mit ihren Lernenden gebildet. Der Teildatensatz, der die Datengrundlage für Beitrag #2 und #3 bildet, besteht

aus einer 45-minütigen DaZ-Unterrichtsstunde, die mit zwei Videokameras so aufgenommen wurde, dass gleichzeitig Lehrkraft und Lernende erfasst wurden. Weiterhin wurden Befragungsdaten der Lehrkraft über die wahrgenommenen sprachlichen DaZ-Fähigkeiten der einzelnen Lernenden erfasst. Die Items hierfür wurden im Projekt entwickelt und sind angelehnt an Spinath (2005). Weiter wurden Informationen darüber erhoben, wie lange die Lehrkräfte die einzelnen Lernenden bereits unterrichten und wie hoch der eingeschätzte Punktscore in einem DaZ-Leistungstest (C-Test, Grotjahn, 2014) sein wird, den die Lernenden im Anschluss an die videographierte Unterrichtsstunde absolviert haben. Somit liegen standardisierte Leistungsdaten über die Fähigkeitswerte im Bereich DaZ der einzelnen Lernenden vor. Gleichzeitig können diese Daten um die von der Lehrkraft wahrgenommenen DaZ-Fähigkeiten der einzelnen Lernenden ergänzt werden. Letzteres ermöglicht es z.B., die Anpassung von Lehrkräften an sprachlich heterogene Lerngruppen zu untersuchen, auch wenn dabei noch nicht gewährleistet ist, dass Lehrkräften stets eine treffende Diagnostik der sprachlichen Leistungsfähigkeit eines jeden einzelnen Lernenden gelingt.

5. Fragestellungen der empirischen Studien

In den bisherigen Kapiteln konnte die zentrale Bedeutung der Sprache aufgezeigt werden und es wurde deutlich gemacht, welche große Rolle die Beherrschung der Sprache des Ziellandes, in dem jemand lebt, auf nahezu alle Bereiche des gesellschaftlichen Lebens hat. Ebenfalls wurde gezeigt, wie durch organisierte formale Bildungsprozesse in Form von Integrationskursen, Vorbereitungsklassen und Alphabetisierungskursen versucht wird, Spracherwerbsprozesse zu ermöglichen. Im Anschluss wurde unter anderem anhand des Angebots-Nutzungs-Modells wirksamen Unterrichts die Bedeutung aufgezeigt, dass sowohl die Sprache als Medium und Inhalt als auch die Lehrkräfte bei der Gestaltung der Lehr-/Lernprozesse von zentraler Bedeutung sind. Es wurde auch dargelegt, welche Bedeutung die Zone of Proximal Development (Zone der nächsten Entwicklung) für erfolgreichen Spracherwerbsprozess hat und wie diese bei sprachlich heterogenen Lerngruppen die Lehrkräfte beim Unterrichten herausfordert. Danach wurden Möglichkeiten aufgezeigt, wie Lehrkräfte durch adaptives Unterrichtshandeln verbal und nonverbal auf sprachliche Heterogenität reagieren können. Das vierte Kapitel verdeutlicht, wie Lehr-/Lernprozesse in der Bildungsforschung anhand von videogestützter Unterrichtsbeobachtung und in der Linguistik anhand gesprochener Lernerkorpora empirisch überprüfbar werden und wie im Forschungsprojekt COLD die Untersuchung von Kompetenzen von Lehrkräften beim Unterrichten von DaZ in sprachlich heterogenen Lerngruppen durchgeführt wurde. Basierend auf diesen bisherigen Erkenntnissen schließen sich die Forschungsfragen der drei folgenden Beiträge an.

Teilfrage 1: Wie wird Spracherwerb beforscht? Was sind aktuelle Forschungsfragen? Welche methodischen Zugänge werden verwendet und welche Erkenntnisse zum Spracherwerb Erwachsener werden gewonnen?

Die erste Studie bildet mit einem breit angelegten Überblick über die derzeitige Beforschung und die methodischen Zugänge sowie Befunde zum Spracherwerb bei Erwachsenen die Basisstudie für das vorliegende Promotionsvorhaben. Im Rahmen eines Literaturreviews wird untersucht, auf welche Art und Weise der Spracherwerb Erwachsener in den letzten Jahren beforscht wurde. Die Literatur wird u.a. daraufhin gesichtet, ob eher theoriegenerierende oder eher theorieprüfende Studien durchgeführt wurden. Von Interesse sind zudem Befunde zu den Lehrenden, den Lernenden, deren Interaktion in Lehr-/Lernprozessen sowie die Rolle didaktischer Materialien,

insbesondere mit Blick auf den Einsatz digitaler Materialien beim Spracherwerb. Mithilfe der ersten Studie wurde die Frage untersucht, welche Themen den wissenschaftlichen Diskurs bestimmen, mit welchen methodischen Zugängen diese Themen bearbeitet werden und welche Forschungsstrategien dabei verfolgt werden. Ebenfalls konnte im Rahmen der ersten Studie gezeigt werden, welche Erkenntnisse zum Spracherwerb Erwachsener bereits gewonnen wurden und welche Forschungsdesiderate weiterhin bestehen.

Teilfrage 2: Wie können Videoaufnahmen dazu genutzt werden, Lehr-/Lernprozesse so zu beforschen, dass Lehr-/Lernprozesse in sprachlich heterogenen Lerngruppen messbar werden?

Die zweite Studie greift eines der in Studie #1 identifizierten Forschungsdesiderate auf. Im Rahmen des Literaturreviews wurde sichtbar, dass es bisher nur unzureichende Evidenz über belastbare Befunde zu den Bedingungen erfolgreicher Lehr-/Lernprozesse im Bereich DaZ gibt. In der zweiten Studie wird untersucht, wie Videoaufnahmen von DaZ-Kursen so aufbereitet werden können, dass die Analyse von Lehr-/Lernprozessen möglich wird. Dafür wurde ein Transkriptionsleitfaden erstellt, der es ermöglicht, verbales und nonverbales Lehr-/Lernhandeln zu transkribieren und annotieren. Die Robustheit, im Sinne von Reliabilität der Transkription und Annotation wird ebenfalls im Rahmen dieser Studie anhand experimenteller Analysen überprüft und berichtet. Die Erkenntnisse der zweiten Studie zeigen, wie eine empirische Untersuchung von Lehr-/Lernprozessen in sprachlich heterogenen Lerngruppen anhand von Unterrichtsvideos im DaZ-Unterricht möglich ist.

Teilfrage 3: Wie gestalten sich Lehr-/Lernprozesse im DaZ-Unterricht am Beispiel nonverbaler Kommunikation der Lehrkraft? Berücksichtigen Lehrkräfte beim Einsatz ihrer Gesten die wahrgenommenen sprachlichen Fähigkeiten einzelner Lerner?

In der dritten Studie wird am Beispiel der nonverbalen Kommunikation der Lehrkraft untersucht, ob Lehrkräfte die Lehr-/Lernprozesse im DaZ-Unterricht adaptiv gestalten und ihre Interaktionen an den sprachlichen Fähigkeiten der Lernenden ausrichten. Konkret wird die Frage untersucht, welche Arten von Gesten von Lehrkräften im Unterricht eingesetzt werden und ob der Einsatz von Lehrkraftgesten zwischen den Lehrkräften und den Lernenden einer Klasse variiert. Basierend auf bisheriger Forschung wird in der dritten Studie ebenfalls untersucht, ob die Lehrkräfte die Gesten

adaptiv einsetzen und ob die Komplexität der eingesetzten Gesten durch die wahrgenommenen sprachlichen Fähigkeiten der Lernenden durch die Lehrkraft zu erklären ist.

6. Beitrag #1 Spracherwerb bei Zugewanderten und funktionalen Analphabeten

Sahlender, M. & Schrader, J. (2017). Spracherwerb bei Zugewanderten und funktionalen Analphabeten - Konzepte, Erfahrungen und Befunde für die Erwachsenenbildung. *Zeitschrift für Weiterbildungsforschung*, 40 (3), 275–293.

© 2022 Springer Nature Switzerland AG. Part of [Springer Nature](#). Dieser Artikel ist nicht die originalgetreue Kopie. Das Original ist verfügbar unter <https://doi.org/10.1007/s40955-017-0098-9>.

Zusammenfassung

Die Fähigkeit, sich in der jeweiligen Landessprache mündlich und schriftlich zu verständigen, gilt als eine unverzichtbare Voraussetzung für Teilhabe am öffentlichen und beruflichen Leben. Die Erwachsenenbildung sieht sich vor diesem Hintergrund auch in Deutschland nicht erst aktuell mit großen Herausforderungen konfrontiert. Denn neben der Vermittlung von Sprachkenntnissen an Migrantinnen und Migranten sowie Geflüchtete und Vertriebene, die Deutsch als Erst- oder Zweitsprache erlernen, wurde und wird auch die Alphabetisierung (zugewanderter) funktionaler Analphabeten bildungspolitisch als eine zentrale Aufgabe angesehen. Der vorliegende Literaturüberblick stellt die Frage, welche Bedingungen die jüngere Forschung für einen erfolgreichen Spracherwerb dieser Adressatengruppen identifiziert hat. Dazu wurde eine systematische Literaturrecherche in einschlägigen nationalen und internationalen Datenbanken (FIS-Bildung, ERIC) durchgeführt. Für die Jahre 2010–2016 wurden insgesamt 314 Veröffentlichungen inhaltsanalytisch ausgewertet, um genauer zu erfassen, wie die Erwachsenenbildung und benachbarte Wissenschaftsdisziplinen den Spracherwerb der hier fokussierten Adressatengruppen in organisierten Lehr-Lernprozessen untersuchen. Der Literaturüberblick zeigt u. a., dass Studien zur Erforschung des Spracherwerbs aus den Bereichen der Erwachsenenbildung, der Psychologie und der Linguistik vorliegen, wobei disziplinspezifisch sehr unterschiedliche Fragen gestellt und Forschungsmethoden angewendet werden. Trotz einer beachtlichen Forschungsaktivität lassen sich derzeit jedoch nur wenige belastbare Befunde zu den didaktischen, medialen, professionellen und individuellen Bedingungen eines erfolgreichen Spracherwerbs benennen.

Schlüsselwörter:

Spracherwerb, Sprachen lernen, Lehr-Lern-Forschung

Abstract

Learning a language is an indispensable prerequisite for participating in both public and working life. Adult education, also in Germany, is currently facing a great challenge in the area of language teaching: In addition to language teaching for migrants and refugees, a key problem is the alphabetization of German and non-German speaking functional illiterates. This paper looks at the conditions for successful language learning by the adults here in focus in organized teaching and learning settings as identified by research across different disciplines. A systematic literature research was conducted, using relevant databases (FIS-Bildung, ERIC), upon which 314 publications were identified and subjected to a content analysis. The literature research focuses on three questions: First, how is adult language learning in organized teaching and learning settings discussed in the field of adult resp. further education and other disciplines. Second, which research questions were pursued during the years 2010 and 2016. And third, what suggestions are proposed to adequately support teachers and learners. To illustrate the wide scope and heterogeneity of the reviewed research on language teaching and learning, this paper presents and discusses examples of different topics. The results indicate that research in language teaching and alphabetization is predominantly conducted within the fields of adult education, psychology and linguistics, applying a broad variety of discipline-specific methods. The review also shows that within the span of the analysis only few studies produced reliable results regarding the effectiveness of successful language acquisition.

Keywords:

language acquisition, language learning, teaching and learning

1. Einleitung

Bereits vor mehr als dreißig Jahren attestierte Hans Tietgens insbesondere den Organisationen der öffentlich anerkannten Erwachsenenbildung eine „Nothelferfunktion“ und kritisierte zugleich ihre Reduktion darauf in Teilen der öffentlichen Wahrnehmung (vgl. Tietgens, 1986). Auch aktuell werden der Erwachsenenbildung Aufgaben zugetragen, die diese Bezeichnung rechtfertigen (vgl. Schrader, 2017) und sie vor mindestens zwei große Herausforderungen stellen: Die erste umfasst – bereits seit längerem – den erfolgreichen Spracherwerb von Migrantinnen und Migranten sowie – gegenwärtig besonders – von Geflüchteten und Vertriebenen. Dabei übersteigt die Nachfrage durch die hohe Zahl der Geflüchteten und Vertriebenen nach Deutschland die vorhandenen Angebotskapazitäten im Bereich Deutsch als Fremd- oder Zweitsprache (vgl. DVV, 2015). Zweitens hat insbesondere die öffentlich verantwortete Erwachsenenbildung die ihr zugeschriebene Aufgabe übernommen, gesellschaftliche Partizipationsmöglichkeiten (funktionaler) Analphabeten zu verbessern. Dies trifft zunächst für deutschsprachige Erwachsene zu, aber auch für Zugewanderte, die in ihrem Heimatland keine hinreichenden schriftsprachlichen Kompetenzen erworben haben. So gilt, dass Lehrende mit einer Heterogenität von Lerngruppen im Hinblick auf das Vorwissen, die Herkunftssprachen, die kognitiven Fähigkeiten und die Motivationen konfrontiert werden, die deutlich größer als im schulischen Bereich ist. Dass die Bedeutung dieser Herausforderung gesellschaftlich wahrgenommen wird, zeigt sich unter anderem in der Ausrufung der Nationalen Dekade für Alphabetisierung und Grundbildung 2016–2026 durch Bund und Länder (vgl. BMBF/KMK, 2016), die von einem breiten Bündnis zivilgesellschaftlicher Akteure gestützt wird. Die zahlreichen Forschungs- und Entwicklungsarbeiten in diesem Feld, die teils bereits aus vorangegangenen Förderprogrammen stammen, sind inzwischen auch in einem viel beachteten Handbuch dokumentiert (vgl. Löffler und Korfkamp, 2016). Der Erwerb der Landessprache, d. h. der in der jeweiligen Gesellschaft favorisierten Verkehrssprache, wird auch in mehrsprachigen Gesellschaften (vgl. Gogolin, 2013) weithin unbestritten als eine elementare Voraussetzung für die Integration in Gesellschaft und Arbeitsmarkt betrachtet. Diese verbreitete Überzeugung stützt sich auf empirische Befunde, die zeigen, dass ein erfolgreicher Spracherwerb positive Auswirkungen auf spätere Erwerbssituationen hat (für Zugewanderte z. B. Deeke, 2011; Isphording und Sinning, 2012), aber auch die Teilhabe an sozialen Aktivitäten befördert (Costa et al., 2014).

Doch was wissen wir über die Voraussetzungen und Gelingensbedingungen eines erfolgreichen Spracherwerbs Erwachsener? Welche Rolle spielen das Vorwissen und die Kompetenzen sowohl der Lernenden als auch der Lehrenden? Unter welchen didaktischen Bedingungen und auf welche Weise werden Lehr-Lernprozesse wirksam? Welche Konzepte, Methoden und Materialien versprechen Erfolg, und wie lassen sie sich angemessen einsetzen? Der folgende Literaturüberblick sichtet den nationalen und internationalen Forschungsstand zum Spracherwerb Erwachsener dahingehend, welche Fragen die Wissenschaft der Erwachsenenbildung und andere Wissenschaftsdisziplinen stellen und zu welchen Ergebnissen sie dabei kommen.

2. Ziel des Beitrags und methodisches Vorgehen

Das Ziel dieses Beitrags ist es, den aktuellen Forschungsstand dahingehend zu sichten und zu analysieren, auf welche Art und Weise der Spracherwerb Erwachsener in den letzten Jahren beforscht wurde. Dabei konzentrieren wir uns auf jene Lehr-Lernprozesse, die Kompetenzen in der jeweiligen Landessprache fördern sollen. Je nach biographischer Lebenssituation der Adressaten handelt es sich also um Prozesse des Zweitspracherwerbs, des Fremdsprachenlernens und/oder der (funktionalen) Alphabetisierung. Nur vereinzelt betrachtet wird die Forschung zu Angeboten, die sich mit dem Fremdspracherwerb im Allgemeinen beschäftigen, d. h. mit dem Erlernen fremder Sprachen durch Personen, die nicht Zugewanderte sind oder die über hinreichende schriftsprachliche Kompetenzen verfügen (vgl. dazu Nolda, 2017). Das Interesse ist darauf gerichtet, relevante Themen, Forschungsstrategien und Erkenntnisse zu sichten und Forschungsdesiderata zu benennen.

Für den Bereich Alphabetisierung haben Grosche und Schroeder (2013) in einem Überblicksartikel den aktuellen Forschungsstand sowie die aktuellen Forschungsdesiderata bereits aufgezeigt. Dieser Überblick ist allerdings weniger auf organisierte Lehr-Lernprozesse des Spracherwerbs (funktionaler) Analphabeten als vielmehr auf allgemeine Fragen der Alphabetisierungsforschung hin ausgerichtet. Für die Alphabetisierung von Migrantinnen und Migranten verweisen Feick und Schramm (2016) auf ein erst langsam steigendes Forschungsinteresse und stellen fest, dass bislang oft methodisch-didaktische und materialbezogene Fragen fokussiert werden. Diesen Adressaten widmet sich auch das so genannte „LESSLA“-Netzwerk (Literacy Education and Second Language Learning for Adults), das 2005 mit dem Ziel gegründet wurde, die weniger gebildeten Zweitsprachenlernende und Teilnehmenden an

Alphabetisierungskursen in den Fokus wissenschaftlicher Forschung zu stellen. Die Aktivitäten umfassen sowohl Reviews zu Effekten experimenteller Spracherwerbsforschung (Kurvers et al., 2015) als auch empirische Arbeiten zu Fragen des Lehrens und Lernens bzw. zu didaktischen Materialien sowie zu Einstellungen und Überzeugungen verschiedener Akteure (z. B. Huettig, 2015; Feldmeier, 2015; Young-Scholten et al., 2015). Die Perspektiven der hier begonnenen Arbeit werden als positiv eingeschätzt (vgl. van de Craats, Kurvers und van Hout, 2015; Feick und Schramm, 2016, S. 215).

Der vorliegende Beitrag bietet einen breiter angelegten Überblick über aktuelle Forschungsfragen, methodische Zugänge und Befunde zum Spracherwerb Erwachsener. Die Literatur wird u. a. daraufhin gesichtet, ob eher theoriegenerierende (qualitative) oder eher theorieprüfende (quantitative) Studien durchgeführt wurden. Von Interesse sind zudem Befunde zu den Lehrenden, den Lernenden, deren Interaktion in Lehr-Lernprozessen sowie die Rolle didaktischer Materialien, insbesondere im Blick auf den Einsatz digitaler Materialien beim Spracherwerb.

Um diese Fragen zu beantworten, wurde eine breit angelegte Literaturrecherche durchgeführt. Dabei lag der Fokus auf empirischen Studien aus der Zeit von Januar 2010 bis Juli 2016. Zusätzlich wurden aber auch relevante Lehrwerke und didaktische Materialien erfasst. Im Kern ging es um Publikationen, die sich mit organisierten Lehr-Lernprozessen für Erwachsene beschäftigen. Die deutlich weiter entwickelte schulbezogene Forschung wurde hier aus Gründen der Arbeitsökonomie zunächst ebenso ausgeklammert wie ältere Forschungsarbeiten. Um das Thema Spracherwerb in allen Facetten zu erfassen und eventuelle Unterschiede zwischen den einzelnen Themenbereichen zu identifizieren, wurden Forschungsarbeiten aus den Bereichen Deutsch als Zweitsprache (DaZ), Deutsch als Fremdsprache (DaF) und Alphabetisierung in die Recherche einbezogen. Wir teilen die verbreitete Auffassung, dass sich Lehr-Lernprozesse im Bereich DaF grundlegend von denen im DaZ-Bereich und in Alphabetisierungskursen unterscheiden. Während im DaF-Bereich Menschen im Heimatland eine Fremdsprache erwerben, ist im DaZ-Bereich das Erlernen der Sprache im Zielland gemeint. Mit dieser eher formalen Unterscheidung geht einher, dass in der Regel sehr unterschiedliche Adressatengruppen angesprochen sind, im ersten Bereich häufig gut qualifizierte, im zweiten Bereich auch gering qualifizierte Personen. Besonders für Geflüchtete und Vertriebene, die im Heimatland sozial- und bildungsbenachteiligt waren, ist ein erfolgreicher Spracherwerb ungleich anspruchsvoller, oft verstärkt durch psychische Probleme und Traumata. Für einen Überblicksartikel

scheint es uns dennoch sinnvoll, alle drei Bereiche des Spracherwerbs Erwachsener in den Blick zu nehmen, da die Erkenntnisse aus den verschiedenen Bereichen (DaZ, DaF und Alphabetisierung) durchaus wechselseitig relevant sein können.

Für die Recherche wurden sowohl nationale als auch internationale Veröffentlichungen berücksichtigt. Neben der traditionellen Spracherwerbsforschung im Feld der Erwachsenenbildung, die jüngst von Nolda erneut als ein „vernachlässigtes Feld“ charakterisiert wurde (vgl. Nolda, 1989; 2017), schließt die vorliegende Recherche auch Arbeiten aus benachbarten Disziplinen, wie z. B. der Linguistik, der Unterrichtsforschung und der Psychologie, ein.

Im deutschsprachigen Bereich wurden über das Fachinformationssystem Bildung (FIS-Bildung) alle Veröffentlichungen zum Spracherwerb im Bereich der Erwachsenen- und Weiterbildung gesichtet. Hierbei handelt es sich um 209 Veröffentlichungen, die anhand der mit Platzhaltern versehenen Suchwörter ‚*Spracherwerb*‘ und ‚*erwachsene**‘ sowie ‚**sprach**‘ und ‚*Erwachsene**‘ als auch ‚**sprach**‘ und ‚*Weiterbildung*‘ erschlossen wurden. Für die internationale Forschung wurde die Suche mit Hilfe des Education Resources Information Center (ERIC) durchgeführt. Die Suche ist auf die Zeit von Januar 2015 bis Juli 2016 beschränkt, da die Anzahl der Treffer den Umfang der Analyse andernfalls enorm ausgeweitet hätte. Gesucht wurde nach referierten Beiträgen der Kategorie ‚*second language learning*‘ aus dem Bereich ‚*adult education*‘. Diese Suche ergab 47 Veröffentlichungen. Zusätzlich wurde die Suchmaschine Sagepub des SAGE-Publications Verlag anhand der Suchbegriffe ‚*adult**‘ und ‚*language*‘ durchsucht. Hierdurch wurden weitere 58 Veröffentlichungen zum Spracherwerb Erwachsener identifiziert. Im Folgenden werden die Ergebnisse der Literaturrecherche vorgestellt.

3. Ausgewählte Befunde

Im Rahmen der Literaturrecherche wurden insgesamt 314 Artikel erfasst und auf Basis ihrer Abstracts nach den oben aufgezeigten Kriterien kategorisiert. Unter den 314 Artikeln befanden sich auch Veröffentlichungen, die zwar in den Suchmaschinen als passend identifiziert worden waren, allerdings aus folgenden Gründen aus der weiteren Analyse ausgeschlossen wurden: 17 Artikel befassten sich ausschließlich mit dem Spracherwerb von Kindern und elf Berichte über Tagungen und Konferenzen präsentierten keine eigene Forschung. Ferner wurden 19 Veröffentlichungen nicht berücksichtigt, da sie mit maximal drei Seiten eher eine kurze Zusammenfassung als

einen wissenschaftlichen Beitrag anboten. Acht der ausgeschlossenen Artikel behandelten im Kern soziologische oder andere nicht pädagogische Fragestellungen (z. B. nach den soziodemographischen Merkmalen der Teilnehmenden). Weitere sieben Veröffentlichungen widmeten sich zwar dem Spracherwerb Erwachsener, allerdings ohne Bezug auf organisierte Lehr-Lernprozesse. 16 Beiträge erwiesen sich schließlich aus sonstigen Gründen als nicht passend, da es sich z. B. um bildungspolitische Anfragen oder Stellungnahmen handelte.

Die 236 letztlich als passend zum Themenfeld Spracherwerb in der Erwachsenen- und Weiterbildung identifizierten Veröffentlichungen lassen sich wie folgt systematisieren (Tabelle 1):

Tabelle 1: Übersicht zu Veröffentlichungen und Untersuchungsdesigns (N=236)^a

		Untersuchungsdesign				
		qualita- tiv	quantita- tiv	triangula- tiv	nicht eindeutig/ keine empiri- sche Forschung	ge- samt
Kategorien	Lehrbücher	-	-	-	47	47
	Lehrende	3	14	1	-	23
	Lernende	10	22	1	8	41
	Lehr-Lernpro- zesse	58	12	3	18	91
	Digitale Medien für Lehrende	2	3	-	3	8
	Digitale Medien für Lernende	5	3	-	6	14
	Unterrichts- und Kursmethoden	16	25	2	15	58

^a 236 Veröffentlichungen erfasst (von 314 Veröffentlichungen); ausgeschlossen wurden Veröffentlichungen ohne Bezug zu organisierter Lehr-Lern-Forschung, Berichte über Tagungen und Konferenzen, zu kurze Artikel (≤ 3 Seiten) sowie Artikel über den Spracherwerb von Kindern.

111 der 236 Veröffentlichungen sind in englischer Sprache verfasst. Mit Blick auf die Forschungsmethodik lässt sich feststellen, dass sich 73 Arbeiten einem quantitativen Untersuchungsdesign und 50 Arbeiten einem qualitativen Design zuordnen lassen; weitere sechs Artikel stützen sich auf ein triangulierendes Vorgehen mit

qualitativen und quantitativen Elementen. Des Weiteren finden sich im Analysepool neun systematische Reviews und 15 Überblicksartikel. 32 Veröffentlichungen beschreiben didaktische Methoden oder didaktische Konzepte, mit denen Lehrende den Spracherwerb Erwachsener unterstützen können. Neben einigen wenigen Veröffentlichungen, die keinem genauen Publikationstyp zugeordnet werden können, finden sich außerdem vereinzelt Plädoyers, Diskursaufsätze sowie ein historischer Aufriss zum Thema Spracherwerb.

Die Klassifizierung zeigt, dass mit 47 von 236 Veröffentlichungen mehr als jeder fünfte Artikel thematisch unter die Kategorie der Lehrbücher fällt oder als Handreichung zur Gestaltung und zum Einsatz didaktischer Materialien zu verorten ist (vgl. Tab. 1). 23 Artikel beschäftigen sich explizit mit Lehrenden (und deren subjektiven Überzeugungen). Demgegenüber ist die Anzahl der Publikationen, die sich mit Lernenden (und deren subjektiven Überzeugungen) befassen, fast doppelt so hoch (41 Veröffentlichungen). Die Mehrheit der Studien (91 Artikel) adressiert Fragen, die sich auf Lehr-Lernprozesse des Spracherwerbs beziehen. In Bezug auf den Einsatz digitaler Medien lässt sich feststellen, dass die Zahl der Veröffentlichungen, die digitale Medien und diesbezügliche Unterstützungsmöglichkeiten für Lehrende untersuchen (8), deutlich geringer ist als die Erforschung der Nutzungsmöglichkeiten für Lernende (14). Bemerkenswert ist auch, dass sich 58 Publikationen mit Unterrichts- und Kursmethoden zum Spracherwerb beschäftigen.

Bezüglich der sprachlichen Inhalte zeigt sich, dass 48 der 111 internationalen Veröffentlichungen gezielt den Erwerb von Englisch als Fremd- oder Zweitsprache untersuchen. Bei den 125 deutschen Veröffentlichungen gibt es 40 Veröffentlichungen, bei denen das Thema Alphabetisierung im Vordergrund steht, 32 adressieren explizit Deutsch als Zweitsprache (DaZ) und Deutsch als Fremdsprache (DaF). Weitere 18 Veröffentlichungen befassen sich mit der Frage, wie Sprachstanderhebungen angemessen durchgeführt bzw. gefördert werden können.

Ein erstes Zwischenfazit ergibt, dass der Spracherwerb Erwachsener national sowie international ein Thema mit Konjunktur ist. Gleichwohl fällt es aufgrund der vielen unterschiedlichen Facetten schwer, konkrete Forschungsschwerpunkte zu identifizieren. International geht es vor allem um Englisch als Zweit- oder Fremdsprache, im deutschsprachigen Raum wird insbesondere die Frage der Alphabetisierung sowie die Evaluation von Sprachlernmethoden beforscht. Dabei sind es neben der traditionellen Erwachsenenbildungsforschung Wissenschaftlerinnen und Wissenschaftler aus der Psychologie und der Linguistik, die den Spracherwerb in organisierten Lehr-

Lernprozessen in methodologischer und thematischer Breite untersuchen. Das Spektrum reicht hierbei von einfachen Erfahrungsberichten über Fallstudien bis hin zu experimentellen Studien. Zur Veranschaulichung werden im Folgenden jeweils eine oder mehrere Studien zu thematisch bestimmten Schwerpunkten exemplarisch dargestellt. Die ausgewählten Beiträge repräsentieren typische Publikationen der jeweiligen Kategorie und versuchen, Thematik und Anlage der Forschungsarbeiten kontrastiv oder beispielhaft zu verdeutlichen.

Lehrbücher und die Gestaltung didaktischer Materialien

Einen ersten Forschungsschwerpunkt bilden Studien zur Gestaltung didaktischer Materialien in Lehrbüchern. Ein Beispiel für diesen Forschungsschwerpunkt stellt die Studie von Targonska (2015) dar, in der Lehrwerke für Deutsch als Fremdsprache für Erwachsene mit den Niveaustufen A2 und B1 (nach Gemeinsamem europäischen Referenzrahmen) untersucht wurden. Im Zentrum der qualitativen Analyse stand die Frage, inwieweit Kollokationen (häufig gebrauchte Wortverbindungen) in der gesprochenen und geschriebenen Sprache thematisiert werden. Mittels einer Analyse von Glossaren und lexikalischen Wortverbindungen wurde beurteilt, ob Kollokationen als ein spezifisches fachdidaktisches Thema Teil der Lehrwerke sind und wie sich Übungen hierzu gestalten. Targonska stellt nach Analyse zweier Lehrbücher fest, dass Übungen zu Kollokationen nicht explizit thematisiert werden und auch in den Glossaren kein Bezug darauf genommen wird. Aufgaben zu Kollokationen sind lediglich in andere Übungen eingeflochten. Abschließend fordert sie daher eine stärkere Thematisierung der Kollokationsproblematik in DaF-Lehrwerken.

Andere Untersuchungen beschäftigen sich mit der Bedeutung der didaktischen Gestaltung von Lehrmaterialien für Lehr-Lernprozesse. Hier lässt sich die Studie von LaBrozzi (2015) hervorheben, in der aufgezeigt wird, auf welcher Ebene die Gestaltung von didaktischen Materialien Lernprozesse beeinflussen kann. Aus pädagogischer und linguistischer Perspektive analysiert er, welchen Einfluss verschiedene Formen der Textanreicherung und Textdarstellung auf die Texterkennung und das Leseverständnis beim Zweitspracherwerb haben. In einem Experimental-Kontrollgruppen-Design wurde die Wirkung von zusätzlichen Textmitteln (Unterstreichungen, Schriftgrößen, Markierungen etc.) an $n=109$ Personen untersucht. Die Studie zeigt, dass sich durch den Einsatz von Textmitteln zwar die Textwahrnehmung, nicht aber das Textverständnis der Experimentalgruppe im Vergleich zur Kontrollgruppe verbesserte.

Neben grundständigen Lehr- und Handbüchern zum Spracherwerb, die aufgrund der Kürze des Artikels nicht aufgeführt werden können, findet sich eine Vielzahl an themenspezifischen Veröffentlichungen, deren Typen im Folgenden kurz skizziert werden soll: Inhaltlich verbinden diese häufig Erkenntnisse der Spracherwerbsforschung mit einer weiteren Disziplin. Beispielsweise zeigt Teuchert (2015) in ihrer Veröffentlichung praxisrelevante Ansätze zur Gesprächs- und Redepädagogik auf. Dabei stehen Informationen und Instrumente zur Beurteilung und Förderung mündlicher Kommunikation im Vordergrund. Zielgruppe des Werks sind Lehrende und Lernende, die sich mit mündlicher Kommunikation und Rhetorik beschäftigen. Das Lehrbuch bietet eine systematische Einführung in kommunikative Kompetenzen, welche in drei Teilbereiche unterteilt werden: Eigenkompetenz (z. B. die Wirkung unterschiedlicher Parameter im Sprechstil), Analyse- bzw. Diagnosekompetenz (z. B. die Einschätzung rhetorischer Leistungen) und drittens Lehrkompetenz (z. B. Vermittlungsprozesse). Heidler (2013) verbindet in ihrer Veröffentlichung den Spracherwerb mit Analysen zur Kapazität des Arbeitsgedächtnisses. Abgeleitet werden daraus Diagnostik- und Therapiemöglichkeiten für Therapeutinnen und Therapeuten, aber auch Pädagoginnen und Pädagogen.

Lehrkräften (und ihre subjektiven Überzeugungen)

Blickt man auf die Veröffentlichungen zur Gruppe der Lehrenden, lassen sich die drei Publikationen von Buss (2016), Larrotta, Moon und Huang (2016) sowie Calvert und Sheen (2015) hervorheben. Die Bedeutung der Lehrenden für den Spracherwerb Erwachsener wird in diesen drei Arbeiten auf unterschiedlichen Abstraktionsgraden untersucht. Während Buss die Frage nach dem Einfluss wissenschaftlicher Erkenntnisse auf die Überzeugungen von Lehrenden erforscht, untersuchen Larrotta et al. den Einfluss bestimmter Methoden und Konzepte auf die Überzeugungen Lehrender. Auf mikrodidaktischer Ebene analysieren Calvert und Sheen, wie sich bestimmte Überzeugungen konstituieren und gegebenenfalls verändern lassen.

Buss (2016) geht der Frage nach, inwieweit die intensive Beforschung des Spracherwerbs und die damit einhergehende wachsende Zahl an Veröffentlichungen (z. B. zu Aussprachetrainings) Überzeugungen und Praktiken von Lehrenden beeinflusst. Ausgehend von einer internationalen Beschreibung des aktuellen Forschungsstands zu Überzeugungen von Lehrenden wurde eine online-gestützte Fragebogenstudie durchgeführt. Dabei wurden 60 Lehrende im Feld Englisch als Zweit- und

Fremdsprache zu ihrem Lehrverhalten bei Aussprachetrainings befragt. Mittels Fragebogen konnten deskriptive Befunde zur Häufigkeit der Anwendung unterschiedlicher Aussprachetrainings sowie qualitativ analysierte Einstellungen der Lehrenden zu den verschiedenen Aussprachetrainings identifiziert werden. Buss stellt fest, dass Erkenntnisse aus der Forschung und aus aktuellen Studien den Lehrenden durchaus bekannt waren, diese sich jedoch weitere Unterstützungsmöglichkeiten zur Umsetzung dieser Befunde in der Praxis (z. B. Trainings) wünschten.

Larrotta et al. (2016) untersuchen die Rezeption von Lehrmethoden von Erwachsenenbildnerinnen und -bildnern. Sie erfassen, welchen Stellenwert Lehrende der Erwachsenenbildung aus dem Bereich Englisch als Fremd- oder Zweitsprache dem Konzept „Learning to learn“ beimessen. Unter Berücksichtigung von lerntheoretischen Überlegungen befragten sie dafür 24 Lehrende mittels eines Online-Fragebogens. Die qualitativ ausgewerteten Ergebnisse liefern nun Erkenntnisse darüber, welche Ansichten Lehrende über das Lernen ihrer Kursteilnehmenden haben, wie Learning-to-learn-Konzepte gestaltet werden und welche möglichen Erträge sich daraus für die Lernenden ergeben. Die Ergebnisse zeigen, dass die Mehrheit der Lehrenden das Learning-to-learn-Konzept im Sprachunterricht kennt und anwendet. Ihnen ist die Wirkung des Konzepts bewusst und sie schätzen es als dazu geeignet ein, Lernstrategien zum Spracherwerb mit den Lernenden zu thematisieren.

Auf der Ebene der Einstellungen und Überzeugungen von Lehrenden gegenüber neuen Kursmethoden untersuchen Calvert und Sheen (2015) mittels des Ansatzes der Aktionsforschung, welchen Einfluss Rückmeldungen zum Lernerfolg auf die Ansichten der Unterrichtsgestaltung von Sprachlehrenden haben. Die Autoren wandten im Selbstversuch die Unterrichtsmethode des Task-Based Learning an, indem sie handlungsorientierte Aufgaben für ihren eigenen Sprachkurs entwickelten, implementierten, evaluierten und modifizierten. Aufgrund dessen konnten sie feststellen, dass die Evaluation und anschließende Modifikation selbstentwickelter Aufgaben zum Spracherwerb zum einen den Lernerfolg bei den Lernenden und zum anderen die Ansicht und Einstellung der Lehrenden gegenüber den eingesetzten Methoden positiv beeinflusste.

Lernende (und ihre subjektiven Überzeugungen)

Im Bereich der Studien, die die Lernenden ins Zentrum der Spracherwerbsforschung stellen, dominieren neben traditionellen pädagogischen Zugängen aus der

Alphabetisierungsforschung (z. B. Doberer-Bey, 2013; Egloff und Grotlüschen, 2011) auch psychologische Forschungsansätze, die die kognitiven Aspekte erfolgreichen Lernens adressieren (z. B. Ouellette-Schramm, 2016).

Die Veröffentlichung von Doberer-Bey (2013) lässt sich als typische Publikation aus der Alphabetisierungsforschung beschreiben. Sie beschäftigt sich mit den Faktoren, die dem Erwerb schriftsprachlicher Kompetenzen entgegenwirken. Aufbauend auf den Ergebnissen qualitativer Interviews wurden systemische und biographische Faktoren identifiziert, die zum (funktionalen) Analphabetismus beitragen können. Nachdem die aktuelle Lebenssituation sowie das soziale und schulische Umfeld als mögliche Ursachen identifiziert wurden, untersucht die Autorin, inwieweit der Besuch von Alphabetisierungs- und Grundbildungskursen das Alltagsleben und die weiteren beruflichen und gesellschaftlichen Entwicklungsmöglichkeiten der Teilnehmenden beeinflusst. Hierbei stellt Doberer-Bey fest, dass die Lernenden durch Alphabetisierungskurse zu weiteren Lerngelegenheiten angeregt werden können, insofern die Lehrkraft es schafft, ein lernförderliches und von Vertrauen getragenes Kursklima zu erzeugen. Die Lernenden nehmen sich durch den Besuch von Alphabetisierungskursen oft als souveräner und selbstsicherer im Umgang mit alltäglichen Herausforderungen wahr.

Ähnliche Schwerpunkte setzt auch die sogenannte Verbleibstudie (Egloff und Grotlüschen, 2011), in der in fünf Teilprojekten die aktuelle Lebenssituation von Kursteilnehmenden erhoben und analysiert wurde. Mit qualitativen und quantitativen Zugängen lieferten die Studienergebnisse Erkenntnisse zur Lebenslage von Kursteilnehmenden und zur Akzeptanz diagnostischer Instrumentarien bei Lernenden und Lehrenden. Mit Blick auf den Umgang Lehrender mit Diagnoseinstrumenten im Alphabetisierungsunterricht konnte u. a. gezeigt werden, dass der Diagnostikbegriff bei Lehrenden der Erwachsenenbildung polarisiert und vielfach mit einer selektiven Funktion verbunden und daher negativ besetzt ist. Darüber hinaus werden in weiteren Teilstudien Zugänge zum Feld der Alphabetisierung beschrieben und Herausforderungen bei der Datenerhebung, z. B. in der Studie Alphapanel (von Rosenblatt und Lehmann, 2013), aufgezeigt.

Das Alphapanel erbrachte u. a. Erkenntnisse zum Zusammenhang zwischen der Dauer des Kursbesuchs und der schriftsprachlichen Kompetenz. Von Rosenblatt und Lehmann (2013) untersuchten in einer Längsschnittstudie mit 299 deutschsprachigen funktionalen Analphabeten, die an Kursen der Volkshochschulen teilnahmen, wie sich die Schriftbeherrschung über die Dauer des Kursbesuchs veränderte und wie sich die Lernenden selbst einschätzten. Die Studie zeigte auch einige kontraindutive

Ergebnisse. So stellen die Autoren z. B. fest, dass der Kursbesuch zwar zur sozialen Stabilisierung beiträgt, der durchschnittliche Grad der Schriftbeherrschung auf Kurs-ebene jedoch nicht zu-, sondern abnimmt. Als Grund hierfür vermuten die Autoren, dass vor allem Teilnehmende mit Lernschwierigkeiten die Kurse länger besuchten (vgl. Rosenblatt & Lehmann, 2013).

Doch nicht nur im Bereich der Alphabetisierungsforschung, sondern auch in der klassischen Spracherwerbsforschung Erwachsener finden sich Studien, die sich mit den Lernenden und deren subjektiven Überzeugungen und Ansichten befassen. Ouellette-Schramm (2016) erforscht z. B., welchen Einfluss sogenannte „developmental perspectives“ der Lernenden auf den Spracherwerb im Bereich Englisch als Zweitsprache haben. Ausgehend von Kegan's Constructive-Development-Framework (Kegan, 2003, cop. 1995) wird erfasst, welche unterschiedlichen Entwicklungsperspektiven die Lernenden in Spracherwerbskursen äußern. Kegan unterscheidet dabei drei Perspektiven; auf instrumentelle, auf sozialisierende und auf selbstbestimmte Entwicklungen. Während Lernende mit erstgenannter Perspektive den Lernprozess mittels einer dichotomen Richtig-Falsch-Kategorie als erfolgreich oder nicht erfolgreich bewerteten, orientierten sich Lernende mit sozialisierender Entwicklungsperspektive bei der Beurteilung des Lernprozesses an den Erwartungen anderer (in der Regel der Lehrenden). Lernende mit selbstbestimmten Entwicklungsperspektiven richteten sich beim Lernen primär an eigenen Vorstellungen aus. Für die Studie wurden neun Kurs-Teilnehmende mit 18 teilstandardisierten Interviews und Kursbeobachtungen untersucht. Es konnten unterschiedliche Begründungszusammenhänge für den jeweiligen Spracherwerb identifiziert werden, die es ermöglichten, verschiedene subjektbezogene Entwicklungsperspektiven nachzuvollziehen. Ouellette-Schramm schlussfolgert, dass sich die verschiedenen „developmental perspectives“ erwartungskonform des der Studie zugrunde liegenden theoretischen Modells von Kegan zuordnen lassen. Personen mit einer eher instrumentellen Entwicklungsperspektive verwiesen bei der Frage nach erfolgreichen Lernprozessen auf Situationen, in denen sie eine Übung lösen konnten. Lernende mit sozialisierender Entwicklungsperspektive beschrieben sich selbst als erfolgreich, wenn sie den Anforderungen des Lehrenden gerecht wurden und Fähigkeiten erworben hatten, um Transferprobleme im Sprachbereich erfolgreich zu lösen. Lernende mit einer selbstbestimmten Entwicklungsperspektive orientierten sich dagegen bei der Bewertung erfolgreicher Lernereignisse an ihren eigenen Erwartungen und bewerteten oft den dabei erlebten Lernprozess.

Rüsseler, Gerth und Boltzmann (2011) untersuchen in ihrer Studie schließlich grundlegende Wahrnehmungsfunktionen bei erwachsenen funktionalen Analphabeten im Vergleich zu Kindern mit Lese-Rechtschreib-Schwäche (LRS). Den Ausgang der Untersuchung bildeten psychologische Annahmen darüber, dass funktionaler Analphabetismus durch Wahrnehmungsdefizite bei den Lernenden bedingt sein könnte. Dabei stützten sich die Autoren unter anderem auf Erkenntnisse aus der psychologischen Forschung zur LRS und zur phonologischen Bewusstheit. In der Studie, an der 60 funktionale Analphabeten, 30 Kinder mit LRS und 60 normal lesende Erwachsene teilnahmen, stellten die Autoren fest, dass letztere über deutlich bessere Wahrnehmungsfunktionen verfügten als funktionale Analphabeten. Zwischen funktionalen Analphabeten und Kindern mit LRS gab es hingegen keine Unterschiede. Die Ergebnisse lassen nach Meinung der Autoren eine Schulung der Wahrnehmungsfähigkeiten in Alphabetisierungskursen sinnvoll erscheinen.

Lehr-Lernprozesse

Betrachtet man empirische Studien, die als Untersuchungsgegenstand die Interaktion zwischen Lehrenden und Lernenden adressieren, findet sich eine breite und methodisch heterogene Zahl an Arbeiten, von denen die Folgenden einen exemplarischen Eindruck geben. Während durch die Literaturrecherche 42 Studien identifiziert werden konnten, die die Lernenden in den Fokus rückten, existieren im Bereich der Forschung zu den Lehrenden nur etwa halb so viele Studien (24 Publikationen). Empirische Studien zu Lehr-Lernprozessen dagegen sind mit 92 Veröffentlichungen deutlich häufiger zu finden. Zwei Drittel (63 Publikationen) dieser Studien wurden international veröffentlicht.

Die Studie von Rassaei (2015) ist ein exemplarisches Beispiel für Forschungsfragen, die in diesem Bereich häufig gestellt werden. Rassaei betrachtet, welche Bedeutung das psychologische Merkmal der Feld(un)abhängigkeit für die Fehlerkorrektur in Sprachkursen hat. Ausgehend von Studien, die belegen, dass Menschen sich hinsichtlich ihres Wahrnehmungsstils in zwei Gruppen aufteilen lassen (Personen mit feldabhängigem, d. h. auf die Wahrnehmungsumgebung fixiertem, und feldunabhängigem, d. h. auf den wahrzunehmenden Gegenstand selbst fixiertem Wahrnehmungsstil), untersucht Rassaei die Frage, ob diese beiden Gruppen von unterschiedlichen Formen der Fehlerkorrektur profitieren. Analysiert wurden dabei die Aussprache- und Schreibfähigkeit der Teilnehmenden in der Zielsprache Englisch. Hierzu konnte in

einem Experimental-Kontrollgruppen-Design mit Pre-, Post- und Follow-up-Testung an $n=76$ Teilnehmenden gezeigt werden, dass Lernende mit feldunabhängigem Wahrnehmungsstil von Fehlerkorrekturen der Lehrenden profitierten. Lernende mit feldabhängigem Wahrnehmungsstil sowie die Kontrollgruppen, die keinerlei Fehlerkorrektur im Sprachkurs erhielten, zeigten keinen signifikanten Leistungszuwachs.

Im Bereich der Blended-Learning-Forschung lässt sich die Studie von Frederiksen und Laursen (2015) hervorheben, da sie als klassische Evaluationsstudie der Lehr-Lern-Forschung einzuordnen ist. Die Autoren sind im Bereich Nordische Studien und Linguistik sowie Wissenschaftssprache tätig. Im Rahmen der Studie wurde ein Sprachkurs für Dänisch am Arbeitsplatz entwickelt, der sowohl Präsenz- als auch Online-Formate enthielt. Mit einem Fragebogen wurden 167 Kursteilnehmende zur gelungenen Integration von Informations- und Kommunikationstechnologien befragt. Anschließend wurden 71 Fragebögen deskriptiv ausgewertet. Die Anreicherung des klassischen Kursformats mit Online-Einheiten beurteilten die Lernenden positiv. Besonders die freie Zeiteinteilung zur Bearbeitung der Online-Einheiten wurde als hilfreich eingeschätzt.

Auch Pulker und Vialleton (2015) beschäftigen sich im Bereich Linguistics and Language mit der Frage, wie sich vorhandene Blended-Learning- und Online-Learning-Konzepte für Spracherwerb in der Erwachsenenbildung adaptieren lassen bzw. welche Anforderungen und Schwierigkeiten bei der Kombination von Spracherwerb in Präsenz- und Online-Phasen auftreten können. Hierzu wurden Interviews mit Kursteilnehmenden aus einem Blended-Learning-Seminar geführt, die den Kurs entweder abgebrochen oder keine Prüfungsleistung abgegeben hatten. Die Interviews wurden anschließend inhaltsanalytisch ausgewertet und theoriegestützt kodiert. Pulker und Vialleton identifizieren Faktoren wie fehlende soziale Interaktion und Unklarheiten in der Bedienung und Nutzung der Online-Lern-Umgebung als mögliche Ursachen für gescheiterte Lehr-Lern-Arrangements.

Nutzung digitaler Medien für Lernende

Die Forschung zu digitalen Medien und Spracherwerb Erwachsener ist noch sehr übersichtlich, hat aber in den letzten Jahren zugenommen. Dies lässt sich am Erscheinungsdatum der einschlägigen Studien erkennen.

Ausgehend vom Interaktionsansatz der Zweitspracherwerbsforschung beschäftigen sich Rouhshad, Wigglesworth und Storch (2016) beispielsweise mit der

Frage, wie digitale Medien bei der Methode der Gesprächsführung wirksam eingesetzt werden können. Mit Verweis auf aktuelle Studienergebnisse zeigen sie, dass die Methode der Gesprächsführung den Zweitspracherwerb maßgeblich begünstigt, da sie Raum für Feedback und Korrekturen durch den Lehrenden ermöglicht. Die Autoren untersuchten, inwiefern es einen Unterschied macht, ob Lernende die Gespräche entweder in klassischen Face-to-Face-Situationen (FTF) oder anhand von textbasierten, computergestützten Medien (SCMC = synchronous computer-mediated communication) durchführten. Die Studie wurde mit 24 Lernenden in Zweiergruppen absolviert und zeigte, dass diejenigen Gruppen, die sich in klassischen FTF-Situationen austauschten, deutlich länger interagierten als die Gruppen, die die Unterhaltung mit der textbasierten, computergestützten Variante durchführten. In FTF-Situationen wurden dadurch mehr Möglichkeiten für Feedback und Korrekturen erzeugt und diese insbesondere für zusätzliche inhaltliche Rückmeldungen genutzt.

Ebenfalls mit einem psycholinguistischen Zugang analysiert Hagiwara (2015), wie Spracherwerb durch digitale Medien unterstützt werden kann. Die zentrale Frage seiner Studie war, ob sich Erkenntnisse aus der Kognitionspsychologie zum Arbeitsgedächtnis (bessere Enkodierung mittels audiovisueller gegenüber unimodaler Stimulationen) auch auf den Spracherwerb übertragen lassen. Hierfür wurden den 32 Teilnehmenden mündlich Aufgaben zur Morphem-Verarbeitung jeweils mit und ohne piktoraler Unterstützung dargeboten. Der Autor stellt fest, dass eine Visualisierung des Lerngegenstands die Leistung der Teilnehmenden verbesserte.

Rafieyan et al. (2015) untersuchen, ob der Kontakt von Sprachlernenden mit Muttersprachlerinnen und Muttersprachlern mittels digitaler Medien den Spracherwerbsprozess unterstützen kann. Alle 60 teilnehmenden Erwachsenen der dargestellten Studie, die zu diesem Zeitpunkt einen Englischkurs besuchten, sollten verschiedene Schreibaufgaben erledigen. Teilnehmende der Experimentalgruppe konnten über einen Smartphone-Kommunikationsdienst (WhatsApp) mit Muttersprachlerinnen und Muttersprachlern in Kontakt treten, während die Personen aus der Kontrollgruppe keine Unterstützung bekamen. Die Autoren beobachteten, dass der über digitale Medien moderierte Austausch mit Muttersprachlerinnen und Muttersprachlern zu signifikant besseren Schreibleistungen führte.

Die Studie von Romaña Correa (2015) widmet sich der Frage, ob sich Skype-Konferenzen dazu eignen, die mündliche Kommunikation der Zielsprache zu trainieren. Methodisch wurden 12 Lernende eines Spracheninstituts in Fokusgruppen und mittels Fragebogen zu ihren Kleingruppenarbeiten in Skype-Konferenzen befragt. Der

Autor kommt zu dem Ergebnis, dass ein computergestützter Medieneinsatz (wie z. B. via Skype) für Sprachkurse geeignet ist. Besonders sinnvoll erscheint er für die Verbesserung der mündlichen Kommunikation hinsichtlich der Sprechflüssigkeit und für den Aufbau von Kenntnissen über Kursinhalte.

Eine weitere Veröffentlichung, die sich mit den sprachlichen Fähigkeiten der Kursteilnehmenden beschäftigt, ist der Artikel von Stollhans (2015) über eine Fallstudie zu einer Online-Plattform, die anstelle von Schriftsprache auf Audioaufnahmen der Nutzenden basiert. Dadurch sollen die Lernenden ihre sprachlichen Fähigkeiten der jeweiligen Zielsprache trainieren. Die von Linguisten entwickelte Plattform hat den Anspruch, selbsterklärend und benutzerfreundlich bedienbar zu sein. Das Ziel des Projekts bestand darin, Lernenden eine Möglichkeit zu geben, unabhängig und selbstständig die eigenen Aussprachefähigkeiten zu trainieren und sich dabei gegenseitig zu unterstützen. Die Plattform wurde mit 48 Studierenden im Rahmen eines Auslandsaufenthalts in Deutschland getestet. Im Anschluss wurden diese mit einem Fragebogen zur Akzeptanz und Nutzung befragt. Die deskriptiven Befunde zeigen, dass die Mehrheit der Befragten die Plattform gerne nutzte und die Lernenden zur Reflektion der eigenen sprachlichen Fähigkeiten angeregt wurden.

González Romero (2016) untersucht in ihrer Publikation, welchen Einfluss Probeklausuren bzw. Übungen in einem Online-Sprachkurs auf die Entwicklung der Sprachkompetenz haben. Zwanzig Spanisch sprechende Erwachsene, die einen Business-Englisch-Kurs besuchten, erhielten nach jeder Kursstunde Übungsaufgaben als freiwillige Hausaufgabe. Personen, die diese bearbeiteten, schnitten in der Abschlussprüfung besser ab als alle anderen. Ob dies auf die Bearbeitung der Übungen zurückgeführt werden kann oder z. B. Merkmale der Personen wie ihre Vorkenntnisse ausschlaggebend waren, konnte aufgrund der Anlage der Studie nicht geklärt werden.

Die Recherche förderte neben den empirischen Studien auch eine Vielzahl von Berichten aus der und für die Praxis zutage. Dazu gehören Kurzberichte über den Einsatz von Sprachlernprogrammen. Auerbach (2013) beschreibt beispielsweise den Einsatz einer Lernsoftware zum Spracherwerb für Austauschschülerinnen und -schüler sowie junge Erwachsene und berichtet über ihren Mehrwert. Feick (2015) geht der Frage nach, wie digitale Medien für Lernende in der Schule und der Erwachsenenbildung eingesetzt werden können, um einen sprachlernbezogenen Mehrwert zu erzeugen. Der Beitrag umfasst dabei verschiedene Best-Practice-Beispiele, die Einsatzmöglichkeiten digitaler Medien exemplarisch illustrieren. Reisenleutner (2015) beschreibt in ihrem Artikel verschiedene digitale Werkzeuge, die sich für den Unterricht

in DaF mit Erwachsenen eignen. Neben einer Zusammenfassung verschiedener Tools verortet sie diese auch hinsichtlich ihrer Anwendungsmöglichkeiten und Funktionen für den Deutschunterricht. Und auch Gillies (2012) berichtet über Methoden des digitalen Lernens sowie über Sprachlernprogramme für Lernende.

Nutzung digitaler Medien für Lehrende

Die Unterstützung Lehrender beim Spracherwerb durch digitale Medien spielt in Veröffentlichungen der Erwachsenen- und Weiterbildung derzeit eine eher untergeordnete Rolle. Arbeiten zur Nutzung digitaler Medien durch Lernende überwiegen deutlich. Wenn der Gebrauch digitaler Medien durch Lehrende thematisiert wird, geht es häufig um die Frage, wie der Kursraum durch digitale Medien angereichert werden kann. So begleitete Haines (2015) beispielsweise zwei Lehrende über einen Zeitraum von 14 Monaten, um sie qualitativ zu wahrgenommenen Verwendungsmöglichkeiten digitaler Tools wie Wikis oder Blogs zu befragen. Haines, die selbst in der Lehrerbildung aktiv ist, fragte in ihren beiden Fallstudien darüber hinaus, inwieweit sich die Verwendungsmöglichkeiten für eine erfolgreiche Anwendung digitaler Tools über die Zeit verändern. Das zentrale Ergebnis dieser Untersuchung war, dass die subjektiven Einschätzungen stark von der einzelnen Lehrkraft und ihrer jeweiligen pädagogischen Erfahrung abhingen. Eine Veröffentlichung von Böttcher (2015) beschäftigt sich ebenfalls mit der Frage, wie Lehrende digitale Medien im Spracherwerbsprozess unterstützend einsetzen können. Hierbei wird aus einem konzeptionellen Interesse heraus skizziert, wie Lehrende digitale Medien nutzen können, um individuelle Lernprozesse zu unterstützen. Der Artikel mündet in einem kurzen Überblick über vorhandene digitale Werkzeuge für den Sprachenunterricht.

Angesichts der geringen Zahl von Arbeiten zu digitalen Medien, die Lehrende bei ihren Aufgaben unterstützen, sollen im Folgenden zwei aktuell begonnene Projekte vorgestellt werden. Zum einen hat das Bundesministerium für Bildung und Forschung (BMBF) das Projekt „Kompetenz-adaptive, nutzerorientierte Suchmaschine für authentische Sprachlernertexte (KANSAS)“ bewilligt, das auf einer interdisziplinären Kooperation von Erwachsenenbildung, Computerlinguistik, Kognitionspsychologie und Sprachdidaktik beruht (Schrader, Schröter, Meurers und Becker-Mrotzek, 2017). Ziel dieses Projekts ist die Entwicklung und Evaluation eines digitalen Instruments zur Unterstützung von Lehrkräften bei der Auswahl geeigneter, lebensweltnaher Sprachlernmaterialien für Kurse der Alphabetisierung und Grundbildung sowie DaZ. Das

Instrument soll eine internetbasierte Inhaltsuche unter Berücksichtigung sprachlicher Eigenschaften der Webseitentexte (z. B. globales Anforderungsniveau im Sinne des Gemeinsamen europäischen Referenzrahmens, spezifische sprachliche Konstruktionen) ermöglichen. Zudem wird derzeit am Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache ein an der Universität Tübingen entwickeltes digitales Tool adaptiert (Schrader und Schöb, 2016), welches zukünftig Lehrkräfte in der Kurs- und Unterrichtsvorbereitung bei der Identifikation sprachsensibler Unterrichtsmethoden unterstützen soll.

4. Fazit und Perspektiven

Betrachtet man die vorgestellten Studien, so wird deutlich, dass der aktuelle Forschungsstand sich sehr differenziert, oder treffender: zersplittert darstellt. Die thematische Breite variiert von sehr speziellen Detailfragen zu psychologischen Konstrukten für erfolgreichen Spracherwerb über didaktische und personale Einflussfaktoren bis hin zu Studien über die institutionellen Rahmenbedingungen erfolgreichen Spracherwerbs ganzer gesellschaftlicher Gruppen, wie sie z. B. in Large-Scale-Assessments wie PIAAC (Ramstedt, 2013) behandelt werden. Forschende aus dem Bereich der Linguistik und der Psychologie setzen sich vor allem mit den Inhalten und der Qualität von didaktischen Materialien sowie mit den kognitiven Prozessen beim Spracherwerb auseinander. Dabei ergeben sich überwiegend im DaF-Bereich hilfreiche, aber oft auch sehr kleinteilige Hinweise, so z. B. auf die Gestaltung und Nutzung didaktischer Materialien oder auf das Training kognitiver Kompetenzen. In der Erwachsenenbildung werden Fragen des professionellen Handelns der Lehrkräfte häufiger untersucht. Dazu gehören Studien, die sich mit der Rezeption didaktischer und wissenschaftlicher Konzepte durch Lehrende beschäftigen, oder untersuchen, wie sich Einstellungen zu Unterrichtsmethoden verändern. Diese Erkenntnisse helfen dabei, relevante professionelle Faktoren für die Förderung des Spracherwerbs zu identifizieren. Über die Bedeutung der unterschiedlichen Einflussfaktoren auf erfolgreichen Spracherwerb lässt die bisherige Forschung jedoch oft noch keine Schlüsse zu, da es an generalisierenden, theorieprüfenden Studien fehlt. Auf Seite der Lernenden wurden, besonders im DaZ-Bereich und in der Alphabetisierungsforschung, Einflussfaktoren und Ursachen für (nicht stattfindende) Spracherwerbsprozesse vornehmlich explorativ erforscht. Hier konnten in den letzten Jahren verstärkt subjektive Begründungszusammenhänge rekonstruiert werden. Nach Sichtung des Forschungsstandes wird auch deutlich, dass Lehr-Lernprozesse im Bereich Blended-Learning und Online-Learning noch wenig theoriegestützt untersucht werden. Hier haben Studien derzeit oftmals einen eher explorativen Charakter.

Im Feld der digitalen Unterstützungsmöglichkeiten für Lehrende und Lernende stehen Fragen danach im Vordergrund, wie interaktive Materialien beim Spracherwerb erfolgreich eingesetzt werden können. Zudem werden bereits bewährte Techniken auf ihre Wirksamkeit untersucht und auch neue, in anderen Kontexten entstandene digitale Unterstützungsmöglichkeiten erprobt. Auch macht die große Zahl der Erfahrungsberichte von Praktikerinnen und Praktikern und für Praktikerinnen und Praktiker

deutlich, dass ein reges Interesse der Lehrenden an digitalen Medien zur Unterstützung von Lernenden besteht. Digitale Medien zur Unterstützung von Lehrkräften werden derzeit vergleichsweise selten von der Forschung adressiert.

Abschließend kann für das vergangene Jahrzehnt bilanziert werden, dass eine durchaus breit angelegte (fach-)didaktische, linguistische und empirische Forschung zum Spracherwerb Erwachsener existiert. Es scheint zudem, dass die Forschungsdensität mindestens für einige Themen wie etwa die Nutzung digitaler Medien zunimmt. Gleichzeitig beobachten wir jedoch ein hohes Maß an disziplinärer Vielfalt, die die wechselseitige Rezeption der Forschungsergebnisse erschwert. Es überwiegen zudem Einzelstudien mit oft kleinen Stichproben und deskriptiven Untersuchungsdesigns. Länger angelegte, durch Drittmittelgeber abgesicherte Forschungsprogramme, die auch (feld-)experimentelle oder gar längsschnittliche Studien ermöglichen würden, lassen sich nicht absehen. Auch finden sich noch kaum Studien, die die Bedingungen der Möglichkeit innerer Differenzierung in oft sehr heterogenen Lerngruppen behandeln. Die disziplinäre Vielfalt der Zugänge sowie der singuläre Charakter vieler Studien stellen hohe Hürden für eine handlungsrelevante Rezeption von Forschungsbefunden durch Akteure der Praxis dar. Insbesondere fehlt es noch an belastbaren Befunden zu den Bedingungen erfolgreicher Lehr-Lernprozesse für den Spracherwerb Erwachsener im Zusammenspiel von didaktischen Konzepten und Materialien, professionellen Kompetenzen der Lehrenden und kognitiven Voraussetzungen und Aktivitäten der Teilnehmenden. Angesichts der Relevanz der eingangs beschriebenen praktischen und bildungspolitischen Herausforderungen kann dies nicht befriedigen.

Um die Herausforderungen für die Wissenschaft genauer definieren zu können, bedarf es u. a. systematischer Reviews, die noch differenzierter und umfassender angelegt sind als der hier vorgelegte, auf jüngere Arbeiten beschränkte Überblicksbeitrag. Diese sollten zukünftig auch die schulbezogene Forschung einbeziehen, da die wechselseitige Wahrnehmung disziplinär vielfältiger Forschung zu verschiedenen Bildungsbereichen theoretisch wie empirisch anregend sein kann.

5. Literaturverzeichnis

- Auerbach, S. (2013). Erfolgreiche Fremdsprachen-Methode: Rotary International stellt erfolgreiches Lernsystem für Austauschschüler vor. *LA-Multimedia*, 10(4), 27.
- BMBF/KMK. (2016). *Grundsatzpapier zur Nationalen Dekade für Alphabetisierung und Grundbildung 2016–2026: Den funktionalen Analphabetismus in Deutschland verringern und das Grundbildungsniveau erhöhen*. https://www.alphadekade.de/files/01_Grundsatzpapier%20zur%20Nationalen%20Dekade%20Alphabetisierung%20und%20Grundbildung_final.pdf. Zugegriffen: 14.08.2017.
- Böttcher, R. (2015). Lernen mit digitalen Medien in informellen Situationen und die Verbindung zu formal organisierten Lernprozessen. *Fremdsprache Deutsch*, (53), 9–13.
- Buss, L. (2016). Beliefs and practices of Brazilian EFL teachers regarding pronunciation. *Language Teaching Research*, 20(5), 619–637.
doi:10.1177/1362168815574145
- Calvert, M. & Sheen, Y. (2015). Task-based language learning and teaching: An action-research study. *Language Teaching Research*, 19(2), 226–244.
doi:10.1177/1362168814547037
- Deeke, A. (2011). Berufsbezogene Sprachförderung und berufliche Weiterbildung von Arbeitslosen mit Migrationshintergrund: Eine Verbleibs- und Wirkungsanalyse. In M. Granato, D. Münk, & R. Weiß (Hrsg.), *Berichte zur beruflichen Bildung. Migration als Chance*, (S. 91–112). Bielefeld: Bertelsmann.
- Costa, P., Rodrigues, M., Vera-Toscano, E. & Weber, A. (2014). *Education, adult skills and social outcomes: Empirical evidence from the survey on adult skills (PIAAC 2013)*. *JRC scientific and policy reports: Vol. 26626*. Luxembourg: EUR-OP.
- Doberer-Bey, A. (2013). Alphabetisierung mit Erwachsenen: Von Fehlentwicklungen beim schulischen Erwerb von Schriftsprachlichkeit und Lernerfolg im Erwachsenenalter. In A. Doberer-Bey & A. Hrubesch (Hrsg.), *leben = lesen? Alphabetisierung und Basisbildung in der mehrsprachigen Gesellschaft* (Schulheft 149, S. 16–32). Innsbruck: Studienverlag.

- DVV – Deutscher Volkshochschulverband. (2015, 03. November). Volkshochschulen warnen von Engpass bei Integrationskursen für Flüchtlinge. Bonn.
<https://www.dvv-vhs.de/presse/details/news/News/detail/volkshochschulen-warnen-vor-engpass-bei-integrationskursen-fuer-fluechtlinge/>. Zugegriffen: 14.08.2017.
- Egloff, B. & Grotlüschen, A. (2011). *Forschen im Feld der Alphabetisierung und Grundbildung: Ein Werkstattbuch* (Alphabetisierung und Grundbildung, Band 7). Münster u. a.: Waxmann.
- Feick, D. (2015). Mehr als nur Apps. Mobiles Lernen im DaF-Unterricht. *Fremdsprache Deutsch*, (53), 14–18.
- Feick, D. & Schramm, K. (2016). Alphabetisierung mit Migrantinnen und Migranten. In C. Löffler & J. Korfkamp (Hrsg.), *Handbuch zur Alphabetisierung und Grundbildung Erwachsener* (S. 214–225). Münster u. a.: Waxmann.
- Feldmeier, A. (2015). Encouraging Learner Autonomy: Working with Portfolios, Learning Agreements and Individualized Materials. In I. van de Craats, J. Kurvers, & R. van Hout (Hrsg.), *LESLLA: Vol. 10. Adult literacy, second language and cognition* (S. 89–114). Nijmegen: Centre for Language Studies (CLS).
- Frederiksen, K.-M. & Laursen, K.Å. (2015). Developing and evaluating a multimodal course format: Danish for knowledge workers – labour market-related Danish. In F. Helm, L. Bradley, M. Guarda & S. Thouësny (Hrsg.), *Critical CALL. Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (S. 186–191). doi:10.14705/rpnet.2015.000331
- Gillies, C. (2012). Schlau auf jedem Kanal: Fremdsprachen lernen. *Manager-Seminar*, (170), 58–63.
- Gogolin, I. (2013). Mehrsprachigkeit und bildungssprachliche Fähigkeiten. In I. Gogolin, I. Lange, U. Michel & H.H. Reich (Hrsg.), *Herausforderung Bildungssprache. Und wie man sie meistert* (FörMig-Edition, Band 9, S. 7–19). Münster u. a.: Waxmann.
- González Romero, R. (2016). The implications of business English mock exams on language progress at higher education. In A. Pareja-Lora, C. Calle-Martínez & P. Rodríguez-Arancón (Hrsg.), *New perspectives on teaching and working with languages in the digital era* (S. 293–302). doi:10.14705/rpnet.2016.tislid2014.442

- Grosche, M. & Schroeder, J. (2013). Stand und Perspektiven der empirischen Alphabetisierungsforschung: Diskussion der Beiträge des Themenhefts. *Empirische Sonderpädagogik*, (3), 250–262.
- Hagiwara, A. (2015). Effect of visual support on the processing of multiclausal sentences. *Language Teaching Research*, 19(4), 455–472.
doi:10.1177/1362168814541715
- Haines, K.J. (2015). Learning to Identify and Actualize Affordances in a New Tool. *Language Learning & Technology*, 19(1), 165–180.
- Heidler, M.-D. (2013). *Das Arbeitsgedächtnis: Ein Überblick für Sprachtherapeuten, Linguisten und Pädagogen*. Bad Honnef: Hippocampus-Verlag.
- Huettig, F. (2015). Literacy influences cognitive abilities far beyond the mastery of written language. In I. van de Craats, J. Kurvers & R. van Hout (Hrsg.), *LESLLA: Vol. 10. Adult literacy, second language and cognition* (S. 115–128). Nijmegen: Centre for Language Studies (CLS).
- Isphording, I.E. & Sinning, M. (2012). The Returns to Language Skills in the US Labor Market. *IZA Discussion Paper, No. 7080*. doi:10.2139/ssrn.2200833
- Kegan, R. (2003, cop. 1995). *In over our heads: The mental demands of modern life*. Cambridge, Mass.: Harvard University Press.
- Kurvers, J., van de Craats, I. & van Hout, R. (2015). Footprints for the future: Cognition, literacy and second language learning by adults. In I. van de Craats, J. Kurvers & R. van Hout (Hrsg.), *LESLLA: Vol. 10. Adult literacy, second language and cognition* (S. 7–32). Nijmegen: Centre for Language Studies (CLS).
- LaBrozzi, R.M. (2015). The effects of textual enhancement type on L2 form recognition and reading comprehension in Spanish. *Language Teaching Research*, 20(1), 75–91. doi: 10.1177/1362168814561903
- Larrotta, C., Moon, J.Y.C. & Huang, J. (2016). Learning a New Language is Like Swiss Cheese. *Adult Learning*, 27(4), 168–174.
doi:10.1177/1045159516643946
- Löffler, C. & Korfkamp, J. (Hrsg.). (2016). *Handbuch zur Alphabetisierung und Grundbildung Erwachsener*. Münster u.a.: Waxmann.
- Nolda, S. (2017). *Fremdsprachenlernen Erwachsener*. Bielefeld: Bertelsmann.

- Nolda, S. (1989). *Sprachwissenschaft als Bezugswissenschaft der Erwachsenenbildung. Theorie und Praxis der Erwachsenenbildung*. Bad Heilbrunn: Klinkhardt.
- Ouellette-Schramm, J.R. (2016). Developmental Diversity in the Academic Language–Learning Experiences of Adult English as a Second or Other Language Learners. *Adult Education Quarterly*, 66(3), 219–236.
doi:10.1177/0741713616640882
- Pulker, H. & Vialleton, E. (2015). Learners' perceptions of online elements in a beginners' language blended course – implications for CALL design. In F. Helm, L. Bradley, M. Guarda, & S. Thouëсны (Hrsg.), *Critical CALL. Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (S. 475–479).
doi:10.14705/rpnet.2015.000378
- Rafieyan, V., Rafieyan, A., Rafieyan, N., Rafieyan, S., Rafieyan, P. & Rafieyan, M. (2015). Effect of Developing Pragmatic Competence through Telecollaboration on Improving English as Foreign Language Learners' Writing Proficiency. *Journal of Education and Practice*, 27(6), 121–131.
- Rammstedt, B. (2013). *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich: Ergebnisse von PIAAC 2012*. Münster: Waxmann.
- Rassaei, E. (2015). Recasts, field dependence/independence cognitive style, and L2 development. *Language Teaching Research*, 19(4), 499–518.
doi:10.1177/1362168814541713
- Reisenleutner, S. (2015). Hybride Lernszenarien. *Fremdsprache Deutsch*, (53), 40–45.
- Romaña Correa, Y. (2015). Skype™ Conference Calls: A Way to Promote Speaking Skills in the Teaching and Learning of English. *PROFILE Issues in Teachers' Professional Development*, 17(1), 143–156. doi:10.15446/profile.v17n1.41856
- Rosenblatt, B. v. & Lehmann, R.H. (2013). Grade der Schriftbeherrschung und subjektiver Lernerfolg bei Teilnehmenden an Alphabetisierungskursen. *Zeitschrift für Erziehungswissenschaft*, 16(1), 55–77. doi:10.1007/s11618-013-0342-z
- Rouhshad, A., Wigglesworth, G. & Storch, N. (2016). The nature of negotiations in face-to-face versus computer-mediated communication in pair interactions. *Language Teaching Research*, 20(4), 514–534.
doi:10.1177/1362168815584455

- Rüsseler, J., Gerth, I. & Boltzmann, M. (2011). Basale Wahrnehmungsfähigkeiten von erwachsenen funktionalen Analphabeten und Analphabetinnen. In Projektträger im Deutschen Zentrum für Luft- und Raumfahrt e.V. (Hrsg.), *Lernprozesse in Alphabetisierung und Grundbildung Erwachsener. Diagnostik, Vermittlung, Professionalisierung* (Alphabetisierung und Grundbildung Erwachsener, Band 2, S. 11–27). Bielefeld: Bertelsmann.
- Schrader, J. (2017, in Druck). Institutionelle Rahmenbedingungen, Anbieter, Angebote und Lehr-Lernprozesse der Erwachsenen- und Weiterbildung. In O. Köller, M. Hasselhorn, F.W. Hesse, K. Maaz, J. Schrader, H. Solga & C.K. Spieß (Hrsg.), *Das Bildungswesen in Deutschland. Bestand und Potenziale*. Bad Heilbrunn: Klinkhardt.
- Schrader, J. & Schöb, S. (2016). Die Planung von Lehr-Lern-Einheiten mit digitalen Medien: Konzepte und Befunde. *Zeitschrift für Weiterbildungsforschung - Report*, 39(3), 331–347. doi:10.1007/s40955-016-0078-5
- Schrader, J., Schröter, H., Meurers, D. & Becker-Mrotzek, M. (2017). *Kompetenzadaptive, nutzerorientierte Suchmaschine für authentische Sprachlernertexte*. Bonn: Deutsches Institut für Erwachsenenbildung.
- Stollhans, S. (2015). The e-learning tool Voxopop and its benefits on oral skills: Activities for final year students of German. In K. Borthwick, E. Corradini & A. Dickens (Hrsg.), *10 years of the LLAS elearning symposium. Case studies in good practice* (S. 185–192). Dublin: Research-publishing.net. doi:10.14705/rpnet.2015.000278
- Targonska, J. (2015). Kollokationen in DaF-Lehrwerken - eine empirische Studie zur Erfassung und Widerspiegelung der Kollokativität der Sprache in ausgewählten Lehrwerken. *Zielsprache Deutsch*, 42(3), 3–25.
- Teuchert, B. (Hrsg.). (2015). *Mündliche Kommunikation lehren und lernen: Facetten der Rhetorik in Schule und Beruf* (Sprache und Sprechen, Band 47, 1. Aufl.). Baltmannsweiler: Schneider Hohengehren.
- Tietgens, H. (1986). *Erwachsenenbildung als Suchbewegung: Annäherungen an eine Wissenschaft von der Erwachsenenbildung. Theorie und Praxis der Erwachsenenbildung*. Bad Heilbrunn: Klinkhardt.

van de Craats, I., Kurvers, J. & van Hout, R. (Hrsg.). (2015). *Adult literacy, second language and cognition. LESLLA: Vol. 10*. Nijmegen: Centre for Language Studies (CLS).

Young-Scholten, M., Peyton, J.K., Sosinski, M. & Cabeza, A.M. (2015). LESLLA teachers' views of the knowledge and skills they need: An international study. In I. van de Craats, J. Kurvers & R. van Hout (Hrsg.), *LESLLA: Vol. 10. Adult literacy, second language and cognition* (S. 165–186). Nijmegen: Centre for Language Studies (CLS).

7. Beitrag #2 Investigating spoken classroom interactions

Sahlender*, M., ten Hagen*, I., Knaus, A., Helbert, S., Weiss Z. (eingereicht). Investigating spoken classroom interactions in linguistically heterogeneous learning groups – An interdisciplinary approach to process video-based data in second language acquisition classrooms. *Manuskript eingereicht zur Publikation.*

Die in dieser Dissertationsschrift abgebildete Version wird voraussichtlich nicht mit der publizierten Version übereinstimmen.

*Der Beitrag ist in geteilter Erstautorenschaft erstellt worden.

Abstract

Speaking the local language (e.g., German in Germany) is central for successful integration into society. The teacher's language, for example in German as a Second Language (GSL) classrooms, serves as a crucial tool in language learning. Heterogeneity of learners language proficiency levels challenges teachers to adapt their language and accompanied instructional behavior. We offer an approach to study language acquisition processes and how teachers adapt their instructional language. We present our language-independent guidelines for processing video-based data of classroom interactions and demonstrate their reliability. These guidelines enable transcriptions of spoken language in noisy environments and detailed annotations of non-verbal classroom behavior. We outline research avenues at the intersection of empirical education research and linguistics that become feasible through these resources focusing on studying (non-)verbal adaptation strategies of teachers for learners at different proficiency levels. Our work directly fosters the interdisciplinary study of teacher-student interactions, teacher competencies, and language acquisition.

Highlights

- Interdisciplinary approach to compiling a spoken learner corpus
- Publishing guidelines for processing video-based classroom data
- Reporting innovative evaluation of the robustness of annotations and transcriptions
- Presenting a L2 spoken learner corpus with rich meta-data on teachers and learners

Keywords

Second Language Acquisition, Spoken Learner Corpora, Spoken Classroom Interactions, Non-Verbal Classroom Interactions, Heterogeneous Classrooms, Guidelines for Transcription and Annotation,

1. Relevance of Language Teaching

It is important for non-native speaking residents to acquire the official national language of the country they live in because it promotes their successful integration, both in terms of their identification with society and their social participation (Esser, 2006; Hochman & Davidov, 2014; Martinovic et al., 2009). Moreover, language proficiency is linked to educational (for an overview, see Esser, 2006) and labor market success (e.g., Bleakley & Chin, 2004; Guven & Islam, 2015). To promote successful Second Language Acquisition (SLA), explicit language instruction and comprehensible language input are needed (Ellis, 2005; Krashen, 1985; Loewen & Sato, 2018; Long, 1983) because mere exposure to a language is insufficient to foster second language (L2) learning. Language learners have been shown to benefit most from language input in their Zone of Proximal Development (Vygotsky & Cole, 1978) which challenges them by being slightly above their current level of proficiency while remaining still at an attainable level. Such adaptations to learners' individual linguistic skills do not typically transpire in naturally occurring conversations, especially for learners with low language proficiency. Thus, the classroom is one of the most important places for SLA, not only as a source of explicit instruction but also as a main source of comprehensible spoken language input (cf. Ellis, 2005; Hall, 2022; Hall et al., 2020).

The spoken classroom interactions that teachers facilitate are a crucial tool in making instruction beneficial to language learning (Howe & Abedin, 2013; Johnston et al., 2016; Thoms, 2012), especially because established didactic techniques such as scaffolding (Gibbons, 2002) allow teachers to systematically adapt input to learners' individual competence levels. However, successful input adaptation can be very challenging for teachers in classrooms consisting of learners with different skill levels. In practice, SLA classrooms are often comprised of learners with diverse backgrounds and competence levels, especially outside of university contexts. For example, in Germany German as a Second Language (GSL) is predominantly being taught in schools and institutions of adult education (Otto et al., 2016; Tissot et al., 2019). GSL classes are highly heterogeneous in terms of the cultural background and language biographies of learners and appropriately accommodating these differences between learners has proven to be a particular challenge for school teachers and adult educators (Otto et al., 2016; Tissot et al., 2019). However, it remains unclear to what extent successful language input adaptation actually takes place in heterogeneous SLA classrooms. This is partially due to a lack of data on teachers' language use in authentic, heterogeneous SLA classrooms (Hall et al., 2020; Sahlender & Schrader, 2017).

This type of data could offer research possibilities for various disciplines. For instance, data on spoken (and accompanied nonverbal) classroom interaction is a key resource to gain better insights on language teaching processes (cf. Hall et al., 2020; Lim, 2019; Mercer & Dawes, 2014). In addition, from a linguistic perspective, authentic language use data is crucial in SLA research (Ellis & Bogart, 2007; Gilquin, 2015; Tarone, 1998) as it allows to study the interlanguage system, L2 developmental trajectories and the linguistic properties of scaffolded language in teacher-learner interactions.

Thus, the aim of the present study is to present an approach for processing data on spoken language of teachers and learners in L2 classrooms to make such data usable for multiple research disciplines. The present study is part of the COLD project (*Competencies of School Teachers and Adult Educators in Teaching German as a Second Language in Linguistically Diverse Classrooms*) which aims to investigate GSL teachers' professional competence on various dimensions. For this purpose, video data of GSL lessons in preparatory classes in schools and adult integration courses was collected. We give a detailed overview on how these complex observed classroom data can be transferred into a machine-readable format in form of a linguistic corpus while taking into account aspects that are relevant for interdisciplinary research. Importantly, to ensure the transparency and re-usability, we report the used guidelines and corresponding empirical quality criteria for transcription and annotation. We exemplify the application of our procedure using authentic GSL data. Moreover, we give an outlook on research opportunities rooted in the interdisciplinary approach of the study.

2. Language Teaching in Heterogeneous Classrooms

In language teaching, language is both content and medium of the instruction (Borg, 2006; König et al., 2016). Especially in this educational context, teachers' linguistic and communication skills and the interactions with their learners are of great relevance. They facilitate learners' oral productions and engagement in meaningful communication, which is a fundamental part of language learning (Thoms, 2012). Therefore, the spoken language of both teacher and learners in the classroom plays a central role for language acquisition (Borg, 2006; König et al., 2016).

GSL learning groups consist of learners from a variety of social and linguistic backgrounds with differing cognitive, motivational, and self-regulatory resources which poses pedagogic challenges for teachers (Otto et al., 2016; Tissot et al., 2019). Especially in these heterogeneous groups of learners, teachers' consideration of their

individual learners' prerequisites and resulting instructional adaptations within the group is important for a successful promotion of all learners (Hardy et al., 2019; Smale-Jacobse et al., 2019; Tomlinson et al., 2003). Adaptive teaching, which is generally defined as teachers' adjustment of their instruction to the individual learners' abilities, motivations, and backgrounds (Corno, 2008; Hardy et al., 2019), is an important characteristic of teaching quality that is associated with effective teaching because it has been shown to positively affect learners' outcomes (for a review, see Parsons et al., 2018). Thus, in order to create high-quality classroom interactions in GSL classrooms that promote learners' language skills, teachers need to take into account individual differences in language proficiency within a class and adapt their instructional language to learners' language skills so that learners are enabled to follow the teachers' instruction. However, the extent to which teachers adapt their instruction in SLA has not been studied in any detail.

To date, we lack research on whether teachers' instructional language varies depending on the individual student addressed, especially in authentic language classroom settings. With teachers' instructional language we intend all utterances that teachers direct at learners during language classes which are a special type of instructed setting. Evidence from video studies in mathematics lessons suggest that teachers might address differences in their learners' language proficiencies by adapting their instructional language (e.g., Nathan & Kim, 2009; Prediger & Buró, 2021). For instance, teachers vary the level of cognitive complexity of their elicitations depending on the targeted learner's prior utterance or responses (Nathan & Kim, 2009) or reduce unnecessary language demands to provide appropriate language-learning opportunities (Prediger & Buró, 2021). A case study in L2 language instruction indicates that teachers adapt their used instructional language (turn design, utterance formats) over the school years to meet learners' changing linguistic needs and achieve pedagogical goals such as language production practice (L2 French; Konzett-Firth, 2020). Yet, it remains an open question whether, when, and how different teachers may adapt the complexity of their language within GSL classes, for instance, as a function of the language proficiencies of their heterogeneous learners.

To investigate the extent to which teachers use their language in the classroom adaptively, it is necessary to gather data on teachers' verbal interaction with their individual learners. Furthermore, because the learners' linguistic prerequisites determine the language adaptation needed, it is important to assess not only teachers' instructional language towards individual learners, but also individual learners' levels of

language proficiencies and corresponding teachers' perception of each individual learner's language skills (cf. Brophy & Good, 1970; Denessen et al., 2020). Thus, this type of investigation requires a multi-perspective approach that takes into account teachers' and learners' spoken language in the authentic classroom as well as meta-information about abilities and backgrounds. Furthermore, it is crucial to consider the communicative context of language usage in the L2 classroom because it shapes the way people interact verbally (Thoms, 2012). Contexts vary for instance regarding the educational environment (e.g., number of learners in the classroom), the structure of the lesson (e.g., social form), and non-verbal behavior that accompanies verbal utterances (e.g., teachers' gestures or learners' hand-raising). A suitable approach to gather authentic data on naturally occurring language and teacher-learner interaction patterns in L2 classrooms are video observations because they offer the possibility to analyze the verbal interactions in the classroom along with associated nonverbal interactions and additional contextual variables (Hall, 2022). Furthermore, videos and transcribed verbal interactions are also suitable data for investigating teachers' implemented adaptive teaching behavior in the classroom (cf. Hardy et al., 2019).

3. Research Perspectives on Spoken L2 Data

Data on spoken language in teaching-learning contexts are interdisciplinary in nature and located at the intersection between corpus linguistics, SLA research, teaching pedagogy, and empirical educational research (Mercer & Dawes, 2014; Paquot & Plonsky, 2017; Thoms, 2012; Wisniewski, 2022). For these research disciplines, these data are a crucial resource for the study of language acquisition. For example, from a linguistic research perspective, they allow to track the development of a target language system to gain valuable insights into linguistic systematicities in the language acquisition process (Wisniewski, 2022). Furthermore, from an educational research perspective, these data can serve as important source to inform language teaching practices by providing empirical insights into the learning process and typical acquisition patterns (e.g., Granger, 2002; Meunier, 2002; Seidlhofer, 2002). These two research perspectives differ concerning their methodological approaches to analyze spoken language in the classroom (cf. Mercer, 2010), which we will elaborate in the following sections.

3.1 Perspectives from Linguistic Corpus Research

In linguistic research, structured collections of language data are referred to as corpora. Language corpora illustrate authentic language use in a pre-defined context,

such as news articles for the study of formal media language or twitter data for the study of social media language. These data typically include (predominantly linguistic) annotations and meta-information that provide information about the represented target group and its language use in a machine-readable format. Corpora are one of the central resources for empirical-driven linguistic research and although their elicitation and compilation are often very resource demanding, a well-structured corpus can be re-used for a variety of research purposes.

One important type of corpora are spoken learner corpora, which elicit spoken language productions of learners (e.g., in classroom interactions with teachers) as opposed to written language productions. Previous research has shown that spoken and written language differ in terms of numerous linguistic factors (e.g., Biber et al., 1999; Goulart et al., 2020; Koch & Oesterreicher, 1985). In contrast to written language, spoken language data is more likely to allow linking linguistic insights with social, cognitive, or psycho-linguistic factors and spoken input is central to foster not only first but also second language acquisition (cf. Hall, 2022), especially in early acquisition stages (Wisniewski, 2022). Spoken output also plays a key role for the assessment of L2 competence, for example in models of communicative L2 competence (Canale & Swain, 1980; Salaberry et al., 2019) or in the Common European Framework of Reference (Council of Europe, 2001). Spoken learner corpora can represent various types of language acquisition data such as second or academic language acquisition within or outside of university. The latter includes, for example, integration courses, preparatory classes, or vocational language trainings for foreign language learners. They allow to track the development of a target language system and to compare shared systematics and variations in the so-called inter-language systems (Selinker, 1972) of learners sharing certain group characteristics [such as proficiency, first language(s), or age] to gain valuable insights into systematics of the language acquisition process (Wisniewski, 2022).

Still, the number of spoken learner corpora remains limited (Gilquin, 2015), despite a notable increase in recent years. In their empirical assessment of 378 quantitative learner corpus studies published from 1991 to 2015, Paquot and Plonsky (2017) found that approximately 75% of studies analyzed written language, but that the proportion of spoken learner data grew from 9% in studies from 1991-1999 to 32% in studies from 2008-2015 (p. 73). Only a quarter of the spoken language corpus studies included interactional data of any form and only 2% (7 studies) included interactions between learners as well as between learners and teachers in authentic instructional

contexts. Most spoken learner corpora instead focus on the elicitation of monologues or dialogues (Wisniewski, 2022), often elicited under (quasi-)experimental conditions (e.g., De Clercq & Housen, 2019; Lintunen & Mäkilä, 2014; Pallotti, 2019) which – unlike authentic communicative settings – allow to control for a variety of factors. However, they lack ecological validity. Since findings from controlled settings may not be transferable to real-life settings, it is crucial to also study language use in more authentic contexts. This allows, for example, to investigate to which degree and how interlocutors in general and teachers specifically adapt their language to provide language learners with input in their Zone of Proximal Development. Authentic classroom settings further elicit diverse communicative situations (e.g., group work or plenary teaching) across which adaptivity strategies of teachers in authentic language use can be studied and compared.

One major reason for the scarcity of spoken learner corpora is that their compilation poses several methodological, technical, and legal challenges (Gilquin, 2015). Wisniewski (2022) points out that modern spoken (learner) corpora require an alignment between the transcription layer and the corresponding audio or video data which is not trivial for most corpus formats. Furthermore, the transcription especially of learner data requires considerable resources, which puts a practical limit to the feasible size of these corpora. Additionally, learner language often requires some form of normalization based on target hypothesis, which again can be far from trivial, because the intended meaning often cannot be reconstructed unambiguously (Lüdeling, 2008; Lüdeling et al., 2005; Meurers, 2015). Finally, the publication of audio and video material comes with a variety of legal challenges due to data security and privacy issues. This is particularly true for the elicitation of authentic data from educational institutions, e.g., (video-)recordings of classroom.

Existing spoken learner corpora have several limitations regarding the sample size, target groups and types of linguistic annotations, the lack of appropriate information about the quality of the data, and the availability and validity of associated meta-information (Paquot & Plonsky, 2017). Most learner corpus studies have focused on adults in English foreign language acquisition settings at universities with advanced proficiency levels (Gilquin, 2015; Paquot & Plonsky, 2017) whereas there is a distinct lack of spoken German L2 learner data (see Hirschmann & Schmidt, 2022; Wisniewski, 2022, for two comprehensive and recent overviews). Notable German learner corpora are the Hamburg-Bamberg corpus, which partially consists of classroom interactions in German science classes in elementary and early secondary

school (Köhne et al., 2015), and the GeWiss corpus (Fandrych et al., 2012), which consists of spoken academic L1 and L2 data for German, English, and Polish that was elicited in university. Moreover, there is still a lack of empirical data on specifically German-spoken learner corpora that represent the demographically dominant groups of GSL learners, such as learners with Arabic as a first language (Wisniewski, 2022).

In addition to the limited diversity of samples, the content of linguistic analyses of learner corpora was also mainly concentrated on a few features, namely phonology and production fluency, lexis, or co-occurrences of multi-word expressions (Paquot & Plonsky, 2017). Less corpus linguistic research has been conducted on discourse and pragmatics (cf. Paquot & Plonsky, 2017; Wisniewski, 2022)⁷. These linguistic domains take into account how shared (extra-)linguistic contexts contribute to the interpretation of interactions and the construction of meaning. This makes them particularly suited for interdisciplinary research on classroom interactions in video studies. Moreover, in most spoken learner corpora the quality of the transcription is not reported or evaluated (Paquot & Plonsky, 2017). Wisniewski (2022) points out the importance of transparently reporting and documenting the quality of every corpus compilation step because it immediately leads into the re-usability of the corpus and the robustness of any findings gained from the data. This holds especially for the reliability of transcriptions and annotations. Furthermore, learner corpora often use empirical methods to assess proficiency that might be invalid or unreliable, such as years of study or self-evaluations of learners rather than language testing (Gilquin, 2015; Paquot & Plonsky, 2017). These hinder insights into the link between language proficiency and authentic language performance. However, the validity and reliability of proficiency assessment could be increased by the use of standardized tests and by addressing learners individually to account for intra-group heterogeneity (Paquot & Plonsky, 2017, p.71 f). At the same time, this would allow to further study the link between assessment test performance and spoken or written language performance in authentic language production scenarios. Altogether, these issues of existing spoken learner corpora limit their (re-)usability.

⁷ Despite the lack of pragmatic learner corpora, the study of language learners' ability to construct and encode meaning in their L2 has been an active field of SLA research for several decades known as interlanguage pragmatics (e.g., Kasper & Rose, 1999; Ohta, 2005; Sykes & Cohen, 2018).

3.2 Perspectives from Empirical Educational Research

It is worth mentioning that there is also data of spoken language in authentic language classrooms gathered through video studies in empirical education research. However, to our knowledge, there have been no studies to date that have systematically prepared data on spoken language in German as a Second Language classes. Existing video studies on language classes in Germany targeting secondary school English lessons (DESI, Helmke et al., 2008) or primary school German lessons (VERA, Helmke et al., 2007; PERLE, Lotz et al., 2013). Although these studies created transcripts of the spoken language in the classroom, these transcripts have some limitations that provide challenges for further interdisciplinary analyses. First, the reliability of the transcription process is not reported, so that the robustness of the transcriptions of the spoken language might not be ensured. Second, most transcripts lack important information on verbal interactions such as speech targets. However, the consideration of the targeted learner(s) of teacher utterances is especially important for analyses of teachers' interaction with individual learners in heterogeneous classrooms, for instance, to investigate adaptations of the complexity of teachers' instructional language. And last, these studies focus on didactic and pedagogical-psychological purposes such as the investigation of instructional quality, teaching practices or learner-teacher interactions (e.g., Helmke et al., 2007; Lotz, 2016). Thus, the data are neither prepared nor used for linguistic analyses of spoken language in the classroom. However, linguistic analyses of spoken language in the classroom might also be of interest for educational scientists due to the major role of language in instruction (e.g., Borg, 2006). The amount and quality of the language input learners receive from their teachers contribute to the instructional quality (Ellis, 2005). Therefore, how teachers use language in the L2 classroom is part of their professional competencies (cf. König et al., 2016).

Moreover, not only linguistic data on interactions in the classroom are important for a better understanding of multimodal classroom discourse (Lim, 2019). Beyond the use of spoken and written language, nonverbal behavior, such as teacher gestures, are also of significance for promoting successful language acquisition processes because they can guide the learners' attention, reduce cognitive burden (e.g., Lazaraton, 2004; Valenzeno et al., 2003), and thus, support comprehension of spoken language (Dargue et al., 2019; Goldin-Meadow, 2000; Hostetter, 2011). The combined investigation of teachers' use of spoken language and their gestures seems necessary to gain a deeper understanding of teacher-learner interactions in the language classroom

(cf. Lim, 2019). For instance, teachers might use different types of gestures for different levels of language complexity (e.g., more complex utterances may be accompanied more frequently by gestures that illustrate them; Sato, 2020). While there are already some studies combining gestural and linguistic analyses (e.g., Lim, 2019), they are limited to the analysis of a few selected linguistic properties of language. Especially in the context of (second) language acquisition a broader perspective on the linguistic properties of language is of interest. Moreover, insights from SLA research on how to operationalize and measure language performance can make a valuable contribution.

Besides, given that opportunities to create oral output play a crucial role in language learning (Ellis, 2005), it is an important research question how learners' verbal utterances occur in classroom interactions. In Germany, classroom interactions are usually led by teachers who create and co-regulate opportunities for their learners to engage by calling on different learners (e.g., Böheim et al., 2020). In this context, learners' hand-raising function as a gateway to classroom participation because learners signal through their hand-raising that they want to actively contribute to the lesson (Fredricks et al., 2019). However, teachers also have the possibility to engage different learners' verbal participation by calling on learners who did not raise their hands. Data on naturally occurring language and teacher-learner interaction patterns in L2 classrooms are of great importance for educational research to investigate which learners are involved in the verbal classroom interaction and how opportunities for verbal participation in GSL classrooms are used by the different learners themselves or prompted by teachers (e.g., learner-initiated vs. teacher-initiated participation).

In the present study, we integrate the different perspectives of linguistic learner corpus research and empirical educational research, because we expect that an interdisciplinary approach contributes to a greater understanding of the linguistic structure and educational functions of spoken language and verbal classroom interaction in GSL classes (cf. Mercer & Dawes, 2014).

4. The Present Study

The present study aims to contribute to a variety of existing research desiderata. We address these needs with the COLD corpus by collecting video data from integration courses for adults and preparatory classes for children with a variety of L1 that depict the natural distribution of first language backgrounds in the German

population of GSL learners. For the presenting corpus the videodata of the consenting participants was transcribed and annotated.

The presented COLD corpus has several advantages including the focus on authentic classroom situations, the transparency of the data processing, the interdisciplinary research approach considering various verbal, nonverbal, and contextual characteristics (such as the explicit encoding of speech targets), as well as detailed meta-information about teachers and learners. First, the corpus contains transcriptions and annotations from authentic teaching-learning conversations of GSL classes, a target group that has so far not been included in Spoken Learner Corpora. Thus, the authentic contextual variables (e.g., social forms) can be considered for analysis. Second, by providing our guidelines for transcriptions and annotations, as well as reporting corresponding evaluation of its reliability, we endorse the robustness of any findings gained from the data and the re-usability of the corpus. Third, due to its interdisciplinary approach, the COLD corpus makes it possible to examine current issues from several perspectives and combine research questions across disciplines. Whereas linguistic research typically focuses on the spoken language productions of teachers and learners in the classroom, educational researchers can use the video data to study further teacher-learner interactions beyond the usage of language. To obtain further insights into the interdependency of verbal and nonverbal communicative aspects in authentic learning environments, we make non-linguistic aspects accessible in our data. Furthermore, these data enable the investigation of research questions that link different research areas, such as the adaptation of teachers' instruction to individual learners through adjusted complexity of language or nonverbal behavior. Fourth, the corpus contains rich meta-information about teachers' and learners' educational and cultural backgrounds, language proficiencies, and motivations. Thus, in addition to the observed data on the linguistic heterogeneity of the GSL learners, the meta-data enables extensive analyses of classroom heterogeneity and associated teachers' adaptations to individual learners' prerequisites.

In the following, we give an overview of the structure of the corpus as well as a detailed description of the available meta-data. Furthermore, the transcription guidelines for the spoken language and the annotations of the non-verbal behavior are presented. We then report on the robustness of the data transcription and annotation to ensure the quality of the corpus is transparent. Finally, descriptions of research opportunities point out the value of the corpus, not only for the linguistic but also for a variety of pedagogical-psychological questions.

4.1. The COLD Corpus

Collection

The COLD corpus consists of video recordings of German as a second language classes from the educational sectors of school and adult education. The videos were recorded as part of the COLD project in educational institutions in North Rhine-Westphalia in the Year 2020. The available data of the pilot study contains five videos of lessons from preparatory classes (PC) of secondary level one of secondary schools as well as five videos of lessons from integration courses (IC) from adult education institutions. In both education sectors, the learner groups consist of beginning to low intermediate learners of German. Adult learners who do not yet have written language skills (i.e. persons who are illiterate) usually do not take part in integration courses but attend a literacy course. Two of the five videos from the adult education sector were filmed in a literacy course.

Session structure

Prior to the filmed lessons, teachers were asked to deliver a 45-minute lesson. The content of the lesson was freely chosen. With regard to the social form, the teachers were asked to conduct at least one other social form (group work, partner work, individual work) in addition to a plenary phase (classroom teaching). The filmed lesson duration varied from 44-63 minutes and was on average 48.4 minutes. 59% of the filmed teaching time took place in plenary phases, 18% in group work, 10 % in partner work, and 13% in individual work.

The lesson recordings were made with two video cameras. One camera recorded the teacher from the learners' perspective. Another camera filmed the entire learning group in full frontal view. The teachers were given a lavalier microphone for the audio recordings. The learners' statements were recorded via two room microphones. The recordings were edited and transcribed.

Sample and Meta-data

As part of the COLD study, rich background information on teachers' and learners' demographics, abilities, and motivations was assessed that allows an extensive description of the sample. Table 1 presents an overview of the meta-data on teachers and learners in the COLD study. Beyond the common meta-information for spoken learner corpora such as the learners' language background, age, gender, first language(s), and proficiency level (Wisniewski, 2022), our corpus includes information

about the learners' country of birth, educational background, socioeconomic status, and motivations as well as various information about the teachers' demographics, educational backgrounds, competencies, and motivations.

Table 1*Overview of the meta-data on teachers and learners in the COLD study*

Construct	Meta-data Queries	Type of Assessment	Availability
<i>Information collected from teachers</i>			
Teachers' demographics	Gender, age, country of birth, languages spoken	Questionnaires	yes
Teaching qualification and expertise	Teaching experience, teacher training, educational qualifications	Questionnaires	yes
Work setting	Employment relationship, teaching hours per week	Questionnaires	yes
Teachers' motivation and beliefs	Self-efficacy, teaching enthusiasm, work satisfaction, job stress, burn-out, beliefs about multilingualism	Questionnaires	Aggregated by teacher
Teachers' professional competencies	Pedagogical knowledge, professional competence for teaching GSL, professional vision	(Standardized) tests	Aggregated by teacher
Teachers' assessment of each individual learner	Learners' language proficiency, motivation, classroom behavior	Questionnaires	Aggregated by learner
<i>Information collected from learners</i>			
Learners' demographics	Gender, age, country of birth, languages spoken, educational background, SES,	Questionnaires	yes
Learners' motivation	Motivation to study German, Self-concept for German	Questionnaires	Aggregated by learner
Learners' language proficiency	C-Test (gap text)	Standardized tests	Aggregated by learner

The teachers ($N = 10$) participating in the pilot were on average 46 years old (24-68 years), predominantly female (80.0%), and born in Germany (70.0%). All teachers had at least a Master's degree or state examination and one or more qualification certificates for teaching GSL. As two groups of learners in the school context were each filmed with two different teachers, the pilot study includes eight different classrooms ($n_{PC} = 3$, $n_{IC} = 5$) with a total of 81 learners (8-13 participants in each class, on average 10 per class). Learners in the school classes ($n_{PC\text{-learners}} = 33$) were on average 13 (10-17) years old and 45.5% were female. Learners in the adult classes ($n_{IC\text{-learners}} = 49$) were on average 39 (24-62) years old and 67.3% were female. Further information on the individual teachers and classes contributing to the corpus is shown in Table 2.

Table 2*Description of teachers and learners contributing to the corpus*

Teachers					Learner Groups				
ID	Age	Gender	Country of birth	GLS teaching expertise (years)	<i>n</i> Learners	<i>M</i> (<i>SD</i>) Age	<i>M</i> (<i>SD</i>) Month in Germany	<i>M</i> (<i>SD</i>) School attendance (years)	<i>M</i> (<i>SD</i>) C-Test Performance
PC1	39	F	Spain	4.5	10	12.6 (2.01)	12.7 (8.7)	5.4 (3.3)	37.5 (17.1)
PC2	57	F	France	10.5					
PC3	43	M	Germany	6.1	10	14.1 (2.1)	13.4 (5.1)	7.8 (4.3)	35.2 (8.9)
PC4	43	F	Germany	4.0	13	12.4 (1.0)	12.7 (6.4)	5.2 (4.1)	22.5 (17.6)
PC5	24	F	Germany	4.3					
Mean all PC	41.2	80% F	60% Germany	5.9	11	12.9 (1.8)	12.9 (6.8)	6.0 (4.0)	30.9 (16.4)
IC1	57	M	Germany	33.0	8	36.0 (8.2)	34.4 (21.9)	12.0 (6.3)	21.9 (8.1)
IC2	68	F	Austria	4.0	11	43.7 (12.5)	45.4 (18.2)	5.3 (4.1)	9.7 (11.9)
IC3	53	F	Germany	16.3	8	43.3 (9.7)	96.5 (107.1)	12.3 (3.2)	31.1 (14.6)
IC4	45	F	Germany	3.5	9	35.2 (10.9)	34.0 (23.7)	6.8 (3.2)	7.2 (4.1)
IC5	32	F	Germany	5.9	12	34.2 (7.3)	25.3 (19.4)	11.0 (1.5)	17.3 (7.5)
Mean all IC	51	80% F	80% Germany	12.2	10	38.2 (10.3)	44.9 (53.3)	9.5 (4.6)	17.1 (12.4)

Notes. PC = preparatory classes; IC = integration courses. Teachers PC1 and PC2 as well as PC4 and PC5 were filmed with the same learner group. F = female; M = male. "School attendance" assessed years learners went to school before coming to Germany. Learners' knowledge of German was determined by means of C tests (cloze texts), in which one point is awarded for each correctly completed cloze. The learners completed two level A1 and two level A2 texts; the maximum achievable score is therefore 80.

Corpus statistics

The corpus consists of recordings of ten GSL lessons. Table 3 shows the number of turns, number of tokens, the average turn length in seconds and in words, and the average number of turns and words per session for teachers and learners in the data split by lessons for adults and children. We define *turn* as all continuous utterances by a single speaker which address the same speech target. The data shows that both, in lessons for adults and children, learners contribute more language material than teachers in terms of the number of turns and words. However, the contribution of teachers is considerable given that the number of individual learners per lesson is much larger than the number of teachers per lesson.

Table 3

Corpus statistics

	Speaker	Num. turns	Num. words	Avg. turn length in words	Avg. turn length in seconds	Avg. num. turns per session	Avg. num. words per session
IC	teachers	3,554	18,315	5.15±5.34	1.99±2.14	756±145	3,663±872
	learners	5,611	24,464	4.36±5.24	1.91±3.31	1,135±112	4,893±980
PC	teachers	1,963	17,927	9.13±13.2	4.61±7.11	395±32.4	3,585±505
	learners	3,519	25,101	7.13±10.9	3.92±6.16	709±60	5,020±555

Note. PC = preparatory classes; IC = integration courses. Num. = number; Avg. = average.

4.2. Transcriptions & Annotations

The videotaped lessons were transcribed using the EXMARaLDA transcription editor (www.exmaralda.org; Schmidt & Wörner, 2014). In this tool, the transcription is structured in annotation tiers, each containing events on the timeline. As this compositional structure allows the transcription of spoken language and accompanying behavior from multiple speakers simultaneously while accounting for the timeline, EXMARaLDA is particularly suitable for our research on (verbal) classroom interactions.

We used a multifaceted transcription guide (complete guidelines are presented in the supplemental material) that considers all the features relevant to our research questions. The guidelines were interdisciplinary developed in an iterative process that consisted of targeted specifications of the annotation guidelines following error analysis of previous inconsistencies. Trained transcribers (all authors and their research assistants) prepared the transcripts. We provide detailed instructions on the structure and annotation of the speakers in the EXMARaLDA software, refer to the handling of verbal content (spelling, filler words, utterances in foreign language, read aloud etc.) as well as non-verbal behavior (the recording of calls and hand messages etc.).

Only study participants (both teachers and learners) who gave informed consent were included in the transcription. In the case of underage learners, the consent of the legal guardian was obtained. Utterances from participants who did not give their consent were not transcribed. In order to anonymously link the meta-data from questionnaires and test instruments to the video data, each participant was randomly assigned an ID. Transcripts were completely anonymized by using participants' IDs. To identify the different speaker IDs, a seating plan indicating the ID of each learner in the classroom was used.

Verbal utterances

Each speaker in the classroom (learners, teacher) was assigned a verbal tier where the verbal utterances of this individual person are transcribed. In addition, we also incorporated tiers for verbal utterances of unidentifiable speakers (unknown) and groups of speakers (some, all). Verbal content was transcribed using the standard German orthography as found in a dictionary of German. In plenary teaching, all (intelligible) verbal utterances of all participants in the lesson were transcribed. In group or partner work, however, the transcription focused on the teacher and the learners in the group that the teacher was attending to in that situation. Verbal utterances from other groups and side conversation were not transcribed.

Additionally, we included a verbal tier "audio" which contains verbal utterances from audio recordings in the teaching session. Especially in language instruction, teachers sometimes use audio recordings (e.g., from a CD or a video recorder) to foster listening comprehension.

Speech targets

Each speaker was assigned a target tier in which the ID(s) of the verbally or nonverbally addressed person(s) was annotated. The annotation of the target was aligned with the corresponding event in the verbal or nonverbal tier. For every

intelligible utterance (by any speaker), at least one target had to be specified. Utterances targeting participants who did not consent to be included in the transcription were referenced by a collective anonymous ID (X).

Speech pauses

Pauses in utterances were annotated following cGAT conventions (Schmidt et al., 2015). Pauses longer than or equal to 0.2 seconds were annotated as events in the pause tier. Pauses shorter than 0.2 seconds were marked as (.) in the verbal tier of the person who spoke last before the speech pause.

Teachers' gestures

Teachers' gestures (i.e., spontaneous interactional movements of their hands and arms, with a close relation to their speech; McNeill, 1992) were coded in one, or if two gestures occurred simultaneously, in two gesture tier(s). The coding of the gestures was temporally aligned with the movement of the gestures. For the annotation of gestures, we used the categorization of Sato (2020). This coding system includes seven gesture types based on the widely used hand gesture system of McNeill (2008; *iconic, deictic, metaphoric, beats*) and extended by the body gestures of Wang and Loewen (2015; *emblems, head movements, affect display*). Additionally, the addressed target of each gesture was noted in an extra target tier for gestures.

Initiations of learners' classroom participation

An additional nonverbal tier was assigned to each speaker or group of speakers. The main focus of this tier was on learners' classroom participation, both learner-initiated (i.e., by learners' hand-raising) and teacher-initiated (i.e., by calling on individual learners). The learners' hand-raising (i.e., arm, hand, or finger sign to indicate that they want to contribute to the lesson) was annotated in the learners' nonverbal tiers. The code took into account whether learners' hand-raising was accompanied or followed by a call from the teacher (or other learners). Moreover, learners' classroom participation by speaking without call or permission was also annotated in their nonverbal tier. The teachers' calls of learners (i.e., asking a learner to or giving the learner the right to speak in the classroom) were annotated in the teachers' nonverbal tiers. The annotations of teachers' calls indicated which learner was called (ID) and whether the call was content-related or –unrelated.

In addition, other nonverbal behavior was described in the nonverbal tier if it led to altered or unintelligible language. For instance, nonverbal behavior should be noted when it causes a longer pause in utterances (e.g., when a teacher writes something on the board) or when the speaker talks about an object without naming it (e.g.,

“What is that?” while pointing at the exercise book). However, the nonverbal tier does not provide a complete scene description. Non-verbal behavior without a direct relevance for the understanding of the verbal utterances was not taken into account (e.g., minor body movements such as unrelated head scratching). All annotations in the nonverbal tiers were temporally aligned with the corresponding nonverbal behavior (e.g., start of the hand-raising event when learner raises the hand, end of the event when the learners’ hand is down again).

Social forms (lesson phases)

Social forms in the classroom (plenary phase, group work, partner work, individual work) were annotated at the beginning of a social form or when the social form changed. For partner or group work, the group membership of each learner was noted in the nonverbal tier.

4.3. Inter-Rater Reliability Experiment

Data

We evaluated the robustness of our annotations and the transcriptions of verbal utterances in separate inter-rater reliability (IRR) experiments. For the evaluation of the transcription, annotation of speech targets, duration of utterances and pauses as well as teachers’ calls, we used two full-lesson recordings, one of a PC and one of an IC. Each lesson was divided into eleven partitions of approximately the same length and assigned to a different annotator pair chosen from a pool of overall seven trained annotators.

Different types of gestures and learners’ classroom participations, as well as social forms displayed within a single lesson, did not occur at a sufficient rate in the two lessons. We, therefore, conducted the IRR experiments for these annotations on all ten lesson recordings of approximately 45-minutes, five of PC and five of IC. Videos were coded by two trained observers in an event sampling procedure.

Verbal Utterances

Most spoken learner corpora do not report the reliability of the initial transcription of the speech signal. However, there has been a call for more transparency regarding the reliability of annotations and transcriptions in learner corpus research (Paquot & Plonsky, 2017; Wisniewski, 2022). Thus, we chose to evaluate the robustness of the transcriptions by calculating the turn-wise Levenshtein distance (Levenshtein, 1966) between annotators. The Levenshtein distance is a common

measure for string comparison in computational linguistics and computer science. It quantifies dissimilarities between two strings A and B in terms of the minimal number of character deletions, insertions, or substitutions that are necessary to convert string A into string B. Since the simple Levenshtein distance is sensitive to string length, we normalized it by the maximal distance between strings A and B (i.e., the number of characters in the longer of the two strings). Since the robustness of transcriptions in spoken learner corpora is rarely addressed, there are no established benchmarks for acceptable distances. We propose that a normalized distance < 0.20 is within an acceptable range of deviation for the analysis of meaning and linguistic analyses at or above the word level: This threshold indicates that less than 20% of the characters in two strings deviate from each other which is unlikely to impact such analyses.

To obtain a more robust distance measure within annotator pairs, we calculated the metric turn-wise, defining turns as an uninterrupted utterance by one speaker. Out of overall 1,474 turns, we excluded turns in which at least one annotator had marked parts of the utterance as incomprehensible ($N = 75$) or turns that had been transcribed by one annotator but not by the other ($N = 510$, predominantly learner utterances), because these cannot be meaningfully compared using the Levenshtein distance. We calculated the following distances on the remaining 889 turns.

Overall, we achieve a normalized Levenshtein distance of $\mu = 0.17$ ($SD = 0.20$; $Mdn = 0.10$). The high standard deviation can be explained by a few turns in which one annotator within an annotator pair transcribed some language material that was absent in the other annotators' transcription. However, these are rare instances in the data which is also being highlighted by the median which falls considerably below the mean. Splitting these results by speaker role, we see that the distance is overall lower for teachers ($\mu = 0.13$; $SD = 0.18$; $Mdn = 0.06$) than for students ($\mu = 0.22$; $SD = 0.22$; $Mdn = 0.14$). This can be explained by the higher audio quality for teachers who were equipped with their own microphones, but might also partially be due to the variation in the learner language. We do not observe any meaningful differences between integration courses ($\mu = 0.18$; $SD = 0.21$; $Mdn = 0.10$) and preparatory classes ($\mu = 0.17$; $SD = 0.20$; $Mdn = 0.11$). This is further supported by the results of a two-sided t-test which fails to find evidence for a significant difference between integration courses and preparatory classes ($t = -0.52498$, $df = 820.03$, $p\text{-value} = 0.5997$) While these results indicate a somewhat lower agreement for learners as opposed to teachers, they still show a high robustness across annotators, especially when taking into account the median distances.

Speech Targets

We calculated the IRR for the identification of teachers' speech targets using Cohen's kappa (Landis & Koch, 1977). We only included teacher turns that were transcribed by both annotators and in which both annotators had identified a speech target. In the PC lesson ($N = 556$ turns), our annotator pairs achieved near-perfect agreement (following Landis & Koch, 1977) with an IRR of $\kappa_{\mu} = 0.87$ ($SD = 0.09$). According to our error analysis, 47 out of 60 disagreements arose from annotators disagreeing on whether someone specifically addressed someone or spoke to everyone in the room.

Due to the Covid-19 health regulations in place during the elicitation of our corpus data, the recording of the IC lesson took place while teachers and learners were wearing face masks. For this lesson ($N = 855$ turns), we achieved an average

agreement of $\kappa_{\mu} = 0.72$ ($SD = 0.17$). However, an error analysis revealed that disagreements arose predominantly (101 out of 135) from annotators identifying overlapping subsets of target, e.g., Speaker 5 and 6 vs. Speaker 5, 6, and 7. When considering overlapping targets as sufficient criterion for agreement, our annotators achieve an average agreement of $\kappa_{\mu} = 0.92$ ($SD = 0.08$). This shows that the target annotations are highly robust even when participants are wearing masks.

Duration of Utterances and Speech Pauses

To ensure that our annotators produced robust estimates of speech and pause durations, we compared the difference between turn durations and pause durations in seconds. For the calculation of turn duration, we only included turns that were transcribed by both annotators ($N = 935$). For the calculation of pause duration, we summed up pauses by intervals of 30 seconds to avoid alignment issues between annotators resulting in 179 30-second intervals. To account for the accumulative effect of micro pauses, we included these in the calculation of pause durations assuming a default duration of 0.1 seconds.

For turn duration, we conducted a Wilcoxon rank sum test with continuity correction to compare the durations between annotator pairs, finding no significant difference in either the PC lesson ($W = 123,754$; $p = 0.78$) or the IC lesson ($W = 99,534$; $p = 0.47$). The 935 turns lasted on average 3.43 seconds ($SD = 5.67$). The difference in turn durations between annotators was on average 0.71 seconds ($SD = 1.77$), which we consider to be negligible.

For pause duration, we conducted a Wilcoxon rank sum test with continuity correction to compare the durations between annotator pairs, finding no significant difference in either the PC lesson ($W = 3,296$; $p = 0.27$) or the IC lesson ($W = 3,726$; $p = 0.36$). Across the 179 30-second intervals, the difference in pause durations between annotators was on average 1.99 seconds ($SD = 2.69$), which we consider to be negligible. This shows that the estimates of speech and pause durations in our corpus are robust.

Teachers' Gestures

A total of 4,143 gestures were identified. In 92.9% of the cases (3,848 gestures), both observers identified a gesture. IRR for gesture type identification was sufficiently good as measured by Cohen's kappa ($\kappa = 0.82$). In cases in which the encoded gestures did not match, the authors discussed each gesture individually and determined a gesture.

On average, teachers used 414 gestures per lesson. The number of gestures varied from 228 to 694 gestures per lesson. The most frequent type of gesture in the GSL lessons studied was deictic (1,630). Metaphoric gestures (1,085) and head movements (975) were also frequently used by teachers. Iconic gestures (198), emblems (132) and beat gestures (17) were used rather rarely in the observed GSL lessons.

Initiations of learners' classroom participation

A total of 2,373 learner- or teacher-initiated participations of learners in the classroom were identified across all ten videos. In 94.9% ($N = 2,251$) of the cases, both observers identified an initiation of classroom participation. The coding of the type of learners' participations was measured with an almost perfect agreement ($\kappa = 0.92$). The authors discussed each case in which the annotators did not agree and determined a code.

Teachers' calls of individual learners in plenary teaching phases were identified with an inter-rater agreement of 81.7%. We recommend practicing and controlling the identification of teachers' calls, as our analysis shows that this annotation is somewhat more prone to errors. Nevertheless, the coding of the type of call (i.e., whether the call was content-related or -unrelated) achieved high agreement ($\kappa = 0.90$).

Social forms

Every minute of the 483 minutes was assigned to one of the four social forms by two trained observers. If two social forms occurred within the same minute, the social form with the temporal majority was annotated. In order to ensure that social forms that occur very briefly are also included, we annotated social forms of less than 30 seconds if the two adjacent cells (minutes) had the identical code (e.g., 25 seconds of individual work if plenary teaching was done before and after). With a Cohens' Kappa of $\kappa = 0.92$ the measured agreement can be considered as almost perfect (Landis & Koch, 1977). In addition, all four social forms individually are reliably ascertainable ($K_{\text{plenary teaching}} = .96$; $K_{\text{group work}} = .98$; $K_{\text{partner work}} = .78$; $K_{\text{individual work}} = .85$).

Of the 483 observed minutes of instruction in the COLD corpus, most time (286 min) was spent in plenary teaching (59.2%), 85 minutes were spent in group work (17.6%), 49 minutes in partner work (10.1%), and 63 minutes in individual work (13%).

5. Research Opportunities

The COLD Corpus provides several research opportunities for investigations of teaching and learning processes in GSL classrooms. Importantly, the multimodal transcriptions and annotations facilitated by our guidelines enable the analysis of

interrelations of different aspects that occur simultaneously in the lessons (e.g., verbal and non-verbal behavior). Moreover, a major advantage of our transcription guidelines is the consideration of each individual speaker and the targets of the speakers' verbal and nonverbal behavior. Thus, it allows to analyze within-class heterogeneity and teachers' individualized teaching behavior in relation to the individual learners. In the following, we outline exemplary research possibilities from linguistic and empirical educational perspectives with a focus on interdisciplinary research opportunities. From a linguistic perspective, the COLD corpus is highly valuable because it makes linguistic properties of language in authentic classroom interactions observable and quantifiable. It also gives the opportunity to link them to the extensive meta-data collected about learners and teachers. For example, the corpus allows linking the linguistic properties of learners' spoken language performance in the classroom to their performance in written proficiency assessments such as the C-test. This not only allows to gain insights into the language performance differences between learners of different proficiency levels based on high-quality estimates of language proficiency but also allows to link observations about authentic oral language use with test performance in written assessments. This is not only crucial for SLA research, but also for research on language testing. To this end, the corpus format supports the augmentation of the corpus with additional manually annotated information layers, encoding, for example, specific linguistic constructs such as part-of-speech or syntactic structures. Alternatively, the transcription can be extracted into a machine-readable format that can be used as input for Natural Language Processing pipelines providing automatically generated linguistic annotations. Weiss et al. (in press) use such an approach to study teachers' use of academic language in content-matter teaching across German school types, but state that their analysis is confined to a global comparison of whole teaching sessions because their data does not allow to re-construct who is being addressed on the level of individual utterances. This is necessary, though, to facilitate a fine-grained analysis of linguistic alignment processes in authentic teacher-learner or learner-learner interactions. The annotation of speech targets for teacher and learner utterances in the COLD corpus allows to contrast the language use of teachers within a session depending on whom they are addressing. This allows to study how teachers address individual learners in contrast to addressing the entire classroom and if they adapt their language to their perceived individual learners' language proficiency. We can further zoom in on individual interactive sequences between teachers and learners to study alignment processes between them and obtain an estimate of whether

teachers' language input is in the Zone of Proximal Development of an individual learner and pushes them to produce more elaborate and varied language themselves.

The recorded sessions are segmented into social forms. This not only enables the analysis of the structure of sessions with respect to individual work, partner work, group work, and plenary teaching phases. It also allows to link learner proficiency, teaching phases, and teachers' guidance of learning process in these groups/dyads or the plenary discussion. The annotations also facilitate an interdisciplinary investigation of the linguistic properties of all involved interlocutors in specific teaching phases. This extends the analyses outlined in the previous paragraph by comparing teachers' adaptation strategies at different stages of instruction and with groups of varying language proficiency which form a more homogeneous counterpart to the heterogeneous overall classroom group.

The comprehensive transcripts of the spoken language of all participants in the class facilitate further coding that is based on the learners' or teachers' utterances, such as the complexity of teachers' questions. Teachers' questions play a key role in teaching as they serve to structure and guide the learning process (e.g., Gall, 1970) and thus are of interest for empirical educational research. Evidence suggests that teachers adjust the complexity of questions to the language level of the learning group (Konzett-Firth, 2020). However, it is still unclear whether teachers vary the complexity of their questions between learners within a classroom depending on their perception of learners' language proficiencies. Adaptive questions could provide individualized learning opportunities, especially in heterogeneous groups such as GSL classrooms. Due to the annotation of the target to every verbal utterance as well as rich background data on the learners' language proficiency, our corpus allows analyses of variance in teachers' questions between learners in a classroom.

Furthermore, the systematic annotation of even short speech pauses in the lessons facilitates the investigation of waiting times after teachers' questions. Prior studies indicate that teachers hardly give learners enough time to think before answering the questions (Tobin, 1987). However, an extension of waiting time can enable longer learner responses in language classrooms (Smith & King, 2017). Some evidence indicates variance in the duration of teachers' waiting times between learners depending on teachers' perception of learners' language proficiencies (Shrum, 1985). The complex annotation structure of the COLD corpus furthermore allows for a wide analysis of classroom discourse including variation in wait-time depending on the social form, the complexity of the question, and the learners' language proficiency.

The COLD corpus also provides the opportunity to analyze non-verbal teaching behavior that asides spoken language to analyze potential interrelations. One important part of teachers' instructional communication besides their spoken language is teachers' gestures, which can support learners' comprehension of spoken language and thus the language acquisition (Dargue et al., 2019; Goldin-Meadow, 2000; Hostetter, 2011). It has been repeatedly hypothesized that different types of gestures are conducive to successful language comprehension depending on the complexity of the language (Hostetter, 2011; Lim, 2019). Moreover, there is evidence that different types of gestures (e.g., deictic or iconic gestures) support learners to varying degrees, depending on their language proficiencies. It has been shown that particularly low-performing learners are more likely to benefit from the teacher's use of gestures compared to high-performing learners (Sueyoshi & Hardison, 2005). The COLD corpus and its encoding of teachers' gestures enables to investigate teachers' use of different types of gestures as a function of the spoken language and the target addressed in order to gain a better understanding of the teachers' nonverbal communication and its role in SLA teaching. In addition, the recording of learners' hand-raising and teachers' calls (verbal and non-verbal) makes it possible to analyze classroom participation patterns of learners. On the one hand, the annotations of hand-raising behavior and speaking without prompting enable to examine how often learners participate in a self-directed manner which is an observable indicator of their engagement (e.g., Böheim et al., 2020). On the other hand, the annotations of teachers' calls are an important source to investigate how teachers engage learners' participation, for instance in terms of the frequency with which teachers promote verbal contributions of different learners in a classroom. Since it is also recorded in the corpus whether learners make a statement after a call or a report, it can also be investigated to what extent the teacher controls the teaching-learning processes, allows self-initiated participation, and regulates an equally distributed participation through active calling behavior. In addition, associations between the frequency of classroom participation and the learners' attitudes and competences regarding GSL can be investigated. Evidence from German science classes already indicates systematic associations between learners' individual characteristics and their class participation initiated by themselves or their teachers (cf. Decristan et al., 2020). With regard to the created corpus, a multitude of questions in the field can be answered. With the research opportunities listed in mind, the following exemplary questions can be answered with the transcribed data along with the meta-data:

- How does the teacher distribute speaking time to different learners? Are there learners with particular characteristics (e.g. high level of language proficiency) who have especially large amounts of speaking time in class?
- How is the speaking time distributed among learners with different German language proficiencies?
- Do teachers vary the wait time after their questions depending on how they assess learners in terms of their German language proficiency?
- What is class participation like? Who reports to class, how often, and is this related to German language proficiency? Do the teachers succeed in involving the learners evenly in the lessons?

6. Limitations and Outlook

Systematic verbal interactions in the classroom play an important role for SLA (e.g., Thoms, 2012), which is relevant for successful integration in the country of immigration (e.g., Esser, 2006). To date, however, we lack data on verbal interactions of teachers and learners, especially in authentic language classroom settings (e.g., Hall, 2022). Moreover, it is unclear how teachers deal with the linguistic heterogeneity of their learners. To address these research gaps, we present the COLD corpus, which has several advantages and provides opportunities for further research. We created a corpus of written transcriptions and annotations to facilitate research in the field of GSL acquisition. The Corpus offers new insights into GSL language use and pedagogy, especially due to rich meta-data (e.g., objectively measured language proficiency indicators) and multitudinous annotations (e.g., target track of utterances). In addition, the sample of teachers and learners in GSL classrooms in both school and adult education fills the gap of data on highly relevant demographically dominant groups of GSL learners as requested by Wisniewski (2022).

Furthermore, we provide guidelines for transcriptions and annotations of video data from language classrooms, which have been shown to enable reliable transcriptions of spoken language and annotations of nonverbal behavior in the classrooms. Thus, the present study demonstrates how language usage in the classroom can be transcribed in a reliable way, which offers the opportunity to further develop objective measures of linguistic adaptivity in terms of relation between learners' and teachers' levels of linguistic complexity. Overall, the multimodal COLD corpus provides ample opportunities for in-depth investigations of teachers' verbal and non-verbal interactions with their learners in GSL classrooms. The corpus finally allows to study the extent to

which language input adaptation actually takes place in heterogeneous SLA classrooms. The presented COLD corpus also revealed some limitations that should be considered in future research. First, the COLD corpus data are cross-sectional and do not allow causal inferences. However, the focus of the project was to investigate GSL teachers' competence comprehensively by assessing teachers' actual interaction with their learners and the use of language in the classroom. To realize this research goal, a first step is to provide reliable measures of verbal and nonverbal interactions in the classroom. For further analyses, the existing COLD corpus of ten video recordings will be extended by sixty additional lessons ($n=30$ more videos from each, school PC and adult education IC). Second, when interpreting differences between the video recordings it is necessary to take into account that no standardized lesson content has been prescribed (e.g., as given in the TALIS video study; Klieme et al., 2020). Since the corpus contains teaching content from one region of Germany (North Rhine-Westphalia), no nationwide generalizability can be assumed without further ado. However, at least for the area of integration courses, a nationally uniform curriculum can be assumed in terms of content. However, it cannot be ruled out that the lessons contained in the corpus are characterized by region-specific peculiarities (such as dialects). Third, because the COLD project focused on teachers' professional competencies and teachers' language in the classroom, transcriptions of verbal utterances only include side conversations of learners in which the teacher was involved. Conversations between learners, without teacher participation, are not recorded in the corpus.

Finally, our data collection in the IC took partly place during the covid-19-pandemic. Thus, governmental pandemic regulations such as minimum spacing and fixed seating charts might have affected the observed teaching. In the subsequent interviews of the study, teacher often reported that the regulations greatly affected instruction and framework conditions. For instance, teachers reported using more plenary and less partner work, and that some classes were divided so that the number of learners was significantly reduced. Further research is needed to investigate how these changes in instruction might have affected the language in the classroom, for example in terms of interaction and conversation patterns. Moreover, in some recorded lessons it was mandatory to wear face masks. It has not yet been sufficiently investigated whether these face masks influence language instruction. However, we were able to show that an adequate inter-rater reliability of verbal utterances and identification of the respective targets can still be achieved when participants are wearing masks. These findings are not only relevant for our corpus but might also be indicative

for future research projects in which participants' faces are not fully visible (e.g., due to face masks or due to pixelation).

The presented corpus and guidelines contribute to research on a better understanding of SLA processes and related instructional behavior (such as linguistic adaptivity) in classrooms of linguistically heterogeneous groups of learners. While the corpus allows in-depth analyses for GSL acquisition, the transcription guidelines are potentially applicable to many other languages. The guidelines for transcription and annotation can be found in the online supplemental material and may be freely used and adapted by other researchers. The corpus is available under a Creative Commons Attribution-ShareAlike 4.0 license (CC BY-SA 4.0) starting March 1, 2023.

7. Acknowledgments

Funding

The research leading to these results received funding from the Leibniz-Society in the Joint Initiative for Research and Innovation 'Leibniz Cooperative Excellence' under Grant Agreement No K113/2018 in the project called Competencies of school teachers and adult educators in teaching German as a second language in linguistically diverse classrooms (COLD).

Conflicts of interest/Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethics approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study involving human participants was reviewed and approved by the Local Ethics Committee of the German Institute for Adult Education (No. DIE-LEK 2020-02-07).

Consent to participate

All participants provided their written informed consent to participate in this study. In the case of underage participants, the consent of the legal guardian was obtained.

Availability of data and materials (data transparency)

The datasets and custom codes generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Open Practices Statement

The guidelines for transcription and annotation can be found in the online supplemental material and may be freely used and adapted by other researchers. The corpus is available under a Creative Commons Attribution-ShareAlike 4.0 license (CC BY-SA 4.0) starting March 1, 2023.

8. References

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Longman Publications Group.
- Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *The Review of Economics and Statistics*, *86*, 481-496. <https://doi.org/10.1162/003465304323031067>
- Böheim, R., Urdan, T., Knogler, M., & Seidel, T. (2020). Student hand-raising as an indicator of behavioral engagement and its role in classroom learning. *Contemporary Educational Psychology*, *62*, 101894. <https://doi.org/10.1016/j.cedpsych.2020.101894>
- Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research*, *10*(1), 3-31. <https://doi.org/10.1191/1362168806lr182oa>
- Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, *61*(5), 365. <https://doi.org/10.1037/H0029908>
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*(1), 1-47.
- Corno, L. Y. N. (2008). On teaching adaptively. *Educational Psychologist*, *43*(3), 161-173. <https://doi.org/10.1080/00461520802178466>
- Council of Europe (2001). Common European Framework of Reference for Languages: Learning, teaching, assessment.
- Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin*, *145*(8), 765-784. <https://doi.org/10.1037/bul0000202>
- De Clercq, B., & Housen, A. (2019). The development of morphological complexity: A cross-linguistic study of L2 French and English Special Issue on Linguistic Complexity. *Second Language Research*, *35*(1), 71-97. <https://doi.org/10.1177/0267658316674506>
- Decristan, J., Fauth, B., Heide, E. L., Locher, F. M., Troll, B., Kurucz, C., & Kunter, M. (2020). Individuelle Beteiligung am Unterrichtsgespräch in Grundschulklassen: Wer ist (nicht) beteiligt und welche Konsequenzen hat das für den

- Lernerfolg? [Students' differential participation in classroom discourse in primary schools: Who participates (not), and what are the consequences for student learning?]. *Zeitschrift Fur Padagogische Psychologie*, 34(3-4), 171-186.
- Denessen, E., Keller, A., Bergh, L. V. D., & Broek, P. (2020). Do teachers treat their students differently? An observational study on teacher-student interactions as a function of teacher expectations and student achievement. *Education Research International*, 2020, 1-8. <https://doi.org/10.1155/2020/2471956>
- Ellis, N. C., & Bogart, P. S. H. (2007). Speech and language technology in education: The perspective from SLA research and practice. Speech and Language Technology in Education Conference, Farmington, PA, USA.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33(2), 209-224. <https://doi.org/10.1016/j.system.2004.12.006>
- Esser, H. (2006). *Migration, language and integration*. WZB.
- Fandrych, C., Meißner, C., & Slavcheva, A. (2012). The GeWiss corpus: Comparing spoken academic German, English and Polish. In T. Schmidt & K. Wörner (Eds.), *Multilingual corpora and multilingual corpus analysis* (Vol. 14, pp. 319). John Benjamins Publishing Company
- Fredricks, J. A., Hofkens, T. L., & Wang, M.-T. (2019). Addressing the challenge of measuring student engagement. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge Handbook of Motivation and Learning* (pp. 689-712). Cambridge University Press. <https://doi.org/10.1017/9781316823279.029>
- Gall, M. D. (1970). The use of questions in teaching. *Review of Educational Research*, 40(5), 707-721. <https://doi.org/10.3102/00346543040005707>
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning*. Heinemann Portsmouth, NH.
- Gilquin, G. (2015). From design to collection of learner corpora. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 9-34). <https://doi.org/10.1017/CBO9781139649414.002>
- Goldin-Meadow, S. (2000). Beyond words: The Importance of gesture to researchers and learners. *Child Development*, 71(1), 231-239. <https://doi.org/10.1111/1467-8624.00138>

- Goulart, L., Gray, B., Staples, S., Black, A., Shelton, A., Biber, D., Egbert, J., & Wizner, S. (2020). Linguistic perspectives on register. *Annual Review of Linguistics*, 6, 435-455. <https://doi.org/10.1146/ANNUREV-LINGUISTICS-011718-012644>
- Granger, S. (2002). A Bird's-eye view of learner corpus research. Chapter overview. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 3-36). John Benjamins Publishing. <https://doi.org/10.1075/llt.6.04gra>
- Guven, C., & Islam, A. (2015). Age at migration, language proficiency, and socioeconomic outcomes: evidence from Australia. *Demography*, 52(2), 513-542. <https://doi.org/10.1007/s13524-015-0373-6>
- Hall, J. K. (2022). L2 classroom interaction and its links to L2 learners' developing L2 linguistic repertoires: A research agenda. *Language Teaching*, 55(1), 100-115. <https://doi.org/10.1017/S0261444820000397>
- Hall, J. K., Wang, T., & Khor, S. Y. (2020). The links between the linguistic designs of L2 teacher questions and the student responses they engender. *Iranian Journal of Language Teaching Research*, 8(3), 25-40.
- Hardy, I., Decristan, J., & Klieme, E. (2019). Adaptive teaching in research on learning and instruction. *Journal for educational research online*, 11, 169-191. <https://doi.org/10.25656/01:18004>
- Helmke, A., Helmke, T., Heyne, N., Hosenfeld, A., Kleinbub, I., Schrader, F.-W., & Wagner, W. (2007). Erfassung, Bewertung und Verbesserung des Grundschulunterrichts: Forschungsstand, Probleme und Perspektiven [Assessment, evaluation and improvement of primary school teaching: State of research, problems and perspectives]. In K. Möller, P. Hanke, C. Beinbrech, A. K. Hein, T. Kleickmann, & R. Schages (Eds.), *Qualität von Grundschulunterricht. Entwickeln, erfassen und bewerten* (pp. 17-34). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90755-0_2
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G., & Schröder, K. (2008). Die Videostudie des Englischunterrichts [The video study of English teaching]. In E. Klieme (Ed.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (pp. 345-363). Beltz.

- Hirschmann, H., & Schmidt, T. (2022). Gesprochene Lernerkorpora: Methodisch-technische Aspekte der Erhebung, Erschließung und Nutzung [Spoken learner corpora: Methodological-technical aspects of collection, indexing and use]. *Zeitschrift für Germanistische Linguistik*.
- Hochman, O., & Davidov, E. (2014). Relations between second-language proficiency and national identification: The case of immigrants in Germany. *European Sociological Review*, 30(3), 344-359. <https://doi.org/10.1093/esr/jcu043>
- Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, 137(2), 297-315. <https://doi.org/10.1037/a0022128>
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43(3), 325-356. <https://doi.org/10.1080/0305764X.2013.786024>
- Johnston, P., Dozier, C., & Smit, J. (2016). How language supports adaptive teaching through a responsive learning culture. *Theory Into Practice*, 55(3), 189-196. <https://doi.org/10.1080/00405841.2016.1173992>
- Kasper, G., & Rose, K. R. (2019). Pragmatics and SLA. *Annual Review of Applied Linguistics*, 19, 81-104. <https://doi.org/10.1017/S0267190599190056>
- Klieme, E., McCaffrey, D., & Bell, C. (2020). *OECD Global Teaching InSights / TALIS Video Study policy report*.
- Koch, P., & Oesterreicher, W. (1985). Sprache der Nähe - Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte [Language of Proximity - Language of Distance. Orality and writing in the field of tension between language theory and language history]. In *Romanisches Jahrbuch* (Vol. 36, pp. 15-43). de Gruyter.
- Köhne, J., Kronenwerth, S., Redder, A., Schuth, E., & Weinert, S. (2015). Bildungssprachlicher Wortschatz—linguistische und psychologische Fundierung und Itementwicklung [Educational vocabulary - linguistic and psychological foundations and item development]. *Forschungsinitiative Sprachdiagnostik und Sprachförderung—Ergebnisse 1, 1*, 67-92.
- König, J., Lammerding, S., Nold, G., Rohde, A., Strauß, S., & Tachtsoglou, S. (2016). Teachers' professional knowledge for teaching English as a foreign

- language: Assessing the outcomes of teacher education. *Journal of Teacher Education*, 67(4), 320-337. <https://doi.org/10.1177/0022487116644956>
- Konzett-Firth, C. (2020). Co-adaptation processes in plenary teacher-student talk and the development of L2 interactional competence. *Classroom Discourse*, 11(3), 209-228. <https://doi.org/10.1080/19463014.2019.1597744>
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, 54(1), 79-117. <https://doi.org/10.1111/j.1467-9922.2004.00249.x>
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8), 707-710.
- Lim, F. V. (2019). Investigating intersemiosis: a systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse. *Visual Communication*, 20(1), 34-58. <https://doi.org/10.1177/1470357218820695>
- Lintunen, P., & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: Comparing the incomparable? *Research in Language*, 12(4), 377-399. <https://doi.org/10.1515/rela-2015-0005>
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285-329. <https://doi.org/10.1017/S0261444818000125>
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4(2), 126-141. <https://doi.org/10.1093/applin/4.2.126>
- Lotz, M. (2016). *Kognitive Aktivierung im Leseunterricht der Grundschule: Eine Videostudie zur Gestaltung und Qualität von Leseübungen im ersten Schuljahr* [Cognitive activation in primary school reading lessons: A video study on the design and quality of reading exercises in the first school year]. Springer-Verlag.

- Lotz, M., Berner, N. E., & Gabriel, K. (2013). Auswertung der PERLE-Videostudien und Überblick über die Beobachtungsinstrumente [Evaluation of the PERLE video studies and overview of the observation instruments]. In F. Lipowsky & G. Faust (Eds.), *Dokumentation der Erhebungsinstrumente des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE)*. 3. *Technischer Bericht zu den PERLE-Videostudien* (1 ed., pp. 83-103).
- Lüdeling, A. (2008). Mehrdeutigkeiten und Kategorisierung: Probleme bei der Annotation von Lernerkorpora [Ambiguities and categorisation: Problems in annotating learner corpora]. In M. Walter & P. Grommes (Eds.), *Fortgeschrittene Lernervarietäten: Korpuslinguistik und Zweispracherwerbsforschung* (pp. 119-140). Max Niemeyer Verlag. <http://www.phonetik.uni-freiburg.de/leap/corpus.html>.
- Lüdeling, A., Walter, M., Kroymann, E., & Adolphs, P. (2005). Multi-level error annotation in learner corpora. Proceedings of corpus linguistics, Birmingham.
- Martinovic, B., Tubergen, F., & Maas, I. (2009). Dynamics of interethnic contact: A panel study of immigrants in the Netherlands. *European Sociological Review*, 25. <https://doi.org/10.1093/esr/jcn049>
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought* (Vol. 27). University of Chicago Press. <https://doi.org/10.2307/1576015>
- McNeill, D. (2008). *Gesture and thought*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226514642.001.0001>
- Mercer, N. (2010). The analysis of classroom talk: methods and methodologies. *The British journal of educational psychology*, 80(1), 1-14. <https://doi.org/10.1348/000709909X479853>
- Mercer, N., & Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxford Review of Education*, 40(4), 430-445. <https://doi.org/10.1080/03054985.2014.934087>
- Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching Chapter overview. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 119-142). John Benjamins Publishing. <https://www.ebsco.com/terms-of-use>

- Meurers, D. (2015). Learner corpora and natural language processing core issues, notions, and methods. In *The Cambridge handbook of learner corpus research* (pp. 537-566).
- Nathan, M. J., & Kim, S. (2009). Regulation of teacher elicitations in the mathematics classroom. *Cognition and Instruction, 27*(2), 91-120.
<https://doi.org/10.1080/07370000902797304>
- Ohta, A. S. (2005). Interlanguage pragmatics in the zone of proximal development. *System, 33*(3), 503-517. <https://doi.org/10.1016/j.system.2005.06.001>
- Otto, J., Migas, K., Austermann, N., & Bos, W. (2016). *Integration neu zugewanderter Kinder und Jugendlicher ohne Deutschkenntnisse. Möglichkeiten, Herausforderungen und Perspektiven* [Integration of newly immigrated children and young people without German language skills. Possibilities, challenges and perspectives]. Waxmann.
- Pallotti, G. (2019). An approach to assessing the linguistic difficulty of tasks. *Journal of the European Second Language Association, 3*(1), 58-70.
<https://doi.org/10.22599/jesla.61>
- Paquot, M., & Plonsky, L. (2017). Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research, 3*, 61-94. <https://doi.org/10.1075/ijlcr.3.1.03paq>
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research, 88*(2), 205-242.
<https://doi.org/10.3102/0034654317743198>
- Prediger, S., & Buró, R. (2021). Fifty ways to work with students' diverse abilities? A video study on inclusive teaching practices in secondary mathematics classrooms. *International Journal of Inclusive Education, 1*-20.
<https://doi.org/10.1080/13603116.2021.1925361>
- Sahlender, M. & Schrader, J. (2017). Spracherwerb bei Zugewanderten und funktionalen Analphabeten [Language acquisition among immigrants and functional illiterates]. *Zeitschrift für Weiterbildungsforschung, 40*(3), 275-293.
<https://doi.org/10.1007/s40955-017-0098-9>

- Salaberry, M. R., Kunitz, S., Sandlund, E., & Sundqvist, P. (2019). Doing versus assessing interactional competence. In M. R. Salaberry & S. Kunitz (Eds.), *Teaching and Testing L2 Interactional Competence: Bridging Theory and Practice*. Routledge.
- Sato, R. (2020). Gestures in EFL classroom: Their relations with complexity, accuracy, and fluency in EFL teachers' L2 utterances. *System*, *89*, 102215. <https://doi.org/10.1016/j.system.2020.102215>
- Schmidt, T., Schütte, W., & Winterscheid, J. (2015). *cGAT. Konventionen für das computergestützte Transkribieren in Anlehnung an das Gesprächsanalytische Transkriptionssystem 2 (GAT2)*.
- Schmidt, T., & Wörner, K. (2014). EXMARaLDA. In J. Durand, U. Gut, & G. Kristoffersen (Eds.), *Handbook on Corpus Phonology* (pp. 402-419). Oxford University Press.
- Seidlhofer, B. (2002). Pedagogy and local learner corpora Working with learning-driven data Chapter overview. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 213-234). John Benjamins Publishing. <https://doi.org/10.1075/LLLT.6.14SEI>
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209-232. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Shrum, J. L. (1985). Wait-time and the use of target or native languages. *Foreign Language Annals*, *18*(4), 305-314. <https://doi.org/10.1111/j.1944-9720.1985.tb01808.x>
- Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in Psychology*, *10*, 2366-2366. <https://doi.org/10.3389/fpsyg.2019.02366>
- Smith, L., & King, J. (2017). A dynamic systems approach to wait time in the second language classroom. *System*, *68*, 1-14. <https://doi.org/10.1016/j.system.2017.05.005>

- Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55(4), 661-699. <https://doi.org/10.1111/j.0023-8333.2005.00320.x>
- Sykes, J. M., & Cohen, A. D. (2018). Strategies and interlanguage pragmatics: Explicit and comprehensive. *Studies in Second Language Learning and Teaching*, 8(2 Special Issue), 381-402. <https://doi.org/10.14746/sslit.2018.8.2.9>
- Tarone, E. (1998). Research on interlanguage variation: implications for language testing. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between Second Language Acquisition and Language Testing Research* (pp. 71-89). Cambridge University Press.
- Thoms, J. J. (2012). Classroom discourse in foreign language classrooms: A review of the literature. *Foreign Language Annals*, 45, 8-27. <https://doi.org/10.1111/J.1944-9720.2012.01177.X>
- Tissot, A., Croisier, J., Pietrantuono, G., Baier, A., Ninke, L., Rother, N., & Babka von Gostomski, C. (2019). Zwischenbericht I zum Forschungsprojekt "Evaluation der Integrationskurse (Evlk)": erste Analysen und Erkenntnisse [Interim report I on the re-search project "Evaluation of the integration courses (Evlk)": main results] (Bd. 33). SSOAR, GESIS – Leibniz-Institut für Sozialwissenschaften e.V; Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ).
- Tobin, K. (1987). The role of wait time in higher cognitive level learning. *Review of Educational Research*, 57, 69-95. <https://doi.org/10.2307/1170357>
- Tomlinson, C., Brighton, C., Hertberg, H. L., Callahan, C., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *Journal for the Education of the Gifted*, 27, 119 - 145. <https://doi.org/10.1177/016235320302700203>
- Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28(2), 187-204. [https://doi.org/10.1016/S0361-476X\(02\)00007-3](https://doi.org/10.1016/S0361-476X(02)00007-3)
- Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.

- Wang, W., & Loewen, S. (2015). Nonverbal behavior and corrective feedback in nine ESL university-level classrooms. *Language Teaching Research*, 20(4), 459-478. <https://doi.org/10.1177/1362168815577239>
- Weiss, Z. L., Lange-Schubert, K., Geist, B., & Meurers, D. (in press). Sprachliche Komplexität im Unterricht: Eine computerlinguistische Analyse der gesprochenen Sprache von Lehrenden und Lernenden im naturwissenschaftlichen Unterricht in der Primar- und Sekundarstufe [Linguistic complexity in the classroom: A computational-linguistic analysis of the spoken language of teachers and learners in science education in primary and secondary schools].
- Wisniewski, K. (2022). Gesprochene Lernerkorpora des Deutschen [Spoken German learner corpora. An overview]. *Zeitschrift für Germanistische Linguistik*, 50(1), 1–35. <https://doi.org/10.1515/zgl-2022-2047>

9. Supplementary Material

COLD Corpus - Transcription and Annotation Guidelines

1. Structure of annotation tiers	112
a. Identification of learners	115
b. Division into segments/events	115
2. Verbal content	116
a. Verbal content to transcribe	116
b. Unintelligible content	116
c. Spelling	116
d. Numbers or single letters	116
e. Single words and filler words	117
f. Punctuation	117
g. Speech pause	118
h. Word and sentence break	118
i. Foreign language utterances	119
j. Pronunciation errors	120
k. Linguistic change	121
l. Read aloud	121
m. Audio material	123
n. Indirect speech	123
3. Nonverbal actions	123
a. Calls	123
b. Hand-raising	125
c. Gestures	127
4. Target	129
a. Target switching	129
b. Alignment of target tier to speaker tier	130
5. Coding of the social form	131
6. Other anomalies	132
7. Verification	132

1. Structure of annotation tiers

- A dummy EXMARaLDA file (COLD_dummy.exb) with preset tiers according to these guidelines is available (https://osf.io/qkct4/?view_only=d9d44194fa8a4452aad9b8ac3b8b3968). An example of the annotation tiers in EXMARaLDA for a class of five students is presented below.
- It is important to be systematic when defining the category and display name. In particular, attention should be paid to case sensitivity and type of parentheses.
- The display name of the tiers must be uniform (pay attention to the category in the case of tier properties)

The verbal tier

- Has the category **v**
- The display name of the teacher is **LK [v]**.
- The display name of the learner is **08 [v]** (here as an example for learner 8).
- The display name for some learners is **Some [v]**.
- The display name for all learners is **All [v]**.
- The display name for unidentifiable learners is **Unk [v]**.
- The display name for the camera assistants is **CA [v]**.
- The display name for an audio is **Audio [v]**.
- **X [v]**: all verbal utterances of speaker X (for the teacher the abbreviation is LK; for learners respectively the two-digit number of their ID, e.g., 01, 02, 03...).
- **Unk [v]**: contains verbal utterances of an unidentifiable learner (unknown).
- **Some [v]**: contains verbal utterances of several learners (but not all).
- **All [v]**: contains verbal utterances of (almost) all learners.
- **Audio [v]**: contains verbal utterances from audio recordings in the teaching session, e.g., from a video recorder or a CD.

The nonverbal tier

- Has the category **nv**.
- The display name of the teacher is **LK [nv]**.
- The display name of the learner is **08 [nv]** (here as an example for learner 8).
- The display name for some learners is **Some [nv]**.
- The display name for all learners is **All [nv]**.
- The display name for unidentifiable learner is **Unk [nv]**.
- The display name for the Videographer is **CA [nv]**.
- The display name for an audio is **Audio [nv]**.
- **X [nv]**: all non-verbal utterances of the speaker X.
- **Unk [nv]**: contains non-verbal utterances of an unidentifiable learner; this learner is not differentiated by gender.
- **Some [nv]**: contains non-verbal utterances of several learners (but not all).
- **All [nv]**: contains non-verbal utterances of (almost) all learners.
- **Audio [nv]**: contains sounds from audio recordings in the teaching session (e.g., bang).

Comment
Social
NN
LK [v]
LK [nv]
LK [target]
01 [v]
01 [nv]
01 [target]
02 [v]
02 [nv]
02 [target]
03 [v]
03 [nv]
03 [target]
04 [v]
04 [nv]
04 [target]
05 [v]
05 [nv]
05 [target]
CA [v]
CA [nv]
CA [target]
Audio [v]
Audio [nv]
Audio [target]
All [v]
All [nv]
All [target]
Some [v]
Some [nv]
Some [target]
Unk [v]
Unk [nv]
Unk [target]

The target tier

- Has the category **target**.
- The display name of the teacher is **LK [target]**.
- The display name of the learners is **08 [target]** here as an example for learner 8).
- The display name for some learners is **Some [target]**.
- The display name for all learners is **All [target]**.
- The display name for unidentifiable learners is **Unk [target]**.
- The display name for the camera assistants is **CA [target]**.
- The display name for an audio is **Audio [target]**.
- **X [target]**: ID of the speech target of X's verbal or nonverbal utterance.

The commentary tier

- Has no category.
- The display name is **Comment**.
- **Comment**: Comments of the transcribers concerning uncertainties and open questions (e.g., as an indication for the statement "mhm": "affirmative").

The social tier

- Has no category.
- The display name is **Social**.
- **Social**: Social form of teaching (e.g., group work; see section 5).

The pause tier

- Has no category.
- The display name is **NN**.
- **NN**: contains pauses.

a. Identification of learners

- Use the seating plan of the class to identify the learners.
- Name learners with last two ID digits only (Important: Name complete transcription file with complete class ID!).
- Use the transcription tiers only for actively participating or involved persons.
 - Important: Active participation can also take place non-verbally or as a reaction (e.g., nodding).
 - Delete unused learner tiers from the EXMARaLDA-Dummy file (right click on the tier name -> remove tier).
- Do not transcribe names of participants!
 - When the teacher's name is mentioned, use the abbreviation LK (e.g., "Mrs. Maier" → "Mrs. LK").
 - When a learner's name is mentioned, use the ID number (e.g., "And you are now working with 07.")
 - When someone is talking about a third person/another teacher, systematically replace the name with a number (99, 98, 97, ...).
- Learners who have not given their consent to the videography are positioned outside the field of view of the camera and are not recorded. Their utterances and hand-raising are not transcribed. If these learners are addressed by the teacher, they are named TNX in the (non-)verbal or target tier of the teacher.
- If information is verbalized in the lesson that allows the exact identification of the learning group or a learner (e.g., name of the school), this information is replaced by ((anonymized)).

b. Division into segments/events

- Even if the teacher makes longer utterances: divide the speech into smaller segments for easy transcription. Recommendation: Use segments of approx. 4 seconds.
- For verbal tiers: Always add a space at the end of a segment (also after a period or word), unless this would separate a word because it continues in the next segment (e.g., [partici] [pant]).

2. Verbal content

a. Verbal content to transcribe

- All utterances of the teacher must be transcribed. If they are completely unintelligible: Note this (see following section on unintelligible utterances), code the target and describe the action in the non-verbal tier (e.g., when the teacher corrects texts of the learners and speaks the text quietly to himself).
- In case of plenary lesson phase, all intelligible utterances of learners are transcribed.
- In the case of group or partner work, transcribe only the conversations of the learner of the group with which the teacher is present or which is receiving attention from the teacher at that moment.
- Completely unintelligible side conversations of learners are not transcribed.

b. Unintelligible content

- Unintelligible parts in the utterances are transcribed as ((unintelligible)).
- If ((unintelligible)) appears in the segment, mark the end of the utterance with a full stop at the next possible position, because utterances with ((unintelligible)) cannot be analyzed automatically.
- If conversations become completely unintelligible, it is sufficient to simply stop transcribing them. Please do not note every further statement of the individual learners with ((unintelligible)).
- If the unintelligible content is accompanied by relevant non-verbal behavior, it is to be annotated in the nonverbal tier (see section 3).

c. Spelling

- The correct spelling is always used.
- Orientation to the dictionary, look it up in case of doubt.
 - Example: Transcribe "doin'" as "doing".
 - Example: Transcribe "no probs" as "no problem".
- Do not correct wrong articles, order of words or wrong declension/conjugation.
- Case sensitivity has to be considered.

d. Numbers or single letters

- Spoken numbers are generally written out (e.g., 18 is written as "eighteen", in order to avoid confusion with the learner IDs).

- Single letters are always capitalized (e.g., “Here you have forgotten the N again.”).
 - If several letters follow each other, they are separated by commas (e.g., “H, O, U, S, E” for single spelling of the word “house”).
- e. Single words and filler words
- Single words at the beginning or at the end of a sentence are always followed by a full stop.
 - No comma is placed between multiple or repetitive single words in a sentence (e.g., “I want to go um um um school.”) unless followed by a subordinate clause.
 - Filler words/utterances on vowels (Ah, Uh, Um) etc. are transcribed in the verbal tier and if necessary, described in the commentary tier (agreeing, surprised, ...). Do not use letter repetitions (~~ahhhhh~~).
 - Use the following standardized encodings for filler words (this is a translated example of German filler words that might be adapted to the subtleties of the target language):
 - Uh, um, uh, uh, uh, or similar → transcribe as um
 - Oh, Oha, Oh ha (astonished) → Oh
 - Okay, Oke, Ok → Okay
 - Hm, Hmmm (pondering, hesitating, questioning etc.) → Hm
 - Mhm (agreement) → Mhm
 - Ah (cognition) → Ah
 - Ah so, Ahso, Oh so → Ohso
 - Ah → Ah
 - No, Nay, Ne (negative, rejecting) → No
 - No? → No?
 - Yes, Yap, Yup, Yep, Jo (agreement or similar) → Yes (note exception: other words that are established words according to an adequate dictionary, e.g., “Yeah”, are also transcribed as these words themselves)
 - Further unintelligible expressions (e.g., Urgh) are transcribed as ((unintelligible)) and, if necessary, described in the commentary tier. Importantly, the unintelligible part is then followed by a full stop (see guidelines for unintelligible content).
 - If a gap to be filled in a text/sentence is verbally pronounced by unintelligible content (e.g., “hm hm hm” or “düdüdüm”), this is transcribed as ((gap filler)) (e.g., “I ((gap filler)) home.”; there is a gap in the sentence because the verb *to go* is to be conjugated).
- f. Punctuation
- Use a full stop at the end of each sentence. Place full stop only at the end of sentences, not at the end of each segment.

- A full stop also follows ((unintelligible)) at the end of a sentence (for broken sentences: see section Sentence Break).
- Single and filler words (e.g., "mhm") are also separated by full stops (e.g., "Yes. That's clear. Mhm.")
- In general, place meaningful full stops at the end of sentences even if they are not always audible due to a lack of speech melody and long word strings.
- In general, the distinction between full stops and exclamation marks is not important. In case of doubt, use a full stop.
- Colons are inserted if they are followed by literal speech or individual words in form of indirect speech (e.g. "We have the question: where."). However, quotation marks are not used (see section 2.n on indirect speech).
- Always use mandatory commas, but do not set optional commas.

g. Speech pause

- If the duration of the speech pause is < 0.2 seconds:
 - Note pause as (.) in the verbal utterance tier of speakers, tied without space to the previous word/character (e.g., "This is a sign.(.)" or "This is a sign,(.) but in green." or "This is(.) a sign.").
 - If the pause occurs within a word: bana(.)na (without space).
 - If the pause occurs between two turns (change of speaker), it is transcribed at the end of the first speaker's segment.
- If the duration of the speech pause is ≥ 0.2 seconds:
 - Please use the pause function in EXMARaLDA (Event \rightarrow Insert Pause). Insert it into the NN tier so that the exact duration is calculated (e.g., (0.2) for 0.2 sec.).
 - Insert the pause exactly in time between the spoken words.
- Breathing, throat clearing, laughing, and coughing are not transcribed, but counted as pauses and transcribed in the nonverbal tier if necessary.

h. Word and sentence break

- Word breaks and sentence breaks are encoded with different characters.
- Word and sentence breaks can also occur simultaneously. In this case, first code the word break and then the sentence break.
 - e.g., if the word is broken off and then a new sentence begins: "The side/... Does everyone know what the sidewalk is?".
 - e.g., when someone starts a word and then is interrupted

Word break

- Aborted words are noted with / (e.g., if someone wants to say "banana" but only speaks "bana" and then aborts, write "bana/").
- Distinction between broken words and words that are not pronounced quite correctly: Words are only considered broken if
 - no more words follow in the sentence (Exception: speaker breaks off word and corrects himself directly afterwards).
 - another person speaks afterwards, or something happens that explains the interruption (e.g., disturbance by noise/other learners).
 - more than one syllable is spoken (only a single syllable, e.g., "Uh" or "Tr" is not sufficient and will not be transcribed, unless it is a complete single word or a didactically intended word break).
- In the case of a didactically intended word break, where the teacher begins a word and the learners are to complete the ending, the broken word is ended with a hyphen and the corresponding word parts of the answer are begun with a hyphen (e.g., teacher: "side- " learner: "-walk.").

Sentence break

- Aborted sentences are ended with three periods (...) with no spaces between the word and the "...".
- What is relevant for sentence termination is whether the syntax is complete. Thus, an aborted sentence is any utterance that cannot stand on its own, because there is incomplete syntax due to a missing finite verb or one of its necessary arguments such as subject/object.
 - e.g., "I like..."
 - e.g., "Um. You said... "
 - e.g., "When it's someone's birthday..."
- No sentence break is marked if
 - the sentence is only partially broken off (e.g., by correcting "He owes her some, everything.").
 - the broken sentence can stand alone (e.g., "I think. No. Never mind." Instead of "I think... No. Never mind." or e.g., "I am.").
 - there is only a mere pause in speech between parts of sentences (e.g., "I go (1.7 seconds speech pause) to school.").
 - single words stand alone (e.g., "I.").

i. Foreign language utterances

- If the used language is unknown or not understandable:
 - If the utterance is clearly a foreign language, transcribe the utterance as ((foreign language utterance)) into the speaker's audio tier.
 - If it is not clear whether the utterance is foreign or German, transcribe the utterance normally like (unintelligible) German utterances.
- If the used foreign language is known (e.g., English):

- If the language is English, transcribe the utterance in the spoken English language in the verbal tier and add "((foreign language utterance in English))" in the commentary tier.
- If other known foreign languages are used, transcribe translation in the verbal tier in double parentheses and add "((foreign language utterance in XX [language]))" in the commentary tier.

Comment	((foreign language utterance))	((foreign language utterance in Arabic))
Social		
NN		
LK [v]	((foreign language utterance)).	((Where are)) my scissors?
LK [nv]		
LK [target]	05	07
01.5.1		

j. Pronunciation errors

- A pronunciation error is a non-standard phonological pronunciation of a word.
- Pronunciation errors are marked if they are relevant for communication, i.e. if the dialogue is no longer comprehensible due to the standardized transcription, because, for example, although there are different pronunciations of the same word by the teacher and the learner, the words would be written the same way in the transcript. The pronunciation error becomes relevant for communication when the learner mispronounces a word that the teacher (or possibly a learner) responds to and
 - refers to the error,
 - corrects the error himself (e.g., stressed and/or correct pronunciation of the word),
 - asks the learner to correct him/herself.
- If it is clear which word the learner or the teacher wants to pronounce, for instance by repetition of another person or hints from the teaching material or the context, the word is written correctly and marked with a * after the word. For instance, if the learner pronounces "praticipan" but means "participant" and this becomes clear during the conversation, describe the heard pronunciation in the commentary tier.
 - **Example:** A learner reads a poem to the whole class. He reads: "Here I bret." The teacher correctly recodes the error and additionally stresses the word: "Here I breathe."
 - In the learner's tier, the sentence is transcribed as, "Here I breathe*." In commentary tier, transcribe: "Heard pronunciation: bret".
- Grammatical, syntactical or contextual errors are not marked and are transcribed as spoken (even if this results in new words, e.g., "fighted" instead of "fought").

k. Linguistic change

- Linguistic change is a change in the linguistic variety used (i.e. a change between everyday language and educational language).
 - **Everyday language** is context-bound and refers to the *here* and *now*. It is characterized by incomplete, short, simple (main) sentences, deictic expressions (gestures, supported by adverbs such as *here*, *there*, *there*), repetitions, general and unspecific verbs (e.g., *to be*, *to do*) as well as figures of speech/filler words (e.g., *or so*, *thingy*).
 - **Example:** In chemistry, a learner says, "Now let's dump this in there." This sentence is only understandable for everyone because everyone sees what is happening.
 - **Educational/academic language** is specialized, abstracted language with (precise) technical terms and headings. It includes complex and complete sentence structures, few to no repetitions, filler words or mental leaps.
 - **Example:** After conducting the experiment in chemistry, a learner describes the procedure as follows: "The solution is poured into the flask."
- The linguistic change is noted in the commentary tier as "VW" (Variety Change).
- Language change is only marked if it is **relevant to communication**, for instance if the trainer switches from everyday language (e.g., dialect) to educational/academic language after a learner has had difficulty understanding. We understand "comprehension difficulties" as situations in which the learners do not know what the teacher meant and the teacher repeats what was said, for instance with different words. The linguistic change is therefore relevant to communication when the teacher uses the linguistic change to explain the lesson content to the learners in a more comprehensible way.

l. Read aloud

- Transcribe read aloud in quotation marks, for instance
 - when participants reference the board pictures (e.g., *It does say "jacket."* when "jacket" is written on the board **and** this word is **read off**).
 - when learners read texts aloud (e.g., *"The dignity of the human person is inviolable."*).
 - in correction situations of the teacher with individual learners (e.g., *What did you write here? "Car?"* when "car" is read from the notes of the learner).
 - If the learners **read out (parts of) self-produced sentences**, which they have written by themselves in individual or group work, this read aloud is also written in quotation marks and noted in the commentary tier with the following sign: **(EP)**.
 - It is not annotated as self-produced (EP), if

- completely written words/sentences were written off without any self-production in terms of inflection of sentence components.
 - words dictated by the teacher or heard from an audio were written down (e.g., a listening comprehension task requires to fill in gaps in a cloze text that is read out).
 - However, in the case of mixed forms between self-production and given texts an utterance is only annotated with (EP) if less than half of the words in the utterance were given verbatim, i.e. at least 50% of the utterance contains a learner's own production.
- Dealing with breaks/interruptions:
 - For speech-only pauses in the sentence being read aloud, quotation marks are placed only at the beginning and end (e.g., "I want (1.0) lunch.").
 - For interruptions/speaker changes, quotation marks are placed at the beginning and end of each sequence (e.g., "I want" (interruption) "lunch.").
 - Filler words between what is read aloud are not put in quotation marks (e.g., "*My mother*" *um* "*buys*" *um* "*potatoes*". If the following sentence is to be read aloud: "*My mother buys potatoes.*").
- Dealing with **verbal use** of given sentence components:
 - General approach: If the sentence formation requires the learner's own **verbal** effort beyond the mere reading aloud, **no** quotation marks are placed in the sentence.
 - If several sentence components that are to be primarily inflected are given in writing, they are not enclosed in quotation marks when used **verbally**, since this is where the boundaries between reading aloud and self-production become blurred (e.g., a given list of nouns and a given list of adjectives from which combinations are to be formed: Prince, handsome → the handsome prince). A sorting of completely given sentence components alone does not count as self-production; it only counts as self-production if an inflection of sentence components takes place.
 - If a cloze text is given, in which the learners have to decline or conjugate individual words, in the case of oral formulation neither the given words nor the completed sentences are put in quotation marks (since it is the learners' own production). However, if the formulation is done in writing and the result is read out, quotation marks are used and (EP) has to be annotated.
 - Exception: If a single prompt is given in writing, it is enclosed in quotation marks (e.g., if learners are asked to complete the sentence "In my dream city" (...) from the board).
 - If no decision possible, write two quotation marks (""") in the commentary tier so that these cases can be identified later on.

m. Audio material

- If audio material is used in the lesson (e.g., a CD for listening comprehension tasks), transcribe the audio text to the tier audio [v].
- Transcribe noise from audio material in the tier audio [nv].
- Audio can be written down as one text in the audio tier.
- If there are multiple speakers in the audio, this can be ignored.
- Speech pauses in audio material are not annotated.
- Melodies in the audio without linguistic content can also be ignored.

n. Indirect speech

- Will not be noted further, but transcribed like verbal speech.
- Indirect speech = When a linguistic object means something other than what the wording says, for instance, when talking about words in a continuous text (e.g., "If we have dog here, what article comes before it?" Dog is indirect speech and is not marked; or "On is a preposition." On is indirect speech and is not marked).

3. Nonverbal actions

- Transcribe calls of the teacher (see section on Calls).
- Transcribe other non-verbal behavior (e.g., movements of the teacher in the classroom) only if it is relevant to what is happening in class, for instance because it leads to changed speech behavior or the verbal content cannot be understood without describing the spatial context.

Examples: It should be noted for better understanding if

- the teacher turns off the recorder.
 - the teacher writes something on the blackboard and this causes a longer pause.
 - the teacher goes back and forth between groups (name group).
 - the teacher refers to objects that are not named (e.g., verbal tier: "What does it say?" → Nonverbal tier: "the teacher points to O2's notebook.").
 - the teacher gives non-verbal impulses to the learning group (e.g., the display of questions via beamer, whereupon the teacher possibly expects hand-raising from the learners).
- No comprehensive scene description is required, i.e. do not transcribe head scratching etc.

a. Calls

- *Call = teacher gives the right to speak to the learners by "taking them on" (speech initiation)*

- A distinction is made between content and non-content calls (see Table 1; also for special category A0 for content calls **without a** speech prompt).
- Is annotated in the nonverbal tier of the teacher (**LK [nv]**).
- Is also coded if a learner has not reported or does not respond to the call.
- In the teacher tier, no distinction is made between calls without/after notification.
- A call may also happen non-verbally, by pointing or nodding.
- Each call requires a matching target alignment (also non-verbal calls).
- If both nonverbal behavior and call occur at the same time, code the call first and then describe the nonverbal behavior (e.g., "A1TN01. LK points to picture.").
- If several learners are called at the same time, note analogous to the target tier (in ascending order) and separate by commas (e.g., A2TN05, A2TN07).
- If a dialogue between the teacher and a learner occurs, the call will only be entered once at the beginning, unless the learner raises their hand again and is then called again.

- Does not count as a call:
 - Expressions of the teacher to the whole learning group, i.e. also questions to the whole group where learners have to raise their hand (e.g., "Whose turn is it?").
 - Answers of the teacher to questions of the learners.
 - Announcement by the teacher that it is now the turn of all learners in a clockwise direction. Only explicit looking or pointing at learners when it is their turn is then coded as a call with A1TNXX.
 - Instructions without a voice prompt from the learner, for instance requests from the teacher that learners should sit down or instructions to call on classmates (e.g., when the learners call on each other).

Table 1: Overview of codes for calls by teacher

Event	Code	Description/Example
Content call without speech prompt	A0TNxx (for "xx", enter the ID of the person called)	Refers to learner calls that do not require linguistic activity, but only actions with the learning content (e.g., writing, cutting, pointing, sorting, such as sort word snippets on the overhead project or write something on the board without the learners having to speak).
Call with speech prompt	A1TNxx (for "xx" enter the ID of the called person)	Refers to calls related to the content of the lessons. Code also when whispering, speaking softly and repeatedly, or when starting a new speech in a dialogue (e.g., Teacher: "Form a future tense." Learner: "...". Teacher: "Exactly, read it again." would Also code for questions to ensure understanding (e.g., "02, did you understand that?"). In group work: If the teacher turns to a group in response to a hand signal and subsequently (need not be immediately) turns to the learner, this is coded as a call.
Non-content call with speech prompt	A2TNxx (for "xx" enter the ID of the called person)	Refers to calls that relate to class leadership (e.g., for organizational reasons "Are you ready yet, Lukas?" or "How far are you?").

b. Hand-raising

- *Hand-raising = arm/hand/or finger sign of the learner to indicate that he/she wants to contribute something to the lesson.*
 - does not have to be visible to the teacher
 - the teacher does not need to have asked a question
- For each learner, the appropriate code from the table is assigned in the respective nonverbal tier.
 - If the hand-raising cannot be assigned to a specific learner: Note in commentary tier = M00
 - If both nonverbal behavior and hand-raising occur at the same time, code the hand-raising first and then describe the nonverbal behavior (e.g., "a+m. 02 falls off chair.").
 - Duration of the hand-raising is annotated by the length of the inserted segment. A new segment (Add Event) must be created to not change the length of the already existing one.
 - If a learner interrupts the hand-raising by lowering and raising his arm again, no new hand-raising is coded if the time span is maximum 2 seconds and the teacher has not asked a new question (if the time span is >2 seconds or there is a new question, a new hand-raising is coded).
 - If the question does not change, but there is a new possibility to call (e.g., because the question was answered incorrectly or insufficiently), then a new segment is coded for all hand-raising, also for learners who raised their hand continuously.

Table 2: Overview of codes for learner hand-raising

Event	Code	Description/ Example	Start/End
Hand-raising with call	m+a	The learner raises their hand and is then called by the teacher.	Start: Time of the hand-raising (the second the hand is raised). End: Time at which the hand is down again.
Hand-raising without call	m-a	The learner raises their hand and is not called.	Start: Time of the hand-raising (the second the hand is raised). End: Time at which the hand is down again.
Call without hand-raising	a-m	Teacher calls learner without learner having raised their hand. Includes calls to (non-content-related) questions of the teacher and learners' non-verbal answers such as nodding.	Start and end aligned with the (non-)verbal response of the learner (since no hand-raising is available).
Call - hand-raising not recognizable	ane	The learner is called and it is not visible whether they raised their hand.	Start and end aligned with the verbal response of the learner (since no hand-raising is available).
Hand-raising and call by other learners	atn	E.g., if the learners are to call each other.	Start: Time of the hand-raising (the second the hand is raised). End: Time at which the hand is down again.
	atn-m	If there is no hand-raising from the learner being called by another learner.	Start and end aligned with the (non-)verbal response of the learner (since no hand-raising is available).
Speaking without call	sp-	<p>Verbal utterance of a learner without a call of that learner being present:</p> <ul style="list-style-type: none"> e.g., when the learner says the answer, asks a question, or comments for everyone to hear on other learners' utterances without being called on to do so. e.g., if the teacher addresses all learners and one learner responds without having been called on. e.g., if the teacher responds to a hand-raising, but then asks a question to all learners, and the same learner responds again without being called again. Code also in group work situations (e.g., when the teacher asks a small group a question and then a learner answers unprompted). Code also if the learner raises their hand during the unprompted interjection (the sp- is noted in parentheses after m-a or m+a or atn; the duration of the interjection does not need to be noted, as it is included in the verbal tier). <p>It's not a sp- if:</p> <ul style="list-style-type: none"> the utterance is made during side conversations, for instance with seat neighbors the teacher asks rhetorical question and learners express agreement utterances by learners that serve to make themselves known (e.g., shouting "Mr. LK" or "Excuse me") are coded as a hand-raising, but aligned in time with the statement. However, statements that go beyond this (e.g., "Mrs. LK, is that correct?"), without the learner waiting to be called on, are sp-. if a learner repeatedly speaks in direct succession without being called, sp- is only coded again when the statements refer to a new question by the teacher (or another learner). Rule of thumb: If a hand-raising would be appropriate, code sp-. 	Start and end aligned with the (non-)verbal response of the learner (since no hand-raising is available).

c. Gestures

- gesture = Movement of hand or arm that serves interpersonal communication/interaction.
- For the teacher, the appropriate code from the table is assigned in the respective [gst] tier
 - Do not annotate (uncontrolled) body movements without identifiable meaning/reference (e.g. head scratching).
 - Analysis of speech unit is a speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clauses associated with either of them.
 - Gestures are coded as events. Each gesture is given a code in the [gst] tier (see the following table for gesture types) for the duration of the gesture. Only the teachers' gestures are coded.
 - If gestures appear simultaneously or overlap, create multiple tiers (gst2, gst3 etc.).
 - Each gesture must be aligned with a corresponding target or each gesture must overlap with a target.

LK [gst1]			M		D
LK [gst2]					
LK [v]		So. ((Bonjour)).			08.(.)

Gesture types are based on Sato (2020). Gestures in EFL classroom: Their relations with complexity, accuracy, and fluency in EFL teachers' L2 utterances. *System*, 89, 102215. <https://doi.org/10.1016/j.system.2020.102215>

Table 3: Overview of codes for teacher gestures

Code	Gesture type	Meaning	Possible expression in the classroom
D	Deictic	Speakers point to something or someone with a finger or the palm of their hand, indicating both concrete and abstract entities.	graphic illustrations Deictic gestures are unambiguous and involve "pointing". They do not need to be interpreted to understand them
I	Iconic	Speakers use their hand and/or arms to show images of the actual form of objects and/or actions, which are closely related to semantic content.	Raise the hand with one finger extended, accompanied by the phrase "write a sentence". Teacher makes an underlining gesture (drawing the hand through the air in a horizontal straight line in pen holding position). Clarifying/indicating numbers through gestures is indicative of an iconic gesture.
M	Metaphoric	Speakers use their hand and/or arms to show images of abstract concepts and/or ideas rather than the actual form of objects and/or actions shown by iconics.	teacher rolls her hands over each other, making a backward motion of the hands, accompanied for "Last week" . Likewise, pointing backward over one's own shoulder is a reference to "Past" Closing gesture (both hands in front of body, palms down, pulling apart from center, as if smoothing a tablecloth, accompanied by "that's it" - is a metaphor for "closure"
E	Emblems	Speakers show nonverbal acts that are understood by all members of the same cultural group. For example, putting the palm next to the ear means "I cannot hear you" in Japan, and circling the thumb and index finger means "Okay" in Western cultures.	Emblems are gestures that are uniquely (and beyond cultural circles) unambiguously understandable. (Waving, thumbs up, bowing, spreading arms to welcome, finger pointing for importance and respect, raising arms questioningly/shrugging shoulders like the emoji, bird pointing with finger to forehead).
A	Affect display	Speakers reveal emotions, such as happiness, fear, sadness, anger, distraction, and interest, especially through facial expressions.	Clearly identifiable facial expressions of the LC such as an approving smile, irritated look or similar.
H	Head movements	Speakers tilt their head while speaking to imply their incomprehension or need for more explanation to interlocutors.	Head tilts to the side, nodding, shaking.
B	Beats	Speakers move their hand with a rhythmic pulse. Typically, a beat gesture is a simple flick of the hand or movement of fingers up and down, or back and forth, following the stress peaks of speech.	Interactive processes are meant here. If it can be assumed that the teacher intends a meaning with its head movement. Pure physiological head movement, for example to orient in space or to look around, are not coded. "
99	indefinite	Speaker shows an indefinite gesture. Gesture is not clearly assignable.	e.g. teacher puts her hands together with no apparent connection to what she is saying.
ng	no gesture	no gesture recognizable, although rater nr. 2 has coded a gesture	
nd	not detected	Gesture recognizable, although rater nr. 2 has not coded a gesture	

4. Target

- *Target = the person(s) who is (are) addressed (non)verbally*
- The target is entered in the target tier as ID of the participants with two characters without a period after it (e.g., 01 or LK).
- The target has to be aligned with the corresponding event in the verbal or nonverbal tier (e.g., nonverbal calls).
- In plenary discussions, the target of the learners is usually: LK.
 - If participants specifically address other participants, the target is the ID of these participants.
 - If participants have to present something in front of the whole group, the target is: LK, all
 - The speaker him/herself is usually not the target.
- If the target person is unknown/external, analogous to verbal tier systematically insert digits for the unknown persons: 99 (98, 97, ...).
- When camera assistant is addressed: Abbreviation CA (for camera assistant).
- When learners participate without consent to the videography, so that they are not visible in the video but interact with other participants: TNX.
- If person cannot be identified because it is unclear who is being addressed: unk (for unknown).
- In case of multiple targets, specify the IDs of all persons addressed (e.g., 02, 14, 22).
 - In partner/group work, specify the individual IDs of the learners and not just the number of the group (e.g., 03, 07 instead of P2). If the teacher is standing by a group and listening to this group, the teacher automatically counts as a target for each statement of the learner to this entire group.
 - Targets are listed separated by "," and in ascending order. The teacher is always at the beginning of listings (e.g., LK, 02, 14, 22).
- In case of ambiguities such as unspecified subset of learner group or all learners, use the corresponding IDs (some, all, unk), but explicitly specify maximum plausible audience if possible.
 - Example1: If the teacher addresses several learners at the same time and out of 4 learners only two can be recognized: 02, 14, some.
 - Example2: If the teacher addresses two learners at the same time and one can be detected: 02, unk.

a. Target switching

- Targets often switch in situations where the teacher asks questions to the (whole) learning group, then calls on individual learners and then reacts to the learners' answers. An exact assignment of the targets is important for the further evaluation of the interactions.
- The statement/question of the teacher, before individual learners are named/called, usually refers to the entire group (target is all). Only afterwards the call of individual learners follows (target is the ID of the learner called).

- Feedback from the teacher in the plenary that refers to the utterance of an individual learner and is not clearly addressed either to this learner or to the entire learning group is given the target "ID of the learner, all" (see screenshot).

LK [v]	In the library. Good.	What else?	10.
LK [nv]			A1TN10
LK [target]	02, all	all	10
01 [v]			

b. Alignment of target tier to speaker tier

- Every utterance (from every speaker) must have a target, unless the entire utterance is unintelligible, in which case it does not need a target assignment.
- An event in the verbal tier must be uniquely assigned to an event in the target tier.
- Each verbal segment needs a timely aligned target segment, even if several consecutive verbal segments have the same target. Exception: If there is no interruption in the verbal tier (such as pauses), the successive target segments with the same target can be combined into one segment.

- Alignment examples:

1) Based on timelines:

If the verbal segment stretches across multiple timelines, the timelines must be merged in the target tier.

Wrong:

Library.			
05			

Correct:

Library.			
05			

2) Due to splits:

If the verbal utterances to a target are divided into more than one segment, the target tier also has to be split.

Wrong:

		(0.4)	
Yes.		Library.	
05			

Correct:

		(0.4)	
Yes.		Library.	
05		05	

5. Coding of the social form

- The social forms are recorded in the social tier once at the beginning of a social form and when the social form changes.
- Codes:
 - Plenum = PL
 - Group work = GA
 - Partner work = PA
 - Individual work = EA
- Code the social form that actually happens in class and not the one intended by the teacher, if there is a contradiction with the teachers' statements (e.g., if the teacher introduces partner work, but the learners work de facto individually, the code is EA).
- The beginning of the segment in which the social form is annotated should exactly correspond to the start of the social form.
 - Group and partner work starts as soon as the first group/pair has received its work assignment, even if the other groups still have to wait for precise instructions.
- Social forms should be aligned with a verbal utterances or pause.
 - Transition from one phase to another: Start of individual work is when "Silence" starts (after distribution of tasks).
- In the case of partner and group work, annotate the group in the nonverbal tier of each learner.
 - Partner work (pair1 = P1); (pair2 = P2), (pair3 = P3) etc.
 - Group work (Group1 = G1); (Group2 = G2), (Group3 = G3) etc.
 - Number pairs / groups from left to right from the teachers' perspective.
 - If a group of three is formed in partner work (e.g., due to an odd number of learners), the social form is nevertheless partner work (PA) and the group of three is annotated with "G" instead of "P" in the respective nonverbal tiers.
 - If both partner and group work take place in the learning group at the same time, annotate only GA in the social tier and indicate in the nonverbal tiers which learners are in partner work (P).
 - If the group/partner allocation changes during the group or partner work, the social form is annotated again in parentheses in the social track and the new group/pair number is annotated in the nonverbal tier of the changing participants. The numbering is continued in ascending order for these new group/pair formations.
 - If learners without consent for the study take part in group work (e.g., in a group of three), this group work is coded as group work, even though only two persons are identified.

- Example of the annotation of social forms:

Comment	438 [439 [1251.0] 440]
Social	G4
NN	(1.7)
LK [v]	So.
LK [inv]	
LK [target]	01, 02, 03, 04
01 [v]	
01 [inv]	01
01 [target]	
02 [v]	
02 [inv]	01
02 [target]	
03 [v]	
03 [inv]	01
03 [target]	
04 [v]	
04 [inv]	01
04 [target]	
05 [v]	
05 [inv]	02
05 [target]	
06 [v]	
06 [inv]	02
06 [target]	

6. Other anomalies

- Other conspicuous details or peculiarities about the teacher, the lesson, the lesson group or the learners, which are not apparent from the transcription itself and which require explanation in order to depict the overall picture of the lesson, might be collected in a separate file.
- Examples for conspicuousness: linguistic peculiarities of the teacher, a funny/ incomprehensible situation in the class, learners arriving too late, teachers leaving the class for a short time, another person coming into the room, etc.

7. Verification

At the end of the transcription the following points should be checked again:

- Copy verbal tiers into Word to correct spelling errors.
- We recommend having a second person listen to all unintelligible and foreign-language utterances to check whether the content is understood.
- Listen to external audio tiers for speech alignment (especially for unintelligible utterances).
- Is there a space at the end of each segment?
- Do all verbal segments have a target?
- Do all teacher calls have a target?
- Are there no names (only IDs) in the transcript?
- Are the social forms coded?
- Are all word and sentence breaks, pronunciation errors, language changes marked?

8. Beitrag #3 Do teachers adapt their gestures?

Sahlender, M. & ten Hagen, I. (2024) Do teachers adapt their gestures in linguistically heterogeneous second language teaching to learners' language proficiencies? *Ges-
ture*. Advance online Publication. <https://doi.org/10.1075/gest.22023.sah>

Dieser Artikel ist nicht die originalgetreue Kopie.

Das Original ist verfügbar unter <https://doi.org/10.1075/gest.22023.sah>

Abstract

Teachers' use of gestures in the classroom can support the language acquisition of learners in learning a second language (Stam & Tellier, 2022). Depending on learners' language skills, different dimensions of gestures (e.g., deictic, metaphorical) are considered to facilitate successful language comprehension. This study investigates which gestures teachers use in German as a second language (GSL) classrooms and to what extent teachers adapt their gestures to learners' language proficiency. Teacher gestures in 10 video-recorded integration and preparation classes were analyzed. Two coders reliably identified 4143 gestures. Results show that GSL teachers predominantly used deictic gestures, metaphorical gestures, and feedback by head movements. Moreover, between-learner variability in teachers' use of deictic and metaphorical gestures was explained by teacher-perceived German language proficiency of learners. These results suggest that teachers systematically adapt some dimensions of gestures in GSL classes, thus emphasizing the importance of studying nonverbal interactions for a better understanding of language acquisition processes.

Keywords: second language acquisition; teachers' gestures; adaptive teaching.

1. Introduction

Acquiring the local language of a country is of immense importance for the immigrant inhabitants. It promotes integration, identity development, and social participation in society (Esser, 2006; Hochman & Davidov, 2014; Martinovic et al., 2009). The significance of local language skills is shown in the strong linkage between acquired local language skills and educational success (Esser, 2006) as well as successful participation in the labor market (e.g., Bleakley & Chin, 2004; Guven & Islam, 2015). Successful second language acquisition (SLA) needs comprehensible input and explicit language instruction for learners (Alibali et al., 2013; Ellis, 2005; Krashen, 1989; Loewen & Sato, 2018; Long, 1983). One important place for SLA for immigrants is the second language classroom (Howe & Abedin, 2013; Johnston et al., 2016; Thoms, 2012). Typically, second language classrooms are highly heterogeneous in terms of learners' cultural backgrounds, their language biographies, and their educational background and educational aspirations (e.g., for classrooms of GSL: Otto et al., 2016; Tissot et al., 2019).

One essential component of effective instruction in highly heterogeneous classrooms is teachers' adaption of the input (Hardy et al., 2019; Parsons et al., 2018) to learners' Zone of Proximal Development (ZPD; Vygotsky 1978). This is the case when the linguistic content is challenging for learners by being slightly above their current level of proficiency while remaining still at an attainable level (Vygotsky, 1978). If learners within the classroom hold different zones of proximal development due to their different language proficiencies, teachers will have to adapt their instruction to the individual learners to enable the highest possible learning success in the language acquisition process for each learner. Therefore, teachers shape their communication and usually use didactic techniques such as scaffolding to support individual learners based on their needs (Hammond & Gibbons, 2005; Hardy et al., 2019). Research suggests that speakers do not only adapt their verbal communication depending on interlocutors' different language prerequisites (cf. Ferguson, 1975; Ulichny, 1979), but also underly their verbal scaffolding processes with nonverbal actions such as gestures (Alibali & Nathan, 2006; Tellier et al., 2021). In the present study, we will investigate to which extent teachers adaptively adjust their gestures to the language proficiency of their individual learners in linguistically heterogeneous classrooms.

2. The Role of Gestures in Second Language Learning

The communication part of teaching is a multimodal action (McNeill, 1992). Thus, to understand SLA processes in the classroom, it is important to look at the role of nonverbal action in addition to the linguistic utterances of teachers. Nonverbal behavior such as making gestures plays an important role in communication and the use of gestures in teaching-learning processes has received increased attention in recent SLA research (Gullberg, 2010; McCafferty & Stam, 2008; Sato, 2020; Stam & Tellier, 2022; Stam & Urbanski, 2023). Gestures are defined as movements of hands or arms that usually go together with speech (McNeill, 1992). The appearance of gestures is important for successful language acquisition, as gestures support the comprehension of spoken language (Dargue et al., 2019; Goldin-Meadow, 2000; Hostetter, 2011). Moreover, there is compelling evidence that teachers' gestures in the SLA classroom have positive effects on learning (for a review, see McCafferty & Stam, 2008). For example, the use of gestures can support the learning of new vocabulary and its meanings (Hupp & Gingras, 2016; Lazaraton, 2004; Lazaraton & Ishihara, 2005; Tellier, 2008). Overall, McCafferty and Stam (2008) as well as Stam and Tellier (2022) evaluate the intertwining of language acquisition processes with the use of gestures as substantial and point to the great potential for research in this area. To date, there are two meta-analyses on the question of whether (different) gestures significantly foster listeners' comprehension of the message. Both the meta-analysis by Hostetter (2011) with $N = 63$ studies and the work by Dargue et al. (2019) with $N = 83$ studies concluded that the use of gestures has a moderate beneficial effect on the comprehension of speech. It should be noted that it is not clear from the meta-analyses how many of the included studies took place in the language classroom. Furthermore, there is evidence that the extent to which the use of gesture supports learners' language comprehension varies between different dimensions of gestures (Dargue et al., 2019). However, it should also be mentioned that the use of gestures is not always beneficial. Tellier and Stam (2022) refer to a large number of studies on the question of how second language learners process gestures and show that various studies (Azaoui, 2016; Hauge, 1998, 1999; Sime, 2001, 2006, 2008; Tellier 2008, 2009) suggest that the use of gestures can also be a hindrance if the gestures are used in a culturally specific, too ambiguous or too symbolic way. In addition to the question of how teachers' use of gestures supports learners' language acquisition, there are also frameworks such as the

classification of teacher gestures according to different pedagogical functions (Tellier, 2008). These make it possible to consider the teacher's use of gestures according to whether they provide information about language, assess language levels or manage classroom instruction, thereby providing a framework for understanding the teacher's pedagogical actions.

2.1. Dimensions of Gestures

There are a variety of classification approaches that aim to capture gestures in teaching-learning processes (see review work by Kendon, 2010; McNeill, 1992). For the following study, we draw on a classification system of gestures according to the current and empirically tested approach of Sato (2020), which is based on the work of McNeill (1992, 2005). In addition to the four typical gesture dimensions (deictic, iconic, beats, and metaphoric), this framework considers emblems and other types of non-verbal feedback such as head movements and affect displays. The last two are not gestures but feedback channels for teachers to provide positive or negative feedback to the students. They are the most common non-verbal behaviors in teacher-learner interactions (Wang & Loewen, 2016).

In the following, we will briefly describe each dimension of gesture (see Table 1). *Deictic gestures* are hand/finger pointing movements towards concrete or abstract entities that are often very closely related to the semantic content of the language and whose meaning can be understood without complex language proficiency of the interlocutors. The pointing of teachers can support learners' language comprehension by guiding attention of the learners (Azaoui, 2013, 2020; Dargue et al., 2019). Similarly, *iconic gestures* occur when a speaker uses the hand/arm to create concrete images of the current form of an object or action, which are close to the semantic content of what is being said. For example, a teacher may clarify the word "underline" with a gesture of holding fingers in a pencil pose and drawing a horizontal line through the air. A *beat gesture* is usually a flick of the hand or a movement of the fingers up and down or back and forth that follows the stress peaks of speech, for example, in stressing syllables in vocabulary explanation. Beat gestures also appear when speakers correct themselves while speaking. *Metaphorical gestures* are those in which speakers use their hands or arms to clarify images of abstract concepts or ideas, for example, rotating the hands over each other with index fingers pointing toward each other in front of the body to signal that something should be repeated. The associated speech of the speaker "Please repeat that" is supplemented by the gesture, but not replaced.

In this case, a higher level of speech comprehension is necessary since the meaning cannot be completely deduced from the gesture. Related to metaphorical gestures are *emblems*: emblems usually also refer to metaphoric aspects. However, the gestures of emblematic meanings are characterized by the fact that they are clearly understood by all members of the same cultural group. Gullberg (2006, p. 105) describes emblems as “conventionalised language- and culture-specific movements, fixed form-meaning pairs with standards of well-formedness”. For example, bringing the thumb and index finger together to show a ring can be understood as anything from ‘OK’ or ‘good’, for interlocutors from the same cultural background to ‘worthless’, or ‘money’, or even as an insult to interlocutors from another cultural group, depending on the individual cultural background (Gullberg, 2006; see also Stam 2012; Stam & Urbanski, 2023). *Feedback by head movement* can be sometimes considered extensions of hand gestures (McNeill, 2005). They are also nonverbal behavioral ways for the teacher to imply their feedback, incomprehension, or need for more explanation to the learners. Typically, the feedback is given by nodding or shaking the head. The head is the main articulator for these responses to the students, helping to monitor the unfolding of the interaction between teacher and student. Another way the teacher can provide clear feedback to learners is through the mimicry of emotions, such as happiness, anger, sadness, distraction, and interest. This is known as *Feedback by affect display* (Wang & Loewen, 2016).

Table 1*Categorization of gestures, based on Sato (2020)*

Dimension of Gesture	Explanation/Operationalization
Deictic	Speakers point to something or someone with a finger or the palm of their hand, indicating both concrete and abstract entities.
Iconic	Speakers use their hands and/or arms to show images of the actual form of objects and/or actions, which are closely related to semantic content.
Beat	Speakers move their hands with a rhythmic pulse. Typically, a beat gesture is a simple flick of the hand or movement of fingers up and down, or back and forth, following the stress peaks of speech.
Metaphoric	Speakers use their hands and/or arms to show images of abstract concepts and/or ideas rather than the actual form of objects and/or actions shown by iconic.
Emblem	Speakers show nonverbal acts that are understood by all members of the same cultural group. For example, putting the palm next to the ear means "I cannot hear you" in Japan, and circling the thumb and index finger means "Okay" in Western cultures.
Dimension of Feedback	
Head movement	Speakers tilt their heads while speaking to imply their incomprehension or need for more explanation to interlocutors.
Affect display	Speakers reveal emotions, such as happiness, fear, sadness, anger, distraction, and interest, especially through facial expressions.
Other	
Undefined gesture	Head or arm movements that could not be clearly assigned to one of the previous gesture dimensions.

Evidence on teachers' use of these dimensions of gestures and the relative frequency of each dimension is inconsistent which might result from the different subjects studied, for example mathematics (cf. Alibali et al., 2019) vs. language instruction (cf. Wang & Loewen, 2016). Regarding SLA classrooms, Sato's (2020) study provides frequencies of gestures for the dimensions of gestures listed above from two different teachers. In terms of content, attention was paid to reading, writing, listening and speaking in both courses, although in one group the emphasis was on reading and listening and in the other more on textbook comprehension. The study shows that both teachers differ largely in the total number of gestures (122 vs. 712) made in a fifty-minute lesson. Differences were also notable in the relative frequencies of specific gesture dimensions. For example, the use of affect displays (1% vs. 24.6%), deictic (17.2 vs. 37.4%), and iconic gestures (6.9% vs. 18%) varied between the two teachers. In contrast, beats (31.1% vs. 37.7%) metaphoric (7.4% vs. 10.2%), emblems (0% vs. 3.2%), and head movements (1.6% vs. 3.2%) were similarly distributed proportionally. Whereas the frequency of some dimensions of gestures was comparable across both classrooms, there were substantial differences between other dimensions of gestures.

However, based on the study, it is not clear whether these differences are related to differences in learner characteristics or teacher characteristics, so further research is necessary to explain these differences. Also, existing findings might not be transferable to the instruction of other languages than English as a foreign language. There have been a few studies on the use of gestures in language teaching other than English, for example Azaoui (2013) in French. These data can only be compared to a limited extent due to different evaluation modalities with different gesture dimensions and/or questions. So far, there are no results on teachers' gesture use in GSL classrooms.

In addition to the subject-specific context, characteristics of the learner groups might explain the inconsistent evidence on teachers' gesture frequency. For instance, teachers might use different dimensions of gestures depending on the linguistic prerequisites of the learners in the classroom. Some evidence suggests that speakers/teachers use gestures differently often when talking either to native speakers or non-native speakers (Tellier et al., 2021). However, in these experimentally conducted studies, the level of language proficiency was only recorded dichotomously (native vs non-native speakers) and not broken down further (Tellier et al., 2021). It can be assumed that the composition of the learners in second language classrooms usually is characterized by many different language levels (Otto et al., 2016). If the teachers' use of gestures is adapted to the individual language prerequisites of the learner, the teachers' differential gesture use across learners with individual language proficiencies could play a major role, especially because some gestures, such as metaphorical gestures, require more language comprehension than, for example, deictic gestures, which are often linguistically redundant (Gullberg, 2006).

2.2. Adaption of Gestures and Language Proficiency of the Interlocutors

The issue of adaptive communication has been studied in teacher-learner research for quite some time, focusing on the analysis of verbal communication. The scholarly groundwork on linguistic adjustments that speakers make when speaking with non-native speakers has been scientifically examined by Ferguson (1975), Ulichny (1979), and Derwing (1987). The linguistic adjustments are usually discussed under the term Foreigner Talk (Tellier et al., 2021). However, the linguistic adaptations can also be explained from the perspective of the Communication Accommodation Theory (Giles et al., 1991). This theoretical approach assumes that the speaker and receiver of a message relate to each other in the communication process and speakers

consider the receivers' perspective, level of knowledge, and specific needs, with considerations usually based on stereotyped representations (Brennan & Hanna, 2009; Giles et al., 1991). Tellier et al. (2021) have argued that although this theory has been explicitly designed to cover verbal and nonverbal action, the body of studies has almost exclusively examined verbal communication adaptations, while there is a lack of research on nonverbal communication.

To our knowledge, Tellier et al. (2021) is only one study that investigated (prospective) teachers' use of gestures and its link to the learners' language proficiencies.. Tellier et al. (2021) carried out a semi-controlled quasi-experimental explanation task in which ten prospective teachers had to explain the semantic meaning of twelve words to ten non-native speakers and ten native speakers. Among other things, the authors looked at gesture dimensions, frequency of gestures, and duration of gestures. They hypothesized that the prospective "teachers would produce more illustrative gestures (iconic and deictic gestures) [when they talk to non-native speakers] to make speech more comprehensible" (Tellier et al., 2021, p. 36), as when they talk to native speakers. Additionally, they hypothesized that prospective "teachers would use fewer abstract gestures (metaphorical gestures and emblems) [to non-native speakers]" (Tellier et al., 2021, p. 36), because these gestures have few linguistic references and are more difficult to understand. The results showed that teachers produced more gestures, longer gestures, and more illustrative gestures, in the condition with non-native speakers (Tellier et al., 2021). In contrast, the findings did not confirm the hypothesis that teachers use fewer abstract gestures (metaphoric and emblem gestures). These results on gesture use are in line with the previous findings from a slightly similar study by Adams (1998), which examined gesture use in native speakers rather than teachers and employed a narration task.

Transferred to the educational practice in the classroom, these findings indicate that teachers might consider the needs and understanding of the interlocutors and therefore shape their use of different gesture dimensions. For example, in the study by Azaoui (2013) the same teacher produced the different dimensions of gestures with varying frequency in two different classes (French for native speakers vs. French for non-native speakers). This suggests that teachers might consider the language proficiency levels of the learners and adapt their gesture use to each class (Azaoui, 2013). In addition, there are other studies that have investigated teachers' gesture use in second language classrooms. For example, Smotrova and Lantolf (2013) showed how language teachers enact gestures and thus improve comprehension in the meaning

of words. Video studies in second language teaching have also provided a better understanding of the important role teachers' gestures play in teaching pronunciation (cf. Smotrova, 2017) and how reciprocal interaction between teachers and learners takes place when gesturing accompanies learning (Matsumoto & Dobs, 2017). However, these studies either exclusively looked at native vs non-native interlocutors or examined learners within a class without explicitly looking at the heterogeneous linguistic proficiencies of the learners. In both cases, the naturally occurring variability of learners' language proficiency levels in authentic language classrooms was not considered comprehensively. In particular, in GSL classrooms, all learners are non-native speakers and usually vary greatly in terms of their language proficiencies (Otto et al., 2016; Tissot et al., 2019). The link between teachers' use of gestures and the natural occurring within-class heterogeneity of language proficiency remains an open question. One explanation why teachers might consider this within-class variability of learners' abilities when designing their communication is the ZPD assumption put forth by Vygotsky (1978). It must be noted that the socio-culturally anchored concept of the ZPD, with a variety of learning support activities, includes much more than the mere provision of language support. Considering verbal communication, these linguistic adaptations usually take place by regulating the length of utterances and simplifying lexis and syntax to make the linguistic input more comprehensible for the learners.

In contrast, regarding nonverbal communication, there have been no studies to date that examine whether teachers adapt their actions to the individual language proficiencies of different learners in the same classroom. There is evidence that the potential support of different gesture dimensions (e.g., deictic or iconic gestures) depends on learners' language proficiency. This can also be deduced from the work of Azaoui (2013, 2020), who observed that one and the same teacher uses gestures differently in different classes. Additionally, it has been shown that learners with especially low language proficiency benefit from the teacher's use of gestures (Sueyoshi & Hardison, 2005). Furthermore, there is conjecture that the use of different gestures is also linked to linguistic content (see Stam, 2016, on the differential use of metaphorical and iconic gestures in oral proficiency vs. narrative tasks). Teachers therefore might use more gestures for the linguistically lower performing learners in the classroom (within classroom differences). Thus, the teachers' perception of the within-classroom differences in learners' language proficiency could be meaningful for their use of gestures to learners. Although we expect learners' language proficiency as measured by a standardized test to be fundamental information to address each learner's ZPD

successfully, it is plausible that teachers' subjectively perceived language proficiency is more relevant to the teacher's nonverbal behavior in the classroom than the objective measured language proficiency because teachers might not always be aware of test results. The differences between subjectively perceived language proficiency of teachers and objective language proficiency of learners measured by standardized tests, however, might be rather small since Urhahne and Wijnia (2021) have elaborated in a review that teachers' judgment accuracy concerning learner abilities is relatively high.

So far, we do not know whether or not the teachers' use of gestures varies between learners in the GSL classroom. Moreover, although there has been work examining teachers' gestures in authentic language classrooms,¹ e.g. in English as a Second Language (Smotrova & Lantolf 2013; Smotrova, 2017) and e.g. French as a Foreign Language (Azaoui, 2013, 2020; Tellier, 2016), to our knowledge, so far there has been no study that has looked at teachers' gestures in the teaching of GSL. In the present study, we use video data from ten GSL classrooms to get insights into what dimensions of gestures teachers use in heterogeneous classrooms. Furthermore, we investigate the link between teachers' use of gestures addressed to individual learners and (teacher-perceived) learners' German language proficiencies.

3. Present Study

One area of education in which learners might benefit differently from different gestures is German as Second Language (GSL) instruction. This is due to the typically large heterogeneity of learners' language proficiencies within these learning groups. The extent to which teachers use gestures in GSL instruction and the question whether teachers take learners' proficiency levels into account has not yet been investigated. Thus, comparable to Tellier et al. (2021), we examined the adaptation of teachers' gestures to different interlocutors. The major difference to most previous studies and the advantage of the present study design is that we could draw on authentic video data from GSL teaching situations combined with enriched information on the language proficiency of the learners. Not only did we capture the gestures teachers use during the lessons, but we also considered teacher-rated learners' German language proficiency as well as learners' performance on a standardized German test. This allowed us to empirically test adaptations of teachers' gestures towards different interlocutors.

¹ See Urbanski and Stam (2023) and Tellier and Yerian (2023) for a detailed overview of further studies.

The present study addressed the following questions: First, what dimensions of gestures do teachers use in GSL classes (RQ1)? Second, does the use of gestures vary between different teachers and between learners of a teacher's learning group (RQ2)? Therefore we wanted to investigate whether the number of gestures used by the teacher differs between several teachers. Furthermore, we wanted to investigate whether teachers address different gestures to different learners in the same classroom. Third, can the variance in the use of teachers' gestures with learners in the same classroom be explained by the (teacher-perceived) German proficiency of the learners (RQ3)?

Following the findings of prior research (Azaoui, 2020; Tellier et al., 2021), we expected that teachers would address linguistically weaker learners with fewer complex gestures (i.e., deictic and iconic gestures, which are close to the semantic content of what is being said) and present more linguistically complex gestures (i.e., metaphoric gestures and emblems, which accompany more abstract concepts) to linguistically stronger learners. Thus, we hypothesized that teacher-perceived language proficiency (extent of vocabulary, text comprehension and grammatical skills, German language aptitude) of learners negatively predicts the number of deictic gestures (RQ3-H1) and iconic gestures (RQ3-H2) a learner receives as compared to other classmates and positively predicts the number of metaphorical gestures (RQ3-H3) and emblems (RQ3-H4) a learner receives compared to other classmates. If the teacher's use of gestures is found to be adaptively aligned with perceived learner abilities, we assume that it is also aligned with the test-assessed language proficiency, since teachers' judgment accuracy concerning learner abilities is relatively high (Urhahne & Wijnia, 2021). Therefore, we hypothesized that objectively measured German language proficiency will predict the number of gestures a teacher addresses to individual learners in the same way as teacher-perceived German language proficiency (RQ3-H5).

4. Method

4.1. Sample

To address the research questions, we used the data of ten video recordings of German as second language classes from the educational sectors of school and adult education. These videos were recorded in educational institutions in the federal state of North Rhine-Westphalia, Germany, in the years 2019 and 2020. Five videos were recorded in preparatory classes (PC) in secondary schools (5th to 9th/10th grade) and five videos were recorded in integration courses (IC) from adult education

institutions. In both educational sectors, the learner groups consist of non-native speakers that have still limited knowledge of German. Both, PC and IC have a similar curriculum, as they are designed to empower newly immigrated learners in the target language of German and provide them with the basic techniques to participate in social life through the acquisition of speaking, writing, reading, and listening skills. Which of the two educational sectors the learners attend depends generally on their age (<18 years old = preparatory classes; >18 years old = integration courses). Adult learners who do not yet have written language skills (= illiterates) usually do not take part in integration courses but attend a literacy course. Two of the five videos from the adult education sector were created in a literacy course.

The present sample consists of teachers ($N = 10$) who were on average 46 years old (24-68 years) (Table 2). The majority (80.0%) were female and born in Germany (70.0%). All participating teachers had at least a master's degree or a state examination and one or more qualification certificates to teach GSL. In the school setting (PC), two teachers each participated in the study with the same group of learners. Therefore, the pilot study includes eight different classes ($N_{PC} = 3$, $N_{IC} = 5$) with a total of 82 learners (8-13 participants per class, on average 10 per class). The average age of learners in the school classes (N_{PC} learners = 33) was 13 years (10-17) and 45.5% were female. Learners in adult classes (N_{IC} learners = 49) were on average 39 years old (24-62) and 67.3% were female.

Table 2

Overview of the composition of the sample by numbers of teachers, classes and learners for preparatory classes and integration courses

	PC	IC	Total
Number of Teachers	5	5	10
Number of Classes	3	5	8
Number of Learners	33	49	82
Average Age (years)	13 (10-17)	39 (24-62)	-
Female Percentage	45.5	67.3	-

Only study participants (teachers and learners) who gave informed consent were included in the analysis. In the case of underage participants, consent was obtained from their legal guardians. Thus, teachers' gestures to participants who did not give their consent were not analyzed. The study was reviewed and approved by the

Local Ethics Committee. To anonymously link the data from questionnaires and test instruments to the video data, each participant was randomly assigned an ID. Transcripts were completely anonymized by using participants' IDs. To identify the different speaker IDs, a seating plan indicating the ID of each learner in the classroom was used.

4.2. Materials and Procedure

4.2.1. Videos / Session Structure

Teachers were asked to teach a 45-minute lesson. The content of the lesson was freely chosen by the teachers. This was necessary to increase the willingness to participate and not to interfere with the daily teaching-learning rhythm. The common focus of all ten lessons was on speaking, although the other three skills (reading, listening, writing) were also taught, but were given less time. Since the use of gestures depends on the content and task (Azaoui, 2013; Stam, 2016; Tabensky, 2008), we would like to refer here to the content descriptions of each lesson attached in the online supplementary material in Table 8. Regarding the social form, the teachers were asked to conduct at least one other social form (group work, partner work, individual work) in addition to a plenary phase (classroom teaching). The duration of the ten video-taped lessons varied from 44 to 63 minutes and was on average 48.4 minutes long. Overall, 59% of the filmed teaching time took place in plenary phases, 18% in group work, 10% in partner work, and 13% in individual work. The lesson recordings were made with two video cameras. One camera shows the teacher from the learner's perspective. Another camera filmed the entire learning group. The teachers wore a lapel microphone for the audio recordings. The learners' statements were recorded via two room microphones.

4.2.2. Annotations of Gestures and Targets

The video-taped lessons were transcribed with the EXMARaLDA transcription editor (www.exmaralda.org, last access 17 March 2024; Schmidt & Wörner, 2014). In this tool, the transcription is structured in annotation tiers, each containing events on a timeline. This compositional structure allows the transcription of spoken language and accompanying nonverbal behavior from multiple speakers simultaneously. We annotated the gestures in one, or if two gestures occurred simultaneously, in two gesture tiers. We coded the gestures following the coding system employed by Sato (2020). This coding system includes seven gesture dimensions which combine the widely

used hand gesture system of McNeill (2005; *deictic, iconic, beat, metaphoric, emblem*) with the body gestures of Wang and Loewen (2016; *feedback by head movement, feedback by affect display*). Furthermore, we added coding for unclear gesture dimensions (undefined gesture) to the category system.

The coding of the gestures was temporally aligned with the movement of the gestures. Likewise, for each gesture, we recorded to whom the gesture was directed in the target tier. When gestures were addressed to a single learner, the learner's ID was annotated. In case a gesture was addressed to several learners or even all of them, the respective IDs were noted, or "all" was coded.

4.2.3. Questionnaire Data and Language Proficiency Test

Following the video-taped lesson, teachers were asked to assess the individual German language proficiency of their learners compared to other GSL learners using four Likert-scaled single-item statements (1 = far below average, 2 = below average, 3 = rather below average, 4 = average, 5 = rather above average, 6 = above average, 7 = far above average). The items addressed the learner's level of text comprehension, the scope of vocabulary, grammatical competence, and German language aptitude. Simultaneously, learners answered a questionnaire followed by a standardized gap test (C-test; Eckes & Grotjahn, 2006; Grotjahn, 2019) to assess their German language proficiency. The test consisted of 4 texts with each 20 gaps to fill and covers German skills in the range A1-A2 of the Common European Framework of Reference for Languages. There were two texts for each level. The number of correctly completed gaps across all texts was then calculated for each learner (maximum 80 points). Thereby, we obtained an objective and comparable measurement of the learners' German language proficiency. Table 7 of the supplementary material lists the scatter and location measures (mean values, standard deviation, minimum and maximum values) of the perceived and measured linguistic abilities for two courses as examples to illustrate the degree of heterogeneity.

4.2.4. Reliability of Measurement

To assure the reliability of the gesture measurement, the videos were analyzed independently by two trained coders in an event sampling procedure. A total of 4143 gestures were identified. In 92.9% of the cases (3848 gestures), both coders identified a gesture. Interrater agreement for gesture dimension identification was sufficiently good as indicated by Cohen's kappa ($\kappa = .82$). For the final analysis, gestures that

were identified by only one coder as well as the gestures that were categorized differently ($N = 295$) were reviewed and assigned to a gesture dimension unanimously by the authors.

4.2.5. Analysis

The following analyses were calculated using SPSS Statistics Version 27. To examine what dimensions of gestures teachers use in GSL classes (RQ1), we conducted descriptive frequency analyses of gesture use for each dimension of gesture. To examine whether and how gesture use varies between different learning groups (first part of RQ2), the frequencies and distributions were considered at the teacher level. For the analysis of variability within the learning group (second part of RQ2 and RQ3), mean values per learner were calculated for each gesture dimension over the gestures individually addressed to this learner. Only gestures that were clearly addressed to an individual learner were considered and gestures addressed to multiple learners were excluded for answering these research questions. To examine whether teachers' gesture use is predicted by learners' German proficiency perceived by teachers or test-assessed (RQ3), we conducted multiple regression analyses per gesture dimension, predicting the group-mean centered frequency of the total number of gestures and each gesture separately by the group-mean centered teacher-perceived and test-assessed German language proficiency of the individual learners. Because learners who are addressed more frequently by the teacher are also likely to receive more gestures of a specific dimension compared to their learning group, we included the total number of gestures directed at a learner as a control variable in the regression analysis to capture the relative frequency of the specific gesture dimension. To estimate the effect size of perceived and measured language proficiencies on adaptive gesture use of teachers we reported ΔR^2 which indicates the additional variance explained by the respective independent variable beyond the absolute number of gestures.

There were no missing data in terms of teachers' perceived language proficiency and measured gap test scores. One learner did not take part in the gap test. We accounted for this with a listwise case exclusion.

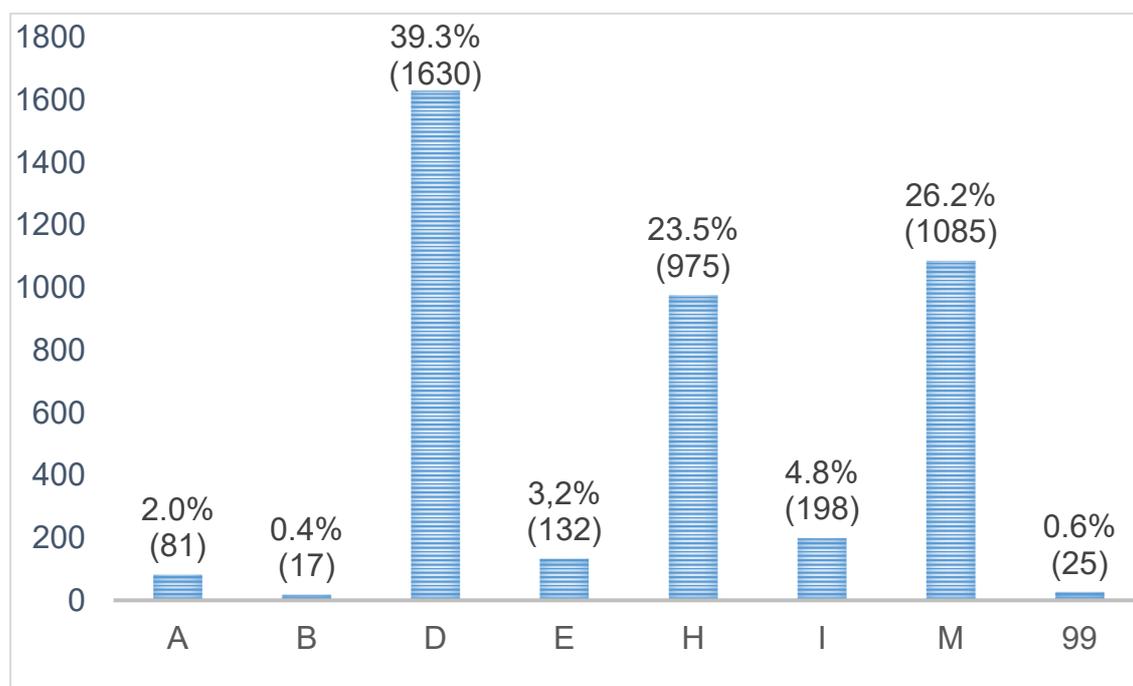
5. Results

RQ1: What dimensions of gestures do teachers use in GSL classes?

On average, teachers used 414 gestures per lesson. The number of gestures varied from 228 to 694 gestures ($SD = 141.27$) per lesson. The detailed distribution of teachers' gestures per dimension can be seen in Figure 1. The most frequent gesture dimension in the GSL lessons studied was deictic ($N = 1630$). Also frequently used by teachers were metaphorical gestures ($N = 1085$) and head movements ($N = 975$). Iconic gestures ($N = 198$), emblem gestures ($N = 132$) and beat gestures ($N = 17$), on the other hand, were used rather rarely in GSL lessons.

Figure 1

Distribution of teachers' gestures



Note. $N = 4143$. Dimension of Gesture: A = Affective Display, B = Beat, D = Deictic, E = Emblem, H = Head Movements, I = Iconic, M = Metaphoric, 99 = undefined gesture.

RQ2: Does teachers' use of gestures vary between teachers and between learners of the same learning group?

Table 3 shows the absolute and relative frequencies of gestures used by the teachers. Both the absolute number and the distribution of the different dimensions of

gestures varied across teachers. Most teachers used predominantly deictic gestures, but some teachers (PC2 and PC5) used head movements equally frequent, and some teachers (IC2 and IC3) used even more metaphorical gestures (IC2: 34.8%; IC3: 42.9%) than deictic gestures (IC2: 32.4%; IC3: 27.7%). A Chi-Square test of independence was performed to assess the relationship between teacher and dimension of gesture. There was a significant relationship between the two variables, $\chi^2(63, N = 4143) = 673.79, p = .000$, showing that the distribution of gestures across the gesture dimensions is related to the teacher. The topics and content of the ten lessons themselves were very different. For example, there were poem analyses with a lot of explanation by the teacher (in PC4 and PC5) or very conversation-intensive lessons in preparation for oral exam situations (IC3). It was also noticeable that within a class, except for PC3, all learners received the same tasks. There was hardly any internal differentiation. For further analysis of the linkage of the different frequencies of gesture use in each lesson, as listed in Table 3 in absolute figures and illustrated in Figure 2 as a percentage, there is a description of the tasks and content of the lesson studies in Table 8 in the Appendix.

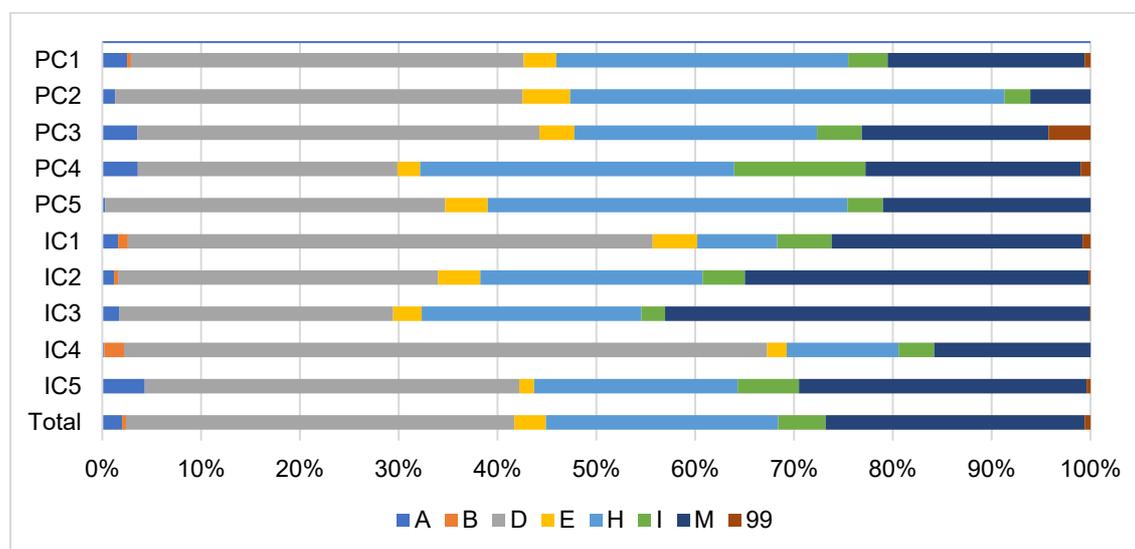
Table 3

Absolute frequencies of gestures by the teacher

ID Teacher	Dimension of Gesture								99	Total
	A	B	D	E	H	I	M			
PC1	12	2	190	16	141	19	95	3	478	
PC2	3	0	94	11	100	6	14	0	228	
PC3	6	0	116	10	70	13	54	12	281	
PC4	11	0	81	7	98	41	67	3	308	
PC5	1	0	105	13	111	11	64	0	305	
IC1	6	4	203	17	31	21	97	3	382	
IC2	6	2	160	21	111	21	172	1	494	
IC3	12	0	192	20	154	17	298	1	694	
IC4	1	9	288	9	50	16	70	0	443	
IC5	23	0	201	8	109	33	154	2	530	
Total	81	17	1630	132	975	198	1085	25	4143	

Note. ID Teacher: PC = preparatory class, IC = integration course.

Dimension of Gesture: A = Affective Display, B = Beat, D = Deictic, E = Emblem, H = Head Movements, I = Iconic, M = Metaphoric, 99 = undefined gesture.

Figure 2*Relative frequencies of gestures by the teacher*

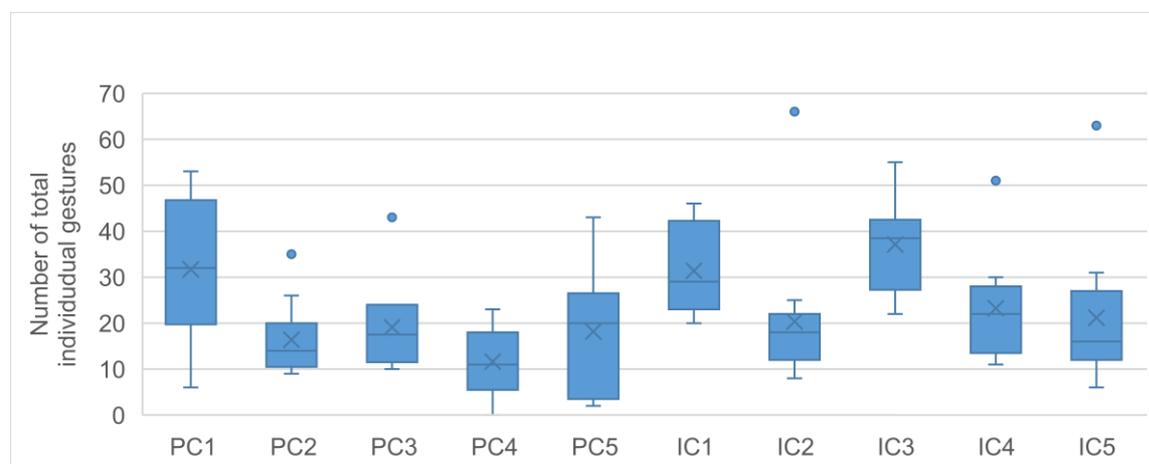
Note. ID Teacher: PC = preparatory class, IC = integration course.

Dimension of Gesture: A = Affective Display, B = Beat, D = Deictic, E = Emblem, H = Head Movements, I = Iconic, M = Metaphoric, 99 = undefined gesture. Percentages are rounded to one decimal place.

Looking only at the gestures directed at individual learners, we found that different learners in the same classroom were also addressed with varying frequency. We identified a total of 2315 gestures that were exclusively directed to one learner. Thus, on average, each learner received 22 gestures per lesson ($SD = 13.71$), with a minimum of zero gestures and a maximum of 66 individually addressed gestures. The broad interquartile range for every single learning group (see Figure 3) illustrates that there was considerable variability within learner groups. We report further boxplots of the distributions separately for each recorded gesture dimension in the online supplementary material. The distributions in Figure 3 show that the teachers' use of gestures varies greatly between teachers and between learners of the same classroom.

Figure 3

Boxplot² of number of total individual gestures a learner received by the teacher



RQ3: Can the variance in the use of teachers' gestures be explained by the teacher's perceived German proficiency of the learners?

To better understand the variance found in teacher gesture use (compare RQ2, Figure 3), we conducted analyses to determine whether teachers' perceived German language proficiency and measured gap test scores could systematically explain gesture use. Table 4 indicates the descriptive statistics of the teacher's perception of each learner's language proficiency and the measured gap test score. The gap-test score was highly correlated with teachers' perception of the learners' language proficiencies ($r = .49 - .63$, all $p < .01$; see Table 4).

² In the boxplot the box represents the middle 50% of the data set, from the first (Q1) to the third quantile (Q3); mean value is represented by cross, median value is represented by a horizontal line; whiskers represent the minimum and maximum values, that are not outliers; outliers are more than 1.5 x interquartile range below Q1 or above Q3 and represented by circles.

Table 4

Descriptive statistics and correlations of the teacher-perceived and test assessed language proficiencies and measured gap test score

Variables	N	M	SD	1	2	3	4	5
1. Text comprehension	105	4.53	1.70	—				
2. Vocabulary	105	4.50	1.59	.84**	—			
3. Grammatical competence	105	4.36	1.62	.89**	.83**	—		
4. German language aptitude	105	4.55	1.59	.87**	.85**	.87**	—	
5. Gap test score	80	22.79	15.64	.58**	.50**	.63**	.49**	—

Note. Two teachers each participated in the study with the same group of learners. Therefore, there are 105 teacher assessments for 82 learners since there are duplicate assessments for learners in two classes ($N = 10$ and $N = 13$). 2 learners did not take part in the gap test. Variable 1-4 is teacher-perceived language proficiency of individual students, variable 5 is test assessed.

** $p < .01$.

Table 5 displays the standardized beta-coefficients of several multiple regressions in which the individual language-related proficiency and the gap test score, respectively, were included as independent variables to predict the number of certain gestures relative to the other gestures a learner receives. The total number of gestures a learner received within a lesson was included as a control variable in the first step of the analyses.

Table 5

Regression results of gesture dimensions predicted by (teacher-perceived) language proficiencies

Variables	Dimension of Gesture													
	A		B		D		E		H		I		M	
	β	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2
Text comprehension	.17	.03	-.10	.01	-.17**	.03	-.05	.00	.13	.02	.04	.00	.09	.00
Vocabulary	.15	.02	-.14	.02	-.16**	.03	-.06	.00	.10	.00	.00	.00	.07	.00
Grammatical competence	.16	.02	-.07	.00	-.21**	.04	-.02	.00	.14*	.02	-.01	.00	.15*	.02
German language aptitude	.13	.02	-.07	.00	-.20**	.04	-.05	.00	.11	.01	.01	.00	.14*	.02
Gap-test score	.17	.03	-.00	.00	-.19**	.04	-.03	.00	.20*	.04	-.09	.01	.04	.00

Note. $N = 105$; Dimension of gesture: A = Affective Display, B = Beat, D = Deictic, E = Emblem, H = Head Movements, I = Iconic, M = Metaphoric.

Predicting the number of individual dimensions of gestures while controlling for the absolute number of individual gestures a learner received as compared to the learning group. Reported ΔR^2 indicates the additional variance explained by the respective predictor beyond the number of absolute gestures a learner received.

* = $p < .05$; ** = $p < .01$.

First, the total amount of gestures the learner received was not significantly predicted by the teacher-perceived German language proficiencies, nor by the gap test scores. However, looking at the gestures separately, teachers' use of certain gestures to individual learners was significantly related to the teacher-perceived German language skills. As Table 5 illustrates, learners, who were rated as weaker (vs. stronger) in terms of text comprehension ($\beta = -.17$, $p = .004$), vocabulary ($\beta = -.16$, $p = .008$), grammatical competence ($\beta = -.21$, $p = .004$), and German language aptitude ($\beta = -.20$, $p = .001$) received relatively more deictic gestures by their respective teacher compared to other learners in the classroom. Thus, our findings support hypothesis RQ3-H1, that teacher-perceived language proficiency of learners negatively predicts the number of deictic gestures a learner receives compared to its classmates.

In contrast, teachers tended to use more metaphorical gestures toward learners to whom they ascribed higher (vs. lower) German language proficiencies. Learners rated stronger (vs. weaker) in terms of grammatical competence ($\beta = .15$, $p = .025$) and German language aptitude ($\beta = .14$, $p = .039$) received relatively more metaphorical gestures. Learners' language proficiency in terms of text comprehension ($\beta = .09$, $p = .178$) and vocabulary ($\beta = .07$, $p = .296$) was not significantly related to the number of metaphorical gestures a learner received compared to the other classmates. Thus,

our findings only partially support hypothesis RQ3-H3 that teacher-perceived language proficiency of learners positively predicts the number of metaphoric gestures a learner received as compared to its classmates.

The number of deictic gestures was negatively predicted also by learners' gap test scores. Learners who scored lower vs. higher in the gap test ($\beta = -.19, p = .001$) received relatively more deictic gestures compared to other learners in the classroom. In contrast, teachers' use of metaphorical gestures was not significantly related to the measured gap test score. Thus, the hypothesis RQ3-H5 that the number of gestures teachers address to individual learners is predicted by German language proficiency both assessed via a standardized test in the same way and perceived by teachers is partly supported.

Beyond deictic and metaphorical gestures, the frequency of head movements was the only other gesture dimension that was significantly related to the language proficiency of learners perceived by teachers or assessed via gap test. Learners rated as grammatically stronger (vs. weaker; $\beta = .14, p = .033$) or learners who scored higher (vs. lower) in the gap test ($\beta = .20, p = .003$) received relatively more head movements from the teacher. There was no significant relation between the language proficiency of the learners and the relative frequency of iconic gestures or emblems a learner received compared to its classmates. Therefore, the hypotheses RQ3-H2 and RQ3-H4 are not supported by the present findings.

6. Discussion and Conclusion

The purpose of this study was to gain insight into teachers' use of gestures in linguistically diverse GSL classrooms. There are three key findings of the present research. First, teachers used different dimensions of gestures unequally often. Deictic and metaphoric gestures and head movements dominated, whereas the other dimensions of gestures recorded in our study were less frequently used by the teachers. Second, the frequency of teachers' gesture in the lessons studied differed substantially between teachers and between different learners of the same teacher. Third, the results show that variance in teachers' gesture use across learners in the same classroom can be partially explained by the teacher-perceived German language proficiency of the individual learners. The perceived relatively lower linguistic ability of a learner compared to the classmates (in terms of text comprehension, vocabulary, grammatical competence, and German language aptitude) was associated with receiving relatively more deictic and less metaphorical gestures. Therefore, this study

provides preliminary evidence that teachers may systematically adapt certain dimensions of gestures in GSL classes to the perceived German proficiency of their learners. These results are preliminary, as the content, which was not examined in the study, may also affect the dimensions of the gestures used. This should be taken into account in future studies.

We compared the descriptive frequencies of teachers' gesture use (RQ1) with a study by Sato (2020), that also investigated gestures of two language teachers in authentic language classroom situations using an identical gesture classification. Similar to the results by Sato, we observed a broad range of absolute numbers of gestures used by the different teachers in a lesson. The high proportion of deictic gestures in our sample also resembles the findings of Sato. In contrast to the study of Sato (2020), other gesture dimensions differ in terms of their relative frequencies. For instance, in Sato's study, both teachers used substantially more beat gestures (31.1% and 37.7%) than was the case in our study (0.4% on average). An explanation for the differing frequencies of the gesture dimensions in existing studies could be the different language of instruction (German vs. English) or the content of the speech produced. This assumption is supported by evidence, that the idiosyncrasy of a language can influence the frequency of gesture use (Gullberg, 2010). Therefore, the frequencies of teachers' use of gestures and gesture dimensions might be related to the German context and the results may not be fully transferable to SLA processes in countries with other languages. Furthermore, the frequencies of gestures or different gesture dimensions could be linked to specific tasks, for instance, gesture rates can be higher for picture descriptions than for translation tasks (Hadar et al., 2002). Also, Tabensky's (2008) study indicates that metaphorical gestures are often associated with expository tasks (such as short presentations) and that iconic gestures are often used when answering questions. In the present study, teachers were free to choose the topic of the lesson. The number of gestures used by teachers might therefore be related to the fact that different teaching-tasks were carried out by the teachers in the video-taped classes. Some additional evidence for this assumption is provided by the study of Stam (2016), who showed that different gestures are used with different frequency depending on the task (in narratives, iconic gestures are used more frequently, in interview-type oral proficiency tasks, metaphorical gestures are used more frequently). The finding is therefore interesting, as it was observed that the interviewers sometimes started to mirror the learners' gestures and thus the gesture use was determined by the content.

The total number of gestures a learner received compared to the classmates was not predicted by teacher-perceived language proficiency. This shows that teachers did not nonverbally interact more often with higher or lower achievers. Therefore, other factors than teacher-perceived language proficiency might be related to the differences in total numbers of gestures which teachers address towards different learners. One explaining factor might be that the actual frequency of interactions with individual learners varies within the classroom. For example, the teachers could spend more time with some learners than with others in individual work phases which might go along with more gestures for these learners. It must be critically noted that the pure comparison of absolute frequencies of gesture use in lessons is insufficiently robust and may appear arbitrary in parts, as the number and gestures are strongly determined by content and task (Stam, 2016; Tabensky, 2008). It nevertheless allows us to draw similarities to Sato's (2020) study and the frequent use of deictic gestures are consistent with work from L2 classrooms by Azaoui (2013, 2020). To answer the further research questions, we chose to calculate the group-centered means to be able to interpret the ratios of the number of gestures per class in which the contents were the same.

Importantly, our results related to RQ3 show that teachers adaptively adjust the composition of gesture dimensions to learners' language proficiency. Learners with lower language proficiency receive relatively more deictic gestures than learners with higher language proficiency. In contrast, learners with lower language proficiency receive relatively fewer metaphorical gestures than learners with higher language proficiency. Our results are in line with previous experimental studies showing that speakers and prospective teachers use more deictic gestures when telling stories and explaining vocabularies to interlocutors with lower language proficiency in the target language (nonnative speakers) than to interlocutors with higher language proficiency in the target language (native speakers; Adams, 1998; Tellier et al., 2021). This indicates that the language proficiency of the respective interlocutor is of relevance for the speakers' use of deictic gestures. Whereas past researchers have found no significant differences in the use of metaphoric gestures addressed to learners with different language proficiency, the present study shows that the relative number of metaphoric gestures received is higher for learners with stronger language proficiency. As the speech content was not examined in detail in the present study, it cannot be ruled out at this point that the language teachers directed to high-proficiency learners was more abstract and therefore included more metaphors (and accompanying more

metaphorical gestures). This could also be the case, as this group of learners may have a more conceptual understanding of language, making the content of the conversation more complex. And it can be assumed that a more complex conversation contains more metaphorical gestures (and hence more metaphorical gestures). For iconic gestures and emblems there was no significant association between a learner's language proficiency and the relative frequency of gestures addressed to this learner. This may be due to the fact, that most teachers in the sample used very few iconic gestures and emblems (less than 5% iconic gestures or emblems, see Table 3). Specifically, in our culturally heterogeneous groups of learners it might be plausible, that the teachers refrained from using emblems too often due to the culture-specific meaning of many emblems (Gullberg, 2006). On the other hand, teachers might not necessarily be aware of the gestures they use. For example, Tellier and Cadet (2013) surveyed French teachers and found that most teachers are not familiar with the culture-specificity of gestures, especially emblems.

The different findings regarding the use of metaphorical gestures in relation to the teacher-perceived language proficiency might be explained by differences in the study designs. First, our sample consisted of trained teachers with professional experience in communication. In contrast, the sample of Tellier et al. (2021) consisted of teachers in training and only some of them had task related experience. Therefore, the teachers in our sample might be more aware of the meaning of gestures and different gesture dimensions in communication. This assumption is in line with the findings from the observational study with a French Language teacher by Azaoui (2013). He found that the teacher used less metaphorical gestures when teaching nonnative speakers as compared to more with the native speakers. Similarly, Azaoui argues that teachers might consider the learners' language proficiencies and their needs for gesture-based support of language comprehension. Second, we used video recordings of authentic classroom situations, which better represent teachers' communication behavior to learners in the context of SLA than controlled laboratory studies. Second, we captured learners' language proficiency in a more differentiated way to reflect the natural variability of learners' language proficiency levels in language classrooms more closely. This extensive measurement of learner's levels of text comprehension, the scope of vocabulary, grammatical competence, and German language aptitude is due to the peculiarities of SLA, as it has been shown that learners have to meet different requirements depending on the task in second language acquisition (Stam, 2016). Third, we assessed teachers' perception of learners' language proficiency which could

inform their behavior in the classroom more than learners' actual abilities. We assume that for the use of gestures, the teacher's perceived language proficiencies are of greater importance than the externally measured language proficiency. Our study design enabled us to show that teachers unequally distribute the number of gesture dimensions between individual learners and that this distribution partially depends on teacher-perceived learner language proficiencies.

Our findings that learners with rather low language proficiency received relatively more deictic and less metaphorical gestures, can also be interpreted in terms of the functions of gestures. Deictic gestures might often be used by teachers to direct a learners' attention, and the accompanying language used might be simpler. In contrast, teachers might be more likely to use metaphorical gestures when engaging in fluent dialogues with more language-proficient learners. The observed relationship between teachers' use of metaphorical gestures on the one hand and teacher-perceived grammatical competence and German language aptitude on the other hand support this interpretation, as these two facets of learners' language proficiency could influence the complexity of the teacher-learner dialogue. However, it must be noted again that we have not considered the actual content of the speech although this content might affect the gestures produced. Another way to better understand our empirical findings is to consider the different functions of gestures (Tellier, 2008). The distinction between gestures to inform, to assess, and to manage the class could provide an explanation for why teachers more frequently address head movements to learners who they perceive as linguistically stronger. Teachers might nod their heads in agreement more often as learners use the language more grammatically correctly in conversation. In contrast, it may be that learners who are perceived to be linguistically weaker receive fewer head-shaking movements when grammatical errors are made, because teachers refrain from using this rather demotivating function. To investigate these assumptions in more detail, it seems promising to examine the linguistic statements associated with the gesture with respect to their correspondence to the learners' language proficiency level in future studies. Here, further analyses could help to understand how certain dimensions of gestures are used. The analysis of the relationship between the adequacy of the individual linguistic statements (in terms of appropriateness of the fitting of the complexity of the linguistic statements and the learners' language proficiency level) and the accompanying gestures could also shed more light on the interplay between language and gestures (Hostetter, 2011).

Some limitations must be considered when interpreting the results of the present study. First, we have a rather small sample of ten teachers in GSL classrooms limiting the generalizability of the results. It should be noted, however, that small sample sizes are rather common in prior studies on teachers' gesture in authentic classroom settings (e.g. Alibali & Nathan, 2006: $N = 1$; Alibali et al., 2013; $N = 4$; Sato, 2020: $N = 2$). Importantly, we accounted for the nested data structure (learners within teachers/classrooms) by calculating deviations from the group mean. Moreover, our sample still enabled us to investigate 4143 gestures in the observed 480 minutes of German language instruction. Second, the small sample did not allow for an analysis whether or not specific teacher characteristics influence the use of gestures. In this regard, it might be meaningful to investigate whether teachers' adaptive use of gestures is consciously controlled or rather is a concomitant of the accompanied language and the instructional design. Thus, we also cannot completely rule out the possibility that teachers used language of varying complexity with different learners. This could mean that learners who were perceived to have a higher language proficiency might not only received more metaphorical gestures, but also more abstract language. However, an initial analysis of the ten lessons, showed no differences in linguistic complexity between the teacher and the different learners. Another factor influencing the adaptive gestures use could be teachers' expertise. Following expertise research, it is reasonable to assume that teaching experience, different training, and professional development can promote the ability to use gestures adequately (Alibali et al., 2013). Third, the different facets of perceived German proficiency were analyzed separately because of multicollinearity. Therefore, further studies are needed to investigate the interrelation of different facets of learners' language proficiency and their incremental roles in teachers' gesture use. Furthermore, as only text but not oral comprehension was tested due to time constraints, the present study employed a rather rough estimator for learners' language proficiency.

Despite these limitations, the present results expand our understanding of teachers' gesture use in heterogeneous second language classrooms. Our findings support the assumption that gestures and nonverbal behaviors are important for teachers' dealing with the heterogeneity of learners in the classroom. Moreover, the present study suggests that teachers are indeed able to adapt gestures to the perceived language proficiency of learners. Therefore, it might be worthwhile to educate teachers about use of gestures and its potential implications in the language classroom. There is empirical evidence from studies on mathematic instruction that teachers' gesture

use can be trained and that teachers can learn how to effectively apply gestures to increase the learning success of the learners (Alibali et al., 2013). In the training and further education of language teachers, more emphasis should be placed on the conscious use of gestures and their adaptation to the language level of the individual learners. We therefore strongly endorse the much-requested recommendation to train teachers on the importance of the body, and in particular the use of the hands for gestures as a tool for designing effective teaching-learning processes (Stam & Tellier, 2022). For a better understanding of adaptive teaching processes (Hardy et al., 2019), future research could also investigate whether scaffolding processes as described by Hammond and Gibbons (2005) are aligned with gestures and adequately adapted to the individual language level. Further detailed analyses are necessary to understand the function and impact of teachers' gesture actions in detail, for example, whether or not the adaptive use of gesture dimensions by the teachers is conducive to the learner's language comprehension in the specific teaching situation.

The present research can be seen as one more step towards integrating two lines of research, namely gestures in the language classroom and adaptive teaching (e.g. Lim, 2021, Stam & Tellier, 2022; Urbanski & Stam, 2023). It contributes to the research field on teacher gestures in language acquisition in authentic classroom settings, which is constantly professionalizing and establishing comparable standards (cf. Tellier & Yerian, 2023). In conclusion, our findings highlight the importance of studying nonverbal interactions in SLA classrooms for a better understanding of teacher-learner-processes.

7. Acknowledgement

Funding

This research was conducted with a part of the data of the COLD project. The COLD project is funded by the Leibniz Association and investigates the competencies of school teachers and adult educators in teaching German as a second language in linguistically diverse classrooms (K113/2018).

8. Supplementary Material

Gestures

- gesture = Movement of hand or arm that serves interpersonal communication/interaction.
- For the teacher, the appropriate code from the table is assigned in the respective [gst] tier
 - Do not annotate (uncontrolled) body movements without identifiable meaning/reference (e.g. head scratching).
 - Analysis of speech unit is a speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clauses associated with either of them.
 - Gestures are coded as events. Each gesture is given a code in the [gst] tier (see the following table for gesture dimensions) for the duration of the gesture. Only the teachers' gestures are coded.
 - If gestures appear simultaneously or overlap, create multiple tiers (gst2, gst3 etc.).
 - Each gesture must be aligned with a corresponding target or each gesture must overlap with a target.

LK [gst]			M		D
LK [gst2]					
LK [v]		So. ((Bonjour)).			08.(.)

Gesture dimensions are based on Sato (2020). Gestures in EFL classroom: Their relations with complexity, accuracy, and fluency in EFL teachers' L2 utterances. *System*, 89, 102215. <https://doi.org/10.1016/j.system.2020.102215>

Table 6

Overview of codes for teacher gestures

Code	Dimension	Meaning	Possible expression in the classroom
D	Deictic	Speakers point to something or someone with a finger or the palm of their hand, indicating both concrete and abstract entities.	Graphic illustrations Deictic gestures are unambiguous and involve "pointing". They do not need to be interpreted to understand them.
I	Iconic	Speakers use their hand and/or arms to show images of the actual form of objects and/or actions, which are closely related to semantic content.	Raise the hand with one finger extended, accompanied by the phrase "write a sentence". Teacher makes an underlining gesture (drawing the hand through the air in a horizontal straight line in pen holding position). Clarifying/indicating numbers through gestures is indicative of an iconic gesture.
M	Metaphoric	Speakers use their hand and/or arms to show images of abstract concepts and/or ideas rather than the actual form of objects and/or actions shown by iconics.	Teacher rolls her hands over each other, making a backward motion of the hands, accompanied for "Last week" . Likewise, pointing backward over one's own shoulder is a reference to "Past". Closing gesture (both hands in front of body, palms down, pulling apart from center, as if smoothing a tablecloth, accompanied by "that's it" - is a metaphor for "closure".
E	Emblems	Speakers show nonverbal acts that are understood by all members of the same cultural group. For example, putting the palm next to the ear means "I cannot hear you" in Japan, and circling the thumb and index finger means "Okay" in Western cultures.	Emblems are gestures that are uniquely (and beyond cultural circles) unambiguously understandable (Waving, thumbs up, bowing, spreading arms to welcome, finger pointing for importance and respect, raising arms questioningly/shrugging shoulders like the emoji, bird pointing with finger to forehead).
A	Affect display	Speakers reveal emotions, such as happiness, fear, sadness, anger, distraction, and interest, especially through facial expressions.	Clearly identifiable facial expressions of the LC such as an approving smile, irritated look or similar.
H	Head movements	Speakers tilt their head while speaking to imply their incomprehension or need for more explanation to interlocutors.	Head tilts to the side, nodding, shaking. Interactive processes are meant here. If it can be assumed that the teacher intends a meaning with its head movement. Pure physiological head movements, for example to orient in space or to look around, are not coded. "
B	Beats	Speakers move their hand with a rhythmic pulse. Typically, a beat gesture is a simple flick of the hand or movement of fingers up and down, or back and forth, following the stress peaks of speech.	
99	indefinite	Speaker shows an indefinite gesture. Gesture is not clearly assignable.	e.g. teacher puts her hands together with no apparent connection to what she is saying.
ng	no gesture	no gesture recognizable, although rater nr. 2 has coded a gesture	
nd	not detected	Gesture recognizable, although rater nr. 2 has not coded a gesture	

Table 7

Descriptive statistics of the teacher-perceived and test assessed language proficiencies and measured gap test score exemplary for two groups (PC1 and IC8)

	PC1 (n=10 learners)				IC8 (n=8 learners)			
	Min	Max	Mean	SD	Min	Max	Mean	SD
1. Text comprehension	1	7	4.20	1.62	2	7	4.63	1.69
2. Vocabulary	1	7	3.90	1.79	3	6	5.12	1.36
3. Grammatical competence	1	7	3.90	1.85	2	7	4.50	1.77
4. German language aptitude	1	7	4.00	1.89	3	6	4.63	1.51
5. Gap test score	10	67	37.50	17.08	11	47	31.13	14.54

Table 8

Content of the lessons of the preparatory classes (PC) and integrations courses (IC)

ID Teacher	Description of content
PC1	<p>At the beginning of the lesson, the teacher asks how the individual learners are doing and has a kind of open social conversation. Sometimes the teacher additionally asks the learners what their favorite leisure activities are. This is a transition to the topic of the lesson. The teacher then presents pictures of different leisure activities on the Smartboard, which the learners are asked to name. Then there is a short review of the modal verb "to be able" and the present tense conjugation. The learners are then asked to each formulate a sentence ("I can do well..."). Afterwards, the sentence position of the verbs in this sentence is discussed on the board. Thereafter, the learners have to complete tasks on a worksheet, which are explained in plenary. Afterwards, the results are discussed in plenary. The lesson ends with a short partner work in which a conversation/interview about any topic in German is to be conducted (with modal verbs).</p>
PC2	<p>The lesson begins with the learners saying "Good morning" in their languages of origin. Then there is the introductory ritual in which the teacher shows questions (e.g., What would you do with one million euros?) on the Smartboard and the learners take it in turns to throw a question, which they then have to read out and answer. The answer is then repeated by another person from the learning group. Learners can take turns and help each other with translations.</p> <p>Next, learners work in groups on grammar. The groups are divided according to their learning level and are given different tasks in the book. The learners work mostly alone and the teacher gives individual help.</p>

-
- PC3 To start the lesson, students work individually to create a mind map on the topic of road traffic. Afterwards, the found terms with articles are collected on the board in plenary. Then the teacher divides the group into 3 small groups, which are given different tasks based on hidden pictures. Group A gets a board game where they have to form sentences with prepositions to the picture ("The king is on the stone."). Group B has to write a dialogue about a scene on a picture. Group C has to write sentences on their own and attach them to the picture with Post-Its, but the group members are given different sentence structures (past tense, passive, perfect, relative) and example sentences. The teacher supervises the groups individually and corrects and checks in turn. As soon as the groups have finished, they get new tasks. Group C gets a card game next, in which they have to form sentences/stories to pictures. Group A is next allowed to listen to how group B presents the dialogue. Afterwards, group B joins group C in the card game and group A has to write post-its with sentences with prepositions, which they then have to distribute appropriately in the classroom (e.g. "The radio is on the cupboard."). The teacher continues to give individual help. The lesson ends with the completion of the group work.
-
- PC4 At the beginning, the teacher shows a picture on the Smartboard and the learners have to say one after the other something they can see in the picture. Then the teacher reads a poem to the picture and the learners have to remember places from the poem, which are then discussed. Then the teacher fades in the poem text and reads the poem again slowly so that the learners in the plenary can read along aloud. In another repetition of the poem, the teacher shows different gestures and movements to the text, which the learners should also do. Then the learners read the poem to each other in groups/partners. Afterwards, some learners try to recite the poem from memory in plenary. Next, prepare for the learners to write a poem themselves. To do this, first collect in plenary words and sentence components with prepositions that the learners use in the poem.
-
- PC5 The filmed lesson builds on the lesson from PC4 but with a different teacher. It is also based on the same poem. At the beginning of the lesson, the poem is read aloud again by some learners in plenary. Then the learners have to rewrite the poem individually using the terms collected in the previous lesson and presented on the board. The teacher gives individual help. Those who have finished earlier may also change verbs in their own version of the poem in addition to the changed sentence components. Later, learners who have finished may also give help to other learners in the class, as the teacher does. Then the learners read their poems in plenary. The listening learners should remember places from the poem and name them. At the end of the lesson, the teacher gives a short preview of the next lesson.
-
- IC1 The group has already had a lesson before the recording and the video starts in the middle of an instruction by the teacher. The learners are asked to ask each other questions in plenary about a worksheet with pictures of different objects ("Who owns X?"). For this purpose, the possessive pronouns are visible to the learners on the board. Next, the teacher instructs a partner work with a new worksheet. Both partners have different information on their worksheets and ask for the missing information ("Do you know what X gives to Y?"). The teacher goes around during the exercise and gives individual help. Afterwards, this exercise is discussed in plenary, with the learners taking turns to ask the questions and answers and correcting each other if necessary. The teacher writes the correct solutions on the board. Finally, the teacher explains the homework.

IC2

At the beginning of the video, some time is spent organising the seating arrangements of the learners. Then the teacher briefly reviews the previous lesson and explains what the content of the lesson will be. She writes the topic "Planning a party" on the whiteboard and discusses the meaning with the learners in plenary. Then she instructs the following group work, in which the learners have to plan the birthday of a fictitious grandmother together. They are given a few key points to plan on a worksheet. Two learners are given another individual task during the group work. The teacher goes around during the group work and gives individual support. Afterwards, the teacher explains the homework and the learners have to start with an exercise in the workbook and the teacher goes around and makes sure that the learners have understood the task.

IC3

At the beginning, the teacher writes the learning objectives and topics of the lesson on the whiteboard and discusses them with the learners. At the same time, the previously discussed question words are briefly repeated. Then the teacher instructs the following partner work and the corresponding worksheets, which correspond to an A2 exam. The learners have to formulate questions and answers to pictures. First, however, the necessary grammar (modal verbs) is discussed in plenary. The teacher also writes the task on the whiteboard. The participants should first formulate questions in individual work and then compare them in partner work. The teacher gives individual help and then discusses particular problems in plenary. Afterwards, two people should improvise a dialogue in front of the group. The teacher gives feedback and hints and tips for the exam situation (e.g. compromise, make comparisons...). Finally, there is another exercise in which the learners are asked to form questions based on question words.

IC4

The teacher starts with an outlook on the different tasks and learning objectives of the lesson. Then they play the game "I'm packing my suitcase", for which the teacher has prepared a suitcase and various items of clothing in the middle of the room. The participants take turns to choose an item of clothing and repeat all the previously mentioned items ("I like..."). The teacher then lays out the terms for the items of clothing with articles on slips of paper so that the learners can read them off. At the end of the exercise, all the terms are read again together in plenary. Then the words are sorted according to the article. Afterwards, the learners get a worksheet on which they have to write down the words in columns according to the article. The teacher goes around and checks and gives individual help. Then the words are read together. For this, the teacher has written the words on the flipchart, divided into letters, and the word parts are spoken together and individually one after the other. In the next task, the learners have to use given sentence blocks to ask each other questions about clothes in pictures ("Do you like...?"). The exercise is first practised together in plenary, then practised in partner work and then repeated again in plenary. Then the teacher reads a short text from the book and has the learners read the text individually.

IC5

First, an exercise in the book is used to repeat the times of day and the corresponding prepositions (in the morning = in the morning). In the next task, the learners listen to an audio narration and are asked to connect sentence halves in the book. Afterwards, the learners have to write sentences individually in which given activities (e.g. listening to music) have to be conjugated to the different times of day. The teacher goes around and gives individual feedback and help. Afterwards, the sentences and sentence order are discussed in plenary. The next exercise involves a board with squares with times and activities. In groups, the learners have to roll the dice and form sentences for the squares. At the end of the lesson, there is a short conclusion in the plenary.

Figure 3

Boxplot of number of total individual affective displays a learner received by the teacher in preparatory classes (PC) and integration courses (IC)

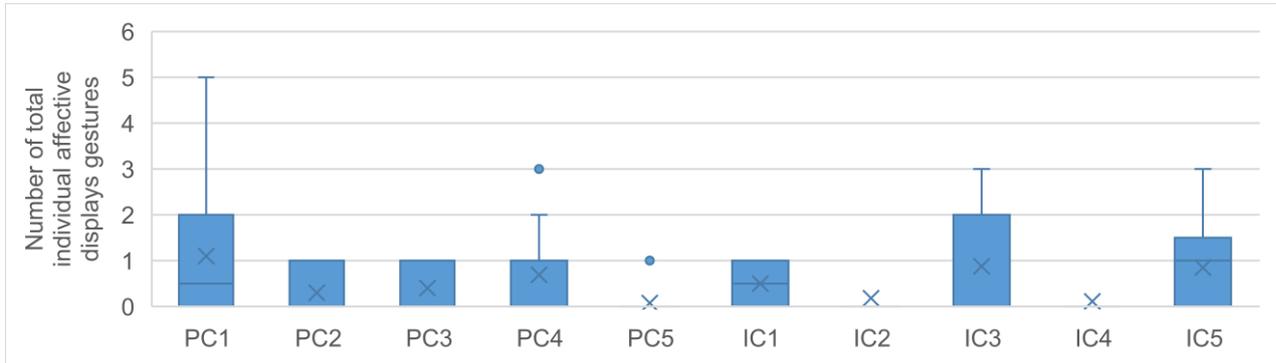


Figure 4

Boxplot of number of total individual beat gestures a learner received by the teacher in preparatory classes (PC) and integration courses (IC)

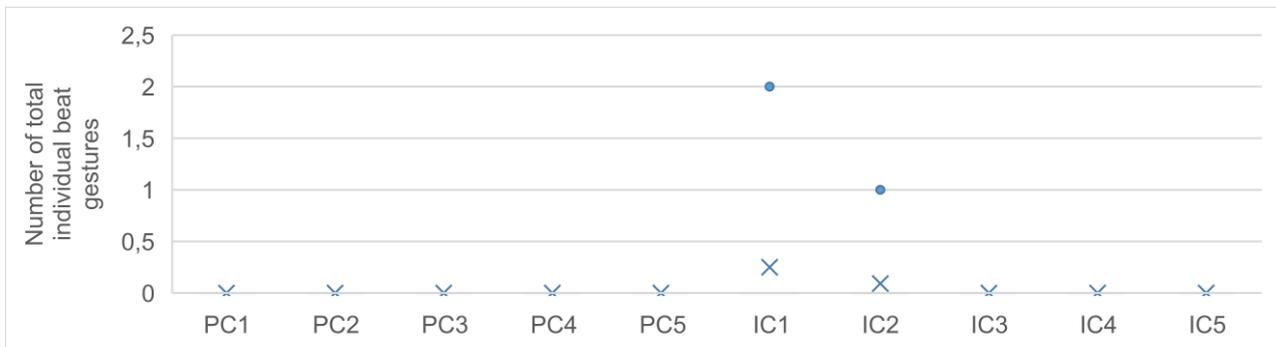


Figure 5

Boxplot of number of total individual deictic gestures a learner received by the teacher in preparatory classes (PC) and integration courses (IC)

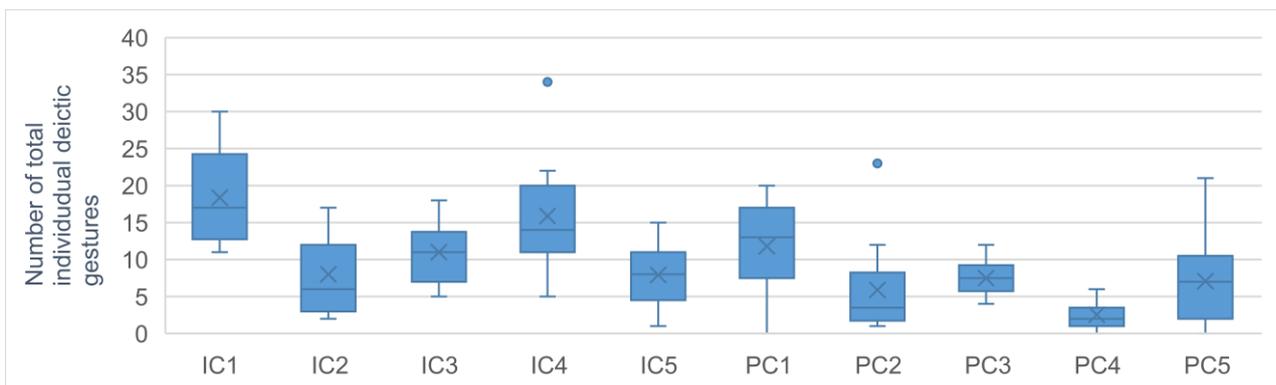
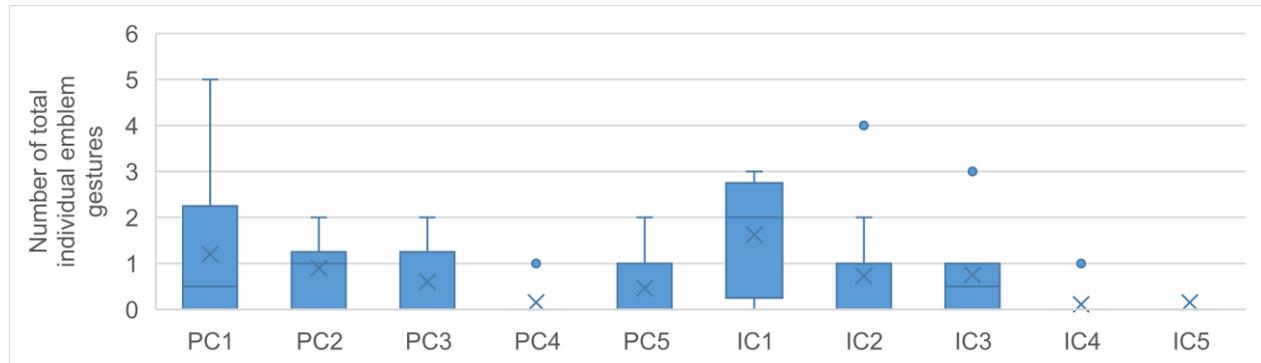
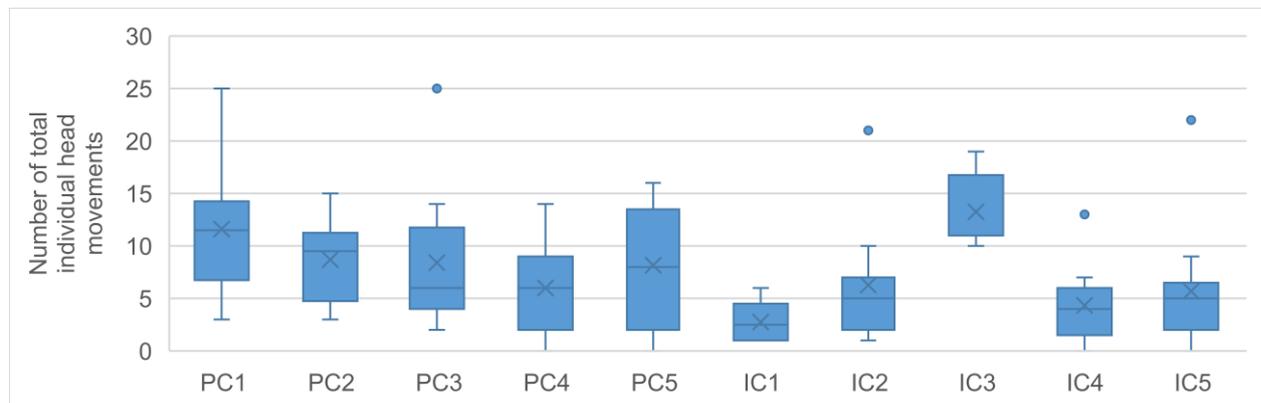


Figure 6

Boxplot of number of total individual emblem gestures a learner received by the teacher in preparatory classes (PC) and integration courses (IC)

**Figure 7**

Boxplot of number of total individual head movements a learner received by the teacher in preparatory classes (PC) and integration courses (IC)

**Figure 8**

Boxplot of number of total individual iconic gestures a learner received by the teacher in preparatory classes (PC) and integration courses (IC)

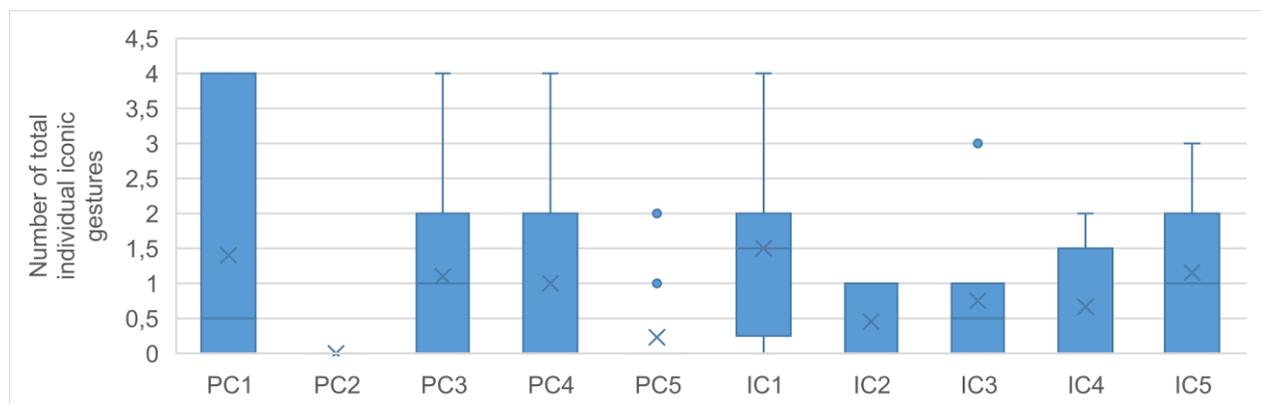
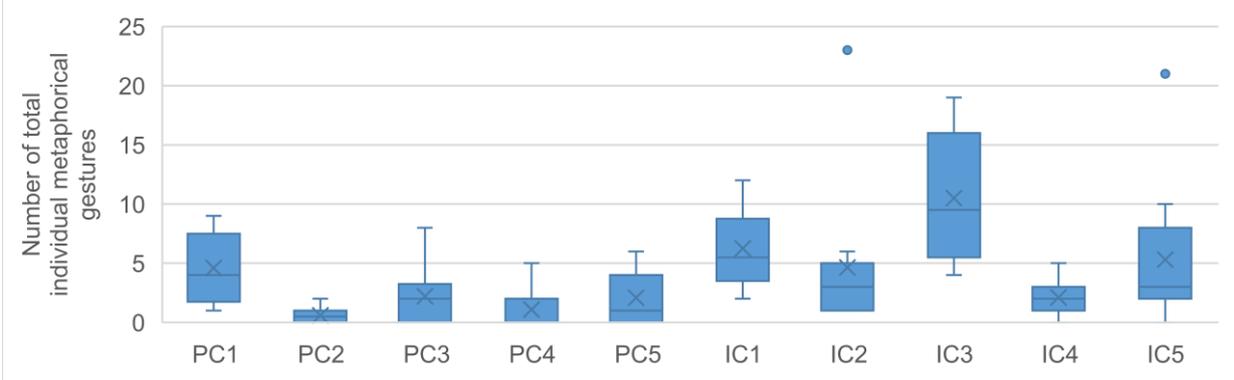


Figure 9

Boxplot of number of total individual metaphorical gestures a learner received by the teacher in preparatory classes (PC) and integration courses (IC)



9. References

- Adams, T. W. (1998). *Gesture in foreigner talk*. [Doctoral Dissertation, University of Pennsylvania.]
- Alibali, M. W., & Nathan, M. J. (2006). Teachers' gestures as a means of scaffolding students' understanding: Evidence from an early algebra lesson. In R. Goldman, R. Pea, B. Barron, & S. J. Denny (Eds.), *Video Research in the Learning Sciences* (pp. 349–365). Lawrence Erlbaum Associates.
- Alibali, M. W., Nathan, M. J., Boncoddò, R., & Pier, E. (2019). Managing common ground in the classroom: Teachers use gestures to support students' contributions to classroom discourse. *ZDM*, *51*(2), 347–360.
<https://doi.org/10.1007/s11858-019-01043-x>
- Alibali, M. W., Nathan, M. J., Church, R. B., Wolfram, M. S., Kim, S., & Knuth, E. J. (2013). Teachers' gestures and speech in mathematics lessons: forging common ground by resolving trouble spots. *ZDM*, *45*(3), 425–440.
<https://doi.org/10.1007/s11858-012-0476-0>
- Azaoui, B. (2013). *One Teacher, Two Instructional Contexts. Same Teaching Gestures?* TiGeR: Tilburg Gesture research meeting, 1-4. <https://hal.science/hal-00833026>
- Azaoui, B. (2016). Mise en abyme des interactions didactiques [Putting didactic interactions in perspective]. *Recherches en Didactique des Langues et des Cultures*, *13*(1). <http://rdlc.revues.org/1472>
- Azaoui, B. (2020). *On pointing in first and foreign language classes*. GESPIN proceedings. Stockholm, Sweden. <https://hal.umontpellier.fr/hal-02912833>
- Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants *Review of Economics and Statistics*, *86*(2), 481–496.
<https://doi.org/10.1162/003465304323031067>
- Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, *1*(2), 274–291. <https://doi.org/10.1111/j.1756-8765.2009.01019.x>
- Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin*, *145*(8), 765–784. <https://doi.org/10.1037/bul0000202>

- Derwing, T. M. (1987). *Individual differences in foreigner talk: factors in successful communication with non-native speakers* [University of Alberta Libraries]. DataCite.
- Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290–325. <https://doi.org/10.1191/0265532206lt330oa>
- Ellis, R. (2005). Principles of instructed language learning. *System*, 33(2), 209–224. <https://doi.org/10.1016/j.system.2004.12.006>
- Esser, H. (2006). *Sprache und Integration: Die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten* [Language and Integration: The Social Conditions and Consequences of Migrants' Language Acquisition]. Campus-Verl.
- Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics*, 17, 1–14.
- Giles, H., Coupland, N., & Coupland, J. (1991). *Accommodation theory: Communication, context, and consequence*. *Contexts of Accommodation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511663673.001>
- Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development*, 71(1), 231–239. <https://doi.org/10.1111/1467-8624.00138>
- Grotjahn, R. (2019). C-Tests. In S. Jeuk & J. Settinieri (Eds.), *DaZ-Handbücher: Band 2. Sprachdiagnostik Deutsch als Zweitsprache: Ein Handbuch* (pp. 585–610). de Gruyter. <https://doi.org/10.1515/9783110418712-024>
- Gullberg, M. (2006). Some reasons for studying gesture and second language acquisition (Hommage à Adam Kendon). *IRAL - International Review of Applied Linguistics in Language Teaching*, 44(2). <https://doi.org/10.1515/IRAL.2006.004>
- Gullberg, M. (2010). Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research*, 26(1), 75–102. <https://doi.org/10.1177/0267658309337639>
- Güven, C., & Islam, A. (2015). Age at migration, language proficiency, and socioeconomic outcomes: Evidence from Australia. *Demography*, 52(2), 513–542. <https://doi.org/10.1007/s13524-015-0373-6>
- Hadar, U., Dar, R., & Teitelman, A. (2002). Gesture during speech in first and second language. *Gesture*, 1(2), 151–165. <https://doi.org/10.1075/gest.1.2.04had>

- Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect*, 20(1), 6–30.
- Hardy, I., Decristan, J., & Klieme, E. (2019). Adaptive teaching in research on learning and instruction. *Journal for Educational Research Online*(11), 169–191. <https://doi.org/10.25656/01:18004>
- Hauge, E. (1998). Gesture in the EFL class: an aid to communication or a source of confusion? In Danielle Killick & Mike Parry (Eds.), *Cross-cultural capability - the why, the ways and the means: New theories and methodologies in language education*, pp. 271-280. Proceedings of the conference at Leeds Metropolitan University Dec. 1997.
- Hauge, E. (1999). Some common emblems used by British English teachers in EFL classes. In Danielle Killick & Mike Parry (Eds.), *Cross-cultural capability - promoting the discipline: Marking boundaries and crossing borders*, pp. 405-420. Proceedings of the conference at Leeds Metropolitan University Dec. 1998.
- Hochman, O., & Davidov, E. (2014). Relations between second-language proficiency and national identification: The case of immigrants in Germany. *European Sociological Review*, 30(3), 344–359. <https://doi.org/10.1093/esr/jcu043>
- Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, 137(2), 297–315. <https://doi.org/10.1037/a0022128>
- Howe, C., & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43(3), 325–356. <https://doi.org/10.1080/0305764X.2013.786024>
- Hupp, J. M., & Gingras, M. C. (2016). The role of gesture meaningfulness in word learning. *Gesture*, 15(3), 340–356. <https://doi.org/10.1075/gest.15.3.04hup>
- Johnston, P., Dozier, C., & Smit, J. (2016). How language supports adaptive teaching through a responsive learning culture. *Theory into Practice*, 55(3), 189–196. <https://doi.org/10.1080/00405841.2016.1173992>
- Kendon, A. (2010). *Gesture: Visible action as utterance*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511807572>
- Krashen, S. (1989). We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *The Modern Language Journal*, 73(4), 440–464. <https://doi.org/10.2307/326879>

- Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, 54(1), 79–117.
<https://doi.org/10.1111/j.1467-9922.2004.00249.x>
- Lazaraton, A., & Ishihara, N. (2005). Understanding second language teacher practice using microanalysis and self-reflection: A collaborative case study. *Modern Language Journal*, 89(4), 529-542. <https://doi.org/10.1111/j.1540-4781.2005.00328.x>
- Lim, F. V. (2021). Investigating intersemiosis: A systemic functional multimodal discourse analysis of the relationship between language and gesture in classroom discourse. *Visual Communication*, 20(1), 34–58.
<https://doi.org/10.1177/1470357218820695>
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285–329.
<https://doi.org/10.1017/S0261444818000125>
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input¹. *Applied Linguistics*, 4(2), 126–141.
<https://doi.org/10.1093/applin/4.2.126>
- Martinovic, B., van Tubergen, F., & Maas, I. (2009). Dynamics of interethnic contact: A panel study of immigrants in the Netherlands. *European Sociological Review*, 25(3), 303–318. <https://doi.org/10.1093/esr/jcn049>
- Matsumoto, Y., & Dobs, A.M. (2017). Pedagogical gestures as interactional resources for teaching and learning tense and aspect in the ESL grammar classroom. *Language Learning*, 67, 7-42.
- McCafferty, S. G., & Stam, G. (Eds.). (2008). *ESL & applied linguistics professional series. Gesture: Second language acquisition and classroom research*. Routledge.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press. <http://www.loc.gov/catdir/description/uchi052/91032575.html>
- McNeill, D. (2005). *Gesture and Thought*. University of Chicago Press.
<https://doi.org/10.7208/chicago/9780226514642.001.0001>
- Otto, J., Migas, K., Austermann, N., & Bos, W. (2016). *Integration neu zugewanderter Kinder und Jugendlicher ohne Deutschkenntnisse. Möglichkeiten, Herausforderungen und Perspektiven* [Integration of newly immigrated children and young

- people without German language skills. Possibilities, challenges and perspectives]. Waxmann. <https://doi.org/10.25656/01:15094>
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research*, 88(2), 205–242. <https://doi.org/10.3102/0034654317743198>
- Sato, R. (2020). Gestures in EFL classroom: Their relations with complexity, accuracy, and fluency in EFL teachers' L2 utterances. *System*, 89, Article 102215. <https://doi.org/10.1016/j.system.2020.102215>
- Schmidt, T., & Wörner, K. (2014). EXMARaLDA. In U. G. Jacques Durand & Gjert Kristoffersen (Eds.), *Handbook on Corpus Phonology*, pp. 402–419. Oxford University Press. <http://ukcatalogue.oup.com/product/9780199571932.do>
- Sime, D. (2001). The use and perception of illustrators in the foreign language classroom. In C. Cave, I. Guaitella, & S. Santi (Eds.), *Oralité et gestualité: Interactions et comportements multimodaux dans la communication* (pp. 582–585). L'Harmattan.
- Sime, D. (2006). What do learners make of teachers' gestures in the language classroom? *International Review of Applied Linguistics*, 44(2), 211–230. <https://doi.org/10.1515/IRAL.2006.009>
- Sime, D. (2008). "Because of her gesture, it's very easy to understand": Learners' perceptions of teachers' gestures in the foreign language class. In S. G. McCafferty & G. Stam (Eds.), *Gesture: Second language acquisition and classroom research*, pp. 259–279. Routledge.
- Smotrova, T. (2017). Making pronunciation visible: Gesture in teaching pronunciation. *TESOL Q*, 51, 59-89. <https://doi.org/10.1002/tesq.276>
- Smotrova, T. & Lantolf, J.P. (2013). The function of gesture in lexically focused L2 instructional conversations. *Modern Language Journal*, 97, 397-416. <https://doi.org/10.1111/j.1540-4781.2013.12008.x>
- Stam, G. (2012). *Second Language Acquisition and Gesture*. In The Encyclopedia of Applied Linguistics, C.A. Chapelle (Ed.). <https://doi.org/10.1002/9781405198431.wbeal1049>

- Stam, G. (2016). Gesture as a window onto conceptualization in multiple tasks: Implications for second language teaching. *Yearbook of the German Cognitive Linguistics Association*, 4(1), 289-314. <https://doi.org/10.1515/gcla-2016-0017>
- Stam, G., & Tellier, M. (2022). Gesture helps second and foreign language learning and teaching. In A. Morgenstern & S. Goldin-Meadow (Eds.), *Gesture in language: Development across the lifespan*, pp. 335–363. De Gruyter Mouton; American Psychological Association. <https://doi.org/10.1037/0000269-014>
- Stam, G., & Urbanski, K. (Buescher) (eds.). (2023). *Gesture and multimodality in second language acquisition: A Research Guide*. Routledge.
- Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55(4), 661–699. <https://doi.org/10.1111/j.0023-8333.2005.00320.x>
- Tabensky, A. (2008). Expository discourse in a second language classroom: How learners use gesture. In S. G. McCafferty & G. Stam (Eds.), *Gesture: Second language acquisition and classroom research*, pp. 298–320. Routledge.
- Tellier, M. (2008). Dire avec des gestes [Say it with gestures]. *Français Dans le Monde, Recherche et Application*, 44, 40–50.
- Tellier, M. (2009). Usage pédagogique et perception de la multimodalité pour l'accès au sens en langue étrangère [Educational use and perception of multimodality for access to meaning in a foreign language]. In R. Bergeron, G. Plessis-Belaire, & L. Lafontaine (Eds.), *La place des savoirs oraux dans le contexte scolaire d'aujourd'hui*, pp. 223–245. Presses de l'Université du Québec.
- Tellier, M., & Cadet, L. (2013). Dans la peau d'un natif: État des lieux sur l'enseignement des gestes culturels [In the skin of a native: State of play on the teaching of cultural gestures in French as a foreign language]. *La Revue Française d'Éducation Comparée*, 9, 111–140.
- Tellier, M. (2016). Prendre son cours à bras le corps. De l'articulation des modalités kinésiques avec la parole [Taking the class by the hand. Articulating kinesic modalities with speech]. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 13(13-1). <https://doi.org/10.4000/rdlc.474>
- Tellier, M., Stam, G., & Ghio, A. (2021). Handling language. *Gesture*, 20(1), 30–62. <https://doi.org/10.1075/gest.19031.tel>

- Tellier, M., & Yerian, K. (2023). How to Study Pedagogical Gesture in Naturalistic Settings. In G. Stam & K. (Buescher) Urbanski, *Gesture and Multimodality in Second Language Acquisition*, pp. 99-123. Routledge.
- Thoms, J. J. (2012). Classroom discourse in foreign language classrooms: A review of the literature. *Foreign Language Annals*, 45(s1), s8-s27.
<https://doi.org/10.1111/J.1944-9720.2012.01177.X>
- Tissot, A., Croisier, J., Pietrantuono, G., Baier, A., Ninke, L., Rother, N., & Babka von Gostomski, C. (2019). *Zwischenbericht I zum Forschungsprojekt "Evaluation der Integrationskurse (Evlk)": erste Analysen und Erkenntnisse* [Interim report I on the research project "Evaluation of the integration courses (Evlk)": main results] (Bd. 33). SSOAR, GESIS – Leibniz-Institut für Sozialwissenschaften e.V; Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ).
- Ulichny (1979). Adult foreigner talk: Input language to L2 learners. *Studi Italiani Di Linguistica Teorica Ed Applicata*, 8, 187.
- Urbanski, K., & Stam, G. (2023). Overview of multimodality and gesture in second language acquisition. In G. Stam & K. (Buescher) Urbanski, *Gesture and Multimodality in Second Language Acquisition*, pp. 1-25. Routledge.
<https://doi.org/10.4324/9781003100683-1>
- Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review*, 32, 100374.
<https://doi.org/10.1016/j.edurev.2020.100374>
- Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes* (M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>.
- Wang, W., & Loewen, S. (2016). Nonverbal behavior and corrective feedback in nine ESL university-level classrooms. *Language Teaching Research*, 20(4), 459–478. <https://doi.org/10.1177/1362168815577239>

9. Gesamtdiskussion – Fragestellungen, Befunde und Perspektiven

Gegenstand der vorliegenden Dissertation ist die Untersuchung von Sprachlernprozessen in heterogenen Lerngruppen wie etwa in DaZ-Kursen (Beiträge #1 – #3) und Alphabetisierungskursen (Beitrag #1) unter besonderer Berücksichtigung von Gesten als Indikator für die Fähigkeit der Lehrkräfte zur adaptiven Unterrichtsgestaltung (Beitrag #3). Der erste Beitrag hatte das Ziel, das Forschungsfeld der Erwachsenenbildung anhand eines breit angelegten Zugangs zu sichten und die Frage zu beantworten, wie Spracherwerbsprozesse in den Bereichen Deutsch als Zweitsprache/Deutsch als Fremdsprache und Alphabetisierung beforscht werden. Weiterhin wurde untersucht, welche methodischen Zugänge in der Forschung verwendet werden und ob dabei eher theoriegenerierende oder theorieüberprüfende Studien veröffentlicht werden. Ebenfalls wurden die im Rahmen der Literaturrecherche identifizierten Artikel daraufhin geprüft, ob verstärkt die Lehrenden oder eher die Lernenden, oder aber die Prozesse zwischen Lehrenden und Lernenden adressiert werden und welche Forschungsdisziplinen an der Produktion wissenschaftlicher Erkenntnisse beteiligt sind. Dabei hat sich gezeigt, dass das Themenfeld maßgeblich von den Disziplinen der Erwachsenenbildung, der Psychologie und der Linguistik bearbeitet wird und die Disziplinen dabei wenig Bezug zueinander herstellen. Die Forschungsaktivitäten sind dabei eher hoch, erlauben aber wenig, empirische Befunde zu den Gelingensbedingungen erfolgreichen Spracherwerbs und der Bedeutung der Lehrkraft in Spracherwerbssituationen zu benennen.

Um die Interaktionen der Lehrkraft, bzw. die Fähigkeit von Lehrkräften zum adaptiven Unterrichten zu erforschen sind empirische Daten nötig. Hierzu wurde im zweiten Beitrag, basierend auf den Ergebnissen des ersten Beitrags, in einem näher eingeschränkten Themenbereich (DaZ-Unterricht) untersucht, wie Videoaufnahmen von Unterrichtssituationen aufbereitet werden können, um einerseits wissenschaftlichen Gütekriterien standzuhalten und andererseits Lehr-/Lernprozesse in sprachlich heterogenen Gruppen messbar zu machen. Dazu wurde die Frage gestellt, ob sich bestimmte, sprachliche Merkmale reliabel erfassen lassen. Ob die Transkription sprachlicher Aussagen und Sprechpausen ohne großen Informationsverlust gelingt und ob die Erfassung des Adressaten oder der Adressatin einer Aussage zuverlässig gelingt, war Teil der empirischen Überprüfung. Weiterhin wurde die Frage gestellt, ob sich auch Elemente der nonverbalen Kommunikation und der Unterrichtsgestaltung, die für die Analyse von Lehr-/Lernprozessen in sprachlich heterogenen Gruppen relevant sein können, reliabel annotieren lassen. Dazu gehörten sprachbegleitende und unterschiedlich komplexe Lehrkraftgesten, Unterrichtsbeiträge

in Form von Wortmeldungen und Aufrufen seitens der Lehrkraft sowie die generelle Sozialform, in der der Unterricht gestaltet wird. Im Rahmen der zweiten Studie konnte die Robustheit der Transkription und Annotation anhand authentischer Unterrichtsdaten präsentiert werden.

Auf dieser Basis wurde in Beitrag #3 unter Zuhilfenahme des Instruments, welches Gegenstand des zweiten Beitrags war, der zentralen Frage nachgegangen, ob Lehrkräfte im DaZ-Unterricht ihre nonverbale Kommunikation adaptiv an den sprachlichen Fähigkeiten der einzelnen Lernenden ausrichten. Dafür wurde zuerst untersucht, welche Arten von Gesten in welcher Häufigkeit von Lehrkräften im DaZ-Unterricht in Vorbereitungsklassen und Integrationskursen verwendet werden. In einer zweiten Frage wurde untersucht, ob sich die Verwendung von Gestenart und -häufigkeit zwischen einzelnen Lehrkräften unterscheidet und ob sich auch die Anzahl der Gesten unterscheidet, die die einzelnen Lernenden jeweils von der Lehrkraft erhalten. In einer dritten Frage wurde schließlich untersucht, ob die wahrgenommene sprachliche Fähigkeit die Anzahl verschiedener Gestentypen, die einzelne Lernende erhalten, erklären kann.

9.1. Zusammenfassung und Einordnung zentraler Ergebnisse

Im ersten Beitrag der vorliegenden Arbeit konnten 236 internationale und nationale Publikationen im Themenfeld Spracherwerb in der Erwachsenen- und Weiterbildung vertretet werden. Die Kategorisierung der Veröffentlichungen nach verschiedenen Untersuchungsdesigns hat verdeutlicht, wie differenziert (bzw. im Beitrag als „zersplittert“ bezeichnet) sich die Forschungslandschaft darstellt. Das liegt mitunter daran, dass die Themenwahl ein großes Spektrum an Inhalten abdeckt und dabei sowohl die Lehrenden als auch die Lernenden in den Blick nimmt. Auffällig ist dabei jedoch auch die fehlende interdisziplinäre Verzahnung der Forschungsgegenstände. So wird aus unterschiedlichen Perspektiven (vorwiegend der Erwachsenenbildungsforschung, der Psychologie und der Linguistik) an Themenbereichen mit unterschiedlicher Tiefe gearbeitet und häufig werden Detailfragen zu Spracherwerbssituationen erforscht, ohne dass Befunde von den benachbarten Disziplinen berücksichtigt oder aufgegriffen werden. Theoretisch fundierte, generalisierbare Informationen über die Gelingensbedingungen für erfolgreichen Spracherwerb sind das zentral identifizierte Forschungsdesiderat des ersten Beitrags. Ebenfalls wurde im Fazit des ersten Beitrags explizit auf die fehlenden Studien zur möglichen inneren Differenzierung in heterogenen Lerngruppen und auf die Notwendigkeit interdisziplinärer Zusammenarbeit hingewiesen.

Im zweiten Beitrag konnte gezeigt werden, wie videogestützte Unterrichtsbeobachtungen im DaZ-Unterricht in Vorbereitungsklassen und in Integrationskursen aufbereitet werden können. Dabei wurde aufgearbeitet, wie das Thema Zweitspracherwerb aus den beiden Perspektiven der (Computer-)Linguistik und der empirischen Bildungsforschung bisher untersucht wurde und welche Überschneidungspunkte es im Hinblick auf die gemeinsame Analyse von Lehr-/Lerninteraktion in sprachlich heterogenen Lerngruppen gibt. Unter Einbezug der DaZ-didaktischen Forschung ist schließlich ein Instrument zur Aufbereitung von videogestützter Unterrichtsbeobachtung entwickelt worden. Bei der Entwicklung wurde auf bereits etablierte Instrumente aus anderen Unterrichtsstudien der empirischen Bildungsforschung (z.B. Lipowsky et al., 2019) und auf Routinen bei der Erstellung von gesprochenen Lernerkorpora zurückgegriffen (z.B. die Transkriptions- und Annotationsgrundlage durch die Software EXMARaLDA, Schmidt & Wörner, 2014). Zentrales Element des Beitrags ist dabei die experimentell überprüfte Beurteilerübereinstimmung, mit der gezeigt werden konnte, dass das entwickelte Transkriptions- und Annotationsmanual geeignet ist, sprachliche Aussagen, Sprechpausen sowie die Erfassung der Adressatinnen und Adressaten einer Aussage/nonverbalen Handlung reliabel zu erfassen. Auch werden Übereinstimmungsmaße berichtet, die zeigen, dass das Manual auch hinsichtlich der Erfassung von Lehrkraftgesten, der Unterrichtseteiligung und Lehrkraftaufrufen sowie der Bestimmung der Sozialform gut geeignet ist.

Im dritten Beitrag wird die Anwendung des Instruments aus Beitrag #2 durchgeführt und die Lehr-/Lernprozesse in sprachlich heterogenen Lerngruppen thematisiert. Dabei konnte gezeigt werden, dass in den untersuchten Integrationskursen und Vorbereitungsklassen überwiegend Zeigegesten, metaphorische Gesten und bejahende und verneinende Kopfbewegungen zum Einsatz kommen. Mit Blick auf die Häufigkeit der eingesetzten Gesten wurde auch deutlich, dass die Zahl der verwendeten Gesten und die Art der Gesten stark zwischen den Lehrkräften variierte. Ein Vorteil, den die Aufbereitungsart des in Beitrag #2 entwickelten Instruments bietet, liegt in der Möglichkeit Lehr-/Lerninteraktion in ihrer Tiefe bis auf die individuelle Lehrenden-Lernenden-Ebene zu erschließen. Dabei zeigte sich, dass auch die Anzahl der Gesten, die die einzelnen Lernenden erhalten, innerhalb einer Lerngruppe stark variierte. Diese Varianz konnte durch die Kombination mit den durch die Lehrkraft wahrgenommenen sprachlichen Deutschfähigkeiten der Lernenden in Teilen erklärt werden. So konnten zwar kleine, aber doch signifikante Effekte der wahrgenommenen sprachlichen Fähigkeiten auf die Anzahl zeigender Gesten und in Teilen auch auf die Anzahl metaphorischer Gesten vorhergesagt werden, die eine Lernende/ein Lernender im Vergleich zu den Mitlernenden erhält. Zusammengefasst wurde

deutlich, dass Lehrkräfte den Lernenden, die sie für sprachlich schwächer halten, tendenziell mehr (deiktische) Gesten präsentieren, die bereits ohne ausgeprägtes Sprachverständnis verständlich sind, als den Lernenden, die sie für sprachlich stärker halten. Ebenfalls deuten die Daten darauf hin, dass Lernende, die von der Lehrkraft hinsichtlich ihrer grammatischen Fähigkeiten und ihrer allgemeinen sprachlichen Begabung für Deutsch eher hoch eingeschätzt werden, auch mehr komplexere (metaphorische) Gesten erhalten, die bereits ein höheres Sprachverständnis benötigen.

Mit dem ersten Beitrag konnte ein Forschungsdesiderat für die Erwachsenenbildung identifiziert werden, das bisher in keiner der benachbarten Disziplinen untersucht wurde. Zwar gibt es im Bereich der empirischen Schulbildungsforschung spätestens seit der DESI-Studie videogestützte Untersuchungen von Spracherwerbsprozessen, allerdings nur im Bereich Deutsch als Erstsprache: Dabei konnten systematisch Analysen zu den Zusammenhängen von beruflichen Kompetenzen und Wissen von Deutschlehrkräften mit Unterrichtshandeln durchgeführt werden (Kunze, 2004). Eine Beforschung von Spracherwerbssituationen im Bereich DaZ, in denen die sprachliche Heterogenität aufgrund der unterschiedlichen Hintergründe der Lernenden besonders hoch ist (Bundesamt für Migration und Flüchtlinge, 2021), gab es bis zur Durchführung der COLD-Videostudie für die Erwachsenenbildung nicht. Der in Beitrag #3 identifizierte fehlende Forschungsstand zu den „[...] Bedingungen der Möglichkeit innerer Differenzierung in oft sehr heterogenen Lerngruppen [...]“ (Sahlender & Schrader, 2017, S. 290) ist ein Thema, das auch in der Schulforschung als hochrelevant erachtet wird (KMK. Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2015). Dabei ist erwähnenswert, dass im Schulbereich die Bedeutung von Heterogenität eher zuzunehmen scheint. Bohl (2017) nennt für die zu erwartende Zunahme verschiedene Gründe, unter anderem neben der demografischen Entwicklung, die Abschaffung der Hauptschule, eine zunehmende Flexibilisierung von Bildungsverläufen und einen normativen Wandel zu einer Pädagogik der Vielfalt. Rückblickend lässt sich konstatieren, dass die Antwort auf die in Beitrag #1 identifizierten Forschungsdesiderate durch die im Projekt COLD erfassten Daten, die auch in Beitrag #2 und Beitrag #3 die Datengrundlage bilden, als Teilantworten auf das Review verstanden werden können. Durch interdisziplinäre Zusammenarbeit wurde ein Instrument zur empirischen Überprüfung von Lehr-/Lernprozessen entwickelt, das Hinweise auf die Gelingensbedingungen erfolgreicher Sprachprozesse ermöglicht. Darüber hinaus wurden im ersten Beitrag auch Schwerpunkte diskutiert, auf die in der Dissertationsschrift nicht weiter eingegangen wird, da sie für die beiden weiteren Beiträge nicht relevant waren.

Im Rahmen der Erstellung von Beitrag #2 konnte neben der Hauptfragestellung, der Gewährleistung einer methodisch reliablen Erfassung des Unterrichtsgeschehens, auch der Transkriptionsleitfaden erstellt, geschärft und die Grenzen der Transkriptionsmöglichkeiten verortet werden. Mit dem zweiten Beitrag wurde ein Instrument geschaffen, das die Beantwortung von vielen Untersuchungsfragen ermöglicht und es somit erlaubt, Forschungslücken⁸ zu schließen. In Kombination mit den zusätzlich erfassten Meta-Informationen erlauben die vorhandenen Daten eine Vielzahl an Analysen von Lehr-/Lernprozessen im Kontext sprachlich heterogener Lerngruppen. Durch die interdisziplinäre Zusammenarbeit wird es somit möglich, mit computerlinguistischen Methoden die Komplexität von Sprache im Kontext Unterricht zu untersuchen. Die technischen Grundlagen hierfür liefert der im Rahmen von Beitrag #2 entstandene Transkriptions- und Annotationsleitfaden, der es erlaubt, sprachliche Aussagen, Sprechpausen, die Adressatinnen und Adressaten von Aussagen und nonverbalen Handlungen, die Lehrkraftgesten, die Unterrichtsbeteiligung und die Sozialform zuverlässig zu erfassen. Die relevanten und von der Linguistik geforderten Robustheitsmaße (Wisniewski, 2022) wurden dabei transparent berichtet. Als Maß für die Transparenz der Transkription gesprochener Sprache wurde eine normalisierte Variante der Levenshtein-Distanz (Levenshtein, 1966) vorgeschlagen und zur Diskussion gestellt. Die Daten aus der Pilotierungsstudie sollen im Rahmen eines gesprochenen Lernerkorpora in vollständig anonymisierter, transkribierter und annotierter Form inklusive verschiedener Meta-Informationen anderen interessierten Wissenschaftlerinnen und Wissenschaftlern zugänglich gemacht werden. Weiterhin wurde der Transkriptions- und Annotationsfaden (als Anhang in Beitrag #2) in die englische Sprache übersetzt und veröffentlicht, um somit die Übertragung auf weitere Zweitspracherwerbs-Situationen auf andere Sprachen zu ermöglichen. Ein direkter Anwendungsbezug zur Erforschung von Lehr-/Lernprozessen ist im zweiten Beitrag bereits im Ausblick und dem Kapitel zu Forschungsmöglichkeiten skizziert: die Analyse von Lehrkraftgesten in sprachlich heterogenen Lerngruppen. An dieser Stelle sei angemerkt, dass es im Schulforschungsbereich bereits früher Forderungen und Versuche gab, die Rolle der nonverbalen Kommunikation stärker in den Fokus der Debatte um Unterrichtsqualität in der videogestützten empirischen Bildungsforschung zu rücken (z.B. bei Gröschner, 2007). Nun existiert mit dem in Beitrag #2 entwickelten Transkriptions- und Annotationsleitfaden die Möglichkeit, diesen Aspekt verstärkt zu beleuchten.

⁸ Eine Auswahl noch zu bearbeitender Desiderate befindet sich in Kapitel 9.2.

Dadurch konnte im dritten Beitrag mit einem Teildatensatz der COLD-Videostudie gezeigt werden, wie videogestützte Unterrichtsdaten genutzt werden können, um Lehr-/Lernprozesse in Integrationskursen und Vorbereitungsklassen so zu analysieren, dass Lehrhandeln besser verstanden werden kann. Die Ergebnisse zeigen, dass Sichtmerkmale, wie der Einsatz von Gesten, von Lehrkräften systematisch an Elemente der wahrgenommenen sprachlichen Fähigkeiten der Lernenden gekoppelt sind. Die Daten zeigen dabei, dass Lehrkräfte zumindest in Teilen die sprachliche Heterogenität ihrer Lerngruppen in ihrem Lehrhandeln berücksichtigten. Die Frage, ob die gewählten Lehrkraftgesten bewusst oder unbewusst von den Lehrkräften getätigt werden, kann aufgrund des Forschungsdesigns an dieser Stelle nicht beantwortet werden. Aus einer eher übergeordneten Perspektive sind die Befunde auch unterstützendes Argument für die videogestützte Unterrichtsbeobachtung als Mittel der Wahl zur Analyse von Lehr-/Lernprozessen. Denn Beitrag #3 hat bereits mit einer kleinen Stichprobe verdeutlicht, dass es unbedingt notwendig ist, nonverbale Kommunikation und ihre Rolle in Spracherwerbsprozessen von Deutsch als Zweitsprache zu berücksichtigen, um den Umgang von Lehrkräften mit sprachlich heterogenen Lerngruppen verstehen zu können.

Die Bedeutung der Ergebnisse und die damit einhergehenden Konsequenzen für die Unterrichtsforschung werden sichtbar, wenn man sie in das in das Angebots-Nutzungs-Modell der Wirkungsweise des Unterrichts (Helmke, 2017) einordnet. Die Wirkung im DaZ-Unterricht ist dabei zweifelsfrei die Verbesserung der DaZ-Fähigkeiten (in einem der in Kapitel 2.1 aufgezeigten Inhaltsbereiche). Diese wird durch die Lernaktivitäten der Lernenden beeinflusst. Deutlich wird aber, wenn man die Ergebnisse des dritten Beitrags verortet, dass die Lernaktivitäten von der Wahrnehmung und Interpretation des unterrichtlichen Angebots abhängen. Unter Berücksichtigung unterschiedlicher Lernpotenziale (Sprachen) und der Kontextvariable Klassenzusammensetzung ist anzunehmen, dass die Lerngruppe aufgrund ihrer sprachlichen Heterogenität im DaZ-Unterricht über unterschiedlich ausgeprägte Fähigkeiten der Wahrnehmung und Interpretation des unterrichtlichen Angebots verfügt. Hier konnte gezeigt werden, dass über alle Lehrkräfte hinweg eine systematische Anpassung der nonverbalen Kommunikation an die unterschiedlichen sprachlichen Fähigkeiten der Lernenden erfolgt ist. Es scheint also, dass Lehrkräfte in der Lage sind, in ihrem Unterricht die unterschiedlichen sprachlichen Leistungsstände der Lernenden differenziert und adaptiv zu adressieren. Gleichwohl die Ursachen für die Fähigkeiten zur adaptiven Unterrichtsgestaltung im Rahmen des dritten Beitrags nicht geklärt werden können, weil unklar ist, ob die Anpassungen bewusst oder unbewusst geschehen, verweisen die Daten

darauf, dass es vielversprechend ist, die nonverbalen Gesten als relevantes Merkmal der Differenzierung in sprachlich heterogenen Lerngruppen zu berücksichtigen.

Im aktuellen Diskurs bezüglich des immer noch prominent genutzten Angebots-Nutzungs-Modells der Wirkungsweise des Unterrichts hat sich in der empirischen Schulbildungsforschung die Unterscheidung in eine Struktur- und eine Prozessperspektive des Modells herausgebildet (Seidel, 2014; Seidel et al., 2021; Seidel & Shavelson, 2007). Der Unterschied besteht nach Seidel et al. (2021) darin, dass man in der Strukturperspektive den Unterricht als natürliches Feld beobachtet, indem man Unterrichtspraxis und Unterrichtsqualität analysiert und auf die Erreichung bestimmter Bildungsziele hin überprüft. In der Prozessperspektive werden „die Qualität von Unterricht über konkrete, im günstigen Fall theorie- und evidenzbasierte Interventionen und experimentell angelegte Manipulationen bestimmt“ (Seidel et al., 2021, S. 298) und damit die Überprüfung der strukturierten Änderung des Lehr- oder des Lernhandelns in den Vordergrund gestellt (Seidel et al., 2021). Die in Beitrag #3 gewonnenen Ergebnisse lassen sich, wie auch die Forschung zu Basisdimensionen der Unterrichtsqualität (Pauli & Reusser, 2006), der Strukturperspektive zuordnen, da Videoaufzeichnungen von authentischen Unterrichtssituationen genutzt wurden, um Unterrichtspraktiken zu beschreiben und die dahinterliegenden Lehr-/Lernprozesse zu verstehen. Somit konnte durch die Analyse nonverbaler Kommunikationsprozesse die Rolle von Gesten im Unterricht messbar gemacht werden. Dieser erste Schritt des Verstehens der Wirkungsweise von Lehrkraftgesten kann und sollte die Grundlage bilden, sich zukünftig auch aus der zweiten Perspektive, dem Prozessparadigma, der Thematik zu nähern und die Qualität von Unterricht durch theorie- und evidenzbasierte Interventionen beim adaptiven Einsatz von Lehrkraftgesten zu untersuchen. Damit könnte ein in Beitrag #1 aufgezeigtes Desiderat zu fehlenden empirischen Befunden über die Gelingensbedingungen wirksamer Spracherwerbsprozesse bearbeitet werden. Das in Beitrag #2 entwickelte Instrument bietet dabei ein Hilfsmittel, um die zu messenden Lehr-/Lernprozesse reliabel zu erfassen.

9.2. Kritische Betrachtung der Studien und Desiderate für zukünftige Forschung

Die drei in dieser Arbeit präsentierten Beiträge haben durch ihre methodische Anlage auch Limitationen, die im Folgenden benannt werden sollen. Die Befunde des ersten Beitrags beschränken sich auf einen eingeschränkten Berichtszeitraum von 2010 bis 2016. Durch diesen zeitlich eingeschränkten Einblick ist die Abbildung der Beforschung

der Untersuchung von Spracherwerbsprozessen nicht in Gänze möglich. Im Rahmen des Reviews konnten auch nur typische Befundmuster für einzelne Schwerpunkte (Lehrende, Lernende, Lehr-/Lernprozesse und digitale Medien) berichtet werden. Die Auswahl typischer Publikationen erfolgte dabei gemeinsam durch beide Autoren. Ebenso muss die Beschränkung auf den deutschsprachigen und englischsprachigen Raum als Einschränkung der Generalisierbarkeit der Befunde benannt werden. So können die Befunde das Feld der Spracherwerbsforschung nur in diesen beiden Sprachbereichen beschreiben. Schlussendlich muss bei der Einschätzung der Befunde berücksichtigt werden, dass der große Bereich der Schulforschung zum Spracherwerb ausgeklammert wurde, da der Spracherwerb bei Erwachsenen explizit im Fokus der Untersuchung stand. Für einen umfassenden Überblick der einzelnen Spracherwerbsbereiche (Deutsch als Zweitsprache, Deutsch als Fremdsprache und Alphabetisierung) ist ein Review, das auch ältere Beiträge und die Schulforschung miteinschließt, unbedingt empfehlenswert. Für die dem Beitrag zugrundeliegende Fragestellung, wie Spracherwerb speziell im Bereich der Erwachsenenbildung beforscht wird, war die methodische Anlage jedoch zielführend und die Erkenntnisse zur Einordnung der Forschungslandschaft hilfreich. So gelang durch das Review eine Momentaufnahme, die es ermöglicht hat, aktuelle Forschungstrends und -desiderate zu identifizieren.

Da Beitrag #2 und Beitrag #3 auf derselben Datengrundlage beruhen, unterliegen sie auch denselben datenspezifischen Limitationen. Die Lehrkräfte und ihre Lernenden, die an der COLD-Pilotstudie teilgenommen haben, entsprechen keiner repräsentativen Stichprobe. Es sei angemerkt, dass dies ist für die Fragestellungen des Projekts jedoch auch nicht zentral erforderlich ist, da eine grundlegende Untersuchung der Komposition verschiedene Kompetenzbereiche bei Lehrkräften im DaZ-Unterricht im Vordergrund stand. Die Teilnahme war freiwillig und ist nur dann erfolgt, wenn sowohl Lehrkraft als auch alle im Unterricht sichtbaren Lernenden mit der Teilnahme einverstanden waren. In den Vorbereitungsklassen wurde hierzu zudem das Einverständnis der jeweiligen Erziehungsberechtigten eingeholt. Da eine Teilnahmeanfrage zur Studie bei einigen angefragten Lehrkräften und ihren Lernenden auf Zurückhaltung und Ablehnung gestoßen ist, muss davon ausgegangen werden, dass die Gelegenheitsstichprobe kein repräsentatives Abbild der Gesamtheit der DaZ-Lehrenden in Vorbereitungsklassen und Integrationskursen darstellt, dazu ist der Umfang der Pilotierungsstudie mit zehn beteiligten Lehrkräften ohnehin zu gering. Da im Rahmen der Instrumentenentwicklung die Qualität der Lehr-/Lernprozesse, die eventuell durch die Teilnahme von eher engagierten, selbstbewussten und kompetenten Lehrkräften überschätzt werden könnte, keine zentrale Rolle spielt, sollte die

nichtrepräsentative Stichprobe keine große Limitation sein. Hinsichtlich der Anpassung der nonverbalen Kommunikation kann nicht ausgeschlossen werden, dass die Fähigkeiten zu adaptivem Unterrichtshandeln bei Lehrkräften durch die Gelegenheitsstichprobe überschätzt werden. Die im Rahmen der Datenerhebung und in Beitrag #2 aufgeführten umfassenden Meta-Informationen und standardisierten Messungen der Kompetenzen der Lehrkräfte können dabei helfen, die Lehrkräfte detailliert zu beschreiben und die Stichprobe näher einzuordnen.

Weiter ist erwähnenswert, dass die Beobachtungen auf einem einzigen Messzeitpunkt, und zwar auf einer 45-minütigen Unterrichtsstunde, beruhen. Wie eingangs erwähnt, ist dieser Zeitraum ausreichend lange, um Klassenmanagement und konstruktive Unterstützung als Dimensionen von Unterrichtsqualität zu untersuchen. Die erfolgreiche Erfassung der kognitiven Aktivierung kann durch nur einen Messzeitpunkt nicht gewährleistet werden (Praetorius et al., 2014). Mit Blick auf die Anlage des COLD-Projekts und auf das Ziel, die Kompetenzen von Lehrkräften beim Unterrichten von DaZ zu erfassen, scheint die Priorität auf einer detaillierten Kompetenzerfassung der Lehrkräfte, zu Lasten einer zusätzlichen Unterrichtsbeobachtung, nachvollziehbar zu sein. Dass eine längsschnittliche Datenerhebung weitere hochrelevante Einblicke in die Kompetenzentwicklung von Lehrenden und Lernenden ermöglicht hätte, steht dabei außer Frage. Jedoch ist anzunehmen, dass mit dem höheren Arbeitsaufwand für die an der Studie teilnehmenden Lehrkräfte auch die Gewinnung von Studienteilnehmenden schwerer geworden wäre. Um die Hürde zur Studienteilnahme zu erleichtern und gleichzeitig einen Einblick in authentischen DaZ-Unterricht zu erhalten, wurde den Lehrkräften die inhaltliche Ausgestaltung der Unterrichtsstunden weitestgehend selbst überlassen, was einen direkten Vergleich der Unterrichtsstunden⁹ erschwert. Wie in Beitrag #2 aufgezeigt, wurde die Lehrkraft lediglich gebeten, neben einer Plenumsphase noch mindestens eine weitere Sozialform durchzuführen.

Durch den Fokus auf die Lehrkräfte und ihre Rolle bei der Gestaltung bei Lehr-/Lernprozessen wurde bei der technischen Realisation auf zusätzliche Mikrofone für jeden einzelnen Lernenden verzichtet. Ebenfalls wurden die Lernenden nur mit einer Kamera gefilmt, so dass sie nur in der Totalen sichtbar sind. Die Lehrkraft hingegen wurde mit einem zusätzlichen Mikrophon und einer eigenen Kamera aufgenommen. Das hat zur Folge, dass alle Interaktionen zwischen der Lehrkraft und den einzelnen Lernenden nahezu

⁹ wie es z.B. in der TALIS Studie anhand von Unterricht zu mathematischen Gleichungen gemacht wurde (Grünkorn et al., 2020).

lückenlos (mit nur wenigen unverständlichen Sätzen) transkribiert werden konnte. Die Interaktionen, die ohne Beteiligung der Lehrkräfte zwischen den einzelnen Lernenden stattgefunden haben, wurden nicht transkribiert, da diese, je nach Anzahl der verschiedenen Lerngruppen im Unterricht, häufig gleichzeitig stattgefunden haben und akustisch häufig so schlecht verständlich waren, dass keine reliable Erfassung der gesprochenen Sprache möglich war.

Die Datenerhebung hat in großen Teilen während der Covid-19-Pandemie stattgefunden. Die im Rahmen der COLD-Studie durchgeführten teilstandardisierten Interviews, die zum Ende der Studie mit den Lehrkräften geführt worden sind, haben ergeben, dass es bei der Durchführung der Unterrichtsstunde Abweichungen hinsichtlich der Interaktionsmuster gegeben hat. So berichteten die Lehrkräfte, dass sie mehr Plenumsphasen und weniger interaktive Gruppen- und Partnerarbeiten durchführten, um die gesetzlich vorgegebenen Mindestabstände einzuhalten. Darüber hinaus waren auch manche der teilnehmenden Lerngruppen kleiner als vor der Covid-19-Pandemie. Das lag zum Teil an Wechselunterrichtsmodellen oder an geänderten Höchstbelegungszahlen für Unterrichtsräume aufgrund des einzuhaltenden Mindestabstands zur Covid-19-Prävention. Die aufgrund der Pandemie von den Lernenden getragenen Gesichtsmasken haben, wie in Beitrag #2 aufgezeigt, kaum Einfluss auf die Identifizierbarkeit der sprechenden Lernenden.

Ob das Tragen von Gesichtsmasken bei Lehrkräften einen Einfluss hat, ist nicht untersucht worden. In der vorliegenden Stichprobe sind bei den Integrationskursen Unterrichtseinheiten enthalten, in denen die Lehrkräfte eine Gesichtsmaske tragen. Es könnte angenommen werden, dass das Tragen einer Gesichtsmaske bei der Lehrkraft die Häufigkeit der Verwendung von Lehrkraftgesten beeinflusst. Zum einen, da die Lehrkraft sich vielleicht dessen bewusst ist, dass ihre Sprache durch den verdeckten Mund nicht mehr so gut verständlich ist, und sie deshalb mehr Gesten verwendet. Zum anderen könnte es auch daran liegen, dass die Lehrkraft berücksichtigt, dass die Lernenden ihren Mund nicht sehen können, und somit die Lehrkraft die Hände zur Verdeutlichung des Gesagten verwendet. Lee und Hart (2022) konnten jüngst empirisch belegen, dass besonders schwächere Lernende im Zweitspracherwerb in ihrer Leistungsentwicklung benachteiligt sind, wenn die Lehrkraft eine Maske trägt. Auch muss erwähnt werden, dass die Analysen in Beitrag #3 aufgrund der geringen Stichprobengröße die den Daten zugrundeliegende genestete Datenstruktur, die eigentlich eine Mehrebenenanalyse (Eid et al., 2013) erfordert, nicht berücksichtigen. Um die gemeinsame Varianz innerhalb der Lerngruppen statistisch dennoch zu berücksichtigen, wurden die Berechnungen mit Abweichungen von den zentrierten Gruppenmittelwerten der Klassen und Kurse durchgeführt. Wie die Befunde zu

Lehrkraftgesten in weitere Unterrichtsfächer oder Kurse übertragen werden können, ist eine Ungewissheit, die eine weitere Grenze der Befunde darstellt. Die Arbeiten von Praetorius et al. (2020) verweisen dabei auf empirische Befunde zur Fachspezifik von Unterrichtsqualität sowie auf die Erfordernis, Elemente des Lehr-/Lerngeschehens im Kontext von Unterrichtsinhalten zu betrachten. Da dies im Moment noch nicht erfolgt ist, ebenso wenig wie die Analyse der Kombination von Lehrkraftgesten und unterschiedlichen Sozialformen im Unterricht, können hierzu noch keine Aussagen getroffen werden. Ebenfalls ist die Qualität der eingesetzten Lehrkraftgesten kein Teil der empirischen Untersuchungen gewesen. Die Analysen beziehen sich auf die anteilige Verteilung verschiedener Gestentypen, die unterschiedliche Lernende erhalten. Aussagen über die Relevanz oder Qualität, im Sinne einer pädagogischen Angemessenheit, sind in der Analyse nicht enthalten. Auch sollte noch untersucht werden, wie der Einsatz von Lehrkraftgesten mit dem Kompetenzerwerb, das heißt hier im konkreten Fall mit der Entwicklung der DaZ-Fähigkeiten der Lernenden zusammenhängt. Hierzu wären longitudinale Daten mit mehreren Messungen zu den DaZ-Fähigkeiten oder gar eine Intervention mit gezielter Variation der Lehrkraftgesten hilfreich.

Ebenso wird die Analyse von Lehr-/Lernprozessen durch die Authentizität der untersuchten Daten erschwert. Denn die Aufnahme von realem Unterricht bringt eine hohe Zahl an potenziellen Störfaktoren und wenig kontrollierbaren, komplexen Unterrichtssituationen mit sich. Da die Hauptfragestellung von Beitrag #2 und Beitrag #3 jedoch explizit das Unterrichtsgeschehen thematisieren, ist eine Untersuchung unter besser kontrollierbaren Bedingungen, z.B. in Laborsituationen, kaum möglich. Ziel des dritten Beitrags war es, die theoretischen Annahmen über die Vermutungen der Wirkungsweisen zu Lehrkraftgesten in ökologisch validen Feldbedingungen zu untersuchen, da in Laborstudien (Adams, 1998; Tellier et al., 2021) bereits nachgewiesen wurde, dass Sprechende ihre Gesten an den sprachlichen Fähigkeiten der Zuhörenden ausrichten. Die Unterschiede wurden aber nur zwischen zwei Gruppen (Personen mit derselben Herkunftssprache vs. Personen mit anderer Herkunftssprache als die Sprechenden) unterschieden. Darauf aufbauend erschien es vielversprechend, die künstliche Einteilung in Experimental- und Kontrollgruppe aufzuheben und das Verhalten im Feld zu betrachten, um den Einsatz von Gesten der Sprechenden in Abhängigkeit der wahrgenommenen sprachlichen Fähigkeiten der Zuhörenden besser zu verstehen. Hierfür halten die Daten noch weitere ergänzende Meta-Informationen bereit, die dabei helfen können, das nonverbale Handeln der Lehrkraft detaillierter zu verstehen.

Perspektivisch ist geplant, die Daten der Pilotierungsstudie im Rahmen eines gesprochenen Lernerkorpora zu veröffentlichen. Dadurch können die transkribierten und annotierten Unterrichtsdaten in anonymisierter Form anderen interessierten Wissenschaftlerinnen und Wissenschaftlern zugänglich gemacht werden. Der interdisziplinäre Ansatz, der bei der Entwicklung des Transkriptions- und Annotationsmanuals verfolgt wurde, hat es ermöglicht, einen Datensatz bereitzustellen, der so aufbereitet wurde, dass er den Anforderungen einer Analyse von Lehr-/Lernprozessen in sprachlich heterogenen Lerngruppen in DaZ in Schule und Erwachsenenbildung aus der Perspektive der empirischen Bildungsforschung, der DaZ-Didaktik und der Computerlinguistik genügt. Die im Rahmen der COLD-Hauptstudie durchgeführte Datenerhebung umfasst weitere Unterrichtsaufnahmen aus Vorbereitungsklassen ($N = 30$) und Integrationskursen ($N = 30$), so dass der Korpus perspektivisch auf $N = 70$ ($N = 10$ Videos aus der Pilotierungsstudie und $N = 60$ Videos aus der Hauptstudie) DaZ-Kurse anwachsen kann. Die größere Stichprobe ermöglicht die Berücksichtigung der genesteten Datenstruktur, so dass die Berechnung von Mehrebenenanalysen möglich wird. Außerdem wird dadurch auch möglich, verschiedene Kurstypen miteinander zu vergleichen.

Die Befunde des dritten Beitrags werfen auch viele Fragen auf, die es zukünftig zu bearbeiten gilt. Denn wenn man die Communications-Accommodation-Theory (Giles, 2016) berücksichtigt und versucht, den Gesteneinsatz theoretisch dadurch zu erklären, dass Lehrkräfte ihre Gesten so auswählen, dass sie zu den sprachlichen Fähigkeiten der jeweiligen Lernenden passen, ist das ein Hinweis auf die Wichtigkeit einer angemessenen Diagnosefähigkeit der sprachlichen Fähigkeiten. So sollte überprüft werden, ob Lehrkräfte, die eine falsche Einschätzung hinsichtlich der sprachlichen Fähigkeit von Lernenden treffen, auch weniger differenziert nonverbal kommunizieren. Ebenfalls legen die Befunde nahe, die Verwendung von Lehrkraftgesten im Kontext der Komplexität sprachlicher Äußerungen zu betrachten. Eine Hauptfragestellung des COLD-Projekts untersucht die Frage, ob Lehrkräfte ihre sprachlichen Äußerungen an die sprachlich heterogenen Lernenden anpassen können. Die Kombination verbaler und nonverbaler Anpassungen könnte Hinweise darauf geben, ob Lehrkräfte Gesten z.B. verstärkend zu sprachlichen Äußerungen einsetzen, so dass sie sprachlich schwachen Lernenden mit einfacher Sprache begegnen und gleichzeitig auch einfache Gesten verwenden. Es könnte aber auch so sein, dass Lehrkräfte ihre Sprache nicht anpassen, aber dafür die Gesten variieren. Wäre Letzteres der Fall, könnten Gesten eine Art kompensatorische Funktion haben.

Schließlich sollten die Lehrkraftgesten auch aus der Perspektive ihrer Relevanz bzw. der Angemessenheit der Geste beurteilt werden. Im dritten Beitrag wurden die

Gesten nicht dahingehend unterschieden, ob sie inhaltlich passend sind und in der fachdidaktischen DaZ-Situation angemessen erscheinen. Mit Hilfe von Experteneinschätzungen könnten hier Einblicke gewonnen werden, welche Lehrkräfte Gesten hochrelevant einsetzen und welche Lehrkräfte Entwicklungspotenzial haben. In Kombination mit den vorliegenden Informationen über bestehende Kompetenzen hinsichtlich allgemein pädagogischem Unterrichtswissen und (fach-)didaktischen Kenntnissen können hier wertvolle Informationen gewonnen werden. Diese erscheinen vielversprechend, weil sie Hinweise darauf geben können, wie Lehrkräfte in ihrem Unterrichtshandeln unterstützt werden können, zum Beispiel im Rahmen einer Fortbildung zu bestimmten pädagogischen Themen. Weiterhin können Analysen der Kombination von Gesteneinsatz und Expertise der Lehrkräfte (im Sinne beruflicher Erfahrung) Hinweise darauf geben, ob der Einsatz von Gesten im Laufe der Berufserfahrung zunimmt und ein angemessener Gesteneinsatz erlernbar ist. Schlussendlich können die Lehrkraftgesten auch noch unter einer gänzlich anderen Perspektive neu betrachtet werden. So wie die Gesten in dieser Studie kategorisiert worden sind, wurde jede Geste auf ihren Wert hinsichtlich der Förderung der DaZ-Fähigkeiten der Lernenden analysiert. Diese Fokussierung auf die angenommene Lernwirksamkeit einer Geste ist nur eine von mehreren Funktionen von Gesten. Wie von Prange bereits beschrieben, hat das Zeigen auch eine strukturierende Funktion im Unterricht (Fuhr, 1999). Man könnte auch aus der Perspektive der Unterrichtsqualitätsdimension z.B. alle Gesten, die ein Feedback enthalten, den Gesten der konstruktiven Unterstützung zuweisen und Gesten, die eher eine Monitoring-Funktion haben (indem die Lehrkraft das Unterrichtsgeschehen durch Redeaufforderungen und Redeunterlassungsaufforderungen nonverbal steuert), zur strukturierten Klassenführung zählen.

9.3. Fazit

Zusammenfassend konnte im Rahmen der vorliegenden Dissertationsschrift das Themenfeld der Lehr-/Lernprozesse im DaZ-Unterricht in sprachlich heterogenen Lerngruppen in den Bildungsbereichen Erwachsenenbildung und Schule differenziert beleuchtet werden. Dabei wurde zunächst der Forschungsstand aufgearbeitet und Forschungsdesiderate identifiziert (Beitrag #1). Das entwickelte Instrument zur Analyse von DaZ-Unterricht (Beitrag #2) hat es ermöglicht, die Bedeutung von Lehrkraftgesten als relevanten Teil des Lehr-/Lernhandelns einzuordnen (Beitrag #3). Mit Hilfe des Angebots-Nutzungs-Modells der Wirkungsweise des Unterrichts konnte die besondere theoretische Bedeutung und Relevanz von Kommunikation im DaZ-Unterricht veranschaulicht werden und

Konsequenzen für Lehrkräfte für die Anpassung der Sprache und der nonverbalen Kommunikation an sprachlich heterogene Lerngruppen abgeleitet werden. Dass auch nonverbale Kommunikationsanpassungen in Form von Lehrkraftgesten im Unterricht stattfinden, ist ein starkes Indiz dafür, dass Lehrkräfte die sprachliche Heterogenität ihrer Lernenden berücksichtigen. Die vorliegende Arbeit und das damit einhergehende Instrument ermöglichen es, DaZ-Unterricht, unter besonderer Berücksichtigung der sprachlichen Heterogenität der Lerngruppe, zu erfassen und zu quantifizieren, um schlussendlich die Wirkungsweise von DaZ-Spracherwerbsprozessen Stück für Stück besser zu verstehen.

10. Literaturverzeichnis

- Ackermann, S. (2011). *Klassengespräch im Mathematikunterricht: Eine Pilotstudie im Rahmen des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern"*. Reihe Studium und Forschung: Bd. 19. Universitätsverlag Kassel.
- Adams, T. W. (1998). *Gesture in foreigner talk*. Ph.D. Dissertation. University of Pennsylvania.
- Bailey, A. L. (Hrsg.). (2007). *The language demands of school: Putting academic English to the test*. Yale University Press.
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, M. Neubrand, U. Klusmann & S. Krauss (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–54). Waxmann.
- Becker-Mrotzek, M. & Roth, H.-J. (2017). Sprachliche Bildung - Grundlegende Begriffe und Konzepte. In M. Becker-Mrotzek & H.-J. Roth (Hrsg.), *Sprachliche Bildung: Bd. 1. Sprachliche Bildung - Grundlagen und Handlungsfelder* (S. 11–36). Waxmann.
- Bleakley, H. & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *The Review of Economics and Statistics*, 86, 481–496.
<https://doi.org/10.1162/003465304323031067>
- Bohl, T. (2017). Umgang mit Heterogenität im Unterricht: Forschungsbefunde und didaktische Implikationen. In T. Bohl, J. Budde & M. Rieger-Ladich (Hrsg.), *UTB Schulpädagogik: Bd. 4755. Umgang mit Heterogenität in Schule und Unterricht: Grundagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen* (S. 257–274). Verlag Julius Klinkhardt.
- Bohl, T. & Kucharz, D. (2010). *Offener Unterricht heute: Konzeptionelle und didaktische Weiterentwicklung. Pädagogik Unterricht: Bd. 22*. Beltz.
- Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research*, 10(1), 3–31. <https://doi.org/10.1191/1362168806lr182oa>
- Bromme, R., Prenzel, M. & Jäger, M. (2017). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. In J. Baumert & K.-J. Tillmann (Hrsg.), *Empirische Bildungsforschung* (S. 129–146). Springer Fachmedien Wiesbaden.
https://doi.org/10.1007/978-3-658-13785-4_8

- Budde, J. (2015). Heterogenitätsorientierung. In J. Budde, N. Blasse, A. Bossen & G. Rißler (Hrsg.), *Edition Erziehungswissenschaft. Heterogenitätsforschung: Empirische und theoretische Perspektiven* (S. 19–37). Beltz.
- Bundesamtes für Migration und Flüchtlinge (2007). *Konzeption für die Zusatzqualifizierung von Lehrkräften im Bereich Deutsch als Zweitsprache*.
<https://www.bamf.de/SharedDocs/Anlagen/DE/Downloads/Infothek/Integrationskurse/Lehrkraefte/konzeption-fuer-die-zusatzqualifikation-von-lehrkraeften-pdf.html>
- Bundesamt für Migration und Flüchtlinge (2020). *Matrix - Zulassungskriterien für Lehrkräfte in Integrationskursen*. https://www.bamf.de/SharedDocs/Anlagen/DE/Integration/Integrationskurse/Lehrkraefte/matrix-zulassung-lehrkraefte-integrationskurse-ab-01102020.pdf?__blob=publicationFile&v=7
- Bundesamt für Migration und Flüchtlinge (2021). *Minas: Atlas über Migration, Integration und Asyl*. https://www.bamf.de/SharedDocs/Anlagen/DE/Statistik/Migrationsatlas/migrationsatlas-2021-11.pdf?__blob=publicationFile&v=4
- Bundesamt für Migration und Flüchtlinge. (2022). *Bericht zur Integrationskursgeschäftsstatistik für das Jahr 2021*. Referat Statistik und Datenanalyse.
https://www.bamf.de/SharedDocs/Anlagen/DE/Statistik/Integrationskurszahlen/Bundesweit/2021-integrationskursgeschaeftsstatistik-gesamt_bund.pdf?__blob=publicationFile&v=2
- Cindark, I., Deppermann, A., Hünlich, D., Lang, C., Perlmann-Balme, M. & Schöningh, I. (2019). *Perspektive Beruf. Mündliche Kompetenz von Teilnehmenden an Integrationskursen und Vorschläge für die Praxis*. Leibniz-Institut für Deutsche Sprache und Goethe-Institut Mannheim. <https://doi.org/10.14618/ids-pub-9202>
- Clark, H. H. & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50(1), 62–81.
<https://doi.org/10.1016/j.jml.2003.08.004>
- Clercq, B. de & Housen, A. (2019). The development of morphological complexity: A cross-linguistic study of L2 French and English. *Second Language Research*, 35(1), 71–97. <https://doi.org/10.1177/0267658316674506>
- Dalehefte, I. M. & Kobarg, M. (2012). Einführung in die Grundlagen systematischer Videoanalysen in der empirischen Bildungsforschung. In M. Gläser-Zikuda, T. Seidel, C. Rohlf's & A. Gröschner (Hrsg.), *Mixed methods in der empirischen Bildungsforschung* (S. 15–26). Waxmann.

- Dargue, N., Sweller, N. & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological bulletin*, 145(8), 765–784. <https://doi.org/10.1037/bul0000202>
- Denn, A.-K., Gabriel-Busse, K. & Lipowsky, F. (2019). Unterrichtsqualität und Schülerbeteiligung im Mathematikunterricht des zweiten Schuljahres. In K. Verrière & L. Schäfer (Hrsg.), *Research. Interaktion im Klassenzimmer: Forschungsgeleitete Einblicke in das Geschehen im Unterricht* (1. Aufl., S. 9–29). Springer VS. https://doi.org/10.1007/978-3-658-23173-6_2
- Demmig, S. (2008). Binnendifferenzierung und Heterogenität. *Deutsch als Zweitsprache*, 4, 34-39.
- Deutscher Bundestag. (2019). *Unterrichtung durch die Bundesregierung: Bericht der Bundesregierung über die Fortschritte der Nationalen Dekade für Alphabetisierung und Grundbildung 2016 bis 2026*. Drucksache 19/14880. https://www.alphadekade.de/alphadekade/shareddocs/downloads/files/drs-1914880.pdf?__blob=publicationFile&v=1
- Dewitz, N. von & Bredthauer, S. (2020). Gelungene Übergänge und ihre Herausforderungen – von der Vorbereitungs- in die Regelklasse. *Informationen Deutsch als Fremdsprache*, 47(4), 429–442. <https://doi.org/10.1515/infodaf-2020-0063>
- Echarti, N., Huntemann, H., Lux, T. & Reichart, E. (2022). *Volkshochschul-Statistik: 59. Folge, Berichtsjahr 2020* (1. Aufl.). *DIE Survey. Daten und Berichte zur Weiterbildung*. wbv Publikation. <https://elibrary.utb.de/doi/book/10.3278/9783763970681>
- Eid, M., Gollwitzer, M. & Schmitt, M. (2013). *Statistik und Forschungsmethoden: Lehrbuch. Mit Online-Materialien* (3. Aufl.). Beltz.
- El-Mafaalani, A., Jording, J. & Massumi, M. (2020). Bildung und Flucht. In U. Bauer, U. H. Bittlingmayer & A. Scherr (Hrsg.), *Springer eBook Collection. Handbuch Bildungs- und Erziehungssoziologie* (S. 1–19). Springer Fachmedien Wiesbaden; Imprint Springer VS. https://doi.org/10.1007/978-3-658-31395-1_67-1
- Esser, H. (2006). *Sprache und Integration: Die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten*. Campus-Verl.
- Evertson, C. M. & Green, J. L. (1986). Observation as inquiry and method. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching: A project of the American Educational Research Association* (3. Aufl., S. 162–243). Macmillan.
- Faltis, C. & Valdés, G. (2016). Preparing teachers for teaching in and advocating for linguistically diverse classrooms: A vade mecum for teacher educators. In C. A. Bell

- & D. H. Gitomer (Hrsg.), *Handbook of research on teaching* (S. 549–592). American Educational Research Association.
- Fuhr, T. (1999). Zeigen und Erziehung. Das Zeigen als "zentraler Gegenstand" der Erziehungswissenschaft. In T. Fuhr (Hrsg.), *Zur Sache der Pädagogik: Untersuchungen zum Gegenstand der allgemeinen Erziehungswissenschaft* (S. 109–121). Klinkhardt.
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom*. Heinemann.
- Giles, H. (2016). Communication accommodation theory. In K. B. Jensen, E. W. Rothenbuhler, J. D. Pooley & R. T. Craig (Hrsg.), *The International Encyclopedia of Communication Theory and Philosophy* (S. 1–7). Wiley.
<https://doi.org/10.1002/9781118766804.wbiect056>
- Gogolin, I. (2008). *Der monolinguale Habitus der multilingualen Schule*. Zugl.: Hamburg, Universität, Habil.-Schr., 1991 (2. Aufl.). *Internationale Hochschulschriften: Bd. 101*. Waxmann.
- Gogolin, I., & Lange, I. (2011). Bildungssprache und durchgängige Sprachbildung. *Migration und schulischer Wandel: Mehrsprachigkeit*, 107-127.
- Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child development*, 71(1), 231–239.
<https://doi.org/10.1111/1467-8624.00138>
- Götz, T., Lohrmann, K. & Ganser, B. (2005). Einsatz von Unterrichtsmethoden - Konstanz oder Wandel? *Empirische Pädagogik*, 19(4), 342–360.
- Granger, S., Gilquin, G. & Meunier, F. (Hrsg.). (2015). *The Cambridge handbook of learner corpus research*. Cambridge University Press.
- Gröschner, A. (2007). Körpersprache im Unterricht. Perspektiven einer kommunikationsorientierten Bildungsforschung mithilfe von Unterrichtsvideos. *Zeitschrift Bildungsforschung e.V.*, 4(2), 1-21. <https://doi.org/10.25656/01:4614>
- Grotlüschen, A. & Buddeberg, K. (Hrsg.). (2020). *LEO 2018: Leben mit geringer Literalität*. wbv.
- Grotlüschen, A., Buddeberg, K., Dutz, G., Heilemann, L. & Stammer, C. (2020). Hauptergebnisse und Einordnung zur LEO-Studie 2018 – Leben mit geringer Literalität. In A. Grotlüschen & K. Buddeberg (Hrsg.), *LEO 2018: Leben mit geringer Literalität* (S. 13–64). wbv.
- Guo, Y., Sun, S., Breit-Smith, A., Morrison, F. J. & Connor, C. M. (2015). Behavioral engagement and reading achievement in elementary-school-age children: A

- longitudinal cross-lagged analysis. *Journal of Educational Psychology*, 107(2), 332–347. <https://doi.org/10.1037/a0037638>
- Guven, C. & Islam, A. (2015). Age at migration, language proficiency, and socio-economic outcomes: evidence from Australia. *Demography*, 52(2), 513–542. <https://doi.org/10.1007/s13524-015-0373-6>
- Hall, J. K. (2022). L2 classroom interaction and its links to L2 learners' developing L2 linguistic repertoires: A research agenda. *Language Teaching*, 55(1), 100–115. <https://doi.org/10.1017/S0261444820000397>
- Hammond, J. & Gibbons, P. (2005). Putting scaffolding to work: the contribution of scaffolding in articulating ESL education. *Prospect*, 20(1), 6–30.
- Hardy, I., Decristan, J., & Klieme, E. (2019). Adaptive teaching in research on learning and instruction. *Journal for Educational Research Online*, 11, 169–191. <https://doi.org/10.25656/01:18004>
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (7. Aufl.). *Schule weiterentwickeln, Unterricht verbessern Orientierungsband*. Klett/Kallmeyer.
- Helmke, A., Helmke, T., Kleinbub, I., Nordheider, I., Schrader, F.-W. & Wagner, W. (2007). Die DESI-Videostudie. Unterrichtstranskripte für die Lehrerbildung nutzen. *Der fremdsprachliche Unterricht. Englisch*, 41(90), 37–45.
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G. & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In E. Klieme (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 345–363). Beltz.
- Hochman, O. & Davidov, E. (2014). Relations between second-language proficiency and national identification: The case of immigrants in Germany. *European Sociological Review*, 30(3), 344–359. <https://doi.org/10.1093/esr/jcu043>
- Horn, H., Lux, T. & Christ, J. (2021). *Weiterbildungsstatistik im Verbund – Ergebnisse für das Berichtsjahr 2019*. wbv Media. <https://doi.org/10.3278/85/0027w>
- Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological bulletin*, 137(2), 297–315. <https://doi.org/10.1037/a0022128>
- Hünlich, D., Wolfer, S., Lang, C. & Deppermann, A. (2018). *Wer besucht den Integrationskurs? Soziale und sprachliche Hintergründe von Geflüchteten und anderen Zugewanderten*. Institut für Deutsche Sprache/Goethe Institut.

- Jeuk, S. (2015). *Deutsch als Zweitsprache in der Schule: Grundlagen - Diagnose - Förderung* (3. Aufl.). Kohlhammer.
- Kade, J., Nolda, S., Dinkelaker, J. & Herrle, M. (Hrsg.). (2014). *Videographische Kursforschung: Empirie des Lehrens und Lernens Erwachsener*. Kohlhammer.
- Karakayali, J. & Heller, M. (2020). *Wie Vorbereitungsklassen zur Chance werden können: Eine Expertise für den Mediendienst Integration*.
https://mediendienst-integration.de/fileadmin/Dateien/Expertise_Mediendienst_Vorbereitungsklassen_2020.pdf
- Kärner, T., Warwas, J. & Schumann, S. (2021). A learning analytics approach to address heterogeneity in the classroom: The teachers' diagnostic support system. *Technology, Knowledge and Learning*, 26(1), 31–52. <https://doi.org/10.1007/s10758-020-09448-4>
- Kendon, A. (2010). *Gesture: Visible action as utterance*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511807572>
- Kita, S. (2004). How representational gestures help speaking. In D. McNeill (Hrsg.), *Language, culture, and cognition: Bd. 2. Language and gesture* (S. 162–185). Cambridge University Press.
- Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Waxmann.
- KMK. Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.). (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring: (Beschluss der 350. Kultusministerkonferenz vom 11.06.2015)*.
- Kniffka, G. & Siebert-Ott, G. (2012). *Deutsch als Zweitsprache: Lehren und Lernen* (3. Aufl.). *StandardWissen Lehramt: Bd. 2891*. Schöningh.
- König, J., Lammerding, S., Nold, G., Rohde, A., Strauß, S., & Tachtsoglou, S. (2016). Teachers' professional knowledge for teaching English as a foreign language: Assessing the outcomes of teacher education. *Journal of Teacher Education*, 67(4), 320-337. <https://doi.org/10.1177/0022487116644956>
- Koordinierungsstelle der AlphaDekade (Hrsg.). (2022). *Monitoring-Bericht: (Zwischen-) Ergebnisse der vom BMBF im Rahmen der AlphaDekade geförderten Projekte für das Jahr 2020*.

- https://www.alphadekade.de/alphadekade/shareddocs/downloads/files/monitoringbericht-projektergebnisse_2020.pdf.pdf?__blob=publicationFile&v=3
- Krashen, S. D. (1989). *The input hypothesis: Issues and implications* (1. publ.). Longman.
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multi-kriteriale Analyse. In M. Kunter, J. Baumert, W. Blum & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 85–113). Waxmann.
- Kunze, I. (2004). *Konzepte von Deutschunterricht: Eine Studie zu individuellen didaktischen Theorien von Lehrerinnen und Lehrern*. 1. Aufl. Zugl.: Hamburg, Universität, FB Erziehungswiss., Habil.-Schr., 2002. *Studien zur Bildungsgangdidaktik#Bd.#1*. VS Verl. für Sozialwissenschaften.
- Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, 54(1), 79–117.
<https://doi.org/10.1111/j.1467-9922.2004.00249.x>
- Lazaraton, A., & Ishihara, N. (2005). Understanding second language teacher practice using microanalysis and self-reflection: A collaborative case study. *Modern Language Journal*, 89(4), 529–542. <https://doi.org/10.1111/j.1540-4781.2005.00328.x>
- Lee, B. J. & Hart, E. T. (2022). Facemask occlusion's impact on L2 listening comprehension. *Speech Communication*, 139, 45–50.
<https://doi.org/10.1016/j.specom.2022.03.005>
- Leutner, D. (1992). *Adaptive Lehrsysteme: Instruktionspsychologische Grundlagen und experimentelle Analysen*. Zugl.: Aachen, Universität, Habil.-Schr., 1991. *Fortschritte der psychologischen Forschung: Bd. 13*. Psychologie-Verlag-Union.
- Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8), 707–710.
- Lintunen, P. & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: Comparing the incomparable? *Research in Language*, 12(4), 377–399. <https://doi.org/10.1515/rela-2015-0005>
- Lipowsky, F., Hess, M. & Denn, A.-K. (Hrsg.). (2019). *Dokumentation der Erhebungsinstrumente des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE). Technischer Bericht zu den PERLE-Videostudien. Band 2: Beobachtungssysteme zur Beschreibung und Qualität von Grundschulunterricht. Materialien zur Bildungsforschung 23/4*. GFPF.

- Lipowsky, F., Rakoczy, K., Pauli, C., Reusser, K. & Klieme, E. (2007). Gleicher Unterricht - gleiche Chancen für alle? Die Verteilung von Schülerbeiträgen im Klassenunterricht. *Unterrichtswissenschaft*, 35(2), 125–147. <https://doi.org/10.25656/01:5489>
- Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching*, 51(3), 285-329. <https://doi.org/10.1017/S0261444818000125>
- Lotz, M., Berner, N. E. & Gabriel, K. (2013). Auswertung der PERLE-Videostudien und Überblick über die Beobachtungsinstrumente. In F. Lipowsky & G. Faust (Hrsg.), *Dokumentation der Erhebungsinstrumente des Projekts "Persönlichkeits- und Lernentwicklung von Grundschulkindern" (PERLE)*. 3. Technischer Bericht zu den PERLE-Videostudien (1. Aufl., S. 83–103).
- Lucas, T., Villegas, A. M. & Freedson-Gonzalez, M. (2008). Linguistically responsive teacher education: Preparing classroom teachers to teach English language learners. *Journal of Teacher Education*, 59(4), 361–373. <https://doi.org/10.1177/0022487108322110>
- Martinovic, B., van Tubergen, F. & Maas, I. (2009). Dynamics of interethnic contact: A panel study of immigrants in the Netherlands. *European Sociological Review*, 25(3), 303–318. <https://doi.org/10.1093/esr/jcn049>
- Massumi, M. & Dewitz, N. von. (2015). *Neu zugewanderte Kinder und Jugendliche im deutschen Schulsystem: Bestandsaufnahme und Empfehlungen*. http://www.mercator-institut-sprachfoerderung.de/fileadmin/Redaktion/PDF/Publikationen/MI_ZfL_Studie_Zugewanderte_im_deutschen_Schulsystem_final_screen.pdf
- McCafferty, S. G. & Stam, G. (Hrsg.). (2008). *ESL & applied linguistics professional series. Gesture: Second language acquisition and classroom research*. Routledge.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- McNeill, D. (2004). *Language and gesture. Language, culture, and cognition: Bd. 2*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511620850>
- McNeill, D. (2008). *Gesture and thought*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226514642.001.0001>
- Mercer, N. (2000). *Words and minds: How we use language to think together*. Routledge.
- Möller, K., Hanke, P., Beinbrech, C., Hein, A. K., Kleickmann, T. & Schages, R. (2007). *Qualität von Grundschulunterricht entwickeln, erfassen und bewerten*. VS Verlag für Sozialwissenschaften.

- Müller, C., Cienki, A., Fricke, E., Ladewig, S. H., McNeil, D. & Bressemer, J. (Hrsg.). (2014). *Handbücher zur Sprach- und Kommunikationswissenschaft = Handbooks of linguistics and communication science: An international handbook on multimodality in human interaction*. De Gruyter Mouton.
- Ninio, A. & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language*, 5(1), 1–15. <https://doi.org/10.1017/S0305000900001896>
- Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S. & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes*, 35(2), 135–198. https://doi.org/10.1207/S15326950DP3502_3
- Ohliger, R., Polat, F., Schammann, H. & Thränhardt, D. (2017). *Integrationskurse reformieren. Steuerung neu koordinieren: Schritte zu einer verbesserten Sprachvermittlung*. Heinrich-Böll-Stiftung.
- Ophardt, D. & Thiel, F. (2017). Klassenmanagement als Basisdimension der Unterrichtsqualität. In M. K. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion* (S. 245–266). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-15083-9_11
- Oser, F. & Baeriswyl, F. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (4. Aufl., S. 1031–1065). American Educational Research Association.
- Pallotti, G. (2019). An approach to assessing the linguistic difficulty of tasks. *Journal of the European Second Language Association*, 3(1), 58–70. <https://doi.org/10.22599/jesla.61>
- Paquot, M. & Plonsky, L. (2017). Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research*, 3, 61–94. <https://doi.org/10.1075/ijlcr.3.1.03paq>
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research*, 88(2), 205–242. <https://doi.org/10.3102/0034654317743198>
- Pauli, C. & Reusser, K. (2006). Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. *Zeitschrift für Pädagogik*, 52(6), 774–798. [//www.pedocs.de/volltexte/2011/4488/](http://www.pedocs.de/volltexte/2011/4488/)
- Petko, D., Waldis, M., Pauli, C. & Reusser, K. (2003). Methodologische Überlegungen zur videogestützten Forschung in der Mathematikdidaktik. *Zentralblatt für Didaktik der Mathematik*, 35(6), 265–280. <https://doi.org/10.1007/BF02656691>

- Praetorius, A.-K. (2014). *Messung von Unterrichtsqualität durch Ratings*. Zugl.: Koblenz-Landau, Universität, Diss., 2012 u.d.T.: Praetorius, Anna-Katharina: Eignung von hoch-inferenten Beobachterratings zur Messung von Unterrichtsqualität (1. Aufl.). *Pädagogische Psychologie und Entwicklungspsychologie: Bd. 90*. Waxmann Verlag GmbH.
- Praetorius, A.-K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B. & Nehring, A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum – zwischen Generik und Fachspezifik. *Unterrichtswissenschaft*, 48(3), 409–446. <https://doi.org/10.1007/s42010-020-00082-8>
- Praetorius, A.-K., McIntyre, N. A. & Klassen, R. M. (2017). Reactivity effects in video-based classroom research: an investigation using teacher and student questionnaires as well as teacher eye-tracking. *Zeitschrift für Erziehungswissenschaft*, 20(S1), 49–74. <https://doi.org/10.1007/s11618-017-0729-3>
- Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K. & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2–12. <https://doi.org/10.1016/j.learninstruc.2013.12.002>
- Reich, H. H. & Krumm, H.-J. (2013). *Sprachbildung und Mehrsprachigkeit: Ein Curriculum zur Wahrnehmung und Bewältigung sprachlicher Vielfalt im Unterricht* (1. Aufl.). Waxmann.
- Rjosk, C., Haag, N., Heppt, B. & Stanat, P. (2017). Zuwanderungsbezogene Disparitäten. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 237–275). Waxmann.
- Sahlender, M. & Schrader, J. (2017). Spracherwerb bei Zugewanderten und funktionalen Analphabeten. *Zeitschrift für Weiterbildungsforschung*, 40(3), 275–293. <https://doi.org/10.1007/s40955-017-0098-9>
- Sato, R. (2020). Gestures in EFL classroom: Their relations with complexity, accuracy, and fluency in EFL teachers' L2 utterances. *System*, 89, Article 102215. <https://doi.org/10.1016/j.system.2020.102215>
- Schmidt, T. & Wörner, K. (2014). EXMARaLDA. In J. Durand, U. Gut & G. Kristoffersen (Hrsg.), *Handbook on Corpus Phonology* (S. 402–419). Oxford University Press.
- Schrader, J. (2015). Large Scale Assessments und die Bildung Erwachsener. Erträge, Grenzen und Potenziale der Forschung. *Zeitschrift für Pädagogik*(3), 410–428.

- Seidel, T. (2003). *Lehr-Lernskripts im Unterricht: Freiräume und Einschränkungen für kognitive und motivationale Lernprozesse; eine Videostudie im Physikunterricht. Pädagogische Psychologie und Entwicklungspsychologie: Bd. 35.* Waxmann.
- Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik* 60(6), 850-866.
<https://doi.org/10.25656/01:14686>
- Seidel, T. & Prenzel, M. (2006). Stability of teaching patterns in physics instruction: Findings from a video study. *Learning and Instruction*, 16(3), 228–240.
<https://doi.org/10.1016/j.learninstruc.2006.03.002>
- Seidel, T. & Prenzel, M. (2010). Beobachtungsverfahren: Vom Datenmaterial zur Datenanalyse. In H. Holling & B. Schmitz (Hrsg.), *Handbuch der Psychologie / hrsg. von Jürgen Bengel u.a: Bd.13. Handbuch Statistik, Methoden und Evaluation* (S. 139–152).
- Seidel, T., Prenzel, M., Duit, R. & Lehrke, M. (Hrsg.). (2003). *BIQUA. Technischer Bericht zur Videostudie "Lehr-Lern-Prozesse im Physikunterricht"*. IPN.
- Seidel, T., Renkl, A. & Rieß, W. (2021). Basisdimensionen für Unterrichtsqualität im Fachkontext konkretisieren: Die Rolle von Unterrichtsartefakten und Bestimmung von Standardsituationen. *Unterrichtswissenschaft*, 49(2), 293–301.
<https://doi.org/10.1007/s42010-021-00108-9>
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499.
<https://doi.org/10.3102/0034654307310317>
- Singer, M. A. & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological science*, 16(2), 85–89.
<https://doi.org/10.1111/j.0956-7976.2005.00786.x>
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19(1/2), 85–95. <https://doi.org/10.1024/1010-0652.19.12.85>
- Stam, G. A. (2013). Second language acquisition and gesture. In C. A. Chapelle (Hrsg.), *The Encyclopedia of Applied Linguistics*. Blackwell Publishing Ltd.
<https://doi.org/10.1002/9781405198431.wbeal1049>
- Stam, G. A. & Tellier, M. (2022). Gesture helps second and foreign language learning and teaching. In A. Morgenstern & S. Goldin-Meadow (Hrsg.), *Gesture in*

- language: Development across the lifespan* (S. 335–363). De Gruyter Mouton.
<https://doi.org/10.1037/0000269-014>
- Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S. & Haag, N. (Hrsg.). (2017). *IQB-Bildungstrend 2016: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Waxmann.
- Statistisches Bundesamt (Hrsg.). (2022). *Wanderungsstatistik (EVAS-Nr. 12711)*.
<https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Wanderungen/Tabellen/wanderungen-alle.html>
- Tellier, M., Stam, G. & Ghio, A. (2021). Handling language. *Gesture*, 20(1), 30–62.
<https://doi.org/10.1075/gest.19031.tel>
- Thoms, J. J. (2012). Classroom discourse in foreign language classrooms: A review of the literature. *Foreign Language Annals*, 45(1), 8–27.
<https://doi.org/10.1111/j.1944-9720.2012.01177.x>
- Tissot, A., Croisier, J., Pietrantuono, G., Baier, A., Ninke, L., Rother, N. & Babka von Gostomski, C. (2019). *Zwischenbericht I zum Forschungsprojekt "Evaluation der Integrationskurse (Evlk)": erste Analysen und Erkenntnisse*. Forschungsbericht / Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ): Bd. 33. SSOAR, GESIS - Leibniz-Institut für Sozialwissenschaften e.V and Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ).
- Tomasello, M. (2010). *Die Ursprünge der menschlichen Kommunikation*. Suhrkamp.
- Trim, J. L. M., Quetz, J., Schieß, R. & Schneider, G. (Hrsg.). (2009). *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. Langenscheidt. <https://www.goethe.de/z/50/commeuro/i4.htm>
- Ulewicz, M. & Beatty, A. (Hrsg.). (2001). *The power of video technology in international comparative research in education*. National Academies Press.
<https://doi.org/10.17226/10150>
- Valenzano, L., Alibali, M. W. & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28(2), 187–204. [https://doi.org/10.1016/S0361-476X\(02\)00007-3](https://doi.org/10.1016/S0361-476X(02)00007-3)
- Vock, M. & Gronostaj, A. (2017). *Umgang mit Heterogenität in Schule und Unterricht*. *Schriftenreihe des Netzwerk Bildung: Bd. 40.2*. Friedrich-Ebert-Stiftung Abt. Studienstiftung.

- Voss, T., Kunter, M. & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology, 103*(4), 952–969. <https://www.doi.org/10.1037/a0025125>
- Vygotsky, L. S. & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*(3), 249. <https://doi.org/10.2307/1170546>
- Wang, M.-T. & Fredricks, J. A. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child development, 85*(2), 722–737. <https://doi.org/10.1111/cdev.12138>
- Weinert, F. E. (1997). Notwendige Methodenvielfalt: Unterschiedliche Lernfähigkeiten erfordern variable Unterrichtsmethoden. *Friedrich-Jahresheft (XV: Lernmethoden)*, 50–52.
- Wisniewski, K. (2022). Gesprochene Lernerkorpora des Deutschen: Eine Bestandsaufnahme. *Zeitschrift für Germanistische Linguistik, 50*(1), 1–35. <https://doi.org/10.1515/zgl-2022-2047>