

Evaluation des Dialogischen Lesens unter Berücksichtigung der Durchführungsqualität

*Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie
des Fachbereich Psychologie und Sportwissenschaften
an der Justus-Liebig Universität Gießen*

*vorgelegt von Nils Hartung
aus Marburg*

2015

1. Gutachter: Prof. Dr. Marco Ennemoser
2. Gutachter: Prof. Dr. Holger Probst

Datum der Disputation: 17.12.2015

für

Maria & Gerhard

Barbara & Wilfried

"Ich hätte auch einen Ferrari aus Ton modellieren können [...] oder einen Film mit vier gespenstischen Ebenen drehen oder ich hätte eine Eliteschule im Alpenvorland gründen können [...], aber Sprache ist nun mal das billigste Material."

Funny van Dannen

Danksagungen

Diese Arbeit wäre ohne die berufliche und private Unterstützung von vielen, mir nahstehenden Kollegen, Freunden und meiner Familie nicht möglich gewesen. Ihnen möchte ich an dieser Stellen meinen tiefen Dank aussprechen.

Zuerst möchte ich meinem Doktorvater Prof. Dr. Marco Ennemoser für die intensive Betreuung dieser Arbeit und für die Möglichkeit zur Mitarbeit an diesem Projekt danken. Aus unseren Gesprächen und gemeinsamen Überlegungen konnte ich stets neue Erkenntnisse zur Weiterentwicklung meiner wissenschaftlichen Ideen gewinnen.

Ebenfalls danke ich Prof. Dr. Holger Probst für seinen – auch im Ruhestand– ungebrochenen fachlichen und freundschaftlichen Zuspruch. Beides war mir immer eine sehr große Hilfe auf dem Weg zur Fertigstellung dieser Arbeit.

Ich könnte mir die Zeit in der diese Arbeit entstanden ist nicht ohne meine lieben Kollegen aus der Arbeitsgruppe „Schulische Prävention und Evaluation“ vorstellen. Ich danke euch für fachlichen Rat, für gute Gespräche bei Kaffee und erlesenen Weinen und eure freundschaftliche Verbundenheit, die ich von Beginn an wahrgenommen habe. Im Besonderen danke ich Dr. Soulemanou Pepouna für eine intensive und fruchtbare Zusammenarbeit an diesem Projekt, für seine hervorragende linguistische Expertise und für viele offene Gespräche auch abseits der Arbeit. Dr. Ana Križan und Armin Vossen danke ich für die tatkräftige Unterstützung bei der Realisierung des Beobachtungsbogens zur Erfassung der Durchführungsqualität, Daniel Sinner und David Haschke für die Beratung in statistischen Fragen. Ebenso möchte ich Christina Czech und meinem Kollegium an der Mosaikschule in Marburg herzlich dafür danken, dass sie mir oftmals den Rücken freigehalten und mich immer wieder darin bestärkt haben, diese Arbeit neben meiner schulischen Tätigkeit fertigzustellen.

Diese Arbeit sei in besonderem Maße allen Kinder gewidmet, die an der Studie teilgenommen haben. Ihren Eltern, den Erzieherinnen und Erziehern im Kindergarten und den Lehrerinnen und Lehrern in der Schule danke ich ausdrücklich für die Bereitschaft zur Teilnahme und das Vertrauen, das sie meiner Arbeit entgegengebracht haben. Ebenso danke ich allen beteiligten Hilfskräften, die durch ihre teils mehrjährige Mitarbeit und mit großem persönlichen Einsatz maßgeblich zum Erfolg des Projekts beigetragen haben. Im Besonderen danke ich den helfenden Hände der ersten Stunde: Kristina Beckermann, Sascha Boy, Lena Briel, Anja Bürger, Barbara Desmytter, Katharina Erb, Swetlana Frei, Dörthe Geers, Katharina Gnadt, Katharina Hasenstab, Carolin Hohmann, Christina Jacker, Diana Klein, Anika Köpge, Lisa Müller, Jessica Nolte, Julia Österle, Stefanie Petrick, Catrin Piechotta, Hubertus Redlich, Karoline Redlich, Anne Rosenkranz, Inga Sangermann, Sandra Severin, Heika Stoßberg, Romy Tischer und Lena Weigel.

Ich konnte mich auf meinem Lebensweg immer auf den Rückhalt, die Unterstützung und die Liebe meiner Eltern und meiner Familie verlassen. Für euren immerwährenden Zuspruch, eure Förderung, euer Interesse, für eure Nachsicht und Geduld danke ich euch von Herzen!

Die vorliegende Arbeit entstand im Rahmen des vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Projekts „Evaluation des Dialogischen Lesens bei Kindern mit Migrationshintergrund“ mit der Förderkennzeichnung 01GJ0901, das unter der Leitung von Prof. Dr. Marco Ennemoser im Zuge der Forschungsinitiative Sprachdiagnostik und Sprachförderung (FISS) an der Justus-Liebig-Universität Gießen durchgeführt wurde.

Zusammenfassung

Die Förderung von sprachlichen Kompetenzen ist eine zentrale Aufgabe der vorschulischen Bildung. Das Angebot an unterschiedlichen Methoden und Trainings ist in den letzten Jahren stark angewachsen. Allerdings liegt selbst für etablierte Sprachfördermaßnahmen bislang kaum wissenschaftliche Evidenz für die Wirksamkeit vor (Jampert et al., 2007). Hinweise auf evidenzbasierte Sprachförderung finden sich zumeist in internationalen Untersuchungen (Justice & Pullen, 2003). Eine vielbeachtete Methode ist hier das Dialogische Lesen (Whitehurst et al., 1988).

Im anglo-amerikanischen Raum konnte das sprachförderliche Potential in verschiedenen Untersuchungsbedingungen nachgewiesen werden (z.B. Mol et al., 2008; 2009; Slavin, 2008). Für den deutschen Sprachraum liegen bislang noch keine umfassenden Studien vor.

Die vorliegende Trainingsstudie befasst sich mit der Frage, ob das Dialogische Lesen auch in deutschen Kindergärten zur Förderung von Risikokindern in Kleingruppen als wirksame Sprachfördermethode eingesetzt werden kann. Hierfür wurden aus einer Ausgangsstichprobe von 660 Kindern 174 sprachauffällige Risikokinder mit Hilfe eines Sprachscreenings (Grimm, 2003) ausgelesen und randomisiert zu einer von drei Untersuchungsbedingungen zugewiesen. Kinder der Trainingsgruppen nahmen entweder an einer Förderung im Sinne des Dialogischen Lesens ($n = 59$) oder an einem Training der Phonologischen Bewusstheit ($n = 58$) teil. Die Kontrollgruppe ($n = 57$) erhielt das konventionelle Sprachförderangebot der Kindergärten. Der Interventionszeitraum erstreckte sich über 3 Monate. Insgesamt wurden 24 Fördersitzungen erteilt. Vor und nach der Förderung, sowie 3 Monate danach, wurden sprachliche und metasprachliche Kompetenzen (Wortschatz, Grammatik, Sprachverständnis, Phonologische Bewusstheit) erfasst. Kontrollvariablen waren die Intelligenzleistung, die Arbeitsgedächtnisleistung und das Erzieherurteil in Bezug auf die Sprachentwicklung der Kinder. Außerdem wurde ein eigenes Instrument entwickelt mit dem die Durchführungsqualität im Dialogischen Lesen erfasst werden konnte.

Die Ergebnisse zeigen, dass Kinder, die am Dialogischen Lesen teilnahmen, im Nachtest signifikant bessere Leistungen im Sprachverständnis und in der Artikelbildung erzielten als Kinder aus der Trainingsgruppe „Phonologischen Bewusstheit“ und aus der Kontrollgruppe. In der morphologischen Entwicklung sind beide Trainingsgruppen der Kontrollgruppe überlegen. Außerdem leistet die Durchführungsqualität im Dialogischen Lesen einen signifikanten Beitrag zur Erklärung der Fördereffekte. Kinder von Trainern, bei denen eine hohe Durchführungsqualität im Dialogischen Lesen festgestellt werden konnte, profitierten im

Wortschatz und im Sprachverständnis mehr, als Kinder von Trainern, bei denen die Durchführungsqualität weniger hoch eingeschätzt wurde.

Die Kinder aus dem Training zur Phonologischen Bewusstheit, sind – wie erwartet – dem Dialogischen Lesen und der Kontrollgruppe in der Phonologischen Bewusstheit überlegen. Zusätzlich ergibt sich für diese Gruppe ein signifikanter Effekt in der syntaktischen Entwicklung. Langfristige Effekte bleiben derweil aus.

Schlüsselwörter: Sprachförderung, Kindergarten, Evaluation, Dialogisches Lesens, Phonologische Bewusstheit, Treatmentvalidität, evidenzbasierte Förderung, best practice

Inhaltsverzeichnis

ABBILDUNGSVERZEICHNIS	III
TABELLENVERZEICHNIS	IV
EINLEITUNG	1
1 SPRACHENTWICKLUNG DES KINDES	4
1.1 PHONOLOGISCH-PHONETISCHE ENTWICKLUNG.....	10
1.2 LEXIKALISCH-SEMANTISCHE ENTWICKLUNG.....	14
1.3 MORPHO-SYNTAKTISCHE ENTWICKLUNG.....	19
1.4 PRAGMATISCHE ENTWICKLUNG.....	23
1.4.1 <i>Motherese und Scaffolding als Vorbild für Dialogisches Lesen.....</i>	<i>25</i>
2 STÖRUNGEN DER SPRACHENTWICKLUNG	27
2.1.1 <i>Kinder mit Spracherwerbsstörungen.....</i>	<i>28</i>
2.1.2 <i>Kinder mit unzureichenden Sprachkenntnissen</i>	<i>29</i>
3 FÖRDERUNG VON SPRACHKOMPETENZEN.....	31
3.1 EVALUATION VON SPRACHFÖRDERMAßNAHMEN IM DEUTSCHEN SPRACHRAUM.....	33
3.2 EVIDENZBASIERTE METHODEN DER SPRACHFÖRDERUNG IN DEUTSCHLAND UND IM INTERNATIONALEN RAUM	39
3.2.1 <i>Dialogisches Lesen.....</i>	<i>39</i>
3.2.2 <i>Phonologische Bewusstheit.....</i>	<i>55</i>
4 EVALUATION VON SPRACHFÖRDERUNG UNTER BERÜCKSICHTIGUNG DER DURCHFÜHRUNGSQUALITÄT	60
5 FRAGESTELLUNG UND HYPOTHESEN.....	66
6 METHODE	69
6.1 STICHPROBE UND VERLAUF DER UNTERSUCHUNG	69
6.2 INSTRUMENTE	72
6.2.1 <i>Auditives Gedächtnis.....</i>	<i>73</i>
6.2.2 <i>Erzieherinnenurteil.....</i>	<i>75</i>
6.2.3 <i>Sprachentwicklung.....</i>	<i>75</i>

6.2.4	<i>Phonologische Bewusstheit</i>	77
6.2.5	<i>Intelligenz</i>	79
6.2.6	<i>Durchführungsqualität</i>	80
6.3	TRAININGSBEDINGUNGEN	85
6.3.1	<i>Dialogisches Lesen</i>	85
6.3.2	<i>Training der Phonologischen Bewusstheit</i>	89
6.3.3	<i>Kontrollgruppe</i>	92
6.4	STATISTISCHES VORGEHEN	94
6.4.1	<i>Prüfung der kurz- und der langfristigen Wirksamkeit</i>	94
7	ERGEBNISSE	97
7.1	PRÜFUNG DER KRITERIUMSVALIDITÄT	97
7.2	DESKRIPTIVE STATISTIKEN UND KORRELATIVE ZUSAMMENHÄNGE	97
7.3	KURZ- UND LANGFRISTIGE TRAININGSEFFEKTE	102
7.3.1	<i>Effekte auf die Phonologische Bewusstheit</i>	102
7.3.2	<i>Effekte auf weitere Sprachentwicklungsbereiche</i>	103
7.3.3	<i>Zusammenfassung der kurz- und der langfristigen Trainingseffekte</i>	113
7.4	EINFLUSS DER DURCHFÜHRUNGSQUALITÄT	114
8	DISKUSSION	117
8.1	PHONOLOGISCHE BEWUSSTHEIT	118
8.2	WEITERE SPRACHENTWICKLUNG	121
8.2.1	<i>Wortschatz</i>	123
8.2.2	<i>Sprachverständnis</i>	125
8.2.3	<i>Grammatik</i>	126
8.3	EINFLUSS DER DURCHFÜHRUNGSQUALITÄT	131
	FAZIT UND AUSBLICK	134
	LITERATURVERZEICHNIS	136
	ANHANG	155

Abbildungsverzeichnis

Abbildung 1	<i>Komponenten der Sprachentwicklung nach Barrett (2001)</i>	<i>5</i>
Abbildung 2	<i>Prinzipien des Dialogischen Lesens für 4-5jährige Kinder in nach Whitehurst et al. (1988).....</i>	<i>41</i>
Abbildung 3	<i>Aufbau des Konstrukts der Phonologischen Bewusstheit</i>	<i>56</i>
Abbildung 4	<i>Studiendesign und Untersuchungsverlauf</i>	<i>69</i>
Abbildung 5	<i>Durchschnittliche Leistungen aller Gruppen in der Phonologischen Bewusstheit im Prä-, Post- und Follow Up-Test.</i>	<i>102</i>
Abbildung 6	<i>Durchschnittliche Leistungen aller Gruppen in „Allgemeine Sprachentwicklung“ im Prä-, Post- und Follow Up-Test.</i>	<i>104</i>
Abbildung 7	<i>Durchschnittliche Leistungen aller Gruppen im Wortschatz im Prä-, Post- und Follow Up-Test.</i>	<i>105</i>
Abbildung 8	<i>Durchschnittliche Leistungen aller Gruppen im Sprachverständnis im Prä-, Post- und Follow Up-Test.</i>	<i>106</i>
Abbildung 9	<i>Durchschnittliche Leistungen aller Gruppen in der Grammatikleistung im Prä-, Post- und Follow Up-Test.....</i>	<i>108</i>
Abbildung 10	<i>Durchschnittliche Leistungen aller Gruppen in „Morphologie“ im Prä-, Post- und Follow Up-Test.</i>	<i>110</i>
Abbildung 11	<i>Durchschnittliche Leistungen aller Gruppen in „Artikelbildung“ im Prä-, Post- und Follow Up-Test</i>	<i>111</i>
Abbildung 12	<i>Durchschnittliche Leistungen aller Gruppen in „Syntax“ im Prä-, Post- und Follow Up-Test..</i>	<i>112</i>

Tabellenverzeichnis

Tabelle 1	<i>Phasen des Wortschatzerwerbs nach Grimm & Weinert (2002) und Weinert (2006).</i>	<i>18</i>
Tabelle 2	<i>Entwicklung der syntaktischen Fähigkeiten nach Tracy (2008).</i>	<i>22</i>
Tabelle 3	<i>Techniken des Dialogischen Lesens nach Ennemoser et al. (2013).</i>	<i>44</i>
Tabelle 4	<i>Erfasste Konstrukte in der Interventionsstudie.</i>	<i>72</i>
Tabelle 5	<i>Struktur des Beobachtungsbogens zur Erhebung der Durchführungsqualität im Dialogischen Lesen.</i>	<i>80</i>
Tabelle 6	<i>Justierte Interklassenkorrelationen der Raterurteile in den Items des Beobachtungsbogens zur Durchführungsqualität.</i>	<i>82</i>
Tabelle 7	<i>Korrelationen zwischen den Gesamtwerten der Beurteilungsmaße Zählen, Rating und Benotung.</i>	<i>83</i>
Tabelle 8	<i>Aufbau des Materialordners.</i>	<i>85</i>
Tabelle 9	<i>Hilfen im Fördermaterial im Sinne der CROWD und PEER-Prinzipien nach Whitehurst et al. (1988).</i>	<i>87</i>
Tabelle 10	<i>Aufbau des Original-Trainingsprogramms „Hören, lauschen, lernen“. ...</i>	<i>89</i>
Tabelle 11	<i>Aufbau der modifizierten Trainingsvariante von „Hören, lauschen, lernen“.....</i>	<i>91</i>
Tabelle 12	<i>Korrelationen zwischen Testwerten des SSV und dem Erzieherurteil zum Sprachentwicklungsstand.</i>	<i>97</i>
Tabelle 13	<i>Deskriptive Statistik der Untertests in den Untersuchungsgruppen.....</i>	<i>98</i>
Tabelle 14	<i>Deskriptive Statistiken der mittleren z-Werte der abhängigen Variablen für alle Untersuchungsgruppen.</i>	<i>101</i>
Tabelle 15	<i>Deskriptive Statistiken der durchschnittlich erreichten Rohpunkte in der erweiterten Auswertung der Satzbildungsaufgabe.</i>	<i>109</i>
Tabelle 16	<i>Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung in „Allgemeiner Sprachentwicklung“ im Dialogische Lesen (n = 55).</i>	<i>114</i>
Tabelle 17	<i>Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung im Wortschatz für Dialogisches Lesen (n = 55).</i>	<i>115</i>

Tabelle 18	<i>Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung im Sprachverständnis für Dialogisches Lesen (n = 55).</i>	<i>116</i>
Tabelle 19	<i>Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung im Bereich Grammatik für Dialogisches Lesen (n = 55).</i>	<i>116</i>
Tabelle 20	<i>Korrelationen zwischen den Subtests zur Erfassung sprachlicher Kompetenzen zu allen Messzeitpunkten.</i>	<i>156</i>
Tabelle 21	<i>Korrelationen der Subtests zur Erfassung der Phonologischen Bewusstheit zu allen Messzeitpunkten.</i>	<i>156</i>
Tabelle 22	<i>Korrelationen zwischen den abhängigen Variablen zu allen Messzeitpunkten.</i>	<i>156</i>
Tabelle 23	<i>Korrelationen zwischen den Skalen der erweiterten Auswertung der Satzbildungsaufgabe über alle Messzeitpunkte.</i>	<i>156</i>
Tabelle 24	<i>Korrelation zwischen den Raterurteilen in den Items zur Erfassung der Durchführungsqualität.</i>	<i>156</i>
Tabelle 25	<i>Korrelationen zwischen den Mittelwerten beider Rater und den Gesamtscores zur Durchführungsqualität.</i>	<i>156</i>

Einleitung

Die frühe Förderung von sprachlichen Kompetenzen spielt ohne Zweifel eine zentrale Rolle für die kindliche Bildungskarriere. Nicht erst mit dem Eintritt in die Grundschule bilden sprachliche Fähigkeiten eine wichtige Basis für den Erwerb von Lese- und Rechtschreibfähigkeiten.

Die Vielfalt an Methoden und Programmen zur sprachlichen Förderung im Elementarbereich ist immens. In ihrer Recherchearbeit „Schlüsselkompetenz Sprache“ fassen Jampert, Best, Guadatiello, Holler und Zehbauer (2007) zusammen, dass sich viele unterschiedliche Methoden und Programme zur Sprachförderung in der Praxis etabliert haben. Sie richten sich an unterschiedliche Zielgruppen. Dazu zählen im Besonderen Kinder mit Sprachentwicklungsstörungen, Kinder mit nicht-deutscher Erstsprache sowie Kinder aus sozial- und bildungsbenachteiligten Verhältnissen. Auch wenn sich in der Vergangenheit eine Vielzahl von Sprachförderkonzepten im Kindergartenalltag etablieren konnte, stellen die Autoren fest, dass für die meisten Programme kaum wissenschaftliche Evidenz in Bezug auf ihre Wirksamkeit vorliegt. Unklar ist, in welchem Maße geförderte Kinder von einer Maßnahme im Allgemeinen profitieren können und in welchen Sprachkompetenzbereichen eine Methode im Speziellen ihre Wirkung zeigt (z.B. Wortschatz, Grammatik, Sprachverständnis). Methodische Debatten befassen sich außerdem mit der Frage, ob sich Sprachkompetenzen nachhaltiger durch formale Übungen (*focus on form*) oder implizite Gesprächsanlässe (*focus on meaning*) verbessern lassen.

Bundesweit wurden in den vergangenen Jahren große finanzielle Anstrengungen unternommen, um Projekte zu unterstützen, die sich im Sinne der evidenzbasierten Wissenschaft mit der Wirksamkeitsprüfung von Sprachfördermaßnahmen befassen. Sie kamen jedoch zu dem ernüchternden Ergebnis, dass sich die Sprachkompetenzen von geförderten und ungeförderten Kindern nach Abschluss der Intervention häufig nicht substanziell unterscheiden (Roos, Polotzek & Schöler, 2007).

Derweil lassen sich, vor allem im anglo-amerikanischen Raum, viele positive Befunde zur Wirksamkeit von Sprachfördermethoden finden (National Early Literacy Panel, 2008). Eine vielfach evaluierte und als evidenzbasiert geltende Methode ist das *Dialogisches Lesen* (Whitehurst et al., 1988; 1991; 1994; 1999). Beim Dialogischen Lesen werden Sprachkompetenzen der Kinder – im Sinne des *focus on meaning* – durch gesprächs-

begleitende Modellierungstechniken gefördert, die Erwachsene beim Vorlesen einer Geschichte einsetzen.

Das sprachförderliche Potential des Dialogischen Lesens konnte bereits über den englischen Sprachraum hinaus als wirksame Fördermethode für Kinder mit Sprachförderbedarf bestätigt werden. Für den deutschen Sprachraum liegen bislang noch keine umfassenden Untersuchungen vor, die die Wirksamkeit des Dialogischen Lesens als Sprachförderangebot für Risikokinder im Kindergarten umfänglich untersuchen.

Ebenfalls fehlt es national wie international an Studien, die den Einfluss der Durchführungsqualität (*treatment validity*) zur Beurteilung der Wirksamkeit von Sprachförderprogrammen in ausreichendem Maße berücksichtigen.

Aus der beschriebenen Forschungssituation ergeben sich für die vorliegende Arbeit mehrere Ziele. So soll geprüft werden, ob sich Dialogisches Lesen auch in deutschen Kindergärten zur Sprachförderung von Risikokindern wirksam als Kleingruppenförderung einsetzen lässt.

Um herauszufinden, ob geförderte Kinder vom Dialogischen Lesen in den Entwicklungsbereichen Wortschatz, Grammatik, Sprachverständnis und Phonologische Bewusstheit gleichermaßen oder unterschiedlich stark profitieren, sollen differentielle Effekte betrachtet und mit einem formfokussierenden Sprachtraining zur Förderung der Phonologischen Bewusstheit verglichen werden.

Zuletzt soll die Arbeit erste Hinweise zum Einfluss der Durchführungsqualität auf die Wirksamkeit des Dialogischen Lesens geben.

Die vorliegende Arbeit umfasst acht Kapitel. Im Anschluss an die Einleitung geben Kapitel 1 bis 4 zunächst einen Überblick über den theoretischen Hintergrund, bevor daraus Hypothesen und Fragestellung für die eigene Untersuchung abgeleitet werden.

In Kapitel 1 werden Meilensteine der kindlichen Sprachentwicklung nachgezeichnet. Ausgangslage für die entwicklungstheoretische Betrachtung ist das Sprachentwicklungsmodell nach Barrett (2001) aus dem sich vier Sprachentwicklungsdomänen ableiten lassen. Dargestellt werden bedeutsame Entwicklungsschritte für den Erwerb von phonetisch-phonologischen, semantisch-lexikalischen, morpho-syntaktischen und pragmatischen Kompetenzen. Gleichzeitig werden Erkenntnisse aus der *motherese*-Forschung, auf denen das Dialogische Lesen basiert, erläutert.

Das zweite Kapitel befasst sich in aller Kürze mit Störungen in der Sprachentwicklung. An dieser Stelle wird definiert, welche Formen von kindlichen Sprachauffälligkeiten gemeint sind, wenn im Verlauf dieser Arbeit von Risikokindern die Rede ist.

Kapitel 3 gibt zunächst einen Überblick über die Ergebnisse aus nationalen Forschungsvorhaben, die sich mit der Evaluation von Sprachfördermaßnahmen im Allgemeinen befassten. Danach werden im Speziellen nationale und internationale Forschungsergebnisse dargestellt, die sich auf die Wirksamkeit des Dialogischen Lesens beziehen. Da in der durchgeführten Studie auch ein Programm zur Förderung der Phonologischen Bewusstheit als Vergleichstraining zum Einsatz kommt, werden zuletzt auch Ergebnisse aus Evaluationsstudien zur Förderung der Phonologischen Bewusstheit berichtet. Theoretische Überlegungen zur Operationalisierung der Durchführungsqualität werden in Kapitel 4 erörtert.

Aus den theoretischen Überlegungen der ersten vier Kapitel werden in Kapitel 5 Hypothesen und Fragestellungen für die eigene Studie abgeleitet. Kapitel 6 gibt einen Überblick über Studiendesign und Methode der eigenen Untersuchung. Außerdem werden die verwendeten Messinstrumente dargestellt und das statistische Vorgehen erläutert. Die Darstellung der Ergebnisse erfolgt in Kapitel 7, bevor sie in Kapitel 8 diskutiert und Implikationen für die pädagogische Praxis abgeleitet werden.

1 Sprachentwicklung des Kindes

Welche Entwicklungsprozesse Kinder schon im Säuglingsalter dazu befähigen Sprache zu verstehen und später selbst zu produzieren, ist eine zentrale und häufig diskutierte Frage in der Psychologie und den Erziehungswissenschaften. Grundsätzlich lassen sich drei Erwerbstheorien unterscheiden: 1) *Inside-Out-Theorien* 2) *Outside-In-Theorien* 3) *interaktionistische Theorien* (Crystal, 2010; Nußbeck, 2007; Owens, 2005; Szagun, 2006).

1) Inside-Out-Theorien

Nativistische Entwicklungsmodelle oder *Inside-Out-Theorien* (Hirsh-Pasek & Golinkoff, 1996) verstehen Spracherwerb als Resultat einer genetischen Disposition (Chomsky, 1986; Pinker, 2000). Die menschliche Fähigkeit eine Sprache zu erwerben wird einem entwicklungsspezifischen Modul zugeschrieben, das nur in sehr geringem Maße an andere Prozesse gekoppelt ist (Nußbeck, 2007). Etwaige Einflüsse der Sprachumwelt auf den Erwerb von Sprache spielen nach dieser Theorie eine untergeordnete Rolle. Sprachlicher Input wie zum Beispiel die kindgerichtete Sprache von Erwachsenen, wird allenfalls dazu benötigt, den universellen Entwicklungsplan des Sprachmoduls zu aktivieren (Chomsky, 1988).

2) Outside-In-Theorien

Dem gegenüber steht die Tradition der soziolinguistischen Erklärungsmodelle oder, um in der Sprache von Hirsh-Pasek und Golinkoff (1996) zu bleiben, die *Outside-In-Theorien*. Sie betonen in besonderem Maß den Einfluss der Sprachumgebung beziehungsweise des sprachlichen Inputs auf den Erwerb von sprachlichem Regelwissen. In dieser Tradition spricht Tomasello (2001) vom gebrauchsbasierten Spracherwerb (*usage-based language acquisition*). Sprache wird hier als Resultat von Erwerbsprozessen verstanden, deren Grundlage die sprachliche Interaktion zwischen Kind und Umwelt bildet. Tomasello geht davon aus, dass kognitive Lernstrategien Kinder befähigen, sprachlichen Input zu abstrahieren. Der Art und Weise, wie Kindern sprachlicher Input präsentiert wird, kommt dabei eine wichtige Steuerungsfunktion zu (Tracy, 2000).

Outside-In- und Inside-Out-Theorien liefern gleichermaßen wichtige Erkenntnisse über die komplexen Prozesse des Spracherwerbs. In bestimmten Fragen stoßen beide Erklärungsansätze indes an ihre Grenzen. Theorien, die ausschließlich dem Outside-In-Ansatz verhaftet sind, vermögen nicht zu klären, wieso manchen Kindern ein ungestörter Spracherwerb auch bei widrigem Kommunikationsverhalten der Umwelt fast problemlos

gelingt. Verfechter der Inside-Out-Theorien geraten in Erklärungsnot, wenn sie Verknüpfungsprozesse einer angeborenen Universalgrammatik mit der zu erwerbenden Zielsprache erklären sollen (*linking problem*). Diese Überlegungen führen zu der Frage, wie sich trotz eines einheitlichen Erwerbsmechanismus eine Vielzahl von Sprachen mit eigenen Regeln und Gesetzmäßigkeiten entwickeln konnte. Unklar bleibt auch, wie im weiteren Entwicklungsverlauf sprachliche Unterschiede bei vermeintlich gleichen Voraussetzungen entstehen können (*continuity problem*; Nußbeck, 2007).

3) Interaktionistische Theorien

Eine Vielzahl von Autoren spricht sich für eine dritte theoretische Sichtweise aus, die gewissermaßen einen Schulterschluss beider zuvor beschriebenen Theorie-Pole darstellt (Grimm & Weinert, 2002; Hennon, Hirsh-Pasek & Golinkoff, 2000; Nußbeck 2007; Ritterfeld, 2005; Szagun, 2006). Gemeint ist die *interaktionistische Sichtweise* (Grimm & Weinert, 2002) oder auch „*radikale Mitte*“ (Hennon, Hirsh-Pasek & Golinkoff, 2000). Vertreter dieses Ansatzes verstehen den kindlichen Spracherwerb als Zusammenspiel aus biologisch bedingten Fähigkeiten und sozialen Umweltfaktoren. Tracy (2000) weist darauf hin, dass bis heute sehr wenig über die genaue Verhältnismäßigkeit kognitiver und sozial-kommunikativer Faktoren und ihren Beitrag zum erfolgreichen Erwerb von Sprache bekannt ist. Deutsch und El Mogharbel (2007a) weisen grundsätzlich darauf hin, dass Sprachentwicklung nicht als einmaliger „Erwerb“ zu verstehen ist, sondern als hierarchisch aufgebauter Entwicklungsprozess. Barrett (2001) nennt drei Säulen der sprachlichen Entwicklung: *Lautentwicklung*, *Semantik* und *Pragmatik* (Abb. 1)

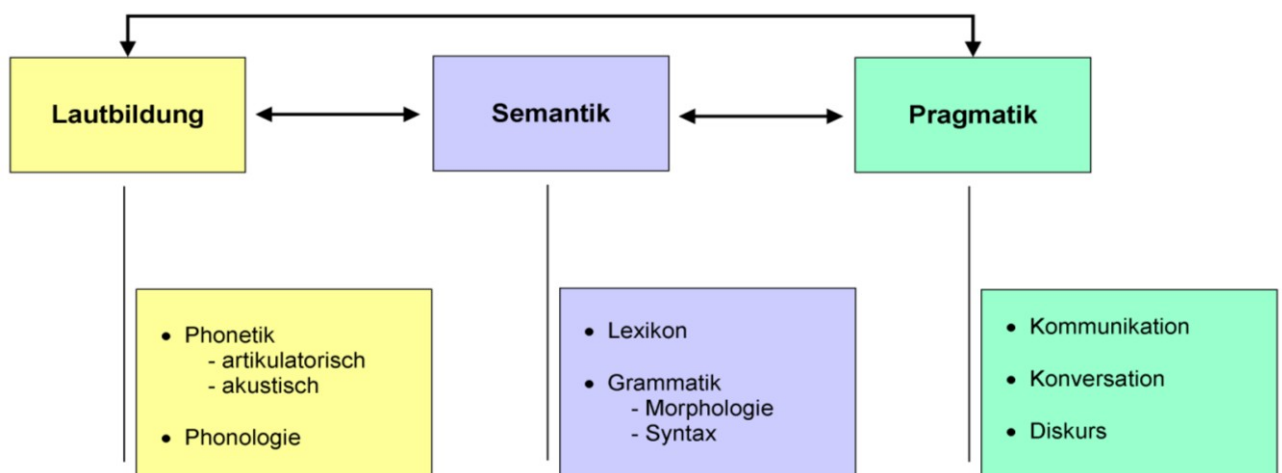


Abbildung 1 Komponenten der Sprachentwicklung nach Barrett (2001)

Einen Grundbaustein jeder gesprochenen Sprache stellt ihre spezifische Lautstruktur dar. Sie setzt sich aus einzelnen Lautelementen (*Phonemen*) zusammen. Die Gesamtheit aller Phoneme und ihre Abfolgen innerhalb eines Sprachsystems werden von der *Phonologie* beschrieben.

Der Teilbereich der *Phonetik* befasst sich mit der physischen Produktion und Perzeption von Lauten. Beschrieben werden Merkmale der Artikulation (zum Beispiel anatomische Voraussetzungen der Lautbildung) und Merkmale des Klangs und der akustischen Eigenschaften (Frequenz, Amplitude, Dauer) von Lauten.

Um ein sprachspezifisches Lautsystem erwerben zu können, verfügen bereits Säuglinge über angeborene Fähigkeiten (Sensitivität für rhythmisch-prosodische Strukturen, Wahrnehmung lautlicher Kontraste, verfügbare Gedächtnisleistung), die zum Teil schon vor der Geburt ausgebildet werden (Weinert, 2006). Diese *suprasegmentalen Kompetenzen* sind dafür verantwortlich, dass bereits Babys aus den Lautströmen ihrer Umwelt eine Struktur ableiten können (Grimm, 2003). Anhand besonders prägnanter oder häufig auftretender Lautmerkmale einer Sprache gelingt es Säuglingen bedeutungsunterscheidende Elemente aus einem anfänglich nicht strukturierbaren Lautstrom aufzufassen und zu abstrahieren. Zahlreiche Studien der Säuglingsforschung liefern Evidenz für die frühe Sensitivität für rhythmisch-prosodische Strukturen (DeCasper, Lecanuet, Busnel, Granier-Deferre & Maugeais, 1994; Höhle, 2004; Jusczyk, Houston & Goodman, 1998). Die besondere Achtsamkeit für auditive Stimuli erleichtert es Säuglingen auch unterschiedliche Umweltsprachen zu differenzieren und spezifische Merkmale eines Sprachcodes zu speichern (Mandel, Jusczyk, Kemler & Nelson, 1994).

Dass es Neugeborenen schon in den ersten Lebensmonaten gelingt die lautliche Struktur ihrer Muttersprache wiederzuerkennen, zeigt eine Untersuchung von Mehler, Jusczyk, Lambertz, Halsted, Bertoncini und Amiel-Tison (1988). Die Forschergruppe konfrontierte Säuglinge aus anglo-amerikanischer Sprachumgebung mit muttersprachlichen und fremdsprachlichen Stimuli. Gleichzeitig wurde ihre Saugrate gemessen. Beim Erklingen bekannter, muttersprachlicher Stimuli produzierten die Neugeborenen höhere Saugraten als es bei fremdsprachlichen Stimuli der Fall war. Offensichtlich gelingt es Säuglingen schon weit vor der Produktion erster bedeutungsvoller Lautstrukturen, häufig gehörte prosodische Muster von unbekannten zu unterscheiden. Diese basale Diskriminationsfertigkeit legt den Grundstein für den Erwerb des phonologischen Systems einer Sprache.

Phoneme sind die kleinsten bedeutungsunterscheidenden Einheiten einer Sprache. Das heißt, die Veränderung einzelner Phoneme innerhalb eines Wortes hat Einfluss auf die Wortbedeutung (z.B. *Mücke* versus *Lücke*). Früh lernen Kinder, dass Phoneme in ihrer Sprache bedeutungsunterscheidend sind. Sie verstehen, dass die Wörter ihrer Muttersprache phonotaktischen Regeln unterliegen (Weinert & Lockl, 2008). Lautkombinationen werden als auditive Zeichen verstanden, mit denen *Bedeutung* transportiert werden kann.

Der Analyse von Bedeutungszusammenhängen innerhalb einer Sprache widmet sich die *Semantik*. Sie beschreibt die zweite Säule in Barrets‘ Sprachentwicklungsmodell (vgl. Abb. 1, S. 5). Bedeutung kann über Sätze und Wörter aber auch nonverbal durch Gestik transportiert werden. Auch nicht-sprachlichen Lauten wie kindlichem Schreien messen wir Bedeutung bei.

Die kleinsten bedeutungstragenden Einheiten stellen die *Morpheme* dar. Sie können als *Derivationsmorpheme* (*schön* - *Schönheit*) oder als *Flexionsmorpheme* mit grammatikalischer Funktion (*gehen* - *du gehst*) in Erscheinung treten (Grimm, 2003; Szagun, 2006; Nußbeck, 2007). Morpheme sind gewissermaßen die Grundbausteine, aus denen sich Wörter konstituieren. Deutsch und El Mogharbel (2007a) sprechen von Wörtern als „kleinste selbständige Bedeutungseinheiten“ (S.12).

Auf welche Weise Wörter mit ihren Morphemen zu Phrasen und Sätzen kombiniert werden dürfen, unterliegt den Regeln der *Syntax*. Syntaktische Konstruktionen folgen einer grammatischen Kongruenz. Das bedeutet, dass die grammatikalische Form eines Wortes die Form von weiteren Wörtern eines Satzes mitbestimmt. So kann die Phrase *die bunten Blumen* nur deshalb als grammatisch korrekt gelten, weil alle Bestandteile (Artikel, Adjektiv und Substantiv) der Pluralmarkierung folgen. Folglich unterliegen grammatische Kompetenzen immer einem Zusammenspiel aus morphologischem und syntaktischem Regelwissen. Es handelt sich um eine äußerst komplexe Domäne der kindlichen Sprachentwicklung.

Der Bereich *Lexikon* bezieht sich auf die Wortsemantik, also die Bedeutungen der Inhalte des Wortschatzes (Grimm & Weinert, 2002). Die Struktur der lexikalischen Unterscheidungen hängt dabei stark von der zu erwerbenden Sprache ab. Theoretisch wäre es vorstellbar, dass eine Sprache keine sprachliche Unterscheidungsmöglichkeit zwischen den Worten „Vogel“ und „Flugzeug“ bereithält, wohl aber ein Wort für „Dinge die am Himmel fliegen“. Das Lexikon eines Muttersprachlers dieser Sprache unterläge dann einer anderen Bedeutungsstruktur als es im Deutschen der Fall ist. Das Beispiel verdeutlicht auch, dass ein Kind im Laufe seiner sprachlichen Entwicklung ein komplex vernetztes Categoriesystem aus Wortbedeutungen aufbaut, das mit zunehmendem sprachlichem Handeln stetige

Umstrukturierung und Erweiterung erfährt. Kinder lernen, dass sich die Bedeutung desselben Wortes im Kontext eines Satzes verschieben kann (Dromi, 2001; Grimm, 2003). Zum Beispiel verknüpfen wir mit dem Wort „leicht“ im Kontext der Aussage „Diese Aufgabe ist *leicht* zu lösen.“ eine andere Bedeutung als bei „Die Feder ist *leicht*.“.

Die dritte Säule der Sprachentwicklung bildet nach Barrett (2001) die *Pragmatik* (vgl. Abb. 1, S. 5). Der Entwicklungsbereich Pragmatik beschreibt das kommunikative System, in dessen Rahmen Sprache gebraucht wird. Grimm (2003) definiert pragmatische Kompetenzen als das Wissen „[...] um die angemessenen kommunikativen Verwendungsweisen von Sätzen in unterschiedlichen Kontexten [...].“ (S.19).

Die Analyse von Sprachhandlungen in der *Kommunikation* ist ein bedeutsamer Aspekt der Pragmatik. Sprachhandlungen können situationsbedingt unterschiedliche Funktionen erfüllen (Barrett, 2001). Abhängig vom Setting und der Beziehung der Sprecher zueinander kann eine bestimmte Aussage zum Beispiel die Intention verfolgen Gefühle auszudrücken, eine Information zu erhalten, eine Handlung zu kommentieren oder mehrere Funktionen zugleich bedienen.

Jedwedes Wissen über gesprächssteuernde Prinzipien und Regeln (zum Beispiel Redeübergabe) ordnet sich dem Feld der *Konversation* unter. Während unter Konversation in besonderer Weise sprachliche Interaktion und Austausch beleuchtet werden, befasst sich der Bereich *Diskurs* in erster Linie mit dem kommunikativen Verhalten des Sprechers. Von einem geübten Sprecher erwarten wir stillschweigend diskursive Kompetenzen. Es muss ihm gelingen einzelne Sätze im Rahmen einer Erzählung in einen angemessenen, kohärenten Zusammenhang zu bringen, sich in die Lage des Zuhörers zu versetzen und seine Sprache an die gegebene kommunikative Situation anzupassen (Barrett, 2001).

In den Erläuterungen zu Barretts Komponentenmodell der Sprachentwicklung ist deutlich geworden, dass die kindliche Sprachentwicklung nicht von einer Komponente allein bestimmt wird. Der Erwerb von Sprache muss als Resultat eines komplexen Zusammenspiels von phonologischen, semantisch-morphologischen und pragmatischen Kompetenzen verstanden werden (Plaut & Kello, 1999; Tracy, 2000).

Ergänzend zu den Entwicklungsprozessen der beschriebenen Sprachdomänen kommt auch der kognitiven Entwicklung eine bedeutsame Rolle im Spracherwerb zu. Viele Studien stellen die Bedeutung des Arbeitsgedächtnis, genauer gesagt der Phonologischen Schleife, für den Spracherwerb heraus (Gathercole & Baddeley, 2004; Grimm 1995; Janczyk, Schöler & Grabowski, 2004). Grundlegenden Gedächtnisleistungen ermöglichen Abstraktions- und Lernprozesse, die eng mit dem Spracherwerb assoziiert sind (Hasselhorn & Werner, 2000; Weinert, 2000, 2006). Auch externale Voraussetzungen, wie das Sprachangebot in der Familie, können die Sprachentwicklung beeinflussen. Für Hof-Ginsberg (2000) sind die Häufigkeit und die Intensität des sprachlichen Angebotes der Eltern maßgebende Einflussfaktoren für ein sprachförderliches Umfeld.

Trotz des komplexen Bedingungsgefüges, gelingt es Kindern im Normalfall recht mühelos immanentes Regel- und Konventionswissen ihrer Muttersprache zu erwerben. Im Folgenden sollen bedeutsame Phasen der kindlichen Sprachentwicklung näher betrachtet werden.

In Anlehnung an das beschriebene Barrett'sche Entwicklungsmodell lassen sich bedeutsame Erwerbsprozesse in vier Entwicklungsdomänen bestimmen: (1) *phonologisch-phonetische Entwicklung* (2) *lexikalisch-semantische Entwicklung* (3) *morpho-syntaktische Entwicklung* und (4) *pragmatische Entwicklung*. Sie sollen im Folgenden vertiefend und mit Bezug auf spezifische Forschungsliteratur dargestellt werden.

1.1 Phonologisch-phonetische Entwicklung

Es wurde bereits erwähnt, dass Säuglinge schon sehr früh die Fähigkeit entwickeln, Lautströme aus ihrer Umwelt in sprach-relevante Einheiten zu gliedern. Schon im ersten Lebenshalbjahr sind Babys in der Lage nach prosodischen Merkmalen ihre Muttersprache von Fremdsprachen zu unterscheiden. Dabei zeigen sie eine Präferenz für ihre Umgebungssprache (Mehler et al., 1988; Nazzi, Bertoncini & Mehler, 1998). Auch gelingt es ihnen verschiedene Betonungsmuster und Wörter nach ihrer Silbenzahl zu differenzieren (Bijeljic-Babic, Bertoncini & Mehler, 1993).

In einer weiteren Experimentalreihe zeigten Christophe, Mehler und Sebastián-Gallés (2001), dass Babys auch Silbenfolgen, die eine Wortgrenze beinhalten (z.B. Ma-Ma-Pa-Pa), und Silbenfolgen ohne Wortgrenze voneinander unterscheiden können. Interessanterweise zeigten die untersuchten Säuglinge diese Fähigkeit nicht nur bei muttersprachlichen Stimuli (Französisch). Auch bei der Vorgabe fremdsprachlicher Stimuli (Spanisch) gelang es ihnen, Unterschiede zwischen Silbenfolgen mit und ohne Wortgrenze wahrzunehmen, sofern die Silben der Muttersprache ähnelten.

Die berichteten Erkenntnisse der Säuglingsforschung liefern Evidenz für die Annahme, dass Neugeborene diese basalen Segmentierungs- und Differenzierungsfähigkeiten aufgrund einer sehr hohen Sensitivität für prosodische Informationen erwerben (Höhle & Weissenborn, 2000; Weinert, 2006; Jusczyk, 2001). Den besonderen Stellenwert der Prosodie für die Perzeptionsleistung sehen die Forschergruppen auch darin bestätigt, dass Säuglinge auch Sprachstimuli, die mit einem *Low-Pass Filter* bearbeitet wurden, diskriminieren können. Ein Low-Pass Filter reduziert den hohen Frequenzbereich einer Sprachaufnahme. Dadurch sind nur die prosodischen Merkmale hörbar, nicht aber einzelne Laute und Silben (Höhle, 2004, 2005; Nußbeck, 2007). Auch unter diesen Bedingungen gelang es Babys ihre Muttersprache von Fremdsprachen zu unterscheiden. Säuglinge schenken den prosodischen Merkmalen ihrer Muttersprache im frühen Stadium des Spracherwerbs offensichtlich viel Aufmerksamkeit, um weitere Informationen über die Beschaffenheit ihrer Muttersprache zu erhalten. Einige Autoren beschreiben den Fokus auf die Prosodie als „Steigbügel“ für den strukturellen Erwerb der Muttersprache (Penner, 2000). International hat sich der Begriff des *prosodic bootstrapping* durchgesetzt (Cutler, 1999; Penner, 2000; Szagun, 2006).

Im Verlauf des ersten Lebensjahres beginnen Babys nicht nur damit lautsprachlichen Input zu diskriminieren. Mit zunehmender Fokussierung auf die lautliche Struktur ihrer Muttersprache

nehmen sie Laute auch ungeachtet ihrer Variation (z.B. durch die väterliche oder mütterliche Stimme) als konstante Kategorien wahr (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Owens, 2005).

Dass Säuglinge schon innerhalb des ersten Lebensjahres in der Lage sind muttersprachliches Lautmaterial in prototypischen Kategorien zu organisieren belegten Eimas, Siqueland, Jusczyk und Vigorito (1971) in einem vielbeachteten Experiment. Die Forschergruppe untersuchte, ob Babys im Alter von 1 bis 4 Monaten zwischen und innerhalb der Lautkategorien „ba“ und „pa“ Unterschiede wahrnehmen können. Um die Diskriminierungsfähigkeit innerhalb einer Lautkategorie zu testen, wurden den Kindern klangliche Varianten der Stimuli „ba“ bzw. „pa“ dargeboten. Die Ergebnisse zeigten, dass Säuglinge den Unterschied zwischen beiden Lautkategorien erhören können. Es gelang ihnen allerdings nicht zwischen Klangvarianten einer Kategorie zu unterscheiden.

Kuhl et al. (1992) beschreiben in der Entwicklung der kategorialen Wahrnehmung von Lauten den sogenannten „Magnet-Effekt“. Sie gehen davon aus, dass jeder Lautkategorie ein phonetischer Prototyp zugrunde liegt, der wie ein Magnet schwächere und unausgeprägte Varianten des gleichen Lauttyps quasi „anzieht“.

Mit der frühen Organisation des muttersprachlichen Lautsystems vollziehen Kinder in den ersten zehn Monaten einen bedeutsamen *shift*. Die generelle Fähigkeit jedwede Sprachmerkmale zu differenzieren weicht einer spezifischen Sensitivität für die phonetisch-phonologische Struktur der Muttersprache. In diesem Entwicklungsstadium sind prosodische Merkmale längst nicht mehr der alleinige Fixpunkt der kindlichen Sprachwahrnehmung. Durch den Wahrnehmungsfokus für die Lautstruktur der Muttersprache gelingt es Babys sehr schnell sequenzielle, regelhaft auftretende Lautfolgen wiederzuerkennen.

Parallel zu den rezeptiven Prozessen der Sprachwahrnehmung entwickeln Kinder schon im ersten Lebensjahr basale Fähigkeiten auf dem Gebiet der Lautproduktion und Artikulation.

In der 6. bis 8. Lebenswoche beginnen Säuglinge erste *Gurr-Laute* zu produzieren und erreichen über das zwischenzeitliche *Expansionsstadium* (gekennzeichnet durch zunehmende Produktion nichtsprachlicher Laute wie Quieken, Brummen, etc.) mit ca. sechs Monaten die *Lallphase* (Penner, 2000; Weinert, 2006). Der Beginn der Lallphase wird bei normalentwickelten Säuglingen durch die kanonische Wiederholung bestimmter Lautsequenzen gekennzeichnet („dadadadada“, „bababa“; Oller & Eilers, 1988). Im weiteren Verlauf der Lallphase ist zu beobachten, dass mehrsilbige Lautketten mit unterschiedlichen Artikulationsstellen (z.B. „dabadaba“) das anfängliche Lallen ablöst.

Mit ca. 10 - 12 Monaten erreicht die Lautproduktion der Kinder eine neue Qualität. Zu beobachten ist ein sogenannter *babbling-drift*. Die Lautproduktion der Kinder nähert sich also immer stärker den Eigenschaften der Zielsprache an (Deutsch & El Mogharbel, 2007b). Daran wird deutlich, dass die Silbenstrukturen, die Säuglinge im ersten Lebensjahr produzieren, keinesfalls universeller Natur sind. Sie konstituieren sich aus Lautclustern, die in der Muttersprache besonders häufig vorkommen (Grimm & Weinert, 2002).

Interessanterweise berücksichtigen Eltern die Entwicklungsvoraussetzungen ihrer Kinder scheinbar intuitiv und passen ihre Sprache in der Interaktion mit Babys an deren Sensibilität für Prosodie, Rhythmus und Betonung an (Weinert, 2006). Noch bevor der Säugling in der Lage ist erste Sprachlaute zu produzieren, wenden Eltern einen stark verlangsamten Sprachstil in hoher Tonlage an. Dieser ist durch viele Pausen und Wiederholungen gekennzeichnet und wird landläufig als *Ammensprache* oder *baby-talk* bezeichnet (Grimm & Weinert, 2002; Meyer, Jungheim & Ptok, 2011). Baby-talk tritt kulturunabhängig in Erscheinung (Ferguson, 1964). Eltern greifen zumeist bis zur Vollendung des ersten Lebensjahres ihres Kindes auf diesen Sprechstil zurück. Baby-talk bietet dem Kind die Möglichkeit lautliche Strukturen seiner Umwelt nachzuahmen. Durch die elterlichen Sprachmodelle erhält der Säugling gleichzeitig eine Art korrekatives Feedback über die „Qualität“ seiner produzierten Laute und wird so an das Lautinventar der Muttersprache herangeführt (Papoušek, 2012). Zu diesem Zeitpunkt entstehen in der Eltern-Kind-Kommunikation schon erste dialogische Strukturen, die für die weitere Sprachentwicklung eine bedeutsame Rolle spielen.

Gegen Ende des ersten Lebensjahres entwickeln Kinder eine Vorstellung davon, welche Lautmuster des Sprachstroms als bedeutungsvolle Einheiten in Frage kommen und produzieren erste *Protowörter*. Gemeint sind damit diejenigen Lautstrukturen, die einen Wortcharakter aufweisen, aber von der Umwelt nicht zwangsläufig als bedeutungstragende Einheiten erkannt werden (Deutsch & El Mogharbel, 2007b; Nußbeck, 2007). Kinder entwickeln zunehmend die Fähigkeit zu erkennen, welche Lautmuster der Muttersprache für bestimmte Wörter stehen.

Aus neurologischer Perspektive untersuchten Friedrich und Friederici (2005) ereigniskorrelierte Potentiale (EKP), die in 12- und 19-Monate alten Säuglingen ausgelöst wurden, wenn sie zu einer bildlichen Darbietung entweder das passende Wort, ein muttersprachnahes Pseudowort oder ein phonetisch fremdes Nichtwort hörten. 19 Monate alte

Kinder zeigten bei der Präsentation von realen und muttersprachlichen Pseudowörtern ein ähnliches EKP wie eine erwachsene Kontrollgruppe. Das heißt: Auch phonotaktisch regelhafte Pseudowörter werden im Alter von 19 Monaten als potentielle neue Wörter wahrgenommen, während nicht zulässige Lautkombinationen keine neuronale Reaktion auslösten. Diese Differenzierung konnte bei Kindern am Ende des ersten Lebensjahres noch nicht nachgewiesen werden.

Es ist deutlich geworden, dass Säuglinge schon kurz nach der Geburt beginnen phonetisch-phonologische Charakteristika ihrer Muttersprache zu erwerben. Die phonetisch-phonologische Entwicklung unterliegt einem komplexen Zusammenspiel aus perzeptiven und produktiven Fertigkeiten. Weiter wirken intuitive Prinzipien der Eltern-Kind-Kommunikation (Verlangsamung, Wiederholen des Gesagten, positive Reaktion auf lautliche Äußerungen) förderlich auf die Entwicklung von phonetisch-phonologischen Kompetenzen. Auch wenn Kinder zu diesem Zeitpunkt nur wenige erkennbare Wörter produzieren, begleiten Eltern den Sprachentwicklungsprozess ihres Kindes intuitiv mit einer stützenden Sprache (*scaffolding*). Das responsive Verhalten der Eltern fördert Sprachverständnis und Sprachproduktion des Kindes gleichermaßen. Um mit seinen Eltern in einen Dialog treten zu können, gewinnt für das Kind die Aneignung und Produktion von Wörtern große Bedeutung (Tomasello, 2000). An dieser Stelle vollzieht das Kind einen bedeutsamen Schritt in der *lexikalisch-semantischen Entwicklung*. Die einzelnen Phasen dieser Entwicklungsdomäne sollen im nächsten Kapitel beschrieben werden.

1.2 Lexikalisch-semantische Entwicklung

Die lexikalisch-semantische Entwicklung beginnt nicht erst mit der Produktion von Wörtern, sondern ist eng mit den phonetisch-phonologischen Kompetenzen und den zeitlich parallelen sozialen, sozial-kognitiven und kognitiv-konzeptuellen Entwicklung assoziiert (Weinert, 2006).

Noch ehe Kinder selbst Wörter produzieren, können sie Lautkombinationen, die sich häufig in der Umwelt Verwendung finden, als eigenständige Einheit wiedererkennen (Juszyk, 2000). Nach Rothweiler und Kauschke (2007) setzt der Beginn der lexikalisch-semantischen Entwicklung dann ein, wenn es Kindern gelingt lautlichen Einheiten eine Bedeutung zuzuweisen. Gleichzeitig bewältigen Kinder die komplexe Aufgabe neuerlernte Wortbedeutungen in ein mentales Kategoriensystem einzuordnen. Dieses Organisationssystem wird häufig als *mentales Lexikon* beschrieben (Kuczaj, 2001). Das mentale Lexikon speichert eine Reihe von Informationen über jede erworbene lexikalische Einheit. Unterschiedliche Informationen werden im mentalen Lexikon schließlich miteinander in Beziehung gesetzt (Rothweiler & Kauschke, 2007).

Kuczaj (2001) nennt vier große Informationskategorien, nach denen Wörter im mentalen Lexikon organisiert werden. Dazu zählen (1) phonetisch-phonologische Informationen (z.B. Aussprache, Wortakzent etc.), (2) morphologisch-syntaktische Informationen (z.B. Genus, Wortart etc.), (3) semantische Informationen (Wortbedeutung) und (4) pragmatische Informationen (Verwendungskonventionen innerhalb der Kommunikation). Einträge im mentalen Lexikon unterliegen ein Leben lang Reorganisationsprozessen, denn der Wortschatzerwerb ist auch im Erwachsenenalter nicht abgeschlossen.¹

Kinder beginnen schon in den ersten Lebensmonaten phonologischen Einheiten eine Bedeutung zuzuweisen. Mit 10 Monaten können Säuglinge bereits über 60 Wörter rezeptiv verstehen (Grimm, 2003; Menyuk, 2000; Weinert 2006). Parallel dazu beginnen sie erste Protowörter zu produzieren und referentiell zu gebrauchen. Das Erlernen der ersten 50 Wörter schreitet verhältnismäßig langsam voran (Dromi, 2001). Aber viele internationale Untersuchungen beschreiben einen sogenannten *Vokabelspurt*, den Kinder nach dem Erwerb der ersten 50 Wörter durchlaufen (Bates, Marchmann, Thal et al., 1994; Goldfield & Reznick, 1990). Kinder erlernen ab diesem Entwicklungspunkt immer schneller neue Wörter und stellen semantische Referenzen her. Häufig treten dabei *Überdiskriminierungen* („Sofa“ nur

¹ Der Wortschatz eines bildungsnahen Erwachsenen umfasst ca. 100.000 Wörter und erweitert sich mit zunehmender Lebensdauer stetig (Nußbeck, 2007; Kuczaj, 2001).

für das Sofa im eigenen Wohnzimmer) und *Übergeneralisierungen* („Hund“ für alle Vierbeiner) auf. Gleichzeitig gelingt es Kindern unbekannte Wörter schon nach dem ersten Hören zu reproduzieren und neue Einträge im mentalen Lexikon miteinander zu verknüpfen (*fast-mapping*; Menyuk, 2000; Rothweiler & Kauschke, 2007; Owens, 2005; Weinert, 2006).

In der Literatur werden Alterspanne und der Verlauf des Vokabelspurts kontrovers diskutiert (Ganger & Brent, 2004). Szagun (2006) vertritt die Ansicht, dass der Vokabelspurt keinem genauen Entwicklungsalter zugeordnet werden kann, da die Entwicklung des Wortschatzes äußerst individuell verlaufe. Andere Autoren (Grimm & Weinert, 2002; Plunkett & Schafer, 2001; Siegmüller & Kauschke, 2011) verorten den Beginn des Vokabelspurts zwischen dem 17. und dem 24. Lebensmonat. Die Begriffe *late-talker* oder günstigenfalls *late-bloomer* beschreiben Kinder, die nach Beginn des zweiten Lebensjahres die 50-Wort-Grenze noch nicht überschritten haben (hierzu ausführlich: Schlesiger, 2009).

Der Einsatz des Vokabelspurts gilt als Indiz für fundamentale Umbrüche in der kognitiven Entwicklung des Kindes. Phonologische Informationen können im Arbeitsgedächtnis nun länger verfügbar gehalten (Gathercole & Baddeley, 2004) und elaboriertere Kategorisierungen und Assoziationen vollzogen werden (Johnson, Jusczyk, Cutler & Norris, 2003).

Tomasello (2000) sieht – wie in Kapitel 1.1 bereits angedeutet – den Ursprung des rasanten Wortzuwachses in Veränderungen des Kommunikationsverhaltens von Eltern und Kind. Er nimmt an, dass es Kindern zunehmend gelingt Handlungsintentionen ihres Gegenübers zu deuten und entsprechend zu handeln. Daraus resultieren längere und intensivere Eltern-Kind-Dialoge, die sich immer häufiger auf einen gemeinsamen Aufmerksamkeitsfokus beziehen (*joint attention*). Der gemeinsame Bezug auf einen Gegenstand (beispielsweise beim Betrachten eines Bilderbuches) erleichtert das Erlernen neuer Wörter (Nußbeck, 2007). Kindern gelingt es, die von nun an neu gelernten Wörter schneller mit Wörtern des bestehenden Wortschatzes in Verbindung zu setzen. Dieser Prozess wird *fast-mapping* genannt.

Die Fähigkeit zum fast-mapping wird häufig auch mit Hilfe sogenannter *constraint-Theorien* erklärt (Markman, 1992). Befürworter der constraint-Theorie nehmen an, dass Kinder eine Reihe von Erwartungen und Vorannahmen (*constraints*) in das Erlernen neuer Wörter (häufig sind dies Substantive) einbringen. „Diese Vorannahmen beschränken den Hypothesenraum

über die Bedeutung neuer Wörter erheblich und machen damit das sogenannte Induktionsproblem des Wortschatzerwerbs lösbar.“ (Weinert, 2006, S. 636).

Es können vier Arten von constraints unterschieden werden (Rothweiler & Kauschke, 2007; Weinert, 2006):

1) Type assumption:

Neue Wörter beziehen sich auf Klassen und nicht auf Individuen.

Beispiel.: Wenn Kinder das Wort „Hund“ hören, nehmen sie an, dass alle Hunde so bezeichnet werden und nicht der Einzelne.

2) Whole Object assumption:

Neue Wörter bezeichnen das ganze Objekt und nicht nur Teile davon.

Beispiel.: Wenn Kinder das Wort „Hund“ hören, nehmen sie an, dass sich das Wort auf das ganze Objekt und nicht nur auf einen Teil (z.B. die Pfote) bezieht.

3) Taxonomic assumption:

Neue Wörter beziehen sich auf das Ausgangs-Objekt und sind häufig gleicher Art.

Beispiel.: Wenn ein Kind „Hund“ kennt, wird es weitere Objekte mit ähnlichen Eigenschaften auch der Kategorie „Hund“ zuordnen.

4) Mutual exclusive assumption / Nameless category bias:

Neue Wörter nehmen Bezug auf eine noch nicht vorhandene Kategorie. Objekten kann jede Wortbezeichnung nur einmal vergeben werden.

Bei den genannten constraints handelt es sich nicht um deterministische Regeln, die die Kinder stereotyp befolgen (Weinert, 2006). Unklar ist, in welcher Häufigkeit und in welchem Entwicklungsstadium Kinder von den einzelnen Strategien Gebrauch machen. Rothweiler und Kauschke (2007) halten dafür, dass das Kind im Lexikonerwerb alle diskutierten Aspekte nutzt und dabei erwerbssteuernde Strategien selbst entwickelt. Kritiker der constraint-Theorien merken an, dass sich constraints zwar auf den Erwerb bestimmter Wortbedeutungen anwenden lassen, nicht aber als allgemeingültige Erwerbsprozesse verstanden werden sollten. Dass Kinder schon sehr früh Worte für Eigenschaften („lieb“), Mengen („alle“, „viel“) oder

auch Details („Pfote“ beim Hund) semantisch korrekt verwenden können, scheint einigen Annahmen der constraint-Theorie zu widersprechen (Kuczaj, 2001).

Mit einem Wortschatzumfang von etwa 100 bis 200 Wörtern steigt allmählich die Komplexität der kindlichen Aussagen. In vielen Sprachen produzieren Kindern in der Ein-Wort-Phase zumeist Nomen (*Noun-bias*). Mit dem Gebrauch von Mehr-Wort-Äußerungen beginnen Kinder syntaktische Bezüge zwischen einzelnen Wörtern herzustellen.

Bates und Kollegen (1994) untersuchten in einer großangelegten Studie die Wortschatzentwicklung von 1803 amerikanischen Kindern zwischen 0;8 und 2;6. Über eine Checkliste² mit 396 Wörtern sollten Eltern einschätzen, ob ihr Kind in seiner aktuellen Entwicklungsphase diese Wörter verstehen oder produzieren kann. Die Autoren fanden signifikante Interaktionseffekte zwischen der Größe des Wortschatzes und den produzierten Wortarten. Kinder mit einem Wortschatzumfang von bis 100 Wörtern produzierten hauptsächlich Substantive. Bei einem Wortschatz zwischen 50 und 200 Wörtern traten vermehrt Verben und Adjektive auf. Ab ca. 400 Wörtern verwendeten die Kinder zusätzlich geschlossene Wortklassen (Adverbien, Konjunktionen, Interjektionen, Präpositionen, Artikel, Pronomen).

Mit ihren Ergebnissen sehen die Autoren die sogenannte *critical mass*-Hypothese bestätigt. Danach wird angenommen, dass Kinder der Erwerb von morpho-syntaktischem Regelwissen erst dann gelingt, wenn sie über einen gewissen „Grundstock“ an Vokabular verfügen. Verfügen Kinder schließlich über morpho-syntaktische Kompetenzen, so können diese ebenso als Steigbügel zum Erwerb neuer Wörter und Wortformen dienen. Gleitman (1990) spricht in diesem Zusammenhang vom *syntactic bootstrapping*. Die lexikalisch-semantische Entwicklung ist also eng mit der Entwicklung morpho-syntaktischer Fähigkeiten verknüpft.

Es wurde bereits erwähnt, dass Eltern die Wortschatzentwicklung ihres Kindes intuitiv durch den Einsatz von Kommunikationsstrategien fördern. Als besonders wortschatzförderlich gelten zum Beispiel Vorlesesituationen mit Bilderbüchern (Grimm, 2003; Kleeck, Stahl & Bauer, 2003; Ritterfeld, 2000; Zevenbergen & Whitehurst, 2003). Dabei unterstützen Eltern den Sprachentwicklungsprozess ihres Kindes, indem sie es unbekannte Wörter nachsprechen lassen und Leerstellen im Wortschatz mit sprachlichem Input füllen. Gleichzeitig fördern Eltern durch das Vorsprechen von komplexeren sprachlichen Strukturen auch die

² Es handelte sich um das MacArthur Communicative Development Inventory (CDI), das Wörter aus verschiedenen semantischen Feldern berücksichtigt.

Entwicklung von morpho-syntaktischen Kompetenzen. Bedeutsame Phasen der morpho-syntaktischen Entwicklung sollen im nächsten Kapitel beschrieben werden. Tabelle 1 fasst abschließend die zentralen Phasen des Wortschatzerwerbs zusammen.

Tabelle 1 Phasen des Wortschatzerwerbs nach Grimm & Weinert (2002) und Weinert (2006).

Phase I	• Erstes Wortverständnis ab dem 10. Lebensmonat • Langsamer Anstieg des produktiven Wortschatzes • Vornehmlich soziale Wörter und kontextgebundene Bezeichnungen • ab ca. 18 Monaten: 50 Wörter produktiv; ca. 200 Wörter rezeptiv
Phase II	• Schneller Zuwachs des nominalen Wortschatzes (<i>Wortschatzspurt</i>) • Schnelle Zuordnung von Wortbedeutungen (<i>fast mapping</i>) • ab 100 - 200 Wörtern erste Wortkombinationen
Phase III	• ab 30 Monaten: Verben und andere relationale Wörter werden erworben • Syntaktische Merkmale als Steigbügel (<i>syntactic bootstrapping</i>)

1.3 Morpho-syntaktische Entwicklung

Die Phase, in der Kinder beginnen erste Sätze mit zwei oder mehr Wörtern zu produzieren, markiert den Einstieg in die produktive Grammatikentwicklung (Dannenbauer, 2002; Grimm & Weinert, 2002).

Rezeptiv gelingt es Kindern weitaus früher, gehörte Sätze aufgrund syntaktischer Eigenschaften voneinander zu unterscheiden. Dies zeigten Laborexperimente von Hirsh-Pasek und Golinkoff (1993, 1996) mit Kindern im Alter von 17 Monaten. Zum Zeitpunkt der Untersuchungen befanden sich alle Kinder in der Einwortphase. Getestet wurde u.a. das Verständnis von Sätzen wie: „*Cookie Monster is washing Big Bird*“. Dafür wurden den Kindern zu jedem Satz zwei Bilder präsentiert. Ein Bild zeigte die Handlung des vorgegebenen Satzes (Cookie Monster wäscht Big Bird), während im zweiten Bild Subjekt und Objekt vertauscht wurden (Big Bird wäscht Cookie Monster). Die Messung der Blickpräferenz ergab, dass die Kinder das korrekte Bild signifikant länger ansahen als den inkongruenten Stimulus. Vermutlich haben Kinder also schon Kenntnisse über die grammatikalischen Strukturen ihrer Muttersprache erworben (hier: die Subjekt-Verb-Objekt Stellung im Englischen), bevor sie in der Lage sind diese selbst zu produzieren.

Welche syntaktischen Strukturen Kinder zuerst präferieren, ist in hohem Maße von der Struktur ihrer Muttersprache abhängig. Somit darf „[...] das Befundmuster aus Studien mit englischsprachigen (oder deutschsprachigen) Studien nicht vorschnell übergeneralisiert werden.“ (Weinert, 2006; S. 657).

Für deutschsprachig aufwachsende Kinder konnte gezeigt werden, dass sie zu Beginn der morpho-syntaktischen Entwicklung häufig Satzformen mit Subjekt-Objekt-Verb Struktur („*Papa Kuchen essen*“) produzieren. Französisch- und englischsprachig aufwachsende Kinder im selben Entwicklungsstadium produzieren indes häufiger Subjekt-Verb-Objekt Kombinationen („*John eats cake*“; Weissenborn, 2000). Weiter gilt für englisch- und deutschsprachige Kinder, dass sie viele Flexionsregeln zunächst auslassen, und der Gebrauch flektierter Verbformen viel später auftritt als beispielsweise bei italienischsprachigen Kindern (Weissenborn, 2000; Weinert, 2006). Deutsche Kinder produzieren zu Beginn der morpho-syntaktischen Entwicklung vornehmlich Inhaltswörter (z.B. „*Ball habe*“). Artikel, Präpositionen und andere Funktionswörter kommen nur selten vor. Internationale Forschungsarbeiten befassen sich häufig mit der Analyse der ersten kindlichen Wortkombinationen (Tomasello, 2003). Es geht dabei um die Frage, auf welchen Wegen Kinder die tiefen Regularitäten ihrer Muttersprache erwerben.

Zwei Annahmen lassen sich nach Weinert (2006) unterscheiden.

Vertreter der *Kontinuitätsannahme* (Pinker, 1984) gehen davon aus, dass Kinder nativ linguistischen Abstraktionen der Erwachsenensprache besitzen. Dieses explizite Regelwissen spiegelt sich daher schon in den frühen Wortkombinationen wider (Nußbeck, 2007). Dagegen nehmen Vertreter der *Diskontinuitätsannahme* (Brain, 1967) an, dass es sich bei den ersten Wortkombinationen um rein semantische Relationen handelt, die keinem Regelwissen unterliegen. Tomasello (2003) konstatiert, dass für beide Annahmen keine überzeugenden empirischen Belege vorliegen. Nativistisch geprägte Autoren vermochten im Sinne der Kontinuitätsannahme nicht zu klären, weshalb sich Kinder höchst unterschiedliche Sprachsysteme aneignen, wenn ihnen gleichzeitig eine universelle Basis an grammatikalischen Kompetenzen (*Universalgrammatik*) zur Verfügung steht.

Gegen die *Diskontinuitätsannahme* spricht, dass Kinder die Wortordnungsregeln ihrer Muttersprache sehr früh beherrschen. Kinder formulieren Sätze wie „*Die schön*“ oder „*Das groß*“. Äußerungen, in denen das Adjektiv an erster Stelle verwendet wird („*Schön die*“) kommen in der Kindersprache so gut wie nicht vor, obwohl auch hier ein enger semantischer Zusammenhang gegeben ist (Grimm & Weinert, 2002).

Karmiloff-Smith (1999) konnte zeigen, dass Kinder in der Lage sind bedeutungslosen Pseudowörtern einen maskulinen oder femininen Artikel korrekt zuzuordnen. Sie orientieren sich dafür vermutlich an bestimmten morphologischen Eigenschaften (z.B. Wortendungen) und folgen unter Berücksichtigung bekannter Regeln der Umgangssprache einer probabilistischen Annahme. Aber möglicherweise sind Kinder auch sehr früh in der Lage strukturelle Informationen über die Grammatik ihrer Muttersprache dialogischen Situationen zu entnehmen, in denen Eltern und Kind miteinander kommunizieren.

Tomasello (2003) sieht in den kommunikativen Aktivitäten zwischen dem Kind und seiner Umwelt den Schlüssel zur Bildung von abstraktem Regelwissen.

„Children build up the abstractness of their item-based constructions using their general skills of intention-reading and pattern-finding. They understand the communicative functions of utterances that embody various syntactic constructions, as well as some of their constituents, by reading the intention of the speaker. They then find patterns across item-based constructions by schematizing and making analogies.” (Tomasello, 2003, S.143).

Die kindgerichtete Sprache von Erwachsenen bietet dem Kind also eine Fülle an morpho-syntaktischen Hinweisen, aus denen es Regelmäßigkeiten der Muttersprache ableiten kann.

Gleichzeitig verwenden Eltern zumeist intuitiv bestimmte Sprechtechniken, mit deren Hilfe sie die morpho-syntaktische Qualität der kindlichen Aussage implizit zurückmelden. Solche rudimentären Dialoge zwischen Eltern und Kind lassen sich ab dem Entwicklungszeitpunkt beobachten, an dem das Kind in der Lage ist, im *telegraphischen Sprachstil* zu kommunizieren (Grimm & Weinert, 2002). Die Äußerungen des Kindes sind zu diesem Zeitpunkt von vielen Leerstellen und Auslassungen geprägt, sodass manche Aussagen nur enge Bezugspersonen im direkten Handlungskontext verstehen können. „*Mama weg*“ könnte z. B. bedeuten, dass die Mutter gerade aus der Tür gegangen ist, oder dass ebendiese eine Handlung beenden soll. Ohne den Handlungskontext zu kennen ist es kaum möglich die intendierte Aussage des Kindes zu entschlüsseln.

Eltern reagieren auf den telegraphischen Sprachstil ihres Kindes meist intuitiv. Sie nehmen Interpretationen vor, indem sie die kindliche Aussage wiederholen und gleichzeitig fehlende Informationen ergänzen („*Ja, die Mama ist weg gegangen*“). Damit greifen Eltern Wortkombinationen auf, die dem Kind bereits bekannt sind und erweitern diese um unbekannte grammatikalische Strukturen. Wenngleich die Komplexität der Kindersprache noch weit von der Erwachsenensprache entfernt ist, erhalten Kinder auf diesem Weg die Möglichkeit, die eignen Äußerungen mit der Reaktion ihrer Umwelt abzugleichen. So können sie ihre linguistischen Kompetenzen weiterentwickeln (Ritterfeld, 2000).

In der Folgezeit steigt die durchschnittliche Länge der kindlichen Äußerungen (MLU)³ auf ca. 4,5 Morpheme bis zum dritten Lebensjahr an (Weinert, 2006). Kindern gelingt es häufiger Verben korrekt zu flektieren. Sie erwerben allmählich das Prinzip der Subjekt-Verb-Kongruenz. Die Subjekt-Verb-Kongruenz beschreibt die morphologische Abhängigkeit des Verbs vom vorausgegangenen Subjekt. Viele Kinder können von nun an Verbendungen auf das Subjekt der Nominalphrase abstimmen (Ich *spiele*, Du *spielst*, usw.; Kany & Schöler, 2010).

Mit zunehmender Spracherfahrung erwerben Kinder typische Wortstellungsmuster der deutschen Sprache (Dannenbauer, 2002). Ein Merkmal ist die Verbzweitstellung. In Hauptsätzen bildet das Verb immer die zweite Konstituente, während bspw. Subjekt oder Objekt flexibel an erster Stelle stehen können. Dieses grammatikalische Regelwissen haben Kinder mit etwa drei Jahren erworben (Clahsen, 1982).

³ Die durchschnittliche Äußerungslänge als wird MLU (*mean length of utterance*) bezeichnet. Sie berechnet sich über die durchschnittliche Anzahl von Wörtern unter Berücksichtigung der Stamm- und Flexionsmorpheme in einer Äußerung. Sie gilt als Maß zur Erfassung der sprachlichen Produktivität eines Kindes.

Gegen Ende des dritten Lebensjahres gelingen Kindern schließlich komplexe Sätze (z.B. Konditionalsätze). Bis zum vierten Lebensjahr versteht das Kind Passivsätze und beginnt diese auch zu produzieren (Kany & Schöler, 2010). Wenn entsprechendes Regelwissen noch nicht vorhanden ist, greifen Kinder häufig auf Interpretationsstrategien zurück um den Inhalt komplexer Sätze zu verstehen. Sie orientieren sich zum Beispiel an der Reihenfolge, in der Personen und Ereignisse in einem Satz erwähnt werden, oder sie ziehen eigenes Weltwissen zum Verständnis heran (Weinert, 2006). Werden bestimmte Satzkonstruktionen indes noch nicht sicher beherrscht, können die genannten Interpretationsstrategien zu Verständnisfehlern führen. Dies zeigt sich in der Sprachdiagnostik häufig in Manipulationsaufgaben zum Sprachverständnis. Mit Hilfe von Figuren sollen Kinder zum Beispiel vorgegebene Temporalsätze wie *„Der Mann springt dreimal in die Luft, nachdem die Frau ihn getragen hat.“* nachspielen (Petermann, Metz & Fröhlich, 2010). Wenn ihnen das Verständnis von Temporalsätzen noch Schwierigkeiten bereitet, bewegen Kindern die Spielfiguren häufig nach der Nennungsreihenfolge im Satz.

Mit zunehmender Spracherfahrung erlernen Kinder grammatikalische Regeln zu dekontextualisieren. Es gelingt ihnen grammatische Formen (z.B. in Texten oder Erzählungen) isoliert und unabhängig vom Inhalt zu betrachten und die Grammatik ihrer Muttersprache zu analysieren (Weinert, 2006). Zusammenfassend seien die Meilensteine der morpho-syntaktischen Entwicklung in Anlehnung an Tracy (2008) tabellarisch dargestellt.

Tabelle 2 Entwicklung der syntaktischen Fähigkeiten nach Tracy (2008).

Alter	Entwicklung
1 bis 1;5	• Einwortäußerungen (meist Nomen, Partikel)
1;5 bis 2	• Erwerb des syntaktischen Prinzips
2 bis 3	• einfache vollständige Sätze • zielsprachliche Wortstellungen (z.B. Verbzweitstellung oder Subjekt-Verb-Kongruenz)
3 bis 4	• komplexe Satzgefüge (z.B. Konditionalsätze) • Nebensätze mit flektiertem Verb am Ende • die meisten Wortklassen sind verfügbar

1.4 Pragmatische Entwicklung

Es ist bereits in den vorangegangenen Abschnitten deutlich geworden, dass der Erwerb einer Sprache nicht allein darin besteht, eine Vielzahl von Wortbedeutungen oder grammatikalischen Regeln zu kennen. Sprachliche Kompetenz äußert sich auch in der Fähigkeit zur Gesprächsführung. Erfahrene Sprecher verfügen über pragmatische Kompetenzen und sind mit Gesprächsregeln vertraut (Nußbeck, 2007).

In den vorangegangenen Abschnitten ist darauf hingewiesen worden, dass Säuglinge die dialogtypischen Wechselmechanismen von Sprechen und Zuhören noch nicht beherrschen. Sie erwerben rudimentäre pragmatische Fertigkeiten zunächst über die sprachliche Interaktion mit den engsten Bezugspersonen (Deutsch & El Mogharbel, 2007b). Bruner (1983) misst dem frühen Dialog zwischen Eltern und ihrem Kind hohe Bedeutung für die weitere Sprachentwicklung bei. Er nimmt an, dass schon Kleinkinder über ein sogenanntes *Language Acquisition Support System* (LASS) verfügen. Bruner geht davon aus, dass Kinder mit Hilfe des LASS in der Lage sind, strukturelle Informationen ihrer Verkehrssprache zu analysieren und auf die eigene Sprachproduktion anzuwenden.

„It [LASS] frames or structures the input of language and interaction to the child's Language Acquisition Device in a manner to 'make the system function'.”
(Bruner, 1983, S. 19).

Im Sinne der Outside-In-Theorie (vgl. Kapitel 1) ist auch Bruner der Ansicht, dass die Sprachentwicklung von Kinder maßgeblich vom Sprachangebot der Eltern (Input) bestimmt wird.

Bei den zuvor beschriebenen Prozessen der phonetisch-phonologischen, semantisch-lexikalischen und morpho-syntaktischen Entwicklung ist bereits darauf hingewiesen worden, dass Eltern durch intuitive Modellierungsstrategien eine Basis für die Sprachförderung des Kindes legen. Die Anpassung der Erwachsenensprache an das Niveau des Kindes ist z. B. beim *baby-talk* zu beobachten. Bei Unterhaltungen mit dem Kind wählen Eltern zumeist eine langsame Geschwindigkeit und sowie eine erhöhte Tonlage für ihre Sprache. Beim *baby-talk* liegt diese zwischen 400 und 600 Hz. Außerdem werden überzogene Intonationsmuster eingesetzt, um dem Kind die Diskriminierung von unterschiedlichen Lautstrukturen zu erleichtern (vgl. Kap. 1.1). Auch längere Pausen zwischen einzelnen Sätzen sind zu beobachten.

Semantik und Syntax der Elternsprache sind geprägt durch häufige Wortwiederholungen, kurze Phrasen, Diminutive („*Na, wo ist das Vögelchen?*“) und eine vereinfachte Wortwahl (z.B. „*Schau, da ist der Wauwau*“). Inhaltlich beziehen sich die Äußerungen der Mütter zumeist auf konkrete Gegenstände und Sachverhalte in der Gegenwart (vgl. Kap. 1.2).

Ritterfeld (2000) weist darauf hin, dass der Einsatz des baby-talks häufig von körperlichen Formen der Kommunikation (zum Beispiel Schaukeln oder Streicheln) begleitet wird. Dadurch werden erste sprachliche Erfahrungen des Kindes in einem engen emotionalen Kontext vermittelt.

Erste dialogartige Strukturen können beim baby-talk dann auftreten, wenn Eltern beginnen, die Emotionen ihrer Kinder zu spiegeln oder Befindlichkeiten zu interpretieren. (El Mogharbel & Deutsch, 2007).

Etwa ab dem zweiten Lebensjahr verändern viele Eltern ihr Sprachverhalten gegenüber ihrem Kind. Der baby-talk weicht einer *stützenden Sprache* (*scaffolding*; Grimm & Weinert, 2002). Sie lässt sich häufig in Situationen beobachten in denen sich die Aufmerksamkeit von Mutter und Kind auf einen gemeinsamen Gegenstand richtet (*joint attention*). Dies ist zum Beispiel beim gemeinsamen Betrachten eines Bilderbuchs der Fall. Viele Mütter bieten ihren Kindern in solchen Situationen einfache, meist hoch ritualisierte Dialogstrukturen (Anrede – Frage – Benennung – Bestätigung) an.

Beispiel einer ritualisierten Dialogstruktur

Anrede: „*Oh, schau, was das ist!*“

Frage: *Was ist das nur?*

Benennung: Das ist ein Hühnchen.

Bestätigung: *Ja, das stimmt, das ist ein Hühnchen.*

(nach Grimm & Weinert, 2002, S. 547)

Mütter verwenden die stützende Sprache vor allem mit dem Ziel den Wortschatz ihrer Kinder zu erweitern. Die stützende Sprache wird von Eltern häufig dann eingesetzt, wenn sie vermuten, dass ihr Kind zu der Erkenntnis gelangt ist, dass alles einen Namen hat und somit jeder Gegenstand benannt werden kann (Ritterfeld, 2000). Kinder erwerben durch stützende Sprache nicht nur linguistische Kompetenzen. Sie erfahren auch, welche Sprachmuster in welchen Situationen angemessen sind und wie sie ihre Gedanken verbalisieren müssen, damit ihre Eltern sie verstehen (Nußbeck, 2007).

Mit zunehmender Komplexität der kindlichen Rede ändert sich auch die Erwartungshaltung der Eltern an die Struktur der kindlichen Aussagen. Der Fokus des Eltern-Kind-Dialogs bezieht sich immer weniger auf das Abfragen von reinem Wortwissen. Eltern gehen von nun an verstärkt auf die morpho-syntaktischen Kenntnisse des Kindes ein. Hierfür nutzen sie implizite Sprachlehrstrategien, die in der Literatur als *lehrende Sprache* oder *motherese* beschrieben werden. Motherese bezeichnet die Strategien, die Eltern anwenden, um sich an das Sprachverständnis des Kindes anzupassen (Zollinger, 1997). Sagt das Kind einen noch nicht korrekten Satz, so wiederholt die Bezugsperson diesen Satz mit identischem Inhalt, aber grammatikalisch richtig, ohne das Kind offensichtlich zu verbessern (Brandl, 2010). Ein Beispiel für diese Modellierungstechnik gibt Grimm (1999). Das Kind sagt: „Hatter put tetangen“ und die Mutter antwortet: „Ja, das ist kaputt gegangen“. Die Mutter bietet dem Kind eine implizite Korrektur an, in dem sie den Satz grammatisch korrekt wiederholt.

In Anlehnung an Grimm und Weinert (2002) lassen sich folgende Merkmale der lehrenden Sprache feststellen:

- Anpassung der durchschnittlichen Äußerungslänge und der Anzahl von Nominalphrasen pro Äußerung an den Entwicklungsstand des Kindes.
- häufiges Stellen von offenen W-Fragen um Sprachanlässe für das Kind zu schaffen.
- Modellierung und Erweiterung der kindlichen Aussagen.
- Wiederholung der kindlichen Äußerung.

1.4.1 Motherese und Scaffolding als Vorbild für Dialogisches Lesen

Im Allgemeinen gilt der Einsatz der oben beschriebenen Modellierungstechniken aus der motherese-Forschung als entwicklungsförderlich (Jungmann, 2007). Zur Prävention von Sprachentwicklungsstörungen erhalten Eltern häufig die Empfehlung scaffolding und Modellierungstechniken bei Unterhaltungen mit ihren Kindern einzusetzen (Warren, Yoder, Gazdag, Kim & Jones, 1993). Allerdings mangelt es noch an gezielten Evaluationsstudien, die das langfristige sprachförderliche Potential dieser Techniken bestätigen (Jungmann, 2007). Einige Studien fanden unmittelbare Effekte bei Kindern, deren Eltern im Rahmen von Elterntrainings sprachförderliche Modellierungstechniken erlernten (Grimm, 2003). Einen ersten experimentellen Nachweis für das sprachförderliche Potential von Modellierungstechniken erbrachten Whitehurst und Kollegen (1988). Ausgehend von den

Erkenntnissen der motherese-Forschung entwickelten sie die Methode des Dialogischen Lesens (DL). Ziel der Methode ist es, die Qualität der elterlichen Inputsprache zu verbessern. Als Fördersetting wählten die Autoren Gesprächssituationen, in denen die Eltern ihren Kindern vorlasen bzw. gemeinsam ein Buch betrachteten. Die Eltern erhielten die Anweisung dem Kind Fragen zu den Inhalten des Buches zu stellen und auf Antworten mit stützender Sprache und Modellierungstechniken zu reagieren. Nachfragen zum Inhalt geben dem Kind die Möglichkeit sich frei zu äußern. Sie regen das Kind dazu an die Geschichte in eigenen Worten nachzuerzählen. Im Gespräch mit dem Kind setzen die Vorleser die beschriebenen Modellierungstechniken ein und wiederholen das Gesagte. Die Autoren konnten Fortschritte in der Sprachentwicklung der DL-Kinder feststellen. Bislang geben hauptsächlich anglo-amerikanische Forschungsarbeiten Hinweise darauf, dass Dialogisches Lesen wirksam zur Förderung der Sprachentwicklung eingesetzt werden kann. Im deutschsprachigen Raum fehlt es derzeit an entsprechenden Befunden.

Die folgenden Kapitel geben zunächst ein Überblick über den nationalen Forschungsstand zur Evaluation von Sprachfördermethoden. Mit Dialogischem Lesen und Trainings zur Förderung der Phonologischen Bewusstheit sind im Anschluss zwei evidenzbasierte Fördermaßnahmen für Vorschulkinder hervorzuheben. Während die Wirksamkeit von Trainings der Phonologischen Bewusstheit in Deutschland und international vielfach bestätigt werden konnte, liegen für das Dialogischen Lesen bislang zumeist Studien aus dem englischsprachigen Raum vor. Für deutschsprachige Kindern fehlen entsprechende Untersuchungen. Einen Überblick über den aktuellen Stand der Forschung zur Evaluation des Dialogischen Lesens geben Kapitel 3.2.1.2 und 3.2.1.3.

Zuvor gibt Kapitel 2 einen Überblick über Varianten von Sprachstörungen. Dabei gilt es herauszuarbeiten, welche Kinder gemeint sind, wenn im Rahmen der eignen Untersuchung von Risikokindern die Rede ist.

2 Störungen der Sprachentwicklung

Kaum ein Bereich der kognitiven Entwicklung ist so häufig von Störungen betroffen wie die Entwicklung der Sprache (Grimm, 1999). Sprachliche Auffälligkeiten können individuell sehr unterschiedlich auftreten.

Studien, die sich mit der Evaluation von Sprachförderung befassen, prüfen die Wirksamkeit einer Maßnahme häufig an Stichproben aus sogenannten *Risikokindern*. Dabei handelt es sich um Kinder, die zu Beginn der Untersuchung Auffälligkeiten in einem Sprachentwicklungstest zeigten. Allerdings sind Sprachstörungen häufig sehr individuell ausgeprägt und können unterschiedliche Ursachen haben. Deshalb soll im Vorfeld der eigenen Studien erörtert werden, welche Art von Sprachauffälligkeiten gemeint sind, wenn im Folgenden von Risikokindern die Rede ist.

Kany und Schöler (2010) schlagen grundsätzlich eine Unterscheidung von fünf Arten der Sprachauffälligkeit vor:

1. Sprachstörung aufgrund kognitiver Beeinträchtigungen
2. Stimm- und Redeflussstörungen (Sprechstörungen)
3. Kommunikationsstörungen
4. Spracherwerbsstörungen (expressive und rezeptive Formen)
5. Unzureichende Sprachkenntnis (fehlende Modelle/ Sprachenregung, Input)

Kinder, bei denen eine kognitive Beeinträchtigung festgestellt wurde und die somit unter die erste Kategorie fielen, wurden von der Untersuchung ausgeschlossen. Gleiches galt für Kinder bei denen der Redefluss durch Stottern oder Poltern gestört war (Kategorie 2). Kinder die eine Kommunikationsstörung aufwiesen (Kategorie 3) waren in der Ausgangsstichprobe nicht vertreten. Wenn im Rahmen der eigenen Untersuchung von Risikokindern die Rede ist, sind Kindern gemeint, bei denen sich das Risiko eine Spracherwerbsstörung zu entwickeln in einem standardisierten Sprachscreening feststellen ließ (Kategorie 4). Auch Migrantenkinder, bei denen sich ein Sprachentwicklungsrisiko aufgrund unzureichender Sprachkenntnis nicht ausschließen ließ, zählten zur Risikogruppe (Kategorie 5). Beide studienrelevante Gruppen werden im Folgenden genauer beschrieben.

2.1.1 Kinder mit Spracherwerbsstörungen

Nach der Internationalen Klassifikation psychischer Störungen der Weltgesundheitsorganisation (ICD-10) werden Sprachentwicklungsstörungen als „Umschriebene Entwicklungsstörungen des Sprechens und der Sprache“ bezeichnet. Dabei handelt es sich um Störungen, die nicht direkt auf neurologische Störungen oder Veränderungen des Sprachablaufs, sensorische Beeinträchtigungen, Intelligenzminderung oder Umweltfaktoren zurückzuführen sind (Dilling & Schulte-Markwort, 2014). Um der genannten Definition gerecht zu werden, hat sich der Begriff *Spezifische Sprachentwicklungsstörung* (SSES) durchgesetzt. Betroffene Kinder fallen zunächst im sprachlichen Bereich durch Artikulationsstörungen, Ausdrucksstörungen (expressive Störungen) oder Störungen im Verstehensprozess (rezeptive Störungen) auf. Es wird zwischen *phonetischen Störungen* und *phonologischen Störungen* unterschieden (Grimm, 1999; Nußbeck, 2007).

Von einer phonetischen Störung kann gesprochen werden, wenn es dem Kind nicht oder nur mit Verzögerung gelingt, sämtliche Laute seiner Muttersprache zu bilden (Kany & Schöler, 2010). Auffälligkeiten können aber auch dann entstehen, wenn dem Kind das Wissen über die sprachsystematische Funktion von Lauten fehlt oder nur teilweise vorhanden ist. In diesem Fall ist von einer phonologischen Störung zu sprechen.

Kany und Schöler (2010) weisen drauf hin, dass sich die diagnostische Unterscheidung zwischen phonetischen und phonologischen Störungen zumeist sehr schwierig gestaltet. Ähnliche Aussprachefehler sind im einen Fall auf fehlende Voraussetzungen und Fertigkeiten zur Lautbildung (phonetische Störung), im anderen Fall auf fehlendes Wissen über die korrekte Verwendung von Lauten (phonologische Störung) zurückzuführen.

In Bezug auf die Diagnostik von Sprachentwicklungsstörungen legen neue Erkenntnisse entgegen der ICD 10-Definition nahe, dass Störungen der Sprachentwicklung oft mit Einschränkungen in Teilleistungen der kognitiven Entwicklung einhergehen (Grimm, 1999; Kany & Schöler, 2010). SSES-Kinder zeigen zum Beispiel Schwächen im verbalen Kurzzeitgedächtnis (phonologische Schleife) und eine daraus resultierende verlangsamte Artikulationsgeschwindigkeit (Grimm, 1999). Kushnir und Blake (1996) konnten in Testaufgaben, die die Nachsprechleistung von Kindern mit Auffälligkeiten in der Sprachentwicklung erfassten, hohe Korrelationen mit produktiven sowie rezeptiven Sprachentwicklungsmaßen feststellen ($r = .68$ - $r = .74$). Auf Grund ihres hohen prädiktiven Wertes werden Aufgaben zur Erfassung der Leistung des phonologischen Arbeitsgedächtnisses und der auditiven Merkspanne in vielen Sprachscreenings zur Diagnose von Risikokindern eingesetzt (Ennemoser, Pepouna & Hartung, 2012; Grimm, 2003). Diese

Aufgabenformate kommen auch in der eigenen Untersuchung zum Einsatz. Wenn also im Zuge dieser Arbeit von „Risikokindern“ die Rede ist, dann sind damit Kinder gemeint, die im Sprachscreening von Grimm und Kollegen (2003) ein unterdurchschnittliches Ergebnis erzielten ($T \leq 40$). Das Screening erfasst die Leistung im Nachsprechen von Pseudowörtern und im Nachsprechen von semantisch korrekten und inkorrekten Sätzen. Ein ausführlicher Überblick über die eingesetzten Instrumente zur Erfassung des Sprachentwicklungsstands in der eigenen Untersuchung findet sich in Kapitel 6.1 dieser Arbeit.

2.1.2 Kinder mit unzureichenden Sprachkenntnissen

Unter diese Kategorie fallen deutschsprachige Kinder, die unter anregungsarmen Entwicklungsbedingungen aufwachsen, aber auch Kinder mit Migrationshintergrund, die Deutsch als Fremdsprache erlernen. Gerade in der Arbeit mit Migrantenkindern ist es eine wichtige diagnostische Aufgabe, Spracherwerbsstörungen aufgrund mangelnder Entwicklungsvoraussetzungen von Auffälligkeiten als Resultat von unzureichender Spracherfahrung zu unterscheiden.

Mangelnde Sprachkenntnisse in der Zweitsprache führen häufig dazu, dass bei Kindern mit Migrationshintergrund in deutschen Sprachentwicklungstests fälschlicherweise eine Sprachentwicklungsstörung diagnostiziert wird. Ennemoser, Pepouna und Hartung (2012) zeigten in einer Untersuchung, dass russische und türkische Muttersprachler deutlich bessere Leistungen beim Nachsprechen von Pseudowörtern erzielten, die auf den phonotaktischen Strukturen ihrer Muttersprache beruhten. In Pseudowortaufgaben auf Grundlage der deutschen Phonologie, die in vielen Sprachentwicklungstests Verwendung finden, sind schlechtere Ergebnisse von Kindern mit Migrationshintergrund eher Ausdruck von mangelnder Spracherfahrung und weniger Hinweis auf eine Sprachentwicklungsstörung.

In der vorliegenden Studie wurden alle Kinder, die ein unterdurchschnittliches Ergebnis im Screening erzielten als Risikokinder klassifiziert – auch jene mit Migrationshintergrund. Da grundsätzlich bei allen Risikokindern die Notwendigkeit für eine sprachliche Förderung besteht, wurde bei der Zusammenstellung der Stichprobe nicht zwischen entwicklungsbedingten und spracherfahrungsbedingten Ursachen des Sprachdefizits unterschieden.

Es ist deutlich geworden, dass die Störung der kindlichen Sprachentwicklung sehr unterschiedliche und individuelle Gründe haben kann. Ganz gleich ob dem Sprachdefizit mangelnde Spracherfahrung oder eine Sprachentwicklungsstörung zugrunde liegen, stellt sich

in der pädagogischen Praxis die Frage nach wirksamen Fördermethoden. Dialogisches Lesen stellt hierfür eine vielversprechende Möglichkeit dar (Whitehurst et al., 1988). Während weltweit zahlreiche Studien das sprachförderliche Potential von Dialogischem Lesen bestätigen (Mol & Bus, 2011), liegen für deutsche Kinder noch keine entsprechenden Evaluationsstudien vor. Ebenfalls ungeklärt ist die Frage, ob sich der Kindergarten als institutionelles Setting für den Einsatz von Dialogischem Lesen eignet. Folglich betrachtet Kapitel 3 Sprachfördermethoden, die aufgrund positiver Befunde in Evaluationsstudien als *evidenzbasiert* gelten können (Slavin, 2008; Justice & Pullen, 2003). Dazu zählen Trainings der Phonologischen Bewusstheit und Dialogisches Lesens. Bevor beide Fördermaßnahmen näher beschrieben werden und eine Übersicht über den Stand der internationalen Forschung erfolgt, fasst das folgende Kapitel zu Beginn Ergebnisse aus deutschen Evaluationsstudien zusammen, die sich in den letzten Jahren mit der Wirksamkeitsprüfung von Sprachfördermaßnahmen befassten.

3 Förderung von Sprachkompetenzen

Die Ergebnisse der internationalen Schulvergleichsstudie PISA lösten in Deutschland um die Jahrtausendwende den vielzitierten „PISA-Schock“ aus. Die Erkenntnis, dass der sozioökonomische Status eines Kindes in Deutschland ein bedeutsamer Prädiktor für den schulischen Werdegang ist, führte zu einer bundesweiten Bildungsdebatte (Baumert et al., 2001; Baumert, Stanat & Watermann, 2006). Trotz positiver Entwicklungstendenzen in den schulischen Leistungsbereichen zeigten die neuerlichen Ergebnisse von PISA 2009, dass Bildungserfolg in Deutschland nach wie vor eng an die soziale Herkunft der Kinder und Jugendlichen gekoppelt ist (Klieme, 2010).

Zwei zentrale Themen dominieren bis heute die Bildungsdebatte: a) *Welche Kompetenzen* müssen zu b) *welchem Entwicklungszeitpunkt* gefördert werden, um die Bindung von Bildungschancen an soziale Herkunft zu verringern.

Zahlreiche Längsschnittstudien der letzten Dekaden verfolgten das Ziel bedeutsame Vorläuferkompetenzen im Bereich des Schriftspracherwerbs zu klassifizieren. Neben kognitiven Faktoren (Arbeitsgedächtnis- und Intelligenzleistung) und metasprachlichen Kompetenzen (Phonologische Bewusstheit) werden stets allgemeine Sprachkompetenzen als bedeutsamer Prädiktor für die Lese-Rechtschreibentwicklung genannt (Marx, Weber & Schneider, 2005a; Schneider, 2004).

Zum Beispiel zeigten die Ergebnisse einer Untersuchung von Ennemoser, Marx, Weber und Schneider (2012), dass die sprachlichen Kompetenzen von Vorschulkindern⁴ bis in die 4. Klasse hinein die Leistungen im Textverständnis und in Rechtschreibung vorhersagen. Die frühzeitige Sprachförderung ist also mit Blick auf die Entwicklung späterer Lese- und Rechtschreibkompetenz eine zentrale Aufgabe der vorschulischen Bildung (Tracy, 2008).

Im Bildungssektor wuchs in den vergangenen Jahren das Angebot an Programmen zur Förderung der sprachlichen Kompetenzen von Kindern mit und ohne Migrationshintergrund. Einen ausführlichen Überblick geben Jampert et. al (2007). Die Autoren stellten fest, dass sich einzelne Förderkonzepte in Bezug auf ihre inhaltliche Ausrichtung und im methodischen Vorgehen zuweilen stark voneinander unterscheiden. Außerdem mangelt es an Evaluationsstudien, die die Wirksamkeit der diversen Förderprogramme belegen. Dies veranlasste in den vergangenen Jahren Kommunen und Länder, aber auch private und kirchliche Stiftungen dazu Forschungsprojekte zu fördern, die sich mit der Entwicklung und Anwendung von Sprachförderprogrammen beschäftigten (Kany & Schöler, 2010).

⁴ Die Autoren sprechen im Rahmen ihrer Untersuchung von linguistischen Kompetenzen.

Auch die Bundesregierung wies in ihrem zweiten Integrationsbericht auf die besondere Bedeutung von empirischer Bildungsforschung im Bereich der Sprachförderung hin. Neben dem Ausbau von Betreuungsangeboten und der Ausbildungsförderung frühpädagogischer Fachkräfte sollten insbesondere Forschungsvorhaben gefördert werden, die fundiertes Wissen über eine angemessene Förderung der Sprache von Kindern mit und ohne Migrationshintergrund bereitstellen (Engels, Köller & Koopman, 2011).

Trotz vielschichtiger Anstrengungen und der gezielten Förderung evaluativer Forschungsvorhaben, steht für die Förderung von Kindern mit geringen Sprachkenntnissen oder mit Sprachentwicklungsstörungen bisher nur eine sehr kleine Anzahl von Sprachförderprogrammen zur Verfügung, die als evidenzbasiert gelten können (Stanat, 2006). Gleichzeitig steigt die Nachfrage nach erprobten Sprachförderkonzepten, deren Wirksamkeit auch wissenschaftlich belegt ist.

Einige Bundesländer sehen den Erfolg von Sprachfördermaßnahmen schon darin bestätigt, dass in den vergangenen Jahren weniger Zurückstellungen auf Grund von sprachlichen Defiziten verzeichnet wurden. In Hessen und Bayern werden zum Beispiel die Einschätzungen von pädagogischen Fachkräften zur Beurteilung der sprachlichen Weiterentwicklung der Kinder nach einer Fördermaßnahme herangezogen (Lisker, 2011). Die empirische Prüfung der Wirksamkeit von unterschiedlichen Sprachförderprogrammen ist eine Aufgabe, der sich in den letzten Jahren einige großangelegte Evaluationsstudien im ganzen Bundesgebiet widmen. Sie sollen im Folgenden beschrieben werden.

3.1 Evaluation von Sprachfördermaßnahmen im deutschen Sprachraum

Das Projekt „Sag mal was“, das 2002 von der Baden-Württemberg Stiftung und dem Landesinstitut für Schulentwicklung initiiert wurde, zählt zu den bundesweit größten und finanzstärksten Vorhaben mit dem Schwerpunkt Evaluation von Sprachfördermaßnahmen.⁵ Ziel des Projekts war es die sprachliche Bildung von Vorschulkindern zu stärken. Hierfür nahmen bis zum Ende des Schuljahres 2009/2010 90.000 Kinder in über 2000 Kindergärten an unterschiedlichen Sprachförderprogrammen teil (Weber & Bacher, 2011). Wissenschaftlich begleitet wurde das Projekt von den Pädagogischen Hochschulen Heidelberg und Weingarten.

An der PH Heidelberg prüften Roos, Polotzek und Schöler (2010) im Rahmen der EVAS-Studie (Evaluation von Sprachförderung bei Vorschulkindern) die Wirksamkeit von drei Sprachfördermaßnahmen bei Kindern mit Förderbedarf. Die Untersuchung folgte einem Prä-Post-Design. Durch den Einsatz von zwei Screeningverfahren (*SISMIK*, Ulich & Mayr, 2004; *HASE*, Brunner & Schöler, 2002) konnte bei 291 Kindern der Gesamtstichprobe sprachlicher Förderbedarf festgestellt werden. 199 Kinder waren nicht förderbedürftig. Alle Kinder befanden sich zum Zeitpunkt der Erhebung im letzten Kindergartenbesuchsjahr und waren im Schnitt 5;3 Jahre alt.

Nach Abschluss des Screenings wurde über alle Messzeitpunkte hinweg die Sprachkompetenz aller Kinder (mit und ohne Förderbedarf) durch vier Untertests des *Heidelberger Sprachentwicklungstests* (H-S-E-T, Grimm & Schöler 1991) erfasst⁶. Außerdem wurde die Wortschatzleistung erhoben. (Ravensburger Prüfmateral). Kontrolliert wurden die Intelligenzleistung und der sozioökonomische Status.

Nach Abschluss des Vortests wurden die Förderkinder einer von drei Trainingsbedingungen ($n = 209$) oder einer Kontrollgruppe ($n = 81$) zugewiesen. Die Kinder der Experimentalgruppen nahmen für sechs Monate (4 bis 5 Förderstunden pro Woche) entweder am *Kon-Lab Programm* (Penner, 2003), an *Deutsch für den Schulstart* (Kaltenbacher, 2008) oder an einer sprachlichen Frühförderung nach Tracy (2003) teil. Ausgebildete Sprachförderkräfte führten die Programme durch. Die Kinder wurden in Kleingruppen mit maximal zehn Teilnehmern gefördert. Die Dauer der Intervention betrug mindestens sechs Monate (4 bis 5 Förderstunden pro Woche). Parallel zu den Experimentalgruppen nahm die

⁵ Für das „Sag mal was“-Projekt stellte die Baden-Württemberg Stiftung ein Projektvolumen von 39 Millionen Euro bereit (Weber & Bacher, 2011).

⁶ Eingesetzt wurden die Untertests *Imitation* und *Verstehen grammatischer Strukturformen* (IS, VS), *Plural-Singular-Bildung* (PS) und *Wortfindung* (WF) (Hoffmann, Polotzek, Roos & Schöler, 2008).

Kontrollgruppe ($n = 81$) am konventionellen Sprachförderangebot der Kindergärten ohne inhaltliche Schwerpunkte teil.

In ihren Analysen betrachten die Autoren die Sprachentwicklung von insgesamt drei Subgruppen der Stichprobe: (1) spezifisch geförderte Kinder mit Sprachförderbedarf, (2) unspezifisch geförderte Kinder mit Sprachförderbedarf und (3) Kinder ohne Sprachförderbedarf.

Wider Erwarten verbesserte sich die Sprachkompetenz von Förderkindern, die an einer spezifischen Sprachförderung teilnahmen, im Vergleich zu unspezifisch geförderten Kindern nicht substanziell. Das heißt, durch eine spezifische sprachliche Förderung, die einem bestimmten Programm folgte, konnten im Vergleich zum konventionellen Förderangebot in den Kitas keine signifikanten Sprachentwicklungseffekte festgestellt werden. Auch konnten geförderte Kinder mit Sprachförderbedarf ihren Leistungsrückstand gegenüber normalentwickelten Kindern nicht bedeutsam verringern.

In einer zweiten Begleitstudie der PH Weingarten untersuchten Gasteiger-Klicpera, Knapp und Kucharz (2011) Bedingungsfaktoren für eine wirksame Sprachförderung in Kindertagesstätten mit dem Ziel, erfolgreiche Vorgehensweisen (*best practice*) für die Sprachförderarbeit zu klassifizieren. In ihrer Längsschnittstudie erhoben die Autoren die Sprachleistungen von 1150 Kindern in zwei Kohorten (606 erste Kohorte; 544 zweite Kohorte). Für die Erhebung der Sprachleistung wurde das *Sprachscreening im Vorschulalter* (SSV; Grimm, 2003) eingesetzt. Außerdem wurden die kognitiven Fähigkeiten der Kinder kontrolliert (CPM; Bulheller & Häcker, 2002). Ca. drei Viertel der Kinder ($n = 867$) nahmen ein Jahr lang an einem Sprachförderangebot ihres Kindergartens teil. Die inhaltliche Gestaltung oblag dem Förderpersonal vor Ort. Informationen zur Konzeption der Sprachförderung, zur Zusammensetzung der Fördergruppe und zum pädagogischen Konzept der Einrichtungen wurden über Fragebögen erfasst, die das Förderpersonal ausfüllte.

Als zentrales Ergebnis stellten die Autoren für die Gesamtstichprobe (Kohorte 1 + Kohorte 2) fest, dass sich Kinder, die eine Förderung erhalten hatten, im Vergleich zu ungefördernten Kindern über die Zeit im Subtest *Satzgedächtnis* verbessern konnten. In weiteren differentiellen Analysen, bei denen ausschließlich sprachförderbedürftige Kinder berücksichtigt wurden, zeigten sich für die erste und zweite Kohorte keine Unterschiede in der Leistungsentwicklung geförderter versus ungeförderter Risikokinder.

Die Ergebnisse der Fragebogendaten des Förderpersonals zeigten starke Unterschiede in Bezug auf die Durchführungsqualität. Auf die Frage, welches Konzept bzw. Programm

mehrheitlich in den Sitzungen umgesetzt wurde, gaben rund 69% der Förderkräfte an kein bestimmtes Programm, sondern Mischformen eingesetzt zu haben. Die Durchführungsintensität variierte abhängig vom Förderer stark zwischen einer und fünf Fördersitzungen pro Woche. Die Fachkräfte unterschieden sich außerdem im Grad ihrer Erfahrung. 23% führten Sprachfördermaßnahmen erst seit weniger als zwei Jahren durch, während 15% auf sechs oder mehr Jahre Erfahrung zurück blicken konnten. Auf Grundlage dieser Ergebnisse ist es nur schwer möglich die Wirksamkeit einzelner Förderprogramme vergleichend zu beurteilen. Die Untersuchung zur Gestaltung der Sprachförderung kam zu dem Ergebnis, dass über 90% der Förderinhalte aus Gesprächen, Wortschatzarbeit, Versen und Liedern sowie Übungen zur phonologischen Bewusstheit bestanden. Vorlesesituationen kamen in 10% der analysierten Videosequenzen vor. Weniger als 5% der Fördersequenzen vermittelten grammatikalische Inhalte. Allerdings geben auch diese Ergebnisse keinen Aufschluss darüber, welche Methoden im Sinne des „best practice“-Gedankens als besonders sprachförderlich gelten können.

Mit der Evaluation des Förderprogramms *Deutsch für den Schulstart* befasste sich auch die DACHS-Studie (Deutsch-Sprachförderung vor der Schule; Sachse, Budde, Rinker & Groth, 2012). Im Rahmen der hessischen Vorlaufkurse⁷ wurde das Förderprogramm ein Jahr lang eingesetzt. 79 Kinder nahmen an der Förderung durch spezifisch geschulte Lehrer teil. Die Kontrollgruppe ($n = 46$) folgte den konventionellen Richtlinien der Vorlaufkurse (Hessisches Kultusministerium, 2001). Rezeptive, expressive und morphologische Sprachleistungen wurden mit drei Untertests des H-S-E-T (Grimm & Schöler, 1991) erhoben. Nach Abschluss der Vorlaufkurse folgte ein Nachtest und am Ende der 1. Klasse eine Follow Up-Erhebung. Wie schon in der EVAS-Studie erwies sich die Trainingsgruppe im Bereich der Sprachkompetenz der unspezifisch geförderten Kontrollgruppe als nicht überlegen. Das Ausbleiben von Effekten begründen die Autoren damit, dass im Studiendesign keine gänzlich ungeforderte Kontrollgruppe vorgesehen war. Außerdem sollten zukünftig mögliche Veränderungen in bestimmten Sprachentwicklungsbereichen (z.B. Semantik) mit Hilfe einer differenzierteren Testauswahl betrachtet werden.

⁷ Seit 2002 werden in Hessen für sprachauffällige Kinder im letzten Kindergartenjahr sogenannte „Vorlaufkurse“ angeboten. Hierbei erhalten Vorschulkinder seitens der Grundschulen in Kleingruppen über 12 Monate 10 bis 15 Wochenstunden additive Sprachförderung. Da die inhaltliche Gestaltung dem Lehrpersonal überlassen ist, variieren die Auswahl der Förderangebote und die Darbietungsqualität zwischen den Vorlaufkursen zum Teil stark (Fiebig & Probst, 2011).

In Brandenburg untersuchten Wolf, Stanat und Wendt (2011) im Rahmen der EkoS-Studie (Evaluation der kompensatorischen Sprachförderung im Jahr vor der Einschulung) die Wirksamkeit eines inhaltspezifischen Sprachförderprogramms (Häuser & Jülisch, 2006). Am Vortest nahm eine Ausgangsstichprobe von 904 Kindern teil. Die sprachlichen Kompetenzen wurden mit dem KISTE (*Kindersprachtest für das Vorschulalter*; Häuser, Kasielke & Scheidereiter, 1994) erfasst. Zusätzlich wurden bei je der Hälfte der teilnehmenden Kinder das BISC (*Bielefelder Screening zur Früherkennung von Lese- und Rechtschreibschwierigkeiten*; Jansen, Mannhaupt, Marx & Skowronek, 2002) oder das HASE (Brunner & Schöler, 2002) eingesetzt. Für die weitere Untersuchung wurden nur Fälle berücksichtigt, für die Ergebnisse aus KISTE und BISC vorlagen ($n = 417$).⁸

Kinder, für die auf Grund der Testergebnisse Sprachförderbedarf feststand, wurden entweder der Experimentalgruppe ($n = 99$) oder der Kontrollgruppe ($n = 170$) zugewiesen. Eine zweite Kontrollgruppe umfasste alle Kinder ohne Sprachförderbedarf ($n = 148$). Alle Kinder befanden sich zum Zeitpunkt der Studie im letzten Kindergartenjahr (Durchschnittsalter 6;3). Die Förderung führten ausgebildete Erzieher in der Experimentalgruppe durch. An einer Fördereinheit (28 Minuten) nahmen im Schnitt 11 Kinder teil. Die Einheiten wurden in der Regel fünfmal pro Woche über einen Zeitraum von 15 Wochen durchgeführt. Im Nachtest war die geförderte Gruppe der Kinder mit Sprachförderbedarf der ungeförderten Gruppe mit Sprachförderbedarf nur im KISTE-Untertest *Satzbildungsfähigkeit* signifikant überlegen. In allen weiteren Bereichen ergaben sich keine signifikanten Leistungsunterschiede zwischen den beiden Gruppen. Außerdem konnten die Kinder der Trainingsgruppe im Nachtest nicht nennenswert zur Gruppe der Kinder ohne Sprachförderbedarf aufschließen. Es gelang also nicht sprachliche Defizite durch ein spezifisches Förderprogramm zu kompensieren.

Insgesamt ist festzuhalten, dass alle Studien zur Wirksamkeit von Sprachförderprogrammen weitgehend ernüchternde Befunde berichten. Als Gründe für die geringe Effektivität der Sprachförderprogramme nennen die Autoren häufig den späten Zeitpunkt der Intervention (Roos et al., 2010), die zu kurze Interventionsdauer (Sachse et al., 2012) oder die fehlende Kontrolle von Moderatorvariablen wie Intelligenz, Herkunftssprache oder sozioökonomischer Status (Esser, 2011). Auch wurde in keiner der genannten Studien die Validität der Treatments, also die Genauigkeit, mit der ausgebildete Fachkräfte eine Fördermaßnahme umsetzen, kontrolliert. Starke Unterschiede in der Durchführungsqualität könnte

⁸Alle HASE-Kinder wurden von der Untersuchung ausgeschlossen. Die Autoren begründen den Ausschluss damit, dass die Ergebnisse aus dem HASE nur für den ersten MZP vorlagen und somit keine längsschnittlichen Analysen durchgeführt werden konnten.

möglicherweise das Ausbleiben von Fördereffekten erklären. Die Bedeutung der Durchführungsqualität für die Wirksamkeit einer Fördermaßnahme soll in Kapitel 4 ausführlich Thema sein.

Sachse et al. (2012) stellten im Zuge der DACHS-Studie fest, dass Lehrer in Vorlaufkursen das durchzuführende Sprachförderprogramm „Deutsch für den Schulstart“ (Kaltenbacher, 2008) sehr unterschiedlich umsetzten. Der Umfang einer Fördersitzung lag zwischen 45 min. bis 120 min ($M = 80.5$, $SD = 25.1$). Ebenso variierte die Anzahl der erfolgten Kurseinheiten über das Schuljahr stark (46 bis 128 Einheiten; $M = 81.5$; $SD = 27.1$). Die Autoren sehen das Ausbleiben von Fördereffekten unter anderem in deutlichen Durchführungsunterschieden begründet.

Im Rahmen der EkoS-Studie wurden die fördernden Erzieher nach der Durchführung des Sprachförderprogramms von Häuser und Jülisch (2006) um eine Selbsteinschätzung zur Durchführungsqualität gebeten. Nach Angaben der Erzieher gelang es ihnen zu 86% die vorgegebenen Übungen im Manual umzusetzen. Um die Durchführungsqualität noch genauer bestimmen zu können, empfiehlt es sich allerdings auf objektivere Maße zurückzugreifen. Eine Möglichkeit ist die Beurteilung der Durchführungsqualität durch geschulte Supervisoren.

Ein weiteres Problem könnte darin bestehen, dass etwaige Veränderungen in der Sprachentwicklung mit den eingesetzten Sprachentwicklungstests nicht valide genug erfasst wurden. Zum Beispiel könnten Kinder durch eine Sprachförderung neue Wörter gelernt haben, die aber in einem allgemeinen Wortschatztests nicht abgefragt wurden. Aus diesem Grund ist das Ausbleiben von Fördereffekten nicht immer ein Beleg für die Unwirksamkeit einer Maßnahme (Esser, 2011).

Ein weiterer Einflussfaktor, der in Wirksamkeitsprüfungen bislang nur selten kontrolliert wird, ist die Durchführungsqualität der Intervention (vgl. Kap. 3.1.2). Auch bei hoch standardisierten Untersuchungsdesigns kann die ungenaue Umsetzung einer Maßnahme ein Grund für das Auftreten von Nulleffekten sein.

Zusammenfassend lässt sich feststellen, dass trotz zahlreicher Forschungsbemühungen in Deutschland weiterhin ein großer Bedarf an Untersuchungen besteht, die sich mit der Wirksamkeitsevaluation von Sprachförderprogrammen beschäftigen. Kuger, Sechting und Anders (2012) weisen darauf hin, dass internationale Studien bereits vielversprechende Ergebnisse in Bezug auf die Wirksamkeit spezifischer Sprachfördermaßnahmen im Kindergarten hervorgebracht haben. Auch wenn internationale Befunde natürlich nur

eingeschränkt auf deutsche Stichproben übertragbar sind, lassen sich dennoch wertvolle Hinweise zur Wirksamkeit dieser Fördermethoden finden.

Eine vielversprechende Sprachfördermethode ist – wie in den vorherigen Kapitel bereits erwähnt – in das Dialogische Lesen (DL). Vor allem in anglo-amerikanischen Studien konnte das sprachförderliche Potential von Dialogischem Lesen bei unterschiedlichen Stichproben und in unterschiedlichen Fördersettings nachgewiesen werden (National Early Literacy Panel, 2008). In Deutschland stehen Untersuchungen bislang noch aus.

Ebenfalls als evidenzbasiert gelten Trainings der Phonologischen Bewusstheit (PB). Die Phonologische Bewusstheit bildet eine Komponente der *metasprachlichen Bewusstheit*. Metasprache kann als allgemeine Fähigkeit verstanden werden, über Sprache zu reflektieren (Metz, Fröhlich, Petermann, 2010). Phonologische Bewusstheit bezieht sich auf die Einsicht in lautliche Strukturen einer Sprache und umfasst Fähigkeiten, wie Reimen, Wörter in Silben zerlegen und die Laute eines Wortes zu diskriminieren. Fähigkeiten auf der phonologisch-phonetischen Ebenen sind demnach wichtige Voraussetzung für den Erwerb der Phonologischen Bewusstheit.

In den folgenden Abschnitten des Kapitels werden das Dialogische Lesen und die Förderung Phonologischen Bewusstheit exemplarisch für evidenzbasierte Sprachförderansätze im Kindergarten vorgestellt. Außerdem sollen Ergebnisse aus Studien berichtet werden, die Aufschluss über die Wirksamkeit und das Effektspektrum beider Maßnahmen geben.

3.2 Evidenzbasierte Methoden der Sprachförderung in Deutschland und im internationalen Raum

3.2.1 Dialogisches Lesen

Dialogisches Lesen (DL) ist eine Methode des Vorlesens, bei der ein Erwachsener verschiedene Frage- und Feedbacktechniken einsetzt, um Kinder zu sprachlicher Aktivität anzuregen. In Kapitel 1.4 wurde bereits darauf hingewiesen, dass die Theorie des Dialogischen Lesens auf den Erkenntnissen der motherese-Forschung beruht. Danach gilt es als sprachförderlich, wenn Eltern die sprachlichen Äußerungen ihres Kindes unterstützen, indem sie bestimmte Modellierungstechniken einsetzen (z.B. Erweiterungen und Wiederholungen) und häufig offene Fragen zu den Inhalten eines Buches stellen. Dazu schreibt Huebner (2000):

„Studies of social interaction between parents and young children have identified many ways in which everyday conversation supports the child's task of language learning. Among the most common examples are parents' use of expansions, repetitions, extensions, responses, and questions that follow the child's interest" (S. 513).

Diese Kommunikationselemente, die Eltern meist intuitiv verwenden, wenn sie mit ihren Kindern sprechen, bilden die Grundprinzipien des Dialogischen Lesens.

Das Dialogische Lesen folgt keinem programmatischen Aufbau. Auch sollen nicht Sprachregeln, wie z.B. die Singular-Plural-Bildung, explizit vermittelt werden. Vielmehr basiert es auf der Annahme, dass sich kindliche Sprachentwicklung durch eine hohe Qualität des Sprachinputs und einen sensitiven Einsatz von Modellierungstechniken bestmöglich fördern lässt. Damit verfolgt Dialogisches Lesen einen *inhaltsbasierten Sprachförderansatz* (*focus on meaning*). Denn grundsätzlich lassen im Bereich von Sprachfördermaßnahmen zwei Konzeptionen unterscheiden. Die angesprochenen inhaltsbasierten Sprachförderansätze basieren auf der Annahme, dass formalsprachliche Regeln durch regelmäßigen Sprachgebrauch implizit erworben werden (Ellis, 1994). Inhaltsbasierte Fördermethoden sind von *formbasierten Förderkonzepten* zu unterscheiden, die sprachstrukturelles Wissen durch konkrete Übungen vermitteln (*focus on form*; DeKeyser, 1995; Darsow, Paetsch, Stanat, Felbrich, 2012; Ellis, 1994). Formbasierte Sprachkonzepte kommen häufig in der Förderung von metasprachlicher Kompetenzen (z. B. Phonologische Bewusstheit; Küspert & Schneider, 2006) und zur Förderung der Grammatik (Penner, 2003) zum Einsatz. Dialogisches Lesen gehört zu den in Kapitel 3 beschriebenen bedeutungsorientierten Sprachförderansätzen (*focus on meaning*).

Anders als in der klassischen Vorlesesituation, in der Kinder zumeist die Rolle des passiven Zuhörers einnehmen, soll Dialogisches Lesen die Kinder unterstützen, selbst Erzähler der Geschichte zu werden. Das Vorlesen gestaltet sich also als Dialog zwischen Kind und Erwachsenen, der sich durch die hohe Qualität der kindgerichteten Sprache auszeichnet (Brannon & Daukas, 2012b). Indem die Vorleser jede Äußerungen des Kindes mit scaffolding und Modellierungstechniken begleiten, nimmt das Kind zunehmend selbst die Rolle des Geschichtenerzählers ein.

Der nächste Abschnitt stellt zunächst die Grundprinzipien von Dialogischem Lesen vor. Darauf folgt ein Überblick über Ergebnisse aus internationalen Evaluationsstudien, die das Förderpotential der Methode prüfen. Dabei werden Ergebnisse für unterschiedliche Zielgruppen (z.B. Kinder ohne Förderbedarf versus Kinder mit Förderbedarf) und für verschiedene Fördersettings verglichen (z.B. häuslicher versus institutioneller Einsatz von Dialogischem Lesen). Einige Studien geben auch Hinweise über das zu erwartende Effektspektrum der Methode. Sie untersuchen differentielle Fördereffekte in den beschriebenen Sprachentwicklungsdomänen (siehe Kapitel 1). Abschließend folgt ein Überblick über Forschungsbemühungen aus dem deutschen Sprachraum.

3.2.1.1 Methode

Die ersten Evaluationsstudien, die sich mit der Dialogischem Lesen befassten führte die amerikanische Forschergruppe um Grover J. Whitehurst in den späten 1980er Jahren durch (Arnold, Lonigan, Whitehurst & Epstein, 1994; Lonigan & Whitehurst, 1998; Whitehurst et al., 1988, 1991, 1994, 1999). Zevenbergen und Whitehurst (2003) schreiben über die Methode:

„This technique [...] is based on the theory that practice in using language, feedback regarding language, and appropriately scaffolded adult-child interactions in the context of picture book reading facilitate young children's language development.” (S. 3).

Abhängig vom Alter der Kinder postulieren Whitehurst und Kollegen zwei unterschiedliche Varianten des Dialogischen Lesens. Sowohl für Kinder im Alter von 2 bis 3 Jahren, als auch für 4 bis 5-jährige liegt eine Zusammenstellung von sprachförderlichen Prinzipien vor, die im Rahmen des dialogischen Vorlesens beachtet werden sollen. Da in der eigenen Untersuchung Dialogisches Lesen bei Vorschulkindern (im Alter zwischen 5 und 6 Jahren) umgesetzt wurde, werden an dieser Stelle die methodischen Prinzipien für Arbeit mit älteren Kindern ausführlich betrachtet. Für die Arbeit mit älteren Kindern nennt Whitehurst insgesamt neun Regeln, die Erwachsene beim Einsatz von DL beachten sollen (siehe Abb. 2).

Fragetechniken CROWD	Modellierungstechniken PEER
• C omplete Prompts: Sätze sollen vervollständigt werden. <i>Bsp.: Der Bär nahm seine Angel mit, um _____.</i> • R ecall Prompts: Inhalte des Buchs sollen erinnert werden. <i>Bsp.: Erkläre mir: Warum musste der Tiger ins Krankenhaus gehen?</i> • O pen-ended Prompts: Aufforderung zum Erzählen in eigenen Worten. <i>Bsp.: Erzähle mal, was passiert auf dieser Seite.</i> • Wh -Prompts Offene und geschlossen W-Fragen stellen. <i>Bsp.: Wen hat der Tiger hier zu Besuch? Was machen der Bär und der Tiger hier?</i> • D istancing Prompts Fragen zum Alltagsbezug. <i>Bsp.: Wart ihr auch schon mal auf einer Schatzsuche wie der Bär und der Tiger?</i>	• P rompt: Stetige Aufforderung zu sprachlichem Handeln. • E xpand: Konsequentes Erweitern der kindlichen Sprache. • E valuate: Jeder Äußerung des Kindes folgt eine Bewertung durch implizite Rückmeldungen des Trainers. • R epeat: Das Kind wird aufgefordert jede erweiterte Äußerung zu wiederholen.

Abbildung 2 Prinzipien des Dialogischen Lesens für 4-5jährige Kinder in nach Whitehurst et al. (1988).

Die ersten fünf Prinzipien fasst Whitehurst unter dem Akronym CROWD zusammen. Beschrieben werden hier unterschiedliche Fragetechniken, die der Vorleser einsetzen soll.

Der Vorleser stellt Fragen, die dem Kind Anlässe bieten Inhalte der Geschichte zu wiederholen (*Weißt Du noch wie die Geschichte weiter ging?, Warum wollten die beiden nach Panama reisen?*).

Um gezielt auf wichtige Inhalte der Geschichte hinzuweisen, kann der Vorleser das Kind auch unvollständige Sätze beenden lassen („Der Bär nahm seine Angel mit, um....“).

Häufiger sollten allerdings offene Sprachanlässe geschaffen werden, die dem Kind die

Möglichkeit bieten, in eigenen Worten die Geschichte nachzuerzählen (z.B. „*Kannst Du mir erzählen, was auf diesem Bild passiert?*“).

Der Vorleser sollte außerdem geschlossene und offene Fragen zu der vorgelesenen Geschichte stellen. Geschlossene Fragen zielen in der Regel auf die Nennung einzelner Wörter (z.B. „*Was hat der Bär in seiner Hand?*“, „*Wo gehen der Bär und der Tiger hin?*“, „*Wer versteckt sich unter dem Tisch?*“). Offene Fragen geben dem Kind die Möglichkeit sich ausführlicher zu äußern. Sie können eine distanzierte Perspektive auf die Handlung einfordern und erfolgen dann, wenn das Kind z.B. nach seiner Meinung gefragt wird oder erraten soll, wie die Geschichte endet (z.B. „*Was glaubst Du, warum ist der Tiger traurig?*“, „*Was macht der Bär wohl als Nächstes?*“). Zum Einsatz geschlossener Fragen (low demand questions) und offener Fragen (high demand questions) merken Whitehurst und Arnold (2004) an, dass Erwachsene geschlossene Fragen vermehrt in den ersten Sitzungen einsetzen sollten, in denen das Kind noch nicht mit der Methode vertraut ist. Sie können aber auch als „Eisbrecher“ zu Beginn einer Fördersitzung hilfreich sein.

Grundsätzlich ist darauf zu achten, dass Erwachsene nach mehreren DL-Sessions den Anteil an geschlossenen Fragen reduzieren und häufiger offene Fragen stellen. Blewitt und Rump (2009) konnten in einer Untersuchung mit Vorschülern im Alter von 3 bis 4 Jahren zeigen, dass Kinder, denen ein Durchführungsleiter beim Vorlesen zunächst häufiger geschlossene und später häufiger offene Fragen stellte, bessere Leistungen in einem Wortdefinitionstest erzielten, als Kinder, denen die Durchführungsleiter über die gesamte Förderphase ausschließlich offene oder geschlossene Fragen stellten.

Auch im Sinne der Entwicklungstheorie nach Wygotsky (1978) kommt den komplexen sprachlichen Äußerungen von Kindern eine große Bedeutung zu. Sie können Erwachsenen Aufschluss geben, auf welchem Entwicklungsniveau sich ihr Kind sprachlich befindet. Durch Fragen, die sich dem geeigneten Niveau des Kindes anpassen, kann es gelingen die „nächste Zone“ der Sprachentwicklung zu erreichen.

Um Eltern eine Hilfestellung zu geben, wie sie mit Hilfe von Modellierungstechniken den sprachlichen Input für ihr Kind förderlich gestalten können, fasst Whitehurst vier Modellierungstechniken unter dem Akronym PEER zusammen. Grundsätzlich sollten Eltern während einer DL-Situation darauf achten, dass sie dem Kind viele Sprechanlässe bieten. Durch Lob sollten komplexe sprachliche Äußerungen positiv verstärkt werden. Besonders wichtig ist, dass Vorleser jede Äußerung aufgreifen und als Möglichkeit zur Modellierung nutzen. Eltern können die Aussage ihres Kindes auch modellieren, indem sie diese erweitern und neue Informationen in ihre Antwort einbauen.

Beispiel: Erweiterung

Kind: „*Der Tiger hat eine Tasche*“

Vorleser: *Richtig, der Tiger hat eine grüne Tasche in seiner rechten Hand*“.

Jede Aussage des Kindes sollte wenn nötig, implizit korrigiert werden.

Beispiel: Implizite Korrektur

Kind: „*Der hat die Eimer 'putt macht.*“

Vorleser: „*Stimmt, der Bär hat den Eimer kaputt gemacht.*“

Außerdem soll das Kind dazu aufgefordert werden korrigierte oder erweiterte Äußerungen zu wiederholen.

Beispiel: Erweiterung + Wiederholung

Kind: „*Der Hase mit den Schuhen.*“

Vorleser: „*Richtig, das ist der Hase mit den schnellen Schuhen.*

Jetzt sag Du es nochmal!“

Kind: „*Das ist der Hase mit den schnellen Schuhen.*“

Grundsätzlich können die DL-Prinzipien in jeder Vorlesesituation, mit jedem Buch zur Anwendung kommen. Vor allem in der Arbeit mit jüngeren Kindern empfiehlt es Bücher auszuwählen, die viele Illustrationen enthalten.

Dass DL einen positiven Beitrag zu kindlichen Sprachentwicklung leisten kann, konnte in diversen Experimentalstudien nachgewiesen werden. Im Folgenden soll ein Überblick über den aktuellen Forschungsstand gegeben werden. Tabelle 3 fasst die sprachförderlichen Interaktionen des Dialogischen Lesens zusammen.

Tabelle 3 Techniken des Dialogischen Lesens nach Ennemoser et al. (2013).

Funktion	Maßnahme / Technik
Anregung der Sprachproduktion	• W-Fragen • offene Fragen • Nachfragen zu Äußerungen des Kindes • Sätze vervollständigen lassen
Modellierung	• korrektive Wiederholung der Äußerungen des Kindes • Erweiterung und Umformulierung • Unterstützung
Verstärkung / Motivation	• Lob und Verstärkung • Orientierung an Interessen und Erfahrungen des Kindes • Spaß haben

3.2.1.2 Internationaler Forschungsstand

Die meisten Untersuchungen, die sich mit der Wirksamkeit von Dialogischem Lesen (DL) befassen, stammen aus dem anglo-amerikanischen Sprachraum. DL wird in der Literatur häufig als eine spezifische Methode unter dem Oberbegriff *storybook-reading interventions* beschrieben (Brannon & Dauksas, 2012a; Marulis & Neuman, 2010; Swanson et al., 2011). Neben dem DL zählen das *Wiederholende Lesen* (repeated reading) und das *konventionelle Vorlesen* (storybook reading) zum Kreis der *storybook-reading interventions*. Allerdings liegen allen genannten Interventionsarten unterschiedliche methodische Überlegungen zugrunde.

Beim *repeated reading* zum Beispiel wird dem Kind dasselbe Buch bei verschiedenen Anlässen mehrfach vorgelesen. Die Modellierungstechniken des Dialogischen Lesens spielen dabei eine untergeordnete Rolle (Swanson et al., 2011). Auch beim konventionellen *storybook reading* kommen die genannten Modellierungstechniken nicht explizit zum Einsatz. In der Literatur firmiert Dialogisches Lesen häufig unter verschiedenen Namen, wie *interactive storybook reading* oder *shared reading*. Die unklar abgegrenzten methodischen Begriffe erschweren es die Wirksamkeit von DL verlässlich einzuschätzen. Um eine möglichst genaue Einschätzung über das sprachförderliche Potential des Dialogischen Lesens zu erhalten, werden im Folgenden vornehmlich Ergebnisse aus Untersuchungen und Metaanalysen berichtet, die sich explizit mit der Evaluation des Dialogischen Lesens nach Whitehurst (1988) befassen.

Die Wirksamkeit von DL wurde über viele Jahre international, bei unterschiedlichen Stichproben und in verschiedenen Settings (Eltern-Kind Förderung vs. ausgebildete Trainer; Einzelförderung vs. Kleingruppenförderung) untersucht. Frühe Studien untersuchten die Wirksamkeit insbesondere bei Müttern, die ihren Kindern zu Hause dialogisch vorlasen.

Whitehurst und Kollegen (1988) führten eine erste Längsschnittuntersuchung mit 29 Kleinkindern im Alter von durchschnittlich 2;1 Jahren durch. Alle Kinder waren normalentwickelt und zeigten keine sprachlichen Auffälligkeiten. Sie wurden randomisiert auf eine Experimentalgruppe (EG; $n = 14$) und eine Kontrollgruppe (KG; $n = 15$) verteilt. In beiden Gruppen wurden die Eltern auf die Bedeutung des Vorlesens für die kindliche Sprachentwicklung hingewiesen. Den Eltern der EG wurde die Anwendung von DL erläutert. Alle Eltern lasen ihren Kindern einen Monat lang 3 bis 4 Mal pro Woche vor. Außerdem bekamen die Eltern die Aufgabe die Vorlesesituationen mit einem Audiogerät aufzunehmen. Vor und nach der Intervention wurden die Leistungen im passiven Wortschatz mittels des Peabody Picture Vocabulary Test (PPVT) erhoben. Der aktive Wortschatz wurde anhand von drei Maßen gemessen: Expressiv One-Word Vocabulary Test (EOWVT), Illinois Test of Psycholinguistic Abilities (ITPA) und der durchschnittlichen Äußerungslänge der Kinder (MLU). Ein Nachtest im Anschluss an die Förderung ergab, dass Kinder, denen dialogisch vorgelesen wurde, signifikant bessere Leistungen im aktiven Wortschatz erzielten als die KG ($d = 1.03$). Im passiven Wortschatz zeigten sich keine signifikanten Gruppenunterschiede. Neun Monate nach der Förderung ergaben sich im passiven wie auch im aktiven Wortschatz zwar tendenzielle Vorteile der DL-Gruppe, jedoch erreichten die Unterschiede nicht das Signifikanzniveau.

Die positiven Ergebnisse der ersten Studie zum Dialogischen Lesen konnten Arnold und Kollegen (1994) anhand einer größeren Stichprobe replizieren. Sie untersuchten den Einfluss verschiedener Coaching-Maßnahmen (persönliches Coaching vs. Video-Coaching), mit denen teilnehmende Eltern die Modellierungsstrategien im Dialogischen Lesen erlernten. Von 64 Kindern im Alter von 2;0 bis 2;10 wurden die Leistungen im passiven und aktiven Wortschatz (PPVT, EOWVT, ITPA) erhoben. Im Anschluss erfolgte eine randomisierte Verteilung auf drei Untersuchungsbedingungen. Eltern der ersten EG nahmen an zwei Coaching-Sitzungen teil, in denen Projektmitarbeiter DL persönlich vermittelten. Eltern der zweiten EG erhielten kein persönliches Coaching. Sie sahen einen Videofilm mit Übungen zum Dialogischen Lesen. Auch die Eltern der Kontrollgruppe erhielten kein spezifisches Coaching. Sie lasen ihren Kindern auf konventionelle Art und Weise vor.

Nach Abschluss der fünfmonatigen Interventionsphase wurde ein Nachtest durchgeführt. Kinder, deren Eltern das Video-Coaching erhalten hatten, erzielten im Vergleich zur KG bessere Leistungen im passiven und aktiven Wortschatz. Kinder, deren Eltern ein persönliches Coaching erhielten, waren der Kontrollgruppe in einem Maß des aktiven Wortschatzes (ITPA) überlegen. Im Vergleich Videocoaching versus persönliches Coaching erzielte die Video-Gruppe signifikant bessere Leistungen im passiven Wortschatz und in einem Maß des aktiven Wortschatzes (EOWVT). Effektstärken werden nicht berichtet. Die Überlegenheit des Videocoaching gegenüber einem persönlichen Coaching sehen die Autoren darin begründet, dass Eltern die DL-Techniken besser erlernen, wenn sie andere Eltern bei der Interaktion mit ihren Kindern beobachten können und die Methode nicht ausschließlich durch Erklärungen und Rollenspiele erlernen.

Die Wirksamkeit des Dialogischen Lesens bei gecoachten Eltern, die ihren Kindern zu Hause vorlasen, konnten zahlreiche Folgeuntersuchungen replizieren. Mol, Bus, de Jong und Smeets (2008) fassen in einer metanalytischen Betrachtung die Ergebnisse aus 16 Studien zusammen ($N = 626$). Die Analyse ergab einen moderaten Zusammenhang zwischen der Methode (DL) und der Sprachentwicklung der geförderten Kinder nach der Intervention ($r = .20$). Verbesserungen in der Sprachentwicklung konnten mit 4% durch die Intervention erklärt werden. Beschränkte sich die Analyse auf die Entwicklung des aktiven Wortschatzes, waren 8% der Varianz im Nachtest durch den Einsatz des Dialogischen Lesens zu erklären.⁹

Weiter konnten die Autoren zeigen, dass jüngere Kinder deutlich stärker vom Dialogischen Lesen profitieren als ältere. Bei 4-5jährigen klärte das Vorlesen durch Eltern weniger als 1% Varianz ($d = .14$) in Bezug auf die weitere Sprachentwicklung der Kinder auf. Bei den 2-3jährigen ergaben sich indes 4 bis 5% Varianzaufklärung. Dies entsprach einer mittleren Effektstärke von $d = .50$.

Weitere Analysen ergaben, dass Risikokinder wider Erwarten weniger vom Dialogischen Lesen profitierten als Kinder ohne Förderbedarf. Die Autoren sehen in diesem Ergebnis den sogenannten Matthäus-Effekt bestätigt, der sinngemäß besagt, dass eine erfolgreiche Entwicklung sehr wahrscheinlich weitere Erfolge nach sich zieht. Bezogen auf die Ergebnisse könnten also Eltern von unauffälligen Kindern die Prinzipien des Dialogischen Lesens möglicherweise besser umgesetzt haben als Eltern von Risikokindern, die zumeist aus Schichten mit niedrigerem Bildungsniveau stammen. Denkbar wäre auch, dass vor allem jüngere Risikokinder weniger von der Förderung profitierten als ältere Risikokinder, weil sie

⁹ Leistungen des aktiven Wortschatzes wurden insgesamt in 9 Studien erfasst.

zusätzlich zu etwaigen Entwicklungsverzögerungen mit größeren sprachlichen Hürden zu kämpfen haben. Gesicherte Befunde, die diese Hypothese stützen könnten, liegen indes noch nicht vor. Insgesamt liefern alle Untersuchungen positive Befunde für das sprachförderliche Potential des Dialogischen Lesens, wenn geschulte Eltern ihren Kindern zu Hause vorlasen. Weitere Untersuchungen befassten sich mit Frage, ob sich DL auch bei Kleingruppen im institutionellen Setting (Schule, Kindergarten) als sprachförderlich erweist. Vorleser waren in diesem Fall nicht die Eltern, sondern geschulte Erzieher, Lehrer oder Studenten.

Erste Hinweise lieferte eine Untersuchung von Valdez-Menchaca & Whitehurst (1992), in der 20 Kinder im Alter zwischen 2;3 und 2;11 Jahren teilnahmen. Alle Kinder sprachen Englisch als Zweitsprache und zeigten Schwächen in der Zweitsprachentwicklung. Die Stichprobe wurde randomisiert auf zwei Untersuchungsbedingungen verteilt. Kinder der EG erhielten Dialogisches Lesen im Rahmen einer Einzelförderung, die von ausgebildeten Studenten durchgeführt wurde. Kinder der KG erhielten eine Einzelförderung im Kunstunterricht. Die Dauer der Intervention erstreckte sich über 30 Tage, wobei die täglichen Sitzungen 10 - 12 Minuten dauerten. Kinder der DL-Gruppe erzielten im Nachtest bessere Leistungen im aktiven wie auch im passiven Wortschatz als die Kinder der KG. Außerdem zeigte sich während einer inszenierten Vorlesesituation, die nach der Förderung mit beiden Gruppen videographiert wurde, dass Kinder der DL-Gruppe sich häufiger zu Wort meldeten sowie längere und komplexere Sätze produzierten, als die Kinder der Kontrollgruppe. Einschränkend muss angemerkt werden, dass die Ergebnisse aufgrund des geringen Stichprobenumfang lediglich als Tendenzen interpretiert werden können. Effektstärken berichten die Autoren nicht.

In einer größer angelegten Untersuchung gingen Whitehurst et al. (1994) der Frage nach, ob sich das DL auch dann als wirksam erweisen würde, wenn es als Kleingruppenförderung in Kindertagesstätten stattfindet. 73 Risikokinder im Alter von drei Jahren nahmen an der Untersuchung teil. Die Stichprobe wurde randomisiert auf drei Untersuchungsgruppen verteilt. In der ersten Gruppe gestalteten ausgebildete Erzieher Vorlesesituationen in einer Kleingruppe mit max. fünf Kindern täglich für ca. 10 Minuten in ihrer Kita. Die zweite Gruppe erhielt ebenfalls Förderung im Sinne des Dialogischen Lesens zu denselben Bedingungen wie die erste Gruppe, aber zusätzlich lasen ihnen die Eltern auch zu Hause vor. Kinder der Kontrollgruppe nahmen an verschiedenen unspezifischen Spielsituationen teil, die ausgebildete Lehrer in Kleingruppen mit maximal fünf Kindern angeleiteten. Die Dauer des

Interventionszeitraums betrug für alle Gruppen sechs Wochen. Im Nachtest zeigten die Kinder beider DL-Gruppen bessere Leistungen im aktiven Wortschatz als die Kontrollgruppe. Auch im passiven Wortschatz waren beide DL-Fördergruppen der Kontrollgruppe tendenziell überlegen. In einer Follow Up-Erhebung nach sechs Monaten blieben die DL-Fördergruppen der Kontrollgruppe im aktiven Wortschatz signifikant überlegen. Effektstärken werden zu keinem Messzeitpunkt berichtet. Die Ergebnisse deuten an, dass mit DL nicht nur als Eins-zu-Eins-Förderung durch Eltern sprachförderliche Effekte erzielt werden können, sondern auch, wenn geschulte Fachkräfte DL in einer Kleingruppenförderung einsetzten. Zudem fanden sich langfristige Fördererfolge.

Diese positiven Befunde konnten Lonigan und Whitehurst (1998) in einer ähnlich angelegten Untersuchung replizieren. Sie wiesen 3- 4jährige Kinder, aus Elternhäusern mit geringem sozioökonomischen Status einer von vier Untersuchungsgruppen zu. In der ersten Gruppe lasen Lehrer den Kindern dialogisch in der Schule vor. Kinder der zweiten Gruppe bekamen ausschließlich zu Hause von ihren Eltern vorgelesen, während die dritte Gruppe zu Hause und in der Schule am Dialogischen Lesen teilnahm. Lehrer und Eltern nahmen an einem Videocoaching teil. Die Intervention dauerte 6 Wochen. Die Ergebnisse zeigten, dass Kinder aus Gruppen, in denen das Dialogische Lesen zu Hause von Eltern angeboten wurde, am meisten von der Förderung profitierten. Dialogisches Lesen als Kleingruppenförderung in der Schule erwies sich ebenfalls als wirksam, allerdings fielen die Effekte im Vergleich zur Einzelförderung durch Eltern geringer aus. Die Autoren berichten Effektstärken im expressiven Wortschatz von $d = .74$ für die kombinierte Trainingsgruppe (Schule + zu Hause) bis $d = .30$ für die Schulgruppe. Alle Trainingsgruppen waren der Kontrollgruppe überlegen.

In Bezug auf das Effektspektrum von Dialogischem Lesen als Kleingruppenförderung im institutionellen Rahmen zeigten Untersuchung von Justice und Ezell (2002) sowie Ukrainetz, Cooney, Dyer, Kysar und Harris (2000), dass sich Kinder nicht nur im Wortschatz verbesserten. Sie profitieren auch in der Buchstabenkenntnis und in der Phonologischen Bewusstheit. Hierzu untersuchte auch Aram (2006) eine Stichprobe von 156 israelischen Risikokindern im Vorschulalter (3-5 Jahre), die randomisiert einer von vier Trainingsbedingungen zugewiesen wurden. Die erste Experimentalgruppe ($n = 37$) nahm an DL teil. Die zweite EG ($n = 38$) trainierte ausschließlich schriftsprachliche Vorläuferkompetenzen (Buchstabenkenntnis und Phonologische Bewusstheit). Eine dritte Gruppe ($n = 40$) erhielt eine kombinierte Förderung, bei der die Trainer in einer dialogischen

Vorlesesituation auch Übungen zur Buchstabenkenntnis und zur Phonologischen Bewusstheit durchführten. Kinder der Kontrollgruppe ($n = 41$) nahmen am konventionellen Vorschulunterricht teil. In allen Experimentalgruppen integrierten geschulte Vorschullehrer die Interventionen über 28 Wochen zwei Mal wöchentlich für 20 - 30 Minuten in ihren Unterricht. Gefördert wurde in Kleingruppen mit maximal sechs Kindern. Im Nachtest waren die Kinder aller Trainingsgruppen der Kontrollgruppe sowohl im Wortschatz als auch in den schriftsprachlichen Vorläuferkompetenzen überlegen.

Differentielle Vergleiche der Trainingsgruppen untereinander zeigten, dass Kinder, die ausschließlich am Schriftsprachtraining teilnahmen, höhere Leistungszuwächse in der Buchstabenkenntnis und der Phonologischen Bewusstheit erzielten als die Kinder der übrigen Interventionsgruppen. Kinder, die an der kombinierten Intervention teilnahmen, profitierten in der Schriftsprachentwicklung mehr als die klassische DL-Gruppe und in der Wortschatzentwicklung mehr als die Schriftsprach-Gruppe. Die klassische DL-Gruppe war der Kontrollgruppe im Wortschatz, in der Buchstabenkenntnis und in Phonologischen Bewusstheit überlegen. Allerdings profitierten Kinder dieser Gruppe im Wortschatz anders als erwartet nicht mehr als die anderen Trainingsgruppen.

Mit der positiven Wirkung interaktiver Vorlesesituationen auf die Schriftsprachentwicklung von Kindern mit und ohne Sprachförderbedarf befassten sich auch Mol, Bus & de Jong (2009) in einer umfangreichen Metaanalyse. Die Autoren fassten die Ergebnisse aus 31 (quasi-) experimentellen Studien ($N = 2049$) zusammen, in denen ausgebildete Fachkräfte die Intervention anleiteten. Davon untersuchten acht Studien das Dialogische Lesen nach Whitehurst et al. (1988). Elf weitere Studien setzten DL-ähnliche sprachförderliche Techniken ein (*interactive reading*), ohne jedoch explizit auf Whitehurst und Kollegen zu verweisen. Weitere 12 Studien verbanden zusätzliche Klassenaktivitäten mit interaktivem Vorlesen (*interactive reading* +).

Unabhängig von der Art der Vorlese-Interventionen konnten die Autoren sprachförderliche Effekte für die Gesamtstichprobe nachweisen ($d = .54$). Eine Teilstichprobe von Risikokindern ($n = 1448$) konnte ihre Sprachkompetenzen unabhängig von der Art des interaktiven Vorlesens in ähnlichem Maße verbessern ($d = .57$). Ähnlich wie in den Studien zum häuslichen Vorlesen profitierten die Kinder der Gesamtstichprobe vor allem im aktiven Wortschatz ($d = .62$). Bei Risikokindern fiel die Effektstärke im aktiven Wortschatz noch etwas höher aus ($d = .72$). Im passiven Wortschatz ergaben sich moderate Effekte für die Gesamtstichprobe ($d = .45$) wie auch für die Stichprobe der Risikokinder ($d = .48$). Obwohl in

den durchgeführten Vorleseinterventionen keine schriftsprachlichen Inhalte vermittelt wurden, berichten die Autoren moderate Effektstärken auf die Schriftsprachentwicklung. Geförderte Kinder der Gesamtstichprobe verzeichneten signifikante Leistungszuwächse in der Buchstabenkenntnis ($d = .39$), der Phonologischen Bewusstheit ($d = .43$) und der orthographischen Bewusstheit ($d = .41$). Die Teilstichprobe der Risikokinder erzielte ähnlich hohe Effektstärken (Buchstabenkenntnis: $d = .40$; Phonologische Bewusstheit: $d = .52$; orthographische Bewusstheit; $d = .36$). Differentielle Analysen ergaben zudem größere Fördereffekte für Interventionen, die von wissenschaftlichen Hilfskräften durchgeführt wurden. Ausgebildete Fachkräfte erzielten geringe Effektstärken in Maßen der Sprach- und Schriftsprachentwicklung. Die Autoren nehmen an, dass es ausgebildeten Fachkräften grundsätzlich schwerer fällt eine neugelernte Methode strikt nach den Vorstellungen ihrer Entwickler umzusetzen.

In Bezug auf die Interventionsform erzielten Studien, die hoch standardisiert nach der Whitehurst-Methode arbeiteten, überraschender Weise die geringsten Effekte. Studien, die das weniger standardisierte interaktive Vorlesen verwendeten, waren dem klassischen Dialogischen Lesen nach Whitehurst (1988) überlegen. Einen Grund sehen Mol, Bus & de Jong (2009) in der Ausbildung der Trainer. In den hoch standardisierten Whitehurst-Studien erhielten die Trainer in Anlehnung an die Ergebnisse von Arnold et al. (1994) ein standardisiertes Coaching mittels Videofilmen, die best practice-Beispiele zeigten. Die Autoren vermuten, dass andere Forschungsgruppen den Trainern intensivere Schulungsmaßnahmen angeboten haben. Um diese Vermutungen zu prüfen sind jedoch verlässliche Daten zur Durchführungsqualität (*treatment validity*) von Nöten. Die Autoren kritisieren, dass keine der berücksichtigten Studien die Qualität der Durchführung in ausreichendem Maß untersuchte. Hier bedarf es weiterer Forschungsarbeiten.

3.2.1.3 Forschungsstand in Deutschland

Verglichen mit den umfangreichen Befunden internationaler Untersuchungen ist die Wirksamkeit von Dialogischem Lesen im deutschen Sprachraum bislang noch nicht hinreichend evaluiert. Hinweise zur positiven Wirkung bei deutschsprachigen Kindern liefern insbesondere Untersuchungen, in denen DL im Rahmen von Elterntrainings eingesetzt wurde. Buschmann (2009) entwickelte das *Heidelberger Elterntraining zur frühen Sprachförderung*. Das Training richtet sich an Eltern, „[...] deren Kinder bei der pädiatrischen Vorsorgeuntersuchung U7 im Alter von 21-24 Monaten durch eine isolierte Verzögerung in der sprachlichen Entwicklung auffallen.“ (Buschmann, 2011, S. 16). Sieben Sitzungen und

eine Nachschulung vermitteln den Eltern den Einsatz von sprachförderlichen Kommunikationstechniken, die sie beim Betrachten von Bilderbüchern einsetzen können (ebd.). Dabei nimmt die Autorin auch Bezug auf das Dialogische Lesen nach Whitehurst (1988). Zur Abbildung der Prozessqualität werden die Teilnehmer am Ende der siebten Sitzung um eine Beurteilung des Elterntrainings anhand eines kurzen Rückmeldebogen gebeten.

Die Wirksamkeit des Trainings wurde im Rahmen einer Evaluationsstudie am Sozialpädiatrischen Zentrum der Universität Heidelberg bei Kindern mit einer isolierten Sprachentwicklungsverzögerung ($N = 65$) ermittelt. Die Studie folgte einem Kontrollgruppen-Design mit Prä-, Post- und Follow Up-Testung. Die Stichprobe wurde randomisiert auf zwei Untersuchungsbedingungen verteilt. Eltern von Kindern der EG ($n = 31$) nahmen am Heidelberger Elterntraining teil. Eltern von Kindern der KG ($n = 34$) nahmen zu einem späteren Zeitpunkt am Trainingsprogramm teil (Wartekontrollgruppe). Die rezeptiven und expressiven sprachlichen Fertigkeiten wurden im Vor- und Nachtest mit dem Sprachentwicklungstest für zweijährige Kinder (*SETK-2*; Grimm, 2000) erfasst. Außerdem wurde die Einschätzung der Eltern zum Sprachentwicklungsstand ihrer Kinder durch den Fragebogen *ELFRA-2* erfasst (Grimm & Doil, 2000). Im Nachtest zeigten Kinder, deren Eltern am Training teilgenommen hatten, im SET-K bessere Leistungen in der Wort- und Satzproduktion. Im Sprachverständnis (Wortverständnis und Satzverständnis) ergaben sich keine Unterschiede zwischen den Gruppen. Effektstärken werden nicht berichtet. Eine Follow Up-Erhebung neun Monate nach der Intervention ergab keine signifikanten Unterschiede zwischen den Gruppen. Maße zur Kontrolle der Durchführungsqualität werden in der Untersuchung nicht eingesetzt. Kritisch anzumerken ist, dass die kleine Stichprobenanzahl sowie das Design der Untersuchung keine abschließende Beurteilung der Wirksamkeit des Trainings zulassen. Es ist nicht ausgeschlossen, dass es sich bei den berichteten Befunden möglicherweise nur um Zuwendungseffekte handelt. Um herauszufinden, ob die Effekte tatsächlich auf Verbesserungen in der Eltern-Kind-Kommunikation zurückzuführen sind, hätte ein anderes Studien-Design gewählt werden müssen, in dem auch die Kontrollgruppe eine elterliche Zuwendung (z.B. durch Spielsituationen) erhalten sollte. Nur auf diesem Weg lässt sich der Mehrwert von Modellierungstechniken beim Vorlesen gegenüber alltäglicher Sprachzuwendung sicher feststellen.

Die Wirksamkeit eines weiteren Elterntrainings „Lass uns lesen!“ (Rückert, Kunze & Schulte-Körne, 2010) untersuchten Rückert, Kunze, Schillert und Schulte-Körne (2010) bei

normalentwickelten Kindern im Alter zwischen 5 und 6 Jahren ($M = 5.5$). Anders als das Heidelberger Elterntraining folgt „Lass uns Lesen!“ einer programmatischen Struktur. In 16 Wochen führen die Eltern mit ihren Kindern täglich 10-15 Minuten vorgegebene Aktivitäten zum Dialogischen Vorlesen und zur Förderung der Phonologischen Bewusstheit durch. Nach abgeschlossener Interventionsphase konnten geförderte Kinder ($n = 37$) ihre Leistungen gegenüber einer Kontrollgruppe ($n = 41$) in der Phonologischen Bewusstheit im engeren Sinne (Bestimmung des Anlauts, Bestimmung des Restworts und Phonemsynthese) signifikant verbessern ($d = .54$). In der Wortschatzentwicklung wie auch im Textverständnis ergaben sich im Nachtest keine signifikanten Unterschiede zwischen den Gruppen. Gleiches traf auf die Buchstabenentwicklung und die Phonologische Bewusstheit im weiteren Sinne zu. Insgesamt konnten die Autoren zwar spezifische Effekte auf den Bereich der Phonologischen Bewusstheit nachweisen. Allerdings geben die Befunde keine Hinweise, inwieweit das dialogische Vorlesen die Effekte beeinflusst. Es ist anzunehmen, dass der formalsprachliche Fokus auf die Phonologische Bewusstheit in erster Linie für die Lernzuwächse verantwortlich ist.

Weitere Untersuchungen befassten sich mit der Wirksamkeit einer Adaption des Heidelberger Elterntrainings (Buschmann, 2009) zur Schulung von Erzieherinnen für den alltagsintegrierten Einsatz von Sprachlehrstrategien im Kindergraten. Buschmann und Joos (2011) evaluierten die Wirksamkeit des Erziehertrainings an ausgelesenen Kindern mit Sprachentwicklungsverzögerung im Alter von 18 bis 30 Monaten. Für Kinder der Interventionsgruppe ($n = 15$) nahmen Erzieherinnen an einer 5-tägigen Veranstaltung zum Heidelberger Training teil. Für Kinder der Kontrollgruppe ($n = 13$) erhielten die Erzieherinnen eine verkürzte eintägige Ausbildung mit den Inhalten des Heidelberger Trainings. Im Anschluss hatten die Erzieherinnen aus beiden Gruppen die Aufgabe mit den Förderkindern über sechs Monate möglichst täglich für zehn Minuten ein Buch anzuschauen. Die Sprachentwicklung vor und nach der Intervention wurde mit dem SET-K 2 erhoben. Die Autoren fanden im Nachtest für Kinder, deren Erzieher die intensivere Schulung erhalten hatten, bessere Leistungen in den SET-K 2-Subtests zur Wort- und Satzproduktion als für die Kinder, deren Erzieher an der kurzen Schulung teilgenommen hatten. Die Autorinnen weisen darauf hin, dass die Untersuchung auf Grund der Stichprobengröße lediglich als pilotierende Untersuchung zu verstehen ist. Es können keine allgemeinen Schlüsse auf die Wirksamkeit des Trainings abgeleitet werden.

Simon und Sachse (2013) untersuchten die kurz- und langfristige Wirksamkeit des Heidelberger Trainings für Erzieher an einer größeren Stichprobe von ausgelesenen Risikokindern ($n = 146$) mit einem Durchschnittsalter von 3;11 Jahren. Eine Experimentalgruppe von 76 Kindern arbeitete mit Erziehern, die nach dem Heidelberger Training geschult wurden. Bei Kindern der Kontrollgruppe ($n = 69$) nahmen die Erzieher an einer eintägigen Schulung zum Zusammenhang zwischen Denken und Sprechen und zu Möglichkeiten der indirekten sprachlichen Förderung durch eine Unterstützung kognitiver Funktionen mit Brett- und Regelspielen teil.

In ihren Ergebnissen fanden die Autorinnen in videographierten Bilderbuchsituationen eine höhere Äußerungslänge und mehr Sprechfreude bei Kindern, deren Erzieherinnen eine Schulung im Sinne des Heidelberger Trainings erhalten hatten. Allerdings zeigten sich in standardisierten Sprachentwicklungstests (SET-K 3-5; AWST-R; HSET und TROG-D) im Nachtest und im Follow-Up keine signifikanten Unterschiede zwischen den Gruppen. Die Autoren merken an, dass die Studie nicht feststellte, ob der Einsatz des Heidelberger Trainings tatsächlich zu einem erhöhten Einsatz von Sprachlehrstrategien im Alltag der geschulten Erzieher führten. Damit ist nicht hinreichend geklärt, welchen Einfluss das Heidelberger Training tatsächlich auf die pädagogische Praxis hat. Bislang mangelt es an Evaluationsstudien, die zunächst die generelle Wirksamkeit von Sprachlehrstrategien in gut kontrollierbaren Fördersituationen – z. B. beim Vorlesen im Kindergarten – untersuchen, bevor Schlussfolgerungen für den Einsatz einer alltagintegrierte Sprachförderung gezogen werden können.

Erste Hinweise, dass sich der Einsatz von Sprachlehrstrategien im Sinne des Dialogischen Lesens in einer kontrollierten Vorlesesituationen als Kleingruppenförderung von Migrantenkindern im Rahmen der hessischen Vorlaufkurse eignet, fanden Ennemoser, Kuhl und Pepouna (2013) in einer pilotierenden Untersuchung. Kindern der EG ($n = 22$) wurde von geschulten Lehramtsstudierenden in acht Sitzungen (à 30 Minuten) dialogisch vorgelesen. Kinder der KG ($n = 23$) nahmen am konventionellen Förderangebot der Vorlaufkurse teil. Die Sprachentwicklung der Kinder wurde mit dem SET-K 3-5 erhoben (Grimm, 2000). Im Nachtest ergaben sich für die DL-Gruppe bei Betrachtung des Gesamtscores größere Leistungszuwächse als für die KG ($d = .47$). Mit Blick auf die einzelnen Untertests waren Kinder, an DL teilnahmen, der KG in der morphologischen Regelbildung ($d = .41$), im Enkodieren semantischer Relationen ($d = .43$) und im Nachsprechen von Pseudowörtern ($d = .43$) überlegen. Ergänzend zu den internationalen Befunden, die zumeist Wortschatz- und

Sprachproduktionseffekte berichten, ist anzunehmen, dass dialogisches Vorlesen möglicherweise auch förderlich auf die Entwicklung von grammatikalischen Kompetenzen wirkt.

Zusammenfassend ist festzustellen, dass im deutschsprachigen Raum bislang keine Untersuchung vorliegt, die das sprachförderliche Potential des Dialogischen Lesens nach Whitehurst (1988) als potentielle Sprachfördermethode im Kindergarten umfänglich untersucht. Das beschriebene Heidelberger Training (Buschmann, 2009) nimmt in seiner theoretischen Konzeption zwar Bezug auf Whitehursts Prinzipien; allerdings wird Dialogisches Lesen nicht isoliert eingesetzt, sondern als Baustein für eine alltagintegrierte Sprachförderung betrachtet. Somit kann nicht eindeutig bestimmt werden, welcher Anteil der Fördererfolge tatsächlich auf Dialogisches Lesen zurückzuführen sind. Weiter untersuchten die Studien zumeist sehr kleine und spezielle Stichproben. Sie lassen keine allgemeine Aussage über die Wirksamkeit von DL als Methode zur Förderung von Kleingruppen zu. Auch fehlen Nachweise zur langfristigen Wirkung. Um das Potential von Dialogischem Lesen als Fördermaßnahme im Kindergarten besser einschätzen zu können, wäre es darüber hinaus interessant, das Effektspektrum der Methode mit einem bereits etablierten, formalsprachlich aufgebauten Sprachförderprogramm zu vergleichen. Die eigene Untersuchung möchte den genannten offenen Fragen zur Wirksamkeit des Dialogischen Lesens Rechnung tragen. Es sollen lang- und kurzfristige Effekte in einem kontrollierten Kleingruppensetting bei Risikokindern mit und ohne Migrationshintergrund untersucht werden, die sich im letzten Kindergartenjahr vor der Einschulung befinden. Um Aussagen über das Effektspektrum von DL treffen zu können, soll ein zweites Sprachtraining zum Einsatz kommen, in dem die Vermittlung von formalsprachlichem Wissen im Vordergrund steht. Im letzten Kindergartenjahr bietet sich der Einsatz eines Trainings der Phonologischen Bewusstheit an, weil sich der Trainingsinhalt auf formalsprachliche Übungen zur Förderung von phonetisch-phonologischen Sprachkompetenz beschränkt. Bei dem Vergleichstraining handelt es sich also um eine Methode der Sprachförderung, die sich klar vom Dialogischen Lesens unterscheiden lässt.

Im folgenden Kapitel wird das Konzept der Phonologischen Bewusstheit vorgestellt und ein Überblick über nationale und internationale Befunde zur Trainingswirksamkeit gegeben.

3.2.2 Phonologische Bewusstheit

Trainingsprogramme zur Förderung der Phonologischen Bewusstheit können als Paradebeispiel für die erfolgreiche Implementierung von evidenzbasierten Erkenntnissen in die pädagogische Praxis gelten. Befunde aus zahlreichen Längsschnittuntersuchungen führten zu der Erkenntnis, dass Kinder bereits vor Schuleintritt spezifische Vorkenntnisse erwerben, die für den Erfolg des späteren Schriftspracherwerbs entscheidend sind (Weber, Marx & Schneider, 2007). Es zeigte sich, dass eine frühzeitige Vermittlung von spezifischem Wissen über die phonologische Struktur einer Sprache höhere Bedeutung für eine erfolgreiche Leserechtschreibentwicklung hat als der Einfluss von allgemeinen Entwicklungsfaktoren wie Gedächtnisleistung, Intelligenz oder visuelle Fähigkeiten (Schneider & Näslund, 1993). Diese Erkenntnisse haben in den letzten Jahren vermehrt Einzug in die Förderpraxis der Kindergärten erhalten. Die Förderung der Phonologischen Bewusstheit (PB) zählt in Deutschland nunmehr zum Standardrepertoire der vorschulischen Förderung. Auch allgemein gefasste vorschulische Sprachförderprogramme greifen häufig Übungen zur Förderung der PB auf (z.B. *Deutsch für den Schulstart*; Kaltenbacher, 2008).

Ziel eines PB-Trainings ist es, Kindern metasprachliches Wissen über die lautliche Struktur ihrer Sprache zu vermitteln (Metz, Fröhlich, Petermann, 2010). Phonetisch-phonologische Kompetenzen, die Kinder im Verlauf ihrer Sprachentwicklung erworben haben (vgl. Kap. 1.1), sollen durch unterschiedliche Übungen (z.B. Reimen, Anlaute bestimmen, Worte in Phoneme zerlegen), zu einem bewusstem Verständnis von formalen Aspekten der sprachlichen Lautstruktur verfestigt werden (Küspert & Schneider, 2006).

Im folgenden Abschnitt wird die Methode der Förderung von PB zunächst näher erläutert. Im Anschluss werden bedeutsame nationale und internationale Befunde zur Wirksamkeit von PB-Trainings dargestellt und diskutiert. Vorweg sei angemerkt, dass die Evaluation von PB-Trainings in Deutschland bereits eine lange Tradition hat. So bezieht sich der folgende Überblick vornehmlich auf Ergebnisse aus dem deutschsprachigen Raum.

3.2.2.1 Methode der Förderung der Phonologischen Bewusstheit

In der Literatur wird das Konzept der PB unter dem Sammelbegriff der *phonologischen Informationsverarbeitung* zusammengefasst. „Damit ist die Nutzung von Informationen über die Lautstruktur bei der Auseinandersetzung mit gesprochener und geschriebener Sprache gemeint“ (Küspert & Schneider, 2006, S.13).

Es würde zu kurz greifen, verstünde man PB als einheitlich strukturierte Kompetenz. Vielmehr setzt sich die Fähigkeit der PB aus einer Reihe von Fertigkeiten zusammen, für

deren Durchführung unterschiedliche kognitive Ressourcen von Nöten sind (ebd.). Skowronek und Marx (1989) unterscheiden daher zwischen *Phonologische Bewusstheit im weiteren* und im *engeren Sinne*. Phonologische Bewusstheit im weiteren Sinne befasst sich mit größeren Einheiten der gesprochenen Sprache, also mit der Untergliederung eines Wortes in Silben oder dem Erkennen von Reimen. Phonologische Bewusstheit im engeren Sinne meint den bewussten Umgang mit dem lautlichen Material einer Sprache. Hierzu zählen klassischerweise Aufgaben zum Anlauterkennen oder zur Phonemsynthese (Zerlegung eines Wortes in seine lautlichen Bausteine). Abbildung 3 fasst die beschriebene Inhaltsbereiche von PB zusammen.

Häufig erwerben Kinder Kompetenzen der PB im weiteren Sinne bereits ohne explizites Training durch Klatsch- oder Singspiele. Fertigkeiten im Bereich der PB im engeren Sinne entwickeln sich häufig erst mit Eintritt in die Schule über erste Erfahrungen mit Lese- und Schreibübungen (ebd.). Um Schwierigkeiten in der Lese-Rechtschreibentwicklung vorzubeugen, kommen PB-Trainings vor allem im Kindergartenalter zum Einsatz. Durch ein PB-Training gewinnen Kinder Einsichten in die lautliche Struktur der gesprochenen Sprache. Da die PB-Trainings formalsprachliche Kenntnisse durch explizite Aufgaben und gerichtete Anleitung vermitteln, sind sie als *formorientierte Förderansätze* (focus on form) zu klassifizieren (vgl. Kap. 3.2.1).

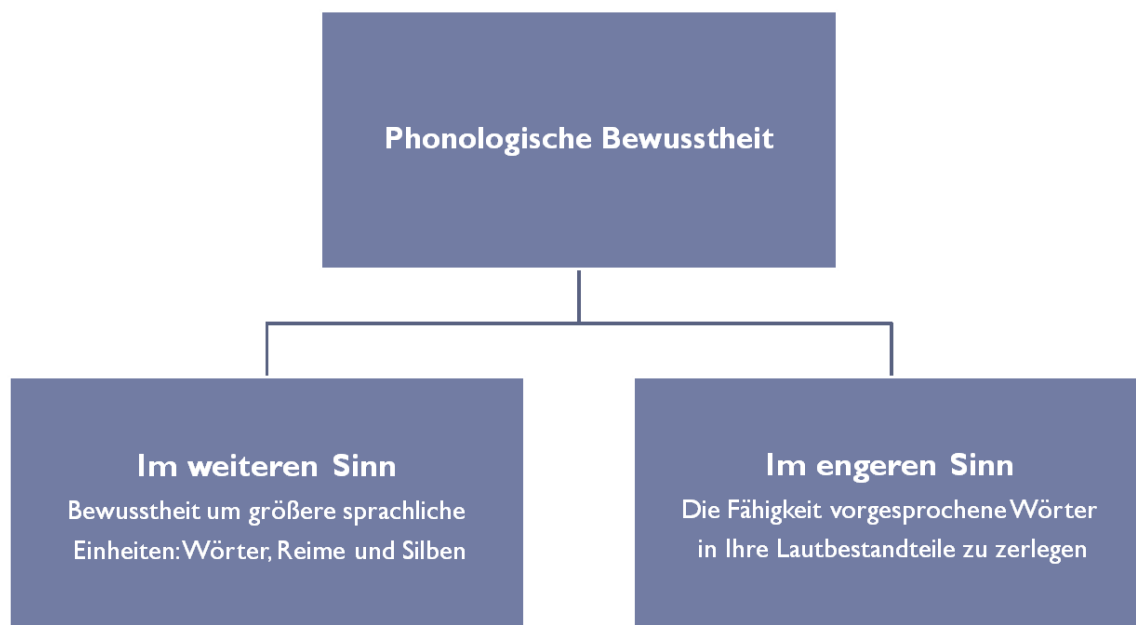


Abbildung 3 Aufbau des Konstrukts der Phonologischen Bewusstheit

3.2.2.2 Wissenschaftliche Evaluation

Die Wirksamkeit von PB-Trainings konnten nationale und internationalen Forschungsarbeiten nachweisen. Eine der ersten Studien führten Lundberg, Frost und Petersen (1988) durch. Sie konnten zeigen, dass sich PB-Leistungen mit Hilfe eines strukturierten Trainingsprogramms für Kinder im Vorschulalter, die keinerlei Schriftsprachkenntnisse besaßen, verbessern ließen. Geförderte Kinder konnten in ihren Lese- und Rechtschreibleistungen bis zum Ende ihrer Grundschulzeit von der Förderung profitieren.

In Deutschland untersuchten Schneider und Kollegen seit Beginn der 1990er Jahre die Wirksamkeit des Würzburger Trainingsprogramms *Hören, lauschen, lernen*. Dabei handelt es sich im Kern um eine deutsche Adaption des Lundberg'schen Trainings.

In einer ersten Interventionsstudie konnten Schneider, Visé, Reimers und Blässer (1994) die kurzfristige Wirksamkeit des Trainings an einer unausgelesenen Stichprobe von Kindergartenkindern ($N = 317$) zunächst nur dann nachweisen, wenn sich Erzieherinnen exakt an die Trainingsvorgaben gehalten hatten. Langfristige Effekte konnten an dieser Stelle noch nicht verzeichnet werden.

In einer zweiten Untersuchung (Schneider, Küspert, Roth, Visé & Marx, 1997; Schneider, Roth & Küspert, 1999) wurde die Qualität der Trainingsdurchführung von den Wissenschaftlern stärker kontrolliert. Die Stichprobengröße war mit der ersten Untersuchung vergleichbar. Die Trainingsgruppe zählte 191 Kinder; 151 Kinder umfasste die Kontrollgruppe. In Nachtest zeigten sich zunächst spezifische Effekte auf die PB-Leistung. Eine Follow Up-Erhebung zeigte zudem positive Transfereffekte auf die Lese- und Rechtschreibentwicklung von geförderten Kinder bis ins 2. Schuljahr.

In einer dritten Untersuchung wurden erstmals ausschließlich Kinder gefördert, deren schriftsprachliche Vorläuferkompetenzen im Risikobereich lagen (unterdurchschnittliche Leistungen in der phonologischen Informationsverarbeitung; Roth, Schneider, 2002; Schneider, Roth & Ennemoser, 2000). Dabei untersuchten die Autoren die Effektivität von drei Trainingsvarianten. Die erste Trainingsgruppe erhielt das klassische PB-Training, die zweite Gruppe ein Buchstaben-Laut-Training und die dritte Gruppe eine Kombination aus beiden Trainingsinhalten. Im Nachtest zeigte sich, dass die PB-Leistungen vor allem durch das klassische PB-Training sowie durch die kombinierte Trainingsversion gefördert wurde. Kinder, die das Buchstaben-Laut-Training erhalten hatten, und Kinder, die am kombinierten Training teilnahmen, zeigten verbesserte Buchstabenkenntnis. Die PB-Effekte blieben bis ins 1. Schuljahr hinein erhalten. Am Ende der ersten Klasse wurden Transfereffekte auf Lese-Rechtschreibkompetenzen untersucht. Es zeigte sich, dass die Risikokinder, die im

Kindergarten an der kombinierten Förderung teilgenommen hatten, am meisten profitierten. Ihre Leistungen unterschieden sich nicht signifikant von denen der unausgelesenen Kontrollgruppe.

Befunde, die auf die Wirksamkeit eines PB-Trainings auch für Kinder mit Sprachentwicklungsstörungen hinweisen, berichten Marx, Weber und Schneider (2005a, 2005b). In einem Vortest wurden neben der PB-Leistung auch die Wortschatz- und Grammatikleistung erhoben, um entsprechende Risikokinder zu klassifizieren. Es zeigt sich zunächst, dass sprachentwicklungsverzögerte Kinder grundlegend schlechtere PB-Leistungen erzielten als unauffällige Kinder. Beide Gruppen konnten jedoch von der Interventionsmaßnahme profitieren und zeigten verbesserte PB-Leistung. Die unmittelbaren Effekte konnten indes nicht langfristig bestätigt werden.

Eine weitere Subgruppen-Analyse führten Weber, Marx und Schneider (2007a, 2007b) an Kindern mit Migrationshintergrund durch. In einem Vortest zeigten Kinder mit Migrationshintergrund deutlich niedrigere Leistungen in der phonologischen Informationsverarbeitung als Kinder mit deutscher Muttersprache. Der Nachtest ergab, dass Migrantenkinder von einem PB-Training im gleichen Ausmaß profitierten wie die muttersprachliche Vergleichsgruppe. Bei der Erhebung von Lese-Rechtschreibfertigkeiten schnitten Kinder mit Migrationshintergrund insgesamt etwas schlechter ab als Muttersprachler.

Die positiven Befunde der in Deutschland durchgeführten Studien zur Wirksamkeit einer PB-Förderung decken sich mit internationalen Forschungsergebnissen, die Bus und van Izedoorn (1999) in einer Metaanalyse zusammenfassten. Sie analysierten die Ergebnisse aus US-amerikanischen Interventionsstudien. Dabei prüften sie die Wirkung von PB-Trainings auf die Entwicklung der PB-Leistung, sowie Transfereffekte auf spätere Leseleistung. Die Autoren berichten, dass Kinder, die an einem PB-Training teilnahmen, unmittelbar nach der Förderung große Lernzuwächse in ihrer PB-Leistung zeigten ($d = 1.04$). Moderate Effekte ergaben sich in der nach Schuleintritt untersuchten Leseleistung der Förderkinder ($d = .44$).

Zusammenfassend lässt sich feststellen, dass PB-Trainings als Fördermaßnahme im Vorschulalter in vielen deutschsprachigen und internationalen Untersuchungen mit verschiedenen Stichproben nachweislich wirksam sind. Im Unterschied zum Dialogischen Lesen, das einem inhaltsbasierten Sprachförderansatz folgt, liegt einem PB-Training ein

formbasierter Förderansatz zugrunde. Das Training beschränkt sich explizit auf die Vermittlung von phonologischen Kompetenzen. Ebenso wie die Dialogisches Lesen ist auch ein PB-Training für den Einsatz im letzten Kindergartenjahr geeignet und kann unter vergleichbaren Bedingungen in Kleingruppen mit maximal fünf Kindern durchgeführt werden. In der vorliegenden Untersuchung soll ein PB-Training als Kontrolltraining eingesetzt werden. Es ist dabei nicht als Konkurrenztraining zum Dialogischen Lesen gedacht, weil sich beide Maßnahmen inhaltlich stark voneinander unterscheiden. Das PB-Training wird in der Untersuchung vielmehr als Repräsentant für ein formbasiertes Sprachfördertraining eingesetzt, das Kompetenzen in einem spezifischen Sprachenentwicklungsbereich fördert. Außerdem ist es zur Kontrolle des Zuwendungseffekts zu verwenden. Dialogisches Lesen hingegen ist eine inhaltsbasierte Fördermethode, von der eine breitere Effekstreuung auf andere Sprachkompetenzbereiche, wie Wortschatz, Sprachverständnis und Grammatik, erwartet wird. Allenfalls im Bereich der phonetisch-phonologischen Entwicklung könnten beide Maßnahmen eine positive Wirkung erzielen. Zu vergleichen ist also das Effektspektrum einer inhaltsbasierten Fördermethode und eines formbasierten Sprachtrainings in Bezug auf die sprachliche Förderung von Risikokindern im Vorschulkindern

Neben der Untersuchung von Wirksamkeit und Effektspektrum möchte die vorliegende Studie auch der Frage nachgehen, welchen Einfluss die Durchführungsqualität (*Treatment Validität*) für die Wirksamkeit des Dialogischen Lesens hat. Da DL keinem vorgegebenen Skript folgt und die Güte der Förderung somit nicht über die Manualtreue erhoben werden kann, gibt Kapitel 4 zunächst einen theoretischen Überblick über Möglichkeiten zur Operationalisierung der Durchführungsqualität. Dabei sollen Impulse für die eigene Untersuchung herausgearbeitet werden.

4 Evaluation von Sprachförderung unter Berücksichtigung der Durchführungsqualität

In Kapitel 3.1 wurden bereits die Ergebnisse der großangelegten Interventionsstudien in Deutschland berichtet, die sich mit Evaluation von Sprachfördermaßnahmen im Kindergarten befassten. Dabei sorgten die Befunde bundesweit für Ernüchterung. Keine der Untersuchungen konnte die Wirksamkeit der untersuchten Fördermaßnahmen umfassend bestätigen. Die naheliegende Schlussfolgerung, dass Sprachfördermaßnahmen generell wirkungslos seien, ist bei der Frage, warum Effekte nicht nachweisbar waren, voreilig. Grundsätzlich unterliegen pädagogische Trainingsstudien einer Vielzahl von Einflussfaktoren, die die Wirksamkeit einer Intervention beeinflussen können. Pressley und Harris (1994) stellen fest, dass elaborierte Interventionsstudien nicht allein danach fragen, ob eine Intervention wirkt, sondern auch, wie eine Wirkung zustande kommt und welche Faktoren für einen wirksamen Einsatz einer Intervention verantwortlich sind. Die Autoren fordern in der Planung von Interventionsstudien eine breitere Berücksichtigung von potentiellen Einflussvariablen, die neben der Intervention Einfluss auf die Wirksamkeit nehmen könnten. Ein einflussreicher Faktor kann die Genauigkeit sein, mit der die fördernde Personen eine Methode oder ein Training nach den Vorgaben der Entwickler umsetzen. Die Güte der Trainingsdurchführung wird häufig mit den Begriffen *Treatment-Validität* (z. B., Pressley & Harris, 1994) oder *Durchführungsqualität* (z.B. Hess, Möller & Probst, 2008) beschrieben. Diese Arbeit verwendet den Begriff *Durchführungsqualität*. Der Prüfung der Durchführungsqualität kommt nicht nur als post hoc Analyse bei ausbleibender Wirkung Bedeutung zu (Bickman & Athay, 2009). Sie ist ebenso Beleg für die Ähnlichkeit gegenüber Gruppenbedingungen während der Interventionsphase (Mowbray, Holter, Teague & Bybee, 2003).

Auch wenn viele Evaluationsstudien die Kontrolle der Durchführungsqualität als wünschenswert und notwendig erachten, finden sich für die Evaluation von Sprachfördermaßnahmen bislang kaum Untersuchungen, die diesen Aspekt angemessen berücksichtigen. Eine der ersten Studien zur Evaluation eines Trainings der Phonologischen Bewusstheit von Schneider, Visé, Reimers und Blässer (1994) zeigte, dass die Durchführungsqualität einen entscheidenden Einfluss auf die Wirksamkeit von Trainings haben kann. In der genannten Untersuchung konnten Effekte des Trainings nur für Gruppen gezeigt werden, in denen sich die fördernden Erzieher genau an die Vorgaben des Manuals

hielten. Die hier durchgeführte Erfassung der *Manualtreue* bietet eine Möglichkeit die Durchführungsqualität eines Trainings zu erfassen. Weitaus schwieriger gestaltet sich die Operationalisierung, wenn die eingesetzte Intervention keinem programmatischen Aufbau folgt, dessen Einhaltung kontrolliert werden kann. Dies ist bei vielen inhaltsbasierten Sprachfördermethoden, wie dem Dialogischen Lesen, der Fall.

Bislang finden sich in der Literatur keine Hinweise, mit welchen Maßen die Durchführungsqualität des Dialogischen Lesens erhoben werden kann. Allerdings finden sich unterschiedliche methodische Möglichkeiten, mit denen die Beurteilung der Durchführungsqualität unabhängig von der inhaltlichen Ausrichtung einer Intervention realisiert werden kann.

Unterschieden wird zwischen 1) Programmtreue, 2) Durchführungsintensität, 3) Vermittlungsqualität, 4) Teilnehmerreaktion und 5) Differenziertheit (Dane & Schneider, 1998; Dusenbury, Brannigan, Falco & Hansen, 2003).

1) Programmtreue (adherence; fidelity; program integrity)

Die Programmtreue ist das geläufigste Maß zur Operationalisierung der Programmgüte (Durlak & DuPre, 2008). Programmtreue wird häufig als Genauigkeit verstanden, mit der sich ein Programmanwender an Instruktionen, Aufbau und Methode der Interventionsmaßnahme hält. Berkel und Kollegen (2011) weisen darauf hin, dass derzeit noch keine einheitlichen Standards zur Erfassung der Programmtreue vorliegen. Dies ist vor allem der Tatsache geschuldet, dass Interventionsprogramme in verschiedenen Einsatzbereichen (z.B. in Medizin, Psychologie oder Pädagogik) unterschiedliche Inhalte abdecken. Um die Programmtreue zu beurteilen, erheben viele Studien das prozentuale Verhältnis zwischen vorgegebenen und tatsächlich erteilten Instruktionen (*Manualtreue*). Auch Zeitmessungen können feststellen, wie häufig die Anwender die vorgegebenen Inhalte eines Trainings in einer Sitzung behandeln (ebd.). Für die Beurteilung von Interventionen, bei denen Trainer kein standardisiertes Manual an die Hand bekommen, kann die Durchführungsqualität über Beobachtungsbögen von externen Ratern erfasst werden. Dabei können unterschiedliche Bewertungsverfahren zum Einsatz kommen. Eine recht einfache Möglichkeit, die Durchführungsqualität zu bewerten, bieten mehrstufige Ratingskalen. Dabei werden externe Rater gebeten, die Stufe der Skala anzukreuzen, die ihrer Einschätzung in Bezug auf das zu beurteilende Merkmal (z. B. Durchführungsqualität) am ehesten entspricht. Ratingskalen werden in den USA z. B. häufig für die Beurteilung von Eltern-Kind-Interaktionen eingesetzt. Einen Überblick geben Neuman & Carta (2011).

Neben Ratingskalen kann die Beurteilung der Durchführungsqualität auch über Auszählungen erfolgen. Gezählt wird z. B., wie häufig ein Trainer ein gewünschtes Verhalten innerhalb einer Fördersitzung zeigt. Dies erfordert allerdings eine umfassende und intensive Ausbildung der externen Beobachter und ist zugleich mit einem hohen Auswertungsaufwand verbunden (Kany & Schöler, 2010). Außerdem müssen die förderlichen Kernelemente der Intervention im Vorfeld eindeutig definiert und voneinander abgrenzbar sein (Bickman, Riemer, Brown, Jones, Flay et al., 2009). Im Vergleich zu einer Ratingskala könnte der Vorteil des Zählens von gewünschten Verhaltensweisen darin liegen, dass es sich um ein sehr sensitives Maß im Vergleich zu den eher grobmaschigen Ratingitems handelt. Dennoch werden Auszählungssitems bislang selten in Beobachtungsbögen eingesetzt (Berkel et al., 2011).

2) Durchführungsintensität (*exposure; dosis*)

Die Durchführungsintensität wird in der Regel über die Anzahl der durchgeführten Fördersitzungen (*Sessions*) operationalisiert (Dusenbury et al., 2003). In den meisten Interventionsstudien ist die Anzahl der Sessions ein streng kontrollierter Faktor, so dass Unterschiede in der „Dosierung“ innerhalb einer Untersuchung zumeist ausgeschlossen werden können. Dane und Schneider (1998) nennen zwei weitere Maßeinheiten der Durchführungsintensität, die zusätzlich kontrolliert werden: a) die Länge der Sitzungen und b) die Frequenz, in welcher die Sitzungen erteilt wurden (z.B. Anzahl der Sitzungen pro Woche).

Innerhalb einer Untersuchung können Unterschiede in der Durchführungsintensität z.B. durch Fehlzeiten der Teilnehmer oder durch ausgefallene Sitzungen entstehen. Die Häufigkeit der Teilnahme können Anwesenheitslisten und Trainerberichte erfassen (Dusenbury et al., 2003). Post hoc-Schätzungen der Trainer oder Hochrechnungen auf Grundlage von zufällig ausgewählten Sessions sind keine probaten Erhebungsmethoden, da Fehlzeiten systematisch unterschätzt werden könnten (Berkel et al., 2011).

Problematisch ist eine divergierende Durchführungsintensität auch dann, wenn die Effektivität eines Programms anhand mehrerer Studienergebnisse beurteilt werden soll, zum Beispiel in Rahmen einer Metanalyse. Um dennoch Vergleichbarkeit in Metaanalysen zu gewährleisten, werden a priori-Bedingungen in Bezug auf das Untersuchungsdesign formuliert. Unzureichende Durchführungsintensität (z.B. Interventionszeitraum < 3 Monate) ist dabei ein häufiges Ausschlusskriterium für Studien (Swanson et al., 2011).

3) Vermittlungsqualität (*Quality of program delivery*)

Dane und Schneider (1998) fassen unter dem Begriff *Vermittlungsqualität* all jene Faktoren zusammen, die unabhängig von Umsetzungsvorgaben die Interventionsqualität positiv oder negativ beeinflussen können. Gemeint sind zum Beispiel die Motivation der Trainer, die Arbeitsatmosphäre, die persönliche Einstellung gegenüber der Intervention, der Umgang mit Störungen oder die Empathie im Verhalten gegenüber den Teilnehmern.

Viele Interventionsmaßnahmen verweisen in ihren Handreichungen auf sogenannte „soft skills“ hin, die Anwender bei der Durchführung des Trainings beachten sollten. Die Autoren des Würzburger Trainings „Hören, lauschen, lernen“ schlagen den Anwendern zum Beispiel vor die Durchführung „in gemütlicher und entspannter Atmosphäre“ sicherzustellen, deutlich und langsam zu sprechen und individuelle Leistungsunterschiede zwischen den Kindern zu berücksichtigen (Küspert & Schneider, 2006, S.21 ff.).

Die Art und Weise der Vermittlung von Trainingsinhalten kann sich selbst bei professionellen Anwendern stark unterscheiden. Dusenbury und Kollegen (2003) schlagen folgende Merkmale zur Operationalisierung der Vermittlungsqualität vor:

„The quality of interaction and the degree to which interactive activities focus on attention on desired elements are thus important to measure.“ (S. 224).

Es ist möglich eine Selbsteinschätzung der Trainer zur Bewertung der Vermittlungsqualität heranzuziehen. Allerdings stellen Beobachtungsbögen, die von externen Ratern bearbeitet werden, vermutlich ein objektiveres Maß dar.

4) Teilnehmerreaktion (*participant responsiveness*)

Einige Untersuchungen erfassen die Durchführungsqualität mit der Reaktion der Trainingsteilnehmer bezüglich der erhaltenen Intervention. Dane und Schneider (1998) definieren die Teilnehmerreaktion als ein Maß „[...] which may include indicators such as levels of participation and enthusiasm.“. Als weitere Indikatoren schlagen Berkel und Kollegen (2011) Messungen der aktiven Teilnahme, Befragungen zur Zufriedenheit und zur Übertragbarkeit der Programminhalte in den Alltag vor. Erstaunlicherweise erfährt die Untersuchung der Teilnehmerreaktion bei Interventionsstudien in pädagogischen Forschungsfeldern bislang nur wenig Beachtung.

5) Differenziertheit (*program differentiation*)

Um die Differenziertheit einer Intervention sicherzustellen, sollte kontrolliert werden, ob die Teilnehmer während des Interventionszeitraums zusätzlich an anderen Maßnahmen mit

ähnlicher Zielrichtung teilnehmen. Dane und Schneider (1998) definieren die Kontrolle der Differenziertheit als „[...] manipulation check that is performed to safeguard against the diffusion of treatments, that is, to ensure that the subjects in each experimental group received only planned intervention.“ (S. 45). Bei der Wirksamkeitsevaluation schulischer und vorschulischer Fördermaßnahmen ist davon auszugehen, dass nahezu alle Kinder neben dem Forschungsprogramm an weiteren Maßnahmen des Schul- und Vorschulcurriculums teilnehmen. Durch Befragungen der pädagogischen Fachkräfte können Inhalte und Intensität zusätzlicher Zuwendungen erhoben werden.

Es ist deutlich geworden, dass Durchführungsqualität für die Beurteilung der Wirksamkeit von Interventionen eine bedeutsame Rolle spielt. Sie bietet eine Möglichkeit die interne Validität einer Intervention zu kontrollieren (Gollwitzer & Jäger, 2007). Neuman und Dickinson (2011) kritisieren, dass viele Untersuchungen den Aspekt der Durchführungsqualität bei der Wirksamkeitsprüfung von Sprachfördermaßnahmen nicht berücksichtigen.

Da die Kontrolle der Durchführungsqualität in vielen Evaluationsstudien von Sprachfördermaßnahmen bislang keine Beachtung fand, möchte die eigene Untersuchung diesen Aspekt mitberücksichtigen. Durch die recht offen gestaltete Förderung ohne explizite Handlungsvorgaben eines Manuals, könnten bei der Umsetzung des Dialogischen Lesens wirksamkeitsbeeinflussende Unterschiede zwischen den Trainern entstehen. Bezogen auf das Dialogische Lesen konnten Swanson et al. (2011) bereits zeigen, dass Effekte in der Sprachkompetenz in Abhängigkeit vom Trainer unterschiedlich ausfallen können. Kinder, die von geschulten studentischen Hilfskräften gefördert wurden, erzielten größere Effekte als Kinder, bei denen Lehrer Dialogisches Lesen im Unterricht einsetzen. Effekte im Dialogischen Lesen können also in Abhängigkeit von der durchführenden Person unterschiedlich ausfallen. Offen bleibt, ob die Durchführungsqualität oder die Durchführungsintensität für diesen Unterschied verantwortlich sind. Um dies genauer zu erfassen fehlen bislang Instrumente mit denen die Durchführungsqualität des Dialogischen Lesens zuverlässig erhoben werden kann. Deshalb wurde im Vorfeld der vorliegenden Untersuchung ein Beobachtungsbogen entwickelt, mit dem externe Rater die Güte der Durchführung beurteilen. Da wenig darüber bekannt ist, welche Maße sich am besten zur Erfassung der Durchführungsqualität des Dialogischen Lesens eignen, berücksichtigt der Beobachtungsbogen Ratingskalen und Auszählungsitems gleichermaßen. Eine ausführliche Beschreibung des Messinstruments erfolgt in Kapitel 6.2.6.

Bevor sich die folgenden Kapitel mit Aufbau und Durchführung der eigenen Untersuchung befassen, werden in Kapitel 5, ausgehend von der aktuellen Befundlage, die Fragestellungen und Hypothesen der eigenen Untersuchung zusammengefasst.

5 Fragestellung und Hypothesen

Es besteht kein Zweifel, dass die sprachlichen Kompetenzen eines Kindes eine hohe Bedeutung für seine allgemeine und schulische Entwicklung haben. Damit zählt die Förderung sprachlicher Kompetenzen zu den zentralen Herausforderungen der vorschulischen Bildung (Redder et. al, 2011).

Trotz der hierzulande großen Vielfalt an Programmen und Methoden zur sprachlichen Förderung im Kindergarten, liefern empirische Studien zur Wirksamkeit von Sprachförderprogrammen eher ernüchternde Befunde.

Die große Vielfalt an Sprachförderprogrammen erschwert die Suche nach erprobten und gut evaluierten Angeboten im Sinne des „Best Practice“-Gedankens. Redder et al. (2011) formulieren in Bezug auf den Forschungsstand:

„Es gibt zahlreiche, jedoch unkoordinierte und im Qualitätsstand unvereinbare Diagnose- und Fördermaßnahmen, die über eine ad-hoc Aktivität nicht hinausführen. [...] Es mangelt an wissenschaftlich verantworteten Interventionen.“ (S.6).

Mit Blick auf internationale Forschungsarbeiten lassen sich allerdings zwei Fördermethoden herausstellen, deren Wirksamkeit für die kindliche Sprachentwicklung in vielen Studien positiv belegt werden konnten. Dazu zählen nach Justice und Pullen (2003) Trainings der Phonologischen Bewusstheit (PB) und die Sprachförderung nach der Methode des Dialogischen Lesens (DL). Allerdings zielen beide Fördermethoden – wie in den Kapiteln 3.2.1 und 3.3.3 beschreiben – auf unterschiedliche Bereiche der Sprachkompetenz.

Während die Wirksamkeit von Trainings der Phonologischen Bewusstheit in einer Vielzahl von Interventionsstudien auch bei deutschen Kindern nachgewiesen werden konnte (vgl. Kap 3.2.2), ist das Potential des Dialogischen Lesens als Sprachfördermaßnahme bei sprachauffälligen Vorschulkindern in Deutschland kaum erforscht. Außerdem mangelt es an Studien, die den Einfluss der Durchführungsqualität (Treatment Validität) auf die Wirksamkeit von Sprachfördermaßnahmen untersuchen (vgl. Kap. 4). Ausgehend von dieser Forschungslage ergeben sich mehrere Fragestellungen, die zu dieser Untersuchung führten:

1. Erweist sich Sprachförderung durch Dialogisches Lesen auch als Kleingruppenförderung in deutschen Kindergärten bei Risikokindern mit und ohne Migrationshintergrund kurz- und langfristig als wirksam?

Aus vorhandenen Forschungsergebnissen geht hervor, dass sich der Einsatz von Dialogischem Lesen vor allem positiv auf die Wortschatzentwicklung der geförderten Kinder auswirkt (Mol, Bus & de Jong, 2009). Viele Befunde deuten darauf hin, dass Dialogisches Lesen ein breites Spektrum an sprachlichen Kompetenzen wirksam fördern kann (Aram, 2006; Justice & Ezell, 2002, Swanson et al., 2011). Im deutschsprachigen Raum liegt bislang nur eine Untersuchung vor, die den isolierten Einsatz des Dialogischen Lesens nach Whitehurst et al. (1988) als Sprachfördermaßnahme zur Förderung von Kleingruppen untersucht (Ennemoser, Kuhl & Pepouna, 2013). Untersucht wurde dabei allerdings ausschließlich eine sehr kleine Stichprobe von Kindern mit Migrationshintergrund, die an hessischen Vorlaufkursen teilnahmen. Fundierte Erkenntnisse zur Wirksamkeit von DL als Kleingruppenförderung im Kindergarten liegen bislang noch nicht vor. Auch fehlen Erkenntnisse über die langfristige Wirkung des Dialogischen Lesens.

2. Auf welche Bereiche der Sprachkompetenz wirkt sich das Dialogische Lesen im Vergleich mit einem formbasierten Training der Phonologischen Bewusstheit aus?

Die vorliegende Untersuchung soll differentielle Effekte einer Kleingruppenförderung mit Dialogischem Lesen bei Risikokindern untersuchen. Es wird erwartet, dass Risikokinder ihre sprachlichen Kompetenzen durch DL verbessern. Dialogisches Lesen soll Sprachkompetenzbereiche breitgefächert fördern. Es sind Effekte im Wortschatz, im Sprachverständnis und in der Grammatik zu erwarten.

Kindern die an einem Vergleichstraining zur Förderung der Phonologischen Bewusstheit teilnehmen, sollten sich in diesem Bereich am stärksten verbessern. Allerdings gibt es auch Hinweise, dass sich Dialogisches Lesen möglicherweise auch förderlich auf die Entwicklung der Phonologischen Bewusstheit auswirkt (Aram, 2006; Mol et al., 2009; Swanson et al., 2011). Verglichen mit zu erwartenden mittleren bis großen Effekten des formbasierten PB-Training in der Phonologischen Bewusstheit sollten mögliche Effekte des Dialogischen Lesens allerdings deutlich geringer ausfallen. Für das PB-Training werden ausschließlich Effekte in der Phonologischen Bewusstheit und nicht in anderen Sprachkompetenzbereichen erwartet.

3. Hat die Durchführungsqualität Einfluss auf die Wirksamkeit des Dialogischen Lesens?

Die dritte Fragestellung beschäftigt sich mit dem Einfluss der Durchführungsqualität auf die Wirksamkeit des Dialogischen Lesens. Einige Studien, in denen Eltern ihren Kinder dialogisch vorlasen, haben gezeigt, dass Kinder mehr von der Methode profitierten, wenn Eltern Modellierungstechniken häufiger beachten und einsetzen (Blewitt et al., 2009). Die Untersuchung soll mit Hilfe eines Beobachtungsbogens die Durchführungsqualität von geschulten Trainer kontrollieren. Es wird erwartet, dass die Durchführungsqualität einen bedeutsamen Einfluss auf die Effektivität von DL hat. Kinder aus Gruppen, in denen eine hohe Durchführungsqualität festgestellt wurde, sollten stärker von der Förderung profitieren, als Kinder aus Gruppen in denen die Durchführungsqualität weniger hoch eingeschätzt war. Die formulierten Hypothesen können wie folgt zusammen gefasst werden:

(1) Spezifität

Dialogisches Lesen

- a. DL verbessert die allgemeinen Sprachentwicklung.
- b. DL verbessert die Leistungen im produktiven Wortschatz.
- c. DL verbessert die Leistungen im Sprachverständnis.
- d. DL verbessert die Leistungen in der Grammatik.
- e. DL verbessert die Leistungen in der Phonologischen Bewusstheit, aber in geringerem Maße als das spezifische Training der Phonologischen Bewusstheit.

Phonologische Bewusstheit

- a. Ein Training der Phonologischen Bewusstheit verbessert die PB-Leistungen.
- b. Es gibt keine Effekte auf die eingesetzten Maße der Sprachentwicklung.

(2) Stabilität

Erwartete Fördereffekte bleiben in einer Follow Up-Erhebung ein halbes Jahr nach Abschluss der Förderung bestehen.

(3) Durchführungsqualität

Die Qualität, mit der geschulte Trainer das Dialogische Lesen umsetzen, hat Einfluss auf die Wirksamkeit der Fördermethode.

6 Methode

6.1 Stichprobe und Verlauf der Untersuchung

Die Erhebung gliederte sich in vier Messzeitpunkte (MZP). MZP 1 umfasste eine Screeningenerhebung, die von Oktober bis Dezember 2009 in 63 mittelhessischen Kindertagesstätten stattfand. 54% der Kitas befanden sich in einem ländlich geprägten Einzugsgebiet, 46% in städtischen Gebieten.

Vor der Untersuchung wurden von allen Eltern, deren Kinder das letzte Kindergartenjahr vor der Einschulung besuchten, das Einverständnis zur Teilnahme erbeten ($N = 714$). Am Ende der Rekrutierungsphase im Oktober 2009 lagen Einverständniserklärungen für 660 Kinder vor. Dies entspricht einem Zusage-Anteil von 92,4%.

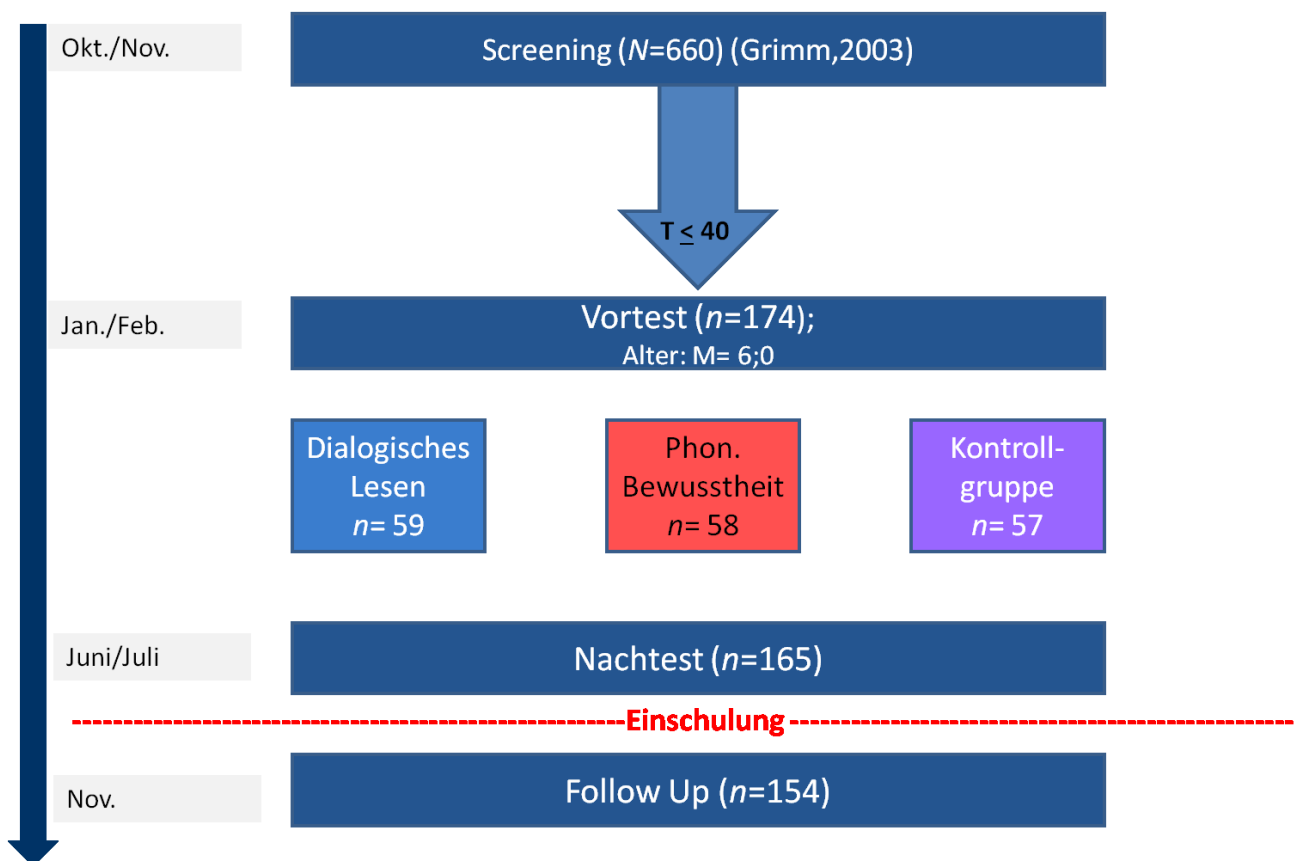


Abbildung 4 Studiendesign und Untersuchungsverlauf

Die Ausgangsstichprobe setzte sich aus 358 Jungen und 302 Mädchen zusammen. Ihr Altersdurchschnitt lag bei 5;9 Jahren (in Monaten: $M = 69.33$; $SD = 4.18$; $MD = 69$).

509 Kinder wurden als deutsche Muttersprachler klassifiziert. 151 Probanden sprachen eine andere Muttersprache.

Zur Identifikation von Risikokindern mit Auffälligkeiten in der Sprachentwicklung wurden die Sprachkompetenzen der Kinder mit dem Sprachscreening für das Vorschulalter (SSV; Grimm, Aktas & Kießig, 2003; vgl. Kapitel 6.2) erfasst. Außerdem schätzte in jeder Kita eine Erzieherin den Sprachentwicklungsstand der teilnehmenden Kinder mit Hilfe eines Beobachtungsbogens ein (Frey, Duhm, Althaus, Heinz & Mengelkamp, 2008). Das Erzieherurteil diente als wichtiges Außenkriterium, um die externe Validität der Screeningergebnisse zu überprüfen.

Nach Abschluss des Screenings erfolgte im Dezember 2009 die Identifikation der Kinder mit sprachlichen Auffälligkeiten. Zur Bestimmung der Risikokinder wurden zunächst die erzielten Rohpunkte in den Subtests des Screenings auf Grundlage der Testnormen in T-Werte umgerechnet. Als kritischer Wert wurde ein $T \leq 40$ festgelegt. Da per Definition ein T-Wert zwischen 60 und 40 den Normalbereich abbildet, gehören Kinder, die diesen Wert unterschreiten, zu den unteren 15% ihrer Altersgruppe. Als Risikokind wurde klassifiziert, wer den kritischen Wert in mindestens einem Subtests unterschritten hatte. Nach Meinung der Testautoren sind diese Kinder in ihrer Sprachentwicklung besonders zu beobachten. Auf diesem Wege wurden 175 Kinder (26.5% der Gesamtstichprobe) als Risikokinder klassifiziert. Sie wurden für die weitere Untersuchung ausgelesen. Ein Kind musste post hoc von der Untersuchung ausgeschlossen werden, da es sich um ein Integrationskind mit geistiger Behinderung handelte. Somit ergab sich für die Prätestphase MZP 2 eine ausgelesene Teilstichprobe von 174 Risikokindern aus 105 Jungen und 69 Mädchen.

Zum Zeitpunkt des Prätests waren die Kinder im Durchschnitt 6;0 Jahre alt ($M = 71.91$; $SD = 4.44$; $MD = 71$). Von der Ausgangsstichprobe verblieben 95 Kinder mit Deutsch als Muttersprache und 79 Kinder mit Deutsch als Zweitsprache in der Untersuchungsstichprobe. Über die Muttersprache der Kinder gaben die Erzieher im eingesetzten Beobachtungsbogen Auskunft. Die ausgelesenen Risikokinder nahmen im Januar und Februar 2010 an einer ausführlichen Prätest-Erhebung (MZP 2) teil.

Zum MZP 2 konnten die Sprachleistungen umfänglicher als zuvor erhoben werden. Erfasst wurden Kompetenzen im Wortschatz, im Sprachverständnis und in der Grammatik. Darüber hinaus wurden die Leistungen in der Phonologischen Bewusstheit erhoben. Da die Intelligenz einen gewichtigen Faktor für die kognitive Entwicklung und somit auch für die Entwicklung von Sprache darstellt, wurde auch dieses Konstrukt erfasst.

Nach Abschluss von MZP 2 nahmen die Kinder an einer dreimonatigen Interventionsphase (März bis Mai 2010) in Kleingruppen teil. Die Kitas wurden zuvor zufällig auf die

Interventionsmaßnahmen oder die Kontrollbedingung verteilt. Die Untersuchung kann somit als experimentelle Studie gelten.

Insgesamt wurden für die Studie drei Untersuchungsgruppen festgelegt: (1) Kleingruppentraining mit Dialogischem Lesen (DL; $n = 59$), (2) Kleingruppentraining der Phonologischen Bewusstheit (PB; $n = 58$), (3) Kontrollgruppe (KG; $n = 57$). Eine ausführliche Beschreibung des Aufbaus und Ablaufs der Kleingruppentrainings folgt in Kap. 6.3.

Direkt nach Beendigung der dreimonatigen Trainingsphase erfolgte im Juni und Juli 2010 eine Posttestung der Untersuchungsstichprobe (MZP 3). Zu MZP 3 kamen dieselben Tests wie zu MZP 2 zum Einsatz – mit Ausnahme des Intelligenztests. Da die Intelligenz ausschließlich als Kontrollvariable diente, und von einer hohen Persistenz der Leistung auszugehen war, wurde sie nur zum MZP 2 erhoben.

Bedingt durch Umzüge konnten die Daten von neun Kindern zum MZP 3 nicht mehr erhoben werden. Grundlage der durchgeführten Prä-Post-Analysen bilden also die Daten von 165 Kindern. Dies entspricht einem Drop-Out von 0.5%.

Vier Monate nach Abschluss des Posttests folgte von November 2010 eine Follow Up-Erhebung zur Feststellung etwaiger Langzeiteffekte (MZP 4). Zu diesem Zeitpunkt befanden sich 121 Probanden der Stichprobe (78.6%) im 1. Halbjahr der ersten Klasse. 33 Probanden (21.4%) erfuhren keine reguläre Einschulung und besuchten weiterhin den Kindergarten oder eine Vorklasse. Die Stichprobe verteilte sich regional weit ausgedehnt auf 51 Schulen und Kindergärten, sodass es nicht vollends gelang alle Probanden noch vor den Weihnachtsferien zu testen. Die Erhebungen waren Anfang Februar 2011 abgeschlossen. Bedingt durch Umzüge und Absagen seitens der Eltern verringerte sich die Stichprobe auf 154 Probanden. Dies entspricht einem Drop-Out von 6.7%.

Im Follow Up wurden dieselben Instrumente wie zuvor zu MZP 2 und MZP 3 eingesetzt. Bei der Auswertung der Grammatikleistung zeigte sich ein Testleitereffekt, der auf einen systematischen Auswertungsfehler in der *Singular-Plural-Bildung* zurückging. Die betroffenen Fälle mussten von der Follow Up-Analyse für den Bereich *Grammatik* ausgeschlossen werden. Betroffen waren 17 Fälle (29.3%) in der PB-Gruppe und 18 Fälle (28.1%) in der Kontrollgruppe. In der DL-Gruppe waren keine Fälle betroffen. In den anderen Bereichen der Sprachentwicklung gab es keine Testleitereffekte.

6.2 Instrumente

Das folgende Kapitel bietet einen Überblick über die eingesetzten Messinstrumente und nennt theoriegeleitet Gründe für deren Auswahl.

Tabelle 4 Erfasste Konstrukte in der Interventionsstudie.

Konstrukt	Messzeitpunkt	Testaufgaben	Testverfahren
Auditives Gedächtnis	Screening	Phonologisches Arbeitsgedächtnis für Nichtwörter (PGN) Satzgedächtnis (SG)	Sprachscreening für das Vorschulalter (SSV; Grimm et al., 2003)
Erzieherinnenurteil	Screening	Sprachentwicklungsskala	Der Beobachtungsbogen für 3-6-jährige Kinder (BBK; Frey et al., 2008)
Intelligenz (IQ)	Prä	Matrizen-Test (MAT) Klassifikation (KLAB)	Hannover-Wechsler-Intelligenztest für Kinder im Vorschulalter – III (HAWIVA – III; Ricken et al., 2007)
Wortschatz (WS)	Prä, Post, Follow Up	Bildbenennung (BB) Kategoriebildung (KB)	Sprachentwicklungstest für Fünf- bis Zehnjährige (SET 5-10; Petermann, Fröhlich & Metz, 2010)
Sprachverständnis (SV)	Prä, Post, Follow Up	Handlungssequenzen (HS) Fragen zum Text (FT)	SET 5-10
Grammatik (GRA)	Prä, Post, Follow Up	Satzbildung (SB) Singular-Plural Bildung (SPB)	SET 5-10
		Morphologie Artikelbildung Syntax	Erweiterte Auswertung des SET-Subtests <i>Satzbildung</i>
Phon. Bewusstheit (PB)	Prä, Post, Follow Up	Reimen Phonemsynthese I (PS) Phonemsynthese II (SYN) Anlautidentifikation (ANID)	Aus der Würzburger Arbeitsgruppe von Wolfgang Schneider (z.B. Ennemoser, Marx, Weber & Schneider, 2012)
Durchführungsqualität des Dialogischen Lesens	Nach Abschluss der Interventionsphase	Beobachtungsbogen für externe Rater	eigenes Instrument

6.2.1 Auditives Gedächtnis

Die auditive Gedächtnisleistung ist eng mit der Entwicklung von Sprachkompetenzen assoziiert (Ennemoser, Pepouna & Hartung, 2012; Hasselhorn, Seidler-Brandler & Körner, 2000). Aufgaben, bei denen sich Kinder Pseudowörter und Sätze merken und nachsprechen müssen, werden häufig zur Auslese von sprachauffälligen Risikokindern eingesetzt (Roos & Schöler, 2007). Um Risikokinder für die eigene Untersuchung zu klassifizieren, wurde das *Screening für das Vorschulalter* (SSV; Grimm et al., 2003) eingesetzt. Bei Kindern von 4;0 bis 5;11 Jahren wird die Leistung im Nachsprechen von Pseudowörtern (PGN) und im Nachsprechen von Sätzen (SG) erhoben.

6.2.1.1 Phonologisches Arbeitsgedächtnis für Nichtwörter (PGN)

Stärker als klassische Arbeitsgedächtnisaufgaben wie Zahl- oder Wortspannenmaße legen Pseudowortaufgaben den Fokus auf die Erfassung der phonologischen Repräsentationsfähigkeit (Grimm et al., 2003). Sie gelten daher als sensitiveres Maß zur Messung von Verarbeitungskapazität und -genauigkeit des phonologischen Arbeitsgedächtnisses (Ennemoser et. al., 2012; Kany & Schöler, 2010). Es liegen Normen für Vorschulkinder bis 5;11 Jahre vor.

Der Altersdurchschnitt der Gesamtstichprobe für die eigene Untersuchung lag zum Zeitpunkt des Screenings bei 5;9 Jahren. Da Screeningaufgaben zur Auswahl von Risikokindern vor allem in den unteren Leistungsbereichen gut differenzieren sollten, wurde die Aufgabe als geeignet eingeschätzt.

Die Schwierigkeit der einzelnen Items ist eng mit der Silbenzahl der verwendeten Pseudowörter verknüpft. Zum Einsatz kommen zwei- bis fünfsilbige Kunstwörter (z.B. *Maluk Skatagurp* oder *Vominlapertust*).

Die einschlägigen Testgütekriterien gelten PGN als erfüllt. Für die PGN-Aufgaben lagen standardisierte Instruktionen und einheitliche Auswertungskriterien vor. Sie stellen hohe Objektivität sicher (Lienert & Raatz, 1998). Für die Altersgruppe von 5;0 bis 5;11 berichten die Autoren eine akzeptable interne Konsistenz (Cronbachs $\alpha = .74$). In puncto Validität konnte eine hohe Korrelation des PGN-Screeningscores mit dem Gesamtergebnis im SETK 3-5 ($r = .75$) festgestellt werden.

6.2.1.2 Satzgedächtnis (SG)

Das Nachsprechen von Sätzen überprüft, ob es Kindern gelingt Sätze korrekt zu reproduzieren, die in unterschiedlicher semantischer und syntaktischer Schwierigkeit dargeboten werden (Grimm, 2003). Dieses Aufgabenformat überprüft, wie gut ein Kind bereits erworbene grammatikalische Strukturen für die geforderte Nachsprechleistung zu nutzen weiß.

Nach Roos und Schöler (2007) ist das Nachsprechen von Sätzen ein Verfahren, das als „[...] Königsweg in der Diagnostik und Differentialdiagnostik von Sprachentwicklungsauffälligkeiten und -störungen gelten kann.“ (S. 541 f.). Es habe sich in zahlreichen Untersuchungen als diskriminativ valide bei Sprach- und Schriftsprachproblemen erwiesen (ebd.).

Das Format der Items lässt sich in zwei „Schwierigkeitsstufen“ gliedern. In einem ersten Teil der Aufgabe sollen die Kinder sechs semantisch korrekte Sätze nachsprechen (z.B. *Lena lacht, nachdem sie gekitzelt wurde.*). In einem zweiten Teil werden neun syntaxkorrekte Sätze ohne semantischen Zusammenhang vorgesprochen (z.B. *Die klatschenden Stühle singen einen Becher.*) Die Schwierigkeit der Items wird von der Satzlänge und der syntaktischen Komplexität determiniert. Auch für den Subtest SG berichten Grimm et al. (2003) ausreichende Reliabilitäts- und Validitätsmaße. Die interne Konsistenz ergab Cronbachs $\alpha = .89$. Die Korrelation zwischen dem SG-Score und dem Gesamtergebnis im SET-K liegt ebenfalls im hohen Bereich ($r = .80$).

Um die prognostische Validität des Screenings zu überprüfen, untersuchten die Autoren den prädiktiven Wert beider Subtests für die Leistungen im SET-K 3-5 an einer Stichprobe von 162 Fünfjährigen (Grimm et al., 2003). PGN und SG konnten gemeinsam einen Varianzanteil von 78% für die Vorhersage der Gesamtleistung im SETK 3-5 aufklären.

Weitere positive Befunde, die die Validität der Screeningergebnisse stützen, liefert eine Korrelationsanalyse zwischen Screening-Score und Erzieherurteil, die auf Grundlage der vorliegenden Stichprobe berechnet wurde (siehe hierzu: Kap. 7.1).

6.2.1.3 Durchführung

Beide Screening-Aufgaben wurden vor Ort in den Kindergärten durchgeführt. Die Testleiter waren geschulte studentische Hilfskräfte aus den Studiengängen Psychologie und Lehramt an Förderschulen. Die Kinder nahmen an einer Einzelsitzung von ca. 20 Minuten teil. Die Antworten der Kinder wurden protokolliert und mit Hilfe eines digitalen Aufnahmegerätes mitgeschnitten.

6.2.2 Erzieherinnenurteil

Zusätzlich zum Screening wurden die Erzieherinnen aller Kinder der Ausgangsstichprobe gebeten, den Sprachentwicklungsstand mit einem Beobachtungsbogens zu beurteilen. Eingesetzt wurden hierfür die Items *Sprachentwicklung* (SE) aus dem Beobachtungsbogen für Drei- bis Sechsjährige Kinder (BBK 3-6, Frey et al., 2008).

SE setzt sich aus zehn Beobachtungseitems zusammen (Bsp.: „*Es [das Kind] erzählt ein Erlebnis/eine Geschichte nahezu wörtlich nach.*“). Dabei schätzen die Erzieher auf einer fünfstufigen Skala ein, wie häufig ein Kind dieses Sprachverhalten zeigt (1 = „selten (nie)“ bis 5 = „sehr oft (immer)“). Die Autoren berichten für die SE-Items hohe interne Konsistenz (Cronbachs $\alpha = .96$). Außerdem konnten Frey, Mengelkamp und Riebel (2010) im Rahmen einer faktorenanalytischen Bewertung zeigen, dass alle Items der SE-Skala einheitlich auf dem Faktor 1 luden. Somit ist ausreichende Konstruktvalidität anzunehmen.

Im Folgenden sollen die Aufgaben zur Erhebung der Sprachentwicklung, der Phonologischen Bewusstheit und der Intelligenz beschrieben werden. Sie wurden vor und nach der Interventionsphase sowie im Follow Up ausschließlich mit den ausgelesenen Risikokindern durchgeführt.

6.2.3 Sprachentwicklung

6.2.3.1 Wortschatz (WS)

6.2.3.1.1 Bildbenennung (BB)

Die Bildbenennungsaufgabe wurde dem SET 5-10 (Petermann et al., 2010) entnommen. BB prüft den aktiven Wortschatz der Probanden mit 40 Bildkarten. Die Kinder sollen dabei sowohl Gegenstände als auch Handlungen benennen. Jede korrekt benannte Bildkarte wird mit 1 Rohpunkt bewertet. Die Autoren berichten eine gute interne Konsistenz (Cronbachs $\alpha = .91$).

6.2.3.1.2 Kategoriebildung (KB)

KB entstammt ebenfalls dem SET 5-10 (Petermann et al., 2010). Sie überprüft die vertikale Struktur des Wortschatzes durch das Nennen von Oberbegriffen. Auf den eingesetzten Bildkarten sind stets vier Gegenstände abgebildet, für die die Kinder einen Oberbegriff nennen sollen (z.B. Tiere, Pflanzen, Möbel).

Bei der Auswertung werden für eine korrekte Antwort zwei Rohpunkte vergeben. Antworten, die nicht dem Zielwort entsprechen, aber den Oberbegriff implizieren (z.B. statt *Himmelskörper: ist am Himmel, sieht man am Himmel*) sind mit einem Rohpunkt zu bewerten. Für KB fällt Conbachs $\alpha = .74$ ebenfalls zufriedenstellend aus.

6.2.3.2 Sprachverständnis (SV)

6.2.3.2.1 Handlungssequenzen (HS)

Bei den HS-Items (Petermann et al. 2010) sollen die Probanden verbale Instruktionen der Testleiter mit Hilfe von Spielfiguren umsetzen. Aufgaben dieser Art finden sich klassischerweise in einer Vielzahl von Sprachentwicklungstests (z.B. *Verstehen grammatikalischer Strukturformen* im HSET; Grimm & Schöler, 1991). Dieses Aufgabenformat kann überprüfen, inwieweit Kinder grammatisches Regelwissen erworben haben, das zum Verständnis der Sätze von Nöten ist (Beispiel: *Der Junge springt über die Bank, weil der Hund ihn jagt.*). Die vorgegebenen Anweisungen variieren hinsichtlich Länge und sprachlicher Komplexität (z.B. Einsatz verschiedener Konjunktionen und/oder unterschiedlichen Haupt- und Nebensatzstrukturen). Die Autoren berichten befriedigende interne Konsistenz (Cronbachs $\alpha = .77$).

6.2.3.2.2 Fragen zum Text (FT)

Der Subtest *Fragen zum Text* überprüft die Leistungen im Textverständnis. Die Kinder werden aufgefordert Fragen zu kurzen Vorlesetexten zu beantworten. Dabei gibt der Testleiter für jede Frage mehrere Antwortmöglichkeiten vor. Die Kinder sollen die Antwort auswählen, die am besten zu den vorgelesenen Inhalten passt. Es lag laut Testmanual ebenfalls eine akzeptable Homogenität vor (Cronbachs $\alpha = .62$).

6.2.3.3 Grammatik

6.2.3.3.1 Singular-Plural-Bildung (SPB)

Die Aufgabe zur SPB überprüft, ob das Kind die Unterscheidung des Numerus regelhaft beherrscht. Um von der Reproduktionsleistung gelernter Wörter in Pluralform zu unterscheiden, bestehen die 18 Items aus Real- und aus Kunstwörtern. Die Kinder sollen nach Vorgabe des Singulars die Pluralform des Zielworts nennen (z. B. *Das ist eine Katze. Das sind viele...? / Das ist eine Giraffe. Das sind viele...?*). Für einzelne Kunstwörter gibt es mehrere Antwortmöglichkeiten, die laut Manual als korrekt gelten (z.B. *Das ist ein Rolz*.

Das sind viele Rölzer/Rolze/Rolzen). Die berichtete interne Konsistenz kann als gut bewertet werden (Cronbachs $\alpha = .84$).

6.2.3.3.2 Satzbildung (SB)

Bei SB soll das Testkind aus vorgegebenen Wörtern einen grammatikalisch und semantisch korrekten Satz bilden. Die dargebotenen Items enthalten entweder zwei oder drei Wörter, die es regelhaft miteinander zu verknüpfen gilt (z.B. **gelb** – **Banane**: *Die Banane ist gelb*; **Hund** – **bellen** – **Katze**: *Der Hund bellt die Katze an*). Für jeden korrekten Satz vergibt der Testleiter einen Rohpunkt. Die Aufgabe besteht aus zwölf Items. Die Autoren berichten eine hohe Homogenität der Items (Cronbachs $\alpha = .91$).

6.2.3.3.3 Erweiterte Auswertung der Satzbildungsaufgabe

Da mit den SET-Subtests *Satzbildung* und *Singular-Pluralbildung* mögliche Entwicklungen in den Bereichen Morphologie, Syntax und Artikelbildung möglicherweise nur sehr undifferenziert erfasst werden, wurde für die Satzbildungsaufgabe eine erweiternde Auswertung entwickelt. Da alle Antworten mit Hilfe von Aufnahmegeräten dokumentiert wurden, war es möglich die Sätze der Kinder post hoc genau zu analysieren. Insgesamt lagen Aufnahmen von 156 Kindern der Untersuchungsstichprobe vor. Geschulte Rater bewerteten jeden produzierten Satz in Bezug auf morphologische und syntaktische Fehler sowie Fehler in der Artikelbildung.

Zum Beispiel erhielt ein Kind für den gebildeten Satz: „*Der Hund bellen die Katze an*“ keinen Rohpunkte im Bereich *Morphologie*, weil *bellen* nicht korrekt konjugiert ist. Da es alle Artikel allerdings korrekt verwendete, vergaben die Rater einen Rohpunkt für die Artikelbildung. Außerdem beachtete das Kind alle Regeln der Satzstellung. Somit konnte dafür ein Rohpunkt in der *Syntax* vergeben werden.

Am Ende der Auswertung wurden die Rohpunkte für jeden Bereich gesondert addiert. So konnten für jedes Kind die Leistungen in der Morphologie, der Syntax und der Artikelbildung bestimmt werden. Für alle drei Grammatikbereiche ließ sich eine hohe interne Konsistenz feststellen (Cronbachs $\alpha = .88 - .90$).

6.2.4 Phonologische Bewusstheit

Die Aufgaben zur Erfassung der Phonologischen Bewusstheit stammen aus unterschiedlichen Testverfahren. Der Subtest Reimen ist dem *Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten* (BISC, Jansen et. al 2002) entnommen. Außerdem kam

eine Aufgabe zur *Phonemsynthese (ohne Bild)* aus dem *Neuropsychologisches Screening für 5 bis 11-jährige* zur Anwendung (BVN/NPS; Bisiacchi, Kaufmann et al. 2008). Zwei weitere Aufgaben (*Anlautidentifikation und Phonemsynthese (mit Bild)*) stammen aus der Würzburger Arbeitsgruppe von Wolfgang Schneider; sie haben sich in zahlreichen Untersuchungen zur Schriftsprachentwicklung bewährt (z.B. Ennemoser, Marx, Weber & Schneider, 2012).

6.2.4.1 Reime

Im Subtest Reimen sollen die Kinder anhand von zwei vorgegebenen Wörtern (Real - und Kunstwörter) schlicht entscheiden, ob ein Reim vorliegt oder nicht (z.B. Alibamm – Läusekamm; Miste – Schinde). Insgesamt bearbeiten die Probanden zehn Items.

Für die Altersgruppe „Zehn Monate vor der Einschulung“ ist eine ausreichende Homogenität der Aufgaben nachgewiesen (Cronbachs $\alpha = .70$). Die Autoren berichten allerdings eine mittlere Itemschwierigkeit von $p = .83$, die als äußerst gering einzustufen ist. Viele Kinder können demnach die vorgegebenen Items richtig lösen.

Um etwaige Deckeneffekte im Nachtest zu vermeiden, wurde die Reimaufgabe für den weiteren Untersuchungsverlauf um vier selbstentwickelte Items erweitert. Hinzu kamen zwei Items mit semantischem Bezug (Wand – Mauer; Fisch – Wasser), ein Item mit lautlicher Ähnlichkeit (Fleisch – Schleif) und ein Pseudowort-Item (Kaliding – Riesenspring). Die Aufgabenkonsistenz blieb dadurch nahezu unverändert (Cronbachs $\alpha = .66$).

6.2.4.2 Phonemsynthese (ohne Bild)

Im Subtest *Phonemsynthese ohne Bild* (Kaufmann et al., 2009) spricht der Testleiter den Kindern einzelne Phoneme eines Wortes vor (z. B. N-A-S-E oder E-L-E-F-A-N-T). Das gesuchte Zielwort soll nun ohne bildliche Hilfestellung genannt werden. Insgesamt besteht der Subtests aus zehn Items, wobei die Phonemanzahl pro Wort im Verlauf der Aufgabe von vier auf sieben ansteigt. Auch hier wurde die interne Konsistenz der Aufgabe auf Grundlage der eigenen Stichprobendaten überprüft. Cronbachs α lag mit .82 in einem hohen Bereich.

6.2.4.3 Phonemsynthese (mit Bild)

Bei dieser Aufgabe legt der Testleiter den Kindern Bildkarten mit je vier Abbildungen (z.B. Eimer – Eis – Glas – Rakete) vor und spricht im Anschluss die Laute eines vorgegebenen Zielwortes segmentiert vor (z.B. Ei - S). Das Kind erhält die Aufgabe, das segmentiert vorgesprochene Wort auf der Bildkarte zu zeigen. Die Aufgabe umfasst acht Items. Da keine statistischen Kennwerte vorlagen, wurde die Homogenität auf Grundlage der

Vortestergebnisse berechnet. Die Berechnungen ergaben eine ausreichende interne Konsistenz (Cronbachs $\alpha = .61$).

6.2.4.4 Anlautidentifikation

Bei der Anlautidentifikation präsentiert der Testleiter dem Kind eine Karte mit vier Abbildungen (z. B. Schnecke – Koffer – Drachen – Kamel). Nachdem der Testleiter alle vier Wörter einmal vorgesprochen hat, soll das Kind entscheiden, welche der genannten Wörter mit dem gleichen Laut beginnen („Zwei dieser Wörter fangen mit dem gleichen Laut an. Versuche mal herauszufinden, welche Wörter das sind...“). Insgesamt werden zehn Items abgefragt. Da auch hier keine statistischen Kennwerte vorlagen, wurde die interne Konsistenz ebenfalls an Vortestdaten berechnet. Die Prüfung der internen Konsistenz ergab mit Cronbachs $\alpha = .70$ ein ausreichendes Ergebnis.

6.2.5 Intelligenz

Die Messung der Intelligenz erfolgte über den *Matrizen-Test* und den Subtest *Klassen bilden* aus dem Handlungsteil des *Hannover-Wechsler-Intelligenztest für das Vorschulalter III* (HAWIVA-III; Ricken & Wechsler, 2007). Beide Tests prüfen die fluide Intelligenz, die nach Carroll (1993) die Verarbeitung von Abstraktionen, Regeln, Verallgemeinerungen und logischen Beziehungen leistet.

6.2.5.1 Matrizentest

Der Matrizentest erfasst in 17 Items Fähigkeiten des problemlösenden und schlussfolgernden Denkens, die als wichtige Indikatoren der fluiden Intelligenz beschrieben werden (Ricken & Wechsler, 2007). Die Kinder erhalten die Aufgabe ein dargebotenes Muster zu vervollständigen. Die richtige Lösung soll aus vier oder fünf Antwortmöglichkeiten ausgewählt werden. Für den Zusammenhang zwischen MT und dem Gesamtergebnis des HAWIVA-III berichten die Autoren einen mittleren Korrelationskoeffizienten von $r_{tc} = .87$. Die Validität als Korrelation zum Gesamttest kann damit als ausreichend betrachtet werden.

6.2.5.2 Klassen bilden

Klassen bilden erfasst die Fähigkeit des abstrakten und kategorialen Denkens. Insgesamt besteht der Subtest aus 17 Items. Jedes Item besteht aus zwei Bilderreihen. Aus jeder Reihe soll das Kind ein Bild so auswählen, dass sich beide Bilder zu einer Kategorie bzw. Klasse zusammenfassen lassen. Die Autoren berichten für das *Klassen bilden* einen mittleren

Korrelationskoeffizienten von $r_{tc} = .82$ mit der Gesamtskala des HAWIVA-III. Die interne Konsistenz ist hier zufriedenstellend. Im Gegensatz zu den zuvor beschriebenen Konstrukten wurde die Intelligenzleistung nur im Prätest gemessen, da keine Veränderungen über die Zeit zu erwarten waren.

6.2.6 Durchführungsqualität

Um die Durchführungsqualität des DL-Trainings zu überprüfen, wurden alle Trainer der DL-Gruppe gebeten in ihren Trainingsgruppen drei Audioaufnahmen zu unterschiedlichen Zeitpunkten der Intervention zu erstellen. Im ersten, zweiten und dritten Monat der Interventionsphase wurde jeweils eine Sitzung mitgeschnitten. Die Trainer benutzen hierfür ein digitales Aufnahmegerät. Zum Ende der Interventionsphase lagen aus allen Kleingruppen 61 Audioaufnahmen vor.

Die Audioaufnahmen dienten als Grundlage zur Beurteilung der Durchführungsqualität. Beurteilt wurde die Durchführungsqualität nach Abschluss der Förderung von zwei externen Ratern, die mit den Prinzipien des Dialogischen Lesens vertraut waren. Dabei handelte es sich um studentische Hilfskräfte mit den Studienschwerpunkten Psychologie und Lehramt an Grundschulen. Sie nahmen im Vorfeld an einer umfassenden Schulung teil. Im Rahmen der Schulung erhielten die Rater zunächst eine kurze theoretische Einführung in Umsetzung des Dialogischen Lesens. Um geschulte Beobachtungsleistungen sicherzustellen, folgten in den weiteren Einheiten praktische Beurteilungsübungen. Geübt wurde das Beurteilen der Trainings anhand kurzer Videosequenzen, die nicht zum Pool der aufgenommenen Trainingssitzungen gehörten.

Der Beurteilungsbogen erfasste sieben Verhaltenskategorien, die bei einer Förderung durch Dialogisches Lesen zum Einsatz kommen sollten (Tabelle 5).

Tabelle 5 Struktur des Beobachtungsbogens zur Erhebung der Durchführungsqualität im Dialogischen Lesen.

Verhaltenskategorien	
1	Geschlossene Fragen
2	Offene Fragen
3	Erweiterung der Aussage
4	Korrektive Wiederholung
5	Nachfragen
6	Sätze vervollständigen
7	Alltagsbezug

Die Verhaltenskategorien wurden mit zwei verschiedenen Kennwerten erfasst:

1) Häufigkeitszählungen und 2) Rating auf einer vierstufigen Rating-Skala (1 = „nie“ - 4 = „immer“). Die Stufe „immer“ galt als erfüllt, wenn Trainer die DL-Techniken in unterschiedlichen Gesprächsphasen immer angemessen einsetzten. Als dritten Kennwert sollten die Rater ein Globalurteil in Form einer Gesamtnote (0 - 15 Punkte) für jede Fördersitzung angeben.

Nach Abschluss des Proberatings wurden die Ergebnisse der Beurteilung gemeinsam mit den beiden Ratern analysiert und etwaige Unterschiede bei der Beurteilung der Probeaufnahmen diskutiert. Im Anschluss an die Schulung beurteilte jeder Rater alle Fördersitzungen einmal hinsichtlich der vorgegebenen Kennwerte, so dass für jede Trainingssitzung zwei Raterurteile vorlagen.

Um die Reliabilität der Raterurteile in Bezug auf die Durchführungsqualität im Dialogischen Lesen zu überprüfen, wurden Intraklassenkorrelationen (ICC) für alle Items des Beobachtungsbogens berechnet. Die ICC schätzt erwartungstreu die Reliabilität der Urteile beliebig vieler Rater. Wird verlangt, dass die von verschiedenen Ratern vergebenen Urteile absolut übereinstimmen, so muss entweder die unjustierte oder die justierte, einfaktorielle ICC berechnet werden (Wirtz, 2004, 2013). Müssen lediglich die relativen Abweichungen vom individuellen Mittelwert der einzelnen Beurteiler ähnlich ausgeprägt sein, weil Mittelwertunterschiede zwischen den Beurteilern die Datenqualität nicht beeinträchtigen, sollte eine justierte ICC bestimmt werden (ebd.). Da die ausgewählten Rater speziell für die Beurteilung der Durchführungsqualität des Dialogischen Lesens ausgebildet waren und beide Rater alle gesammelten Fördersitzungen beurteilten, wurden justierte ICCs berechnet.

Allgemein können ICCs bis $r = .50$ als ausreichend betrachtet werden. Korrelationen unter $r = .50$ gelten als nicht mehr akzeptabel (Wirtz & Casper, 2002).

Da die Rater von jedem Trainer drei bis vier Fördersitzungen beurteilten, wurde für jeden Trainer ein Mittelwert über alle gerateten Sitzungen gebildet. Eine Analyse der Retest-Reliabilitäten ergab durchweg hohe Korrelationen zwischen den bewerteten Sitzungen eines Trainers (alle im Bereich $r = .80$). Die gemittelten Werte waren Grundlage für die Berechnung der ICCs. Tabelle 6 fasst die ICCs für alle Items des Beobachtungsbogens zusammen.

Tabelle 6 Justierte Interklassenkorrelationen der Raterurteile in den Items des Beobachtungsbogens zur Durchführungsqualität.

Beobachtungsitems	Interklassenkorrelation
Zählen	
Geschlossene Fragen	.92
Offene Fragen	.89
Erweiterung der Aussage	.80
Korrektive Wiederholung	.60
Nachfragen	.73
Sätze vervollständigen	.99
Alltagsbezug	.78
Rating	
Geschlossene Fragen	.56
Offene Fragen	.69
Erweiterung der Aussage	.63
Korrektive Wiederholung	.62
Nachfragen	.36
Sätze vervollständigen	.88
Alltagsbezug	.91
Gesamtnote	.86

Insgesamt liegen die ICCs bei fast allen Items in einem guten bis sehr guten Bereich. Offenbar gelangten beide Rater bei der Beurteilung der Durchführungsqualität zu sehr ähnlichen Einschätzungen. Einzig für das Ratingitem „Nachfragen“ war die ermittelte ICC nicht ausreichend reliabel.

Insgesamt zeigte sich, dass die Übereinstimmung der Raterurteile besonders hoch ausfiel, wenn die Häufigkeit der eingesetzten DL-Techniken exakt gezählt werden sollte. Dies spricht nicht zuletzt für eine erfolgreiche Coaching-Phase, in der es gelang beide Rater so auszubilden, dass sie alle DL-Prinzipien eindeutig erfassen konnten.

Die Reliabilitäten für die Ratingitems, bei denen die Qualität der Feedback- und Modellierungstechniken auf einer vierstufigen Skala eingeschätzt werden sollte, fielen deutlich geringer aus. Angesichts der Tatsache, dass Ratingskalen ein deutlich gröberes Beurteilungsmaß darstellen, als exakte Häufigkeitszählungen, ist dieser Unterschied wenig überraschend.

Ebenfalls hohe Übereinstimmungen zwischen beiden Ratern konnten bei der Vergabe einer Gesamtnote im Sinne der Schulbenotung 0 bis 15 Punkte festgestellt werden ($r = .86$).

6.2.6.1 Berechnung der Interkorrelationen

Die Berechnungen der internen Korrelationen aller Beobachtungsitems ergaben erwartungsgemäß bedeutsame Zusammenhänge von Items, denen dasselbe Auswertungsprinzip zu

Grunde lag. Das heißt, Items, bei denen die Rater Verhaltensweisen zählten, korrelierten häufig bedeutsam miteinander. Die Ratingitems korrelierten untereinander deutlich weniger stark als die Zählitems.

Weiter wurden die Korrelationen zwischen Zähl- und Ratingitems betrachtet. Da beide Auswertungsprinzipien denselben Beobachtungsgegenstand (Trainerverhalten) zur Grundlage hatten und somit dasselbe Konstrukt (Durchführungsqualität) erfassten, sollten beide Auswertungsprinzipien auch in einem bedeutsamen korrelativen Zusammenhang zueinander stehen. Diese Annahme konnte bestätigt werden. Der Beobachtungsbogen erfüllt somit die geforderte interne Kriteriumsvalidität. Da für das Dialogische Lesen bislang kein vergleichbarer Beobachtungsbogen zur Erfassung der Durchführungsqualität vorliegt, konnte kein Außenkriterium zur Absicherung der Validität herangezogen werden. Hierzu bedarf es weiterer Forschungsarbeiten.

Die vollständige tabellarische Darstellung der internen Korrelationen zwischen den Items des Beobachtungsbogens ist dem Anhang beigelegt (Tabellen 20 bis 25).

6.2.6.2 Bildung des Qualitätsscores

Tabelle 7 zeigt, dass alle Gesamtwerte in hohem Maße miteinander korrelierten. Es ist anzunehmen, dass alle Werte trotz unterschiedlicher Beurteilungstechniken dasselbe Konstrukt (Durchführungsqualität) erfassen.

Tabelle 7 Korrelationen zwischen den Gesamtwerten der Beurteilungsmaße Zählen, Rating und Benotung.

N= 59	Gesamtwert: Rating	Gesamtwert: Note
Gesamtwert: Zählen	.87**	.92**
Gesamtwert: Rating	--	.90**

Trotzdem ist es möglich, dass mit Rating- und Zählitems sowie der Vergabe einer Gesamtnote unterschiedliche Facetten der Durchführungsqualität erhoben werden.

Während die Zählitems ausschließlich die Häufigkeit der eingesetzten DL-Prinzipien erfassen, könnten über Gesamtnoten weitere förderrelevante Merkmale, wie die motivationalen Fähigkeiten des Trainers, sein genereller Umgang mit den Kindern oder die Struktur der Fördersitzung, implizit – und eventuell auf subjektiven Theorien basierend – mit erfasst worden sein. Da bislang kaum etwas darüber bekannt ist, welches Maß zur Erhebung der Durchführungsqualität als Mittel der Wahl gelten kann, wurden in den hier durchgeführten Analysen alle drei Qualitätsmaße (Zählen + Rating + Benotung) zu einem Gesamtscore „Durchführungsqualität“ verrechnet. Da bei den Einzelscores unterschiedliche

Maßeinheiten zu Grunde lagen, wurden die Werte zunächst z-standardisiert, bevor sie zu einem gemeinsamen Qualitätsscore zusammengefasst werden konnten.

6.3 Trainingsbedingungen

6.3.1 Dialogisches Lesen

6.3.1.1 Material

Zur Förderung nach der Methode des Dialogischen Lesens (DL; vgl. Kap. 3.2.1.1) wurde für die Trainer ein Materialordner mit 127 Bildkarten zu drei Janosch-Geschichten erstellt (Janosch, 2007). In den dargestellten Studien zur Evaluation des Dialogischen Lesens (siehe Kap. 3.2.1.2) war es Eltern und Trainern häufig selbst überlassen, welche Bücher sie zur Förderung verwendeten. Für die hier durchgeführte Studie wurde ein einheitliches Fördermaterial erstellt, um Unterschiede des Materials wie die sprachliche Komplexität der Vorlesetexte und den Anregungsgehalt der Abbildungen zu kontrollieren.

Die Förderung wurde in Kleingruppen mit bis zu fünf Kindern von geschulten Trainern durchgeführt. Die Trainer waren wiederum studentische Hilfskräfte aus den Studiengängen Psychologie und Lehramt an Förderschulen. Alle Trainer nahmen an drei Workshop-Phasen vor dem und während des Interventionszeitraums teil, in denen sie intensiv für den Einsatz der Methode und den Umgang mit dem Material ausgebildet wurden.

Die Reihenfolge der Bildkarten entsprach jeweils der inhaltlichen Abfolge der ausgewählten Bildergeschichten (Janosch, 2007). Bei der Auswahl des Textmaterials wurde darauf geachtet, dass die einzelnen Geschichten wiederkehrende Strukturen (Orte, Figuren, Textinhalte) enthielten. Dieses Kriterium erfüllten in ansprechender Weise die Janosch-Geschichten vom kleinen Bär und kleinen Tiger. Tabelle 8 veranschaulicht den Aufbau des eingesetzten Materialordners.

Tabelle 8 Aufbau des Materialordners.

Bildkarten Nr.	Geschichte	Σ Bildkarten
1-35	Oh, wie schön ist Panama	35
36 - 64	Komm, wir finden einen Schatz	29
65 - 97	Post für den Tiger	33
98 - 127	Ich mach' dich gesund, sagte der Bär	30

Alle Bildkarten zeigten auf der Vorderseite eine farbige Abbildung, zu der sich auf der Rückseite ein entsprechender Vorlesetext fand.

Dem bevorzugten Einsatz von Bildkarten an Stelle von Büchern lagen Überlegungen zum Konzept gemeinsamer Aufmerksamkeit (*joint attention*) zugrunde. Gemeint ist in diesem Zusammenhang der gemeinsame Fokus *auf* und der verbale Austausch *über* ein gemeinsames Referenzobjekt. Auf das sprachförderliche Potential der joint attention weisen verschiedene Untersuchungen hin (Baldwin, 1995; Carpenter, Nagell & Tomasello, 1998; Neuman & Carta, 2011).

Der Vorteil von Bildkarten liegt nicht allein in ihrem flexiblen Einsatz sowohl in Gruppen- als in Einzelsituation. Sie bieten den Trainern auch die Möglichkeit inhaltlich komplexe Sequenzen durch das Nebeneinanderlegen mehrerer Karten zu verdeutlichen. Somit können auch schwerer verständliche Inhalte der Geschichten anschaulich wiederholt und zu gewinnbringenden Sprachanlässen werden.

Da zu jeder Bildkarte der entsprechende Vorlesetext auf der Rückseite abgedruckt ist, bleiben die zu besprechenden Referenzobjekte für alle Kinder visuell präsent. Im Vergleich zu klassischen Büchern ist dies ein überaus bedeutsamer Vorteil. Stets präsente Bilder können den konstanten Erhalt eines gemeinsamen Aufmerksamkeitsfokus erleichtern und bieten sprachschwachen Kindern visuelle Hilfestellungen besonders in eher monodirektional geprägten Vorlesephasen.

Unter den Vorlesetexten fanden sich auf der Rückseite der Bildkarten vorformulierte Fragen, die sich an den in Kapitel 3.2.1 beschriebenen Fragetypen nach Whitehurst (1988) orientieren. Die Trainer konnten sie bei Bedarf einsetzen, um zum Beispiel einen Einstieg in die Fördersitzung zu schaffen oder den Dialog aufrecht zu halten.

Tabelle 9 gibt einen exemplarischen Überblick über einige Fragentypen, die den Trainern am Ende des Vorlesetextes optional zur Verfügung standen.

Tabelle 9 Hilfen im Fördermaterial im Sinne der CROWD und PEER-Prinzipien nach Whitehurst et al. (1988).

Typus	Beispiel
Sätze vervollständigen	<i>Nachdem der Bär die Kiste gefunden hat _____</i> <i>Der kleine Bär hat Angst, dass _____</i> <i>Der Tiger zündet ein Feuer an, weil _____</i>
Inhalte erinnern	<i>Warum möchte die Krähe, dass die beiden auf einen Baum klettern ?</i> <i>Erinnerst Du dich noch, wen sie alles auf ihrem Weg getroffen haben?</i> <i>Warum brauchten der Bär und der Tiger das Floß?</i>
Offene Anregungen	<i>Was könnte in der Flaschenpost stehen?</i> <i>Wen treffen der Bär und der Tiger wohl als nächstes?</i> <i>Könnt ihr mir erzählen, was hier auf dem Bild passiert?</i>
W-Fragen	<i>Woraus bauen sich die beiden einen Wegweiser?</i> <i>Wohin laufen die beiden?</i> <i>Womit graben der Bär und der Tiger das Loch?</i>
Alltagsbezug	<i>Hast Du schon mal einen Brief verschickt? Was hast Du geschrieben?</i> <i>Bist Du schon mal mit einem Boot gefahren? Wie war das?</i> <i>Hast Du auch schon mal einen Strand gesehen? Was gab es da alles?</i>

Wie schon erwähnt, fand vor Beginn der Trainingsphase eine intensive Schulung statt, um sicherzustellen, dass alle Trainer gleichermaßen mit den sprachförderlichen Techniken vertraut waren. Der nächste Abschnitt beschäftigt sich daher mit dem Aufbau und den Inhalten der Trainerschulung.

6.3.1.2 Schulung der Trainer im Dialogischen Lesen

Alle Trainer nahmen vor Beginn der Intervention an einer eintägigen Schulung teil, in der sie die Methode kennen lernten. Die Schulung gliederte sich in vier Abschnitte:

1) Theoretische Einführung

Zu Beginn des Seminars erhielten die Trainer eine theoretische Einführung in die Methode. Weiter bekamen alle Trainer ein kurzes Manual, das die Prinzipien des Dialogischen Lesens erläuterte und wichtige Hinweise zur Durchführung des Trainings zusammenfasste.

2) Präsentation der Fördermaterialien

Nach Abschluss der theoretischen Einführung folgte die Präsentation des Materialordners (vgl. Tabelle 8). Die Workshop-Leiter informierten die Trainer über die Zusammenstellung der Bildergeschichten und den Aufbau des Ordners. Nachdem die Verwendungsprinzipien der Bildkarten erläutert waren, folgten anwendungsbezogene Übungen, die im Folgenden näher beschrieben werden.

3) Übung zur Anwendung der Dialogischen Prinzipien im Rollenspiel

Zentralen Bestandteil der Schulung stellten praktische Übungen dar, in denen die angehenden Trainer sprachförderliches Verhalten im Sinne des Dialogischen Lesens einübten. Hierfür wurden unterschiedliche Rollenspiel-Settings initiiert. Die Rollenspiele waren so konzipiert, dass jeder Trainer die DL-Interaktionen mehrmals in einer gespielten Trainingssitzung vor einer Kleingruppe von Pseudo-Probanden (gespielt durch andere Trainer des Workshops) einsetzen musste. Während des Rollenspiels bekamen die übrigen Teilnehmer die Aufgabe ihre Beobachtungen zu protokollieren und die Qualität der Umsetzung zu beurteilen.

Nach dem Rollenspiel fand eine Feedbackrunde statt, in der die Trainer Rückmeldungen aus dem Plenum bekamen. Jeder Feedbackrunde folgte eine weitere Rollenspielsituation, in der die Trainer noch einmal den Einsatz der DL-Interaktionen unter Beachtung der zurückgemeldeten Hinweise üben konnten.

4) Klärung offener Fragen und Sicherung der Ergebnisse

Eine Fragerunde zum Abschluss der Schulung gab den Trainern die Möglichkeit bis dahin unbeantwortete Fragen zur Anwendung zu stellen. Abschließend fassten die Seminarleiter zentrale Aspekte und Ergebnisse der Schulung zusammen.

Zusätzlich zur Vorbereitung auf die 12-wöchige Interventionsphase fanden innerhalb des Förderzeitraums zwei Begleitseminare jeweils vier und acht Wochen nach Trainingsbeginn statt. Der Schwerpunkt der Folgeseminare lag auf dem Austausch bisheriger Erfahrungen. Außerdem wurden spezifische Problemsituationen Rollenspielen nachgestellt und Lösungsvorschläge diskutiert.

6.3.1.3 Durchführung

Die Förderung fand in einem Kleingruppensetting mit maximal fünf Kindern pro Gruppe statt. Insgesamt nahmen die Probanden an 24 Sitzungen à 30 Minuten teil. Die Sitzungen fanden zwei bis drei Mal in der Woche statt.

Da beim Dialogischen Lesen im Sinne eines inhaltsbasierten Sprachförderansatzes formalsprachliche Regeln durch regelmäßigen Sprachgebrauch zur Anwendung kommen sollten (vgl. Kap. 3.2.1.1), gab es keine standardisierten Vorgaben zur Durchführung der Fördersitzungen. So war nicht festgelegt, wie viele Bildkarten die Trainer mit den Kindern in einer Sitzung bearbeiten sollten. Einzige Vorgabe war eine fünfminütige Einführungsphase zum Einstieg in jede Sitzung, in der die Kinder die Inhalte der vorangegangenen Sitzung in eigenen Worten nacherzählten. Berechnungen nach Abschluss der Förderung zeigten, dass die Trainer im Durchschnitt acht bis zehn neue Bildkarten pro Sitzung mit den Kindern bearbeiteten.

Ebenfalls beachtet wurde ein sinnvoller Wechsel zwischen geschlossenen und offenen Fragestellungen (Blewitt et al., 2009). Geschlossene Fragen sollen die Trainer häufiger zu Beginn der Sitzung stellen. Im weiteren Verlauf sollten häufiger offene Fragen zum Einsatz kommen. Auch unbekannte oder schwere Wörter wurden zunächst in Form von geschlossenen Fragen abgefragt. Offene Fragen sollten neu gelernte Inhalte aufgreifen und wiederholen. Grundsätzlich sollten die Trainer bei jeder Aussage der Kinder die von Whitehurst (1988) postulierten Modellierungstechniken einsetzen (vgl. Kap. 3.2.1.1).

6.3.2 Training der Phonologischen Bewusstheit

6.3.2.1 Klassischer Aufbau des Trainings

Zur Förderung der Phonologischen Bewusstheit diente das Trainingsprogramm *Hören, lauschen, lernen* (Küspert & Schneider, 2006). Insgesamt setzt sich die klassische Trainingsvariante aus 57 Spielen zusammen, die sich auf 200 Kurzeinheiten à 5 Minuten verteilen. Die effektive Trainingszeit beträgt ungefähr 17 Zeitstunden. Tabelle 10 fasst den Programmaufbau zusammen.

Tabelle 10 Aufbau des Original-Trainingsprogramms „Hören, lauschen, lernen“.

	Lauschspiele	Reimen	Sätze und Wörter	Silben	Anlaut	Phon	Σ
Anzahl der Spiele	9	10	8	7	8	15	57
Σ der Einheiten	11	29	18	22	45	75	200

Zentralen Inhalt des Programms bilden Spielvarianten zur Förderung der Phonologischen Bewusstheit im engeren und im weiteren Sinne.

Den Einstieg bilden *Lautspiele*. Hierbei sollen zunächst basale Fertigkeiten wie das Differenzieren, Lokalisieren und Erkennen von Klängen und Geräuschen geübt werden. Die Lauschspiele haben einführenden Charakter und kommen nur in den ersten Sitzungen zum Einsatz.

Nach der Einstiegsphase folgt eine Phase mit vielen *Reimspielen*. Als übergeordnetes Ziel sollen phonologische Ähnlichkeiten bei unterschiedlichen Wörtern und Lautkombinationen als sprachliches Prinzip verstanden werden (Küspert & Schneider, 2006). Die Trainer erarbeiten dieses Ziel sukzessive zunächst durch das Vor- und Nachsprechen vorgegebener Reimpaare. Später lösen Spiele zum freien Reimen die festen Vorgaben ab. Auch Liedtexte und kurze Abzählreime kommen in dieser Phase zum Einsatz.

Als Nächstes folgen Spiele mit ganzen Sätzen. Die Kinder sollen (ähnlich wie beim Dialogischen Lesen) z. B. Sätze auf Grundlage eines vorgegebenen Schlüsselworts reimend ergänzen (Beispiel: *Auf dem Dache breit und fest, hat ein Storchenpaar sein _____*). Gleichzeitig werden erste Erkenntnisse über den Aufbau von *Sätzen* und *Wörtern* vermittelt. Die Zerlegung von Sätzen in einzelne Wörter wird an konkreten Handlungen (Klatschen) und Materialien (farbige Bauklötze) geübt.

Im weiteren Verlauf vertiefen *Silbenspiele* die Analyse- und Synthesefertigkeiten auf Wortebene. Die Kinder bekommen die Aufgabe Silben zu klatschen oder nachzusprechen. Zunächst arbeitet die Gruppe mit vorgegebenen Wörtern, später können sich die Kinder eigene Wörter ausdenken.

Mit dem fünften Inhaltsbereich, der *Identifikation von Anlauten*, erfolgt nunmehr der Brückenschlag zu einem Kernbereich der Phonologischen Bewusstheit. Aufgaben zur Analyse und Synthese auf Phonemebene zählen zur *Phonologischen Bewusstheit im engeren Sinne* (vgl. Kap. 3.2.2.1) und verstehen sich gewissermaßen als „Königdisziplin“ phonologischer Fertigkeiten. Im Fall der Anlautspiele lassen die Trainer Wörter nachsprechen, deren Anlaute bei der Artikulation stark gedehnt werden (*Eeee-lefant, Mmmm-und*). In einem weiteren Bildkartenspiel sollen die Kinder Worte finden, die mit dem gleichen Laut beginnen. Außerdem werden Wörter bearbeitet, bei denen Anlaute vom Rest des Wortes „getrennt“ werden sollen. Durch das Hinzufügen eines neuen Anlauts lassen die Kinder nun neue Wörter entstehen (*Eis wird zu Rrr-eis*).

Der sechste Trainingsbereich konzentriert sich auf die Diskriminierung von Einzellaute innerhalb eines Wortes. Die Kinder beobachten beim Sprechen ihre Lautbildung und markieren diese durch z. B. gleichzeitiges Legen eines Bauklötzchens. Später sollen sie Einzellaute innerhalb eines Wortes ohne Hilfestellungen erkennen.

Die Autoren weisen drauf hin, dass Analyseaufgaben vorerst den Vorrang vor Synthesaufgaben erhalten sollten, da Vorschulkindern dieses Aufgabenformat leichter fällt. Auch sollten die Übungen zunächst kurze Wörter verwenden bevor längere Wörter geübt werden (Küsper & Schneider, 2008).

Da zur Durchführung des klassischen Trainings effektiv ca. 17 Zeitstunden benötigt werden, im Rahmen des Untersuchungssettings aber nur 24 Sitzungen à 30 Minuten (12 Stunden) zur Verfügung standen, mussten die Trainingsinhalte an die Rahmenbedingungen der Untersuchung angepasst werden. Der modifizierte Programmablauf wird im folgenden Abschnitt beschrieben.

6.3.2.2 Modifizierter Aufbau des Trainings

Um vergleichbare Voraussetzungen zur DL- Gruppe zu schaffen, musste die vorgegebene Trainingsstruktur von *Hören, lauschen, lernen* an den zeitlichen Rahmen des Interventionszeitraums der Trainingsstudie angepasst werden. Statt der vorgegebenen 200 Einheiten erlaubte der verfügbare Zeitrahmen nur 144 Kurzeinheiten. Tabelle 11 stellt die Verteilung der modifizierten Trainingsvariante dar.

Tabelle 11 Aufbau der modifizierten Trainingsvariante von „Hören, lauschen, lernen“.

	Lauschspiele	Reimen	Sätze und Wörter	Silben	Anlaut	Phon	Σ
Anzahl der Spiele	-	10	8	7	8	15	48
Σ der Einheiten	-	18	13	13	39	61	144

Die notwendigen Kürzungen setzen aus inhaltlichen Überlegungen bei den Lautspielen an (9 Spiele in 11 Einheiten), da sie lediglich einführenden Charakter haben und keinen Bereich der Phonologischen Bewusstheit gezielt fördern. Die Trainingsvariante berücksichtigte alle verbleibenden 48 Spiele.

Da das Originaltraining häufige Wiederholungen der Spiele vorsieht, wurde auch dieses Prinzip in der Studienvariante beibehalten. Jedes Spiel kam mehrfach zum Einsatz. Durch die

Kürzung der Fördersitzungen auf eine Länge von 30 Minuten ergab sich für manche Sitzungen ein unverhältnismäßig hoher Anteil an Wiederholungen. Zum Beispiel kam es vor, dass vier Einheiten des Silbenballspiels (20 Minuten) in einer Sitzung untergebracht waren. In solchen Fällen wurde die mehrfache Durchführung eines Spiels durch ein inhaltlich ähnliches Spiel ersetzt. Somit konnten inhaltliche Verschiebungen im Vergleich zur klassischen Form vermieden werden.

In der Trainingsvariante deckten 69% der Einheiten weiterhin den Bereich Phonologische Bewusstheit im engeren Sinne ab (sie waren in der klassischen Variante zu 60% vertreten). 31% der Spiele trainierten die Phonologische Bewusstheit im weiteren Sinne (35% in der Originalform).

6.3.2.3 Durchführung und Schulung der Trainer in der Phonologischen Bewusstheit

Die modifizierte Trainingsvariante wurde ebenso wie das Dialogische Lesen in einem Kleingruppensetting à 30 Minuten mit maximal fünf Kindern zwei bis drei Mal pro Woche durchgeführt. Wie im DL übernahmen weitere geschulte Trainer (Studierende der Psychologie und des Lehramts an Förderschulen) die Förderung.

Anders als bei beim Dialogischen Lesen lagen für das Training der Phonologischen Bewusstheit (PB-Training) standardisierte Instruktionen vor. Um die Rahmenbedingungen konstant zu halten nahmen auch die Trainer der PB-Gruppe an einer Schulung vor und während der Interventionsphase teil. Die Struktur des Schulungsseminars war identisch mit der Schulung der DL-Trainer. Jeweils vier und acht Wochen nach Beginn der Interventionsphase fanden auch in der PB-Gruppe zwei Feedbackseminare statt.

6.3.3 Kontrollgruppe

Die Kinder der Kontrollgruppe nahmen am konventionellen Sprachförderangebot in ihren Kitas teil. Sie erhielten kein zusätzliches Trainingsprogramm. Um festzustellen welche Maßnahmen in der Kontrollgruppe zum Einsatz kamen, wurden die Erzieherinnen vor der Interventionsphase befragt. Das Sprachförderangebot war wie erwartet äußerst heterogen. Viele Erzieherinnen arbeiteten mit einem Mix aus verschiedenen Programmen (z.B. *Hören, lauschen, lernen*; *Lobo vom Globo*; *Wuppis Abenteuerreisen durch die Phonologische Bewusstheit*). Auch Vorlesen, Sing- und Bewegungsspiele wurden genannt. Zum Teil wurden gezielte Grammatikspiele eingesetzt. Einige Kitas strukturierten ihr Angebot auf tägliche

Einheiten à 15 Minuten. Andere förderten in wöchentlichem Turnus ein bis zwei Mal für mindestens eine halbe Stunde.

Von einer Vereinheitlichung des Förderangebots in der Kontrollgruppe wurde abgesehen, da sich der Mehrwert erwarteter Trainingseffekte immer an den Gepflogenheiten der Praxis messen sollte.

6.4 Statistisches Vorgehen

6.4.1 Prüfung der kurz- und der langfristigen Wirksamkeit

Um zu prüfen, ob sich die DL-Gruppe im Vergleich zu den anderen Untersuchungsgruppen kurz- und langfristig in den unterschiedlichen Sprachentwicklungsbereichen verbessert und, ob die PB-Gruppe spezifische Effekte in der Phonologischen Bewusstheit erzielen konnte, wurde die Entwicklung in den Untersuchungsgruppen mit Hilfe von Kovarianzanalysen (ANCOVA) untersucht. Im Folgenden werden die statistischen Voraussetzungen des Verfahrens dargestellt.

6.4.1.1 Anwendungsvoraussetzungen für Kovarianzanalysen

Alle ANCOVAs wurden mit dem Statistikprogramm SPSS 20 berechnet. In der Literatur gilt die ANCOVA als robustes und häufig eingesetztes Verfahren zur Analyse der Programmwirksamkeit in einem Vergleichsgruppen-Design mit mindestens zwei MZP (Field, 2009; Hager, 2000).

Die ANCOVA bietet zudem die Möglichkeit, den Einfluss von (Stör-)Variablen zu kontrollieren, die mit der abhängigen Variable konfundiert sein könnten. Die damit einhergehende Reduktion der Fehlervarianz kann die Testpower (Wahrscheinlichkeit, mit der ein Signifikanztest zu Gunsten einer spezifischen Alternativhypothese entscheidet) der Analyse erhöhen (Dimitrov & Rumrill Jr., 2003; Wildt & Ahtola, 1978). Im vorliegenden Fall wurde die Intelligenzleistung der Probanden als zusätzliche Kovariate in den Analysen berücksichtigt.

Grundsätzlich gelten für alle varianzanalytischen Verfahren drei Anwendungsbedingungen (Bortz & Schuster, 2010 S.274):

- (1) *Bedingung der Normalverteilung:* Die Verteilung der Residuen (Abweichung eines Messwertes vom Stichprobenmittelwert) darf sich pro Treatmentstufe nicht signifikant von der Normalverteilung unterscheiden.
- (2) *Bedingung der Varianzhomogenität:* Die Varianzen der Residuen dürfen sich innerhalb der Stichprobe nicht signifikant unterscheiden.
- (3) *Bedingung der Unabhängigkeit:* Die Residuen innerhalb einer Stichprobe müssen unabhängig voneinander sein.

Glass und Kollegen (1972) weisen darauf hin, dass mit geringem Stichprobenumfang häufig Verletzungen der obengenannten Voraussetzungen einhergehen. Bei ausreichender Stichprobengröße und nahezu gleichgroßen Gruppen gelten varianzanalytische Verfahren als robust gegen etwaige Verletzungen. Als ausreichendes Maß wird dabei eine Stichprobengröße zwischen 10 und 20 Probanden pro Treatmentstufe angesehen (Bortz & Schuster, 2010; Stevens, 2007). In der vorliegenden Untersuchung ist dies gewährleistet, sodass die Bedingungen (1) und (2) erfüllt sind.

Nach Bortz und Schuster (2010) kann die letztgenannte Bedingung der unabhängigen Residuen dann angenommen werden, wenn die Zuteilung zu den Trainingsbedingungen randomisiert erfolgte. In dieser Untersuchung ist die zufällige Zuteilung der Kindergärten zu einer von drei Treatmentstufen gegeben, sodass auch das dritte Kriterium erfüllt ist.

Zusätzlich zu den allgemeinen Voraussetzungen für varianzanalytische Verfahren gelten für die Anwendung der ANCOVA zwei spezifische Bedingungen. Gefordert werden (1) Unabhängigkeit zwischen Kovariate und Treatmentfaktor und (2) Homogenität der Regressionssteigung innerhalb der Kovariaten (Bortz & Schuster, 2010; Dimitrov & Rumrill, 2003; Field, 2009, Keselman, Huberty, Lix et al., 1998).

Miller und Chapman (2001) weisen darauf hin, dass viele ANCOVAs häufig fälschlicherweise auch dann angewendet werden, wenn ein substantieller Zusammenhang zwischen Treatmentfaktor und Kovariate besteht. Rogosa (1980) wie auch Hamilton (1977) unterstreichen indes, dass bei randomisierter Zuteilung und gleichgroßen Stichproben etwaige Verletzungen der Bedingungen zu vernachlässigen sind. Selbst bei ungleichgroßen Stichproben gelten Unterschiede zwischen den Regressionssteigungen bis $\alpha = 0.4$ als tolerabel, da sie nur zu unbedeutenden Verzerrungen führen (Bortz & Schuster, 2010; Wu, 1986).

Obwohl beide Voraussetzungen der Kovarianzanalyse in der vorliegenden Untersuchung als erfüllt gelten können, wurden beide Kriterien auch statistisch abgesichert. Eine ANOVA ergab keine signifikanten Zusammenhänge zwischen den Treatmentstufen und den Kovariaten. Homogenität der Regressionssteigungen konnte ebenfalls nachgewiesen werden, da in ergänzend durchgeführten ANCOVAs für alle abhängigen Variablen keine signifikanten Interaktionseffekte zwischen IQ X Gruppe und Prätestleistung X Gruppe gefunden wurden.

Um zu überprüfen, ob Unterschiede in der Durchführungsqualität bedeutsame Einflüsse auf die Effektivität der Förderung haben, wurden multiple Regressionsanalysen berechnet.

Regressionsanalysen werden häufig als Verfahren zur Vorhersage von *Kriteriumsvariablen* durch den Einfluss einer oder mehrerer *Prädiktorvariablen* eingesetzt. Im Fall der vorliegenden Untersuchung dienten die Regressionsanalysen zur Bestimmung des Einflusses der Durchführungsqualität des Dialogischen Lesens (Prädiktorvariable) auf Trainingserfolge der DL-Gruppe (Kriteriumsvariable).

Bortz und Schuster (2010) nennen eine Vielzahl von Anwendungsvoraussetzungen, die zumindest annähernd erfüllt sein sollten. So sollten die Ausprägungen der einzelnen zu untersuchenden Merkmale ebenso normalverteilt sein wie die Arrayverteilung (die Verteilung der zu einzelnen x-Werten gehörenden y-Werte). Außerdem sollten die Mittelwerte der Arrayverteilung auf einer Geraden liegen. Vorausgesetzt wird zudem die *Homoskedastizität* der Arrayverteilung; dann ist Varianzhomogenität gegeben. Schließlich müssen die zu einem x-Wert gehörenden y-Werte unabhängig voneinander sein. Die Prüfung dieser Kriterien erfolgt in der Forschungspraxis per Augenschein über die optische Prüfung der Verteilungsformen der einzelnen Merkmale. Bortz und Schuster (2010) räumen ein, dass geringfügige Abweichungen zu tolerierbaren Verzerrungen führen. Für die durchgeführten Analysen sind die Voraussetzungen erfüllt. Bei Nichterfüllung der Anwendungsvoraussetzungen wird an entsprechender Stelle darauf hingewiesen.

7 Ergebnisse

7.1 Prüfung der Kriteriumsvalidität

Tabelle 12 Korrelationen zwischen Testwerten des SSV und dem Erzieherurteil zum Sprachentwicklungsstand.

N= 587	Satzgedächtnis	Erzieherurteil
Pseudowort	.53**	.36**
Satzgedächtnis	-	.61**

** signifikant auf dem .01-Niveau

Zur Prüfung der Kriteriumsvalidität wurden die Ergebnisse des Screenings mit dem Erzieherurteil zur Einschätzung der Sprachentwicklung der Ausgangsstichprobe korreliert. In den Berechnungen wurden die Daten von 587 Probanden berücksichtigt. Von 73 Probanden lag kein Erzieherurteil vor, da die Eltern keine Zustimmung erteilten, oder weil einige Erzieher unausgefüllte Bögen abgaben.

Es ergab sich eine solide Korrelation zwischen der Satzgedächtnisaufgabe und dem Erzieherurteil. Das heißt, die Leistungen der Kinder in der Satzgedächtnisaufgabe stimmten in hohem Maße mit den Erzieher-Einschätzungen des Sprachentwicklungsstandes überein. Die Satzgedächtnisaufgabe erweist sich somit als geeigneter Marker zur Beurteilung des Sprachentwicklungsstandes.

Die Korrelationen zwischen der Pseudowortaufgabe und dem Erzieherurteil fällt deutlich geringer aus ($r = .36$). Dies ist nicht verwunderlich, da es sich bei der Pseudowortaufgabe in erster Linie um einen Test der Arbeitsgedächtnisleistung handelt und diese über einen Beobachtungsbogen zur Sprachentwicklung kaum korrekt eingeschätzt werden kann. Insgesamt zeigen die Ergebnisse, dass mit dem SSV ein valides Instrument zur Auswahl von Risikokindern eingesetzt worden ist.

7.2 Deskriptive Statistiken und korrelative Zusammenhänge

Die Prä-, Post- und Follow Up-Ergebnisse in den einzelnen Untertests fasst Tabelle 13 für alle Untersuchungsgruppen zusammen. Die Intelligenzleistung wurde nur einmalig im Prätest erhoben, da es sich um ein relativ stabiles Konstrukt handelt, bei dem keine bedeutsamen Veränderungen zu erwarten waren (s. Kap. 6.2.5).

Tabelle 13 Deskriptive Statistik der Untertests in den Untersuchungsgruppen

		EG1: Dialogisches Lesen						EG2: Phonologische Bewusstheit						Kontrollgruppe					
		Prä		Post		Follow Up		Prä		Post		Follow Up		Prä		Post		Follow Up	
		n = 59		n = 58		n = 53		n = 58		n = 54		n = 49		n = 57		n = 56		n = 52	
	max	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
<u>Wortschatz</u>																			
Bildbenennung	40	22.90	8.7	27.04	6.0	27.64	6.6	23.29	7.7	27.15	7.2	28.08	6.6	22.21	8.2	26.13	6.0	28.58	6.6
Kategoriebildung	30	11.54	6.0	14.98	6.2	14.28	5.0	11.97	4.9	15.17	4.8	15.39	5.1	11.63	4.9	14.64	4.5	15.88	6.0
<u>Sprachverständnis</u>																			
Handlungssequenzen	12	5.03	2.7	6.73	2.6	6.15	3.0	5.14	2.2	6.33	2.4	6.55	2.2	4.81	2.4	6.25	2.3	6.67	2.8
Fragen zum Text	6/10 ^a	2.25	1.6	3.25	1.7	4.92	2.8	2.43	1.8	3.00	1.8	4.48	2.5	2.40	1.7	2.80	1.8	4.90	2.7
<u>Grammatik</u>																			
Satzbildung	12	3.14	2.9	5.11	4.0	4.85	3.7	2.48	3.0	3.48	3.0	4.55	3.3	2.88	3.0	3.43	3.5	6.08	4.2
Singular-Plural Bildung	18	7.12	3.9	8.15	4.3	8.28	4.2	7.64	3.7	8.20	3.6	9.25	3.2	7.40	4.6	8.55	3.7	8.61	4.0
<u>Phonologische Bewusstheit</u>																			
Reimen	10/12 ^b	7.34	1.8	10.27	2.8	11.11	2.1	7.34	2.1	11.00	2.0	11.06	2.2	7.63	1.8	10.38	2.5	11.60	1.9
Phonemsynthese (m. Abbildung)	8	4.61	2.0	5.88	1.9	6.72	1.5	5.03	1.8	6.69	1.4	7.06	1.3	4.07	2.1	5.55	2.0	7.12	1.2
Phonemsynthese (o. Abbildung)	42	4.61	2.0	8.82	10.3	17.51	13.2	5.03	1.8	14.46	9.2	21.02	12.9	2.86	4.3	9.45	9.6	20.00	12.9
Anlautidentifikation	10	2.86	2.5	3.85	2.6	5.64	2.7	2.67	2.4	5.93	2.6	6.76	2.9	2.74	1.9	3.88	2.7	7.19	2.68
<u>Intelligenz</u>																			
Matrizentest	17	6.92	2.9	--	--	--	--	7.05	2.7	--	--	--	--	7.25	2.4	--	--	--	--
Klassen bilden	17	9.49	3.9	--	--	--	--	10.40	3.0	--	--	--	--	10.02	3.1	--	--	--	--

^a Max. 6 Rohpunkte konnten im Prä- und Posttest erzielt werden, max.10 Rohpunkte im Follow Up.^b Max. 10 Rohpunkte konnten im Prätest erzielt werden, max.12 Rohpunkte im Post- und Follow Up-Test.

Zwischen den Untertests der einzelnen Sprachdomänen ergaben sich zu allen drei MZP hohe Korrelationen. (vgl. Tabelle 23 im Anhang). Für den Bereich *Wortschatz* lagen die Korrelationen der Subtests *Bildbenennung* und *Kategoriebildung* abhängig vom Messzeitpunkt zwischen $r = .72$ und $r = .75$. Die Subtests *Handlungssequenzen* und *Fragen zum Text* korrelierten im Bereich *Sprachverständnis* zwischen $r = .46$ und $r = .55$. Die geringe Korrelation ist möglicherweise auf die unterschiedliche Komplexität der Aufgabenformate zurückzuführen. Während die Kinder zur Umsetzung der Handlungssequenzen immer nur einen Satz verstehen mussten, wurden ihnen bei *Fragen zum Text* mehrere Sätze vorgelesen. Um die Fragen richtig zu beantworten war eine höhere Verständnisleistung von Nöten.

Im Bereich Grammatik korrelierten die Subtests *Singular-Plural-Bildung* und *Satzbildung* zwischen $r = .45$ und $r = .55$. Alle Korrelationen zeigen, dass die ausgewählten Subtests innerhalb ihrer Inhaltsbereiche deutliche Zusammenhänge aufweisen. Das heißt, die Wortschatzleistung wurde besonders gut über die Subtests *Bildbenennung* und *Kategoriebildung* erfasst; die Grammatikleistung konnte am besten durch die Aufgaben *Satzbildung* und *Singular-Plural-Bildung* abgebildet werden, und das Sprachverständnis wurde am besten durch die Aufgaben *Handlungssequenzen* und *Fragen zum Text* erfasst.

Um mit den Ergebnissen der Untersuchung eine Aussage über die Entwicklung von größer gefassten Sprachkompetenzbereichen zu treffen, anstatt vereinzelte Subtestleistungen zu untersuchen und Redundanzen zu befördern, wurden die Subtestergebnisse inhaltlich zusammengefasst. Das heißt, die Leistungen in *Bildbenennung* und *Kategoriebildung* wurden als *Wortschatz (WS)* zusammengefasst. Die Subteststs *Handlungssequenzen* und *Fragen zum Text* erfassten zusammen das *Sprachverständnis (SV)*. *Grammatik (GRA)* setzte sich aus *Singular-Plural-Bildung* und *Satzbildung* zusammen. Für die statistische Zusammenfassung der Daten wurden die einzelnen Subtestwerte zunächst z-transformiert. Danach wurden die Testergebnisse über gemittelte z-Werte zu den genannten Sprachentwicklungsbereichen zusammengefasst. Die Ergebnisse aus den Subtests zur Phonologischen Bewusstheit wurden ebenfalls z-transformiert und zu einem eigenen Score *Phonologische Bewusstheit (PB)* zusammengefasst.

Wie erwartet ergaben sich zwischen den Scores der zusammengefassten Sprachentwicklungsbereiche WS, SV und GRA hohe korrelative Zusammenhänge (r zwischen $.68$ und $.76$), die sich über alle Messzeitpunkte als stabil erwiesen (r zwischen $.60$ und $.71$; vgl. Tabelle 24 im Anhang). Dies deutet darauf hin, dass alle drei Scores inhaltlich äußerst eng mit einander assoziiert sind. Somit erschien es sinnvoll aus allen drei Sprachentwicklungsbereichen einen Gesamtscore zur Feststellung der *Allgemeinen*

Sprachentwicklung (AS) zu bilden. Hierfür wurden die Werte aus WS, SV und GRA erneut zu einem gemittelten z-Wert zusammengefasst.

Die Leistung in Phonologischer Bewusstheit wurde getrennt von den anderen Sprachkompetenzen betrachtet, um die Wirksamkeit des formbasierten PB-Kontrolltrainings feststellen zu können. Auch theoriegeleitet erschien eine isolierte Betrachtung der PB-Leistung sinnvoll, da hier möglicherweise mehr metasprachliche Kompetenzen erfasst wurden, als in den anderen Sprachentwicklungsbereichen (siehe. Kap. 3.2.2). Die geringen Korrelationen zwischen PB und den anderen drei Sprachscores ($r = .35$ bis $r = .37$) untermauern diese Vermutungen.

Die nonverbale Intelligenzleistung korrelierte moderat mit WS, SV GRA und PB (r zwischen .30 und .44).

Bevor die Ergebnisse der statistischen Analysen ausführlich berichtet werden, fasst Tabelle 14 die deskriptiven Statistiken auf Basis der einzelnen Trainingsgruppen zusammen.

Tabelle 14 Deskriptive Statistiken der mittleren z-Werte der abhängigen Variablen für alle Untersuchungsgruppen.

	Prä-Test ^a				Post-Test		Follow Up	
	Min	Max	M	SD	M	SD	M	SD
Dialogisches Lesen	<i>n</i> = 59				<i>n</i> = 55		<i>n</i> = 53	
Sprachkompetenzen (SK)	-1.55	1.59	0.00	0.84	0.56	0.94	0.67	0.97
<i>Wortschatz (WS)</i>	-2.44	2.07	-0.01	1.04	0.57	1.04	0.54	0.88
<i>Sprachverständnis (SV)</i>	-1.52	2.31	-0.02	0.90	0.62	0.94	1.03	1.32
<i>Grammatik (GRA)</i>	-1.38	1.68	0.02	0.84	0.48	1.13	0.45	0.97
Phonol. Bewusstheit (PB)	-1.11	2.29	0.05	0.72	0.88	0.91	1.66	1.07
Phonologische Bewusstheit	<i>n</i> = 58				<i>n</i> = 54		<i>n</i> = 49	
Sprachkompetenz (SK)	-1.73	2.14	0.03	0.77	0.42	0.76	0.69	0.63
<i>Wortschatz (WS)</i>	-2.08	2.19	0.05	0.86	0.59	0.83	0.67	0.83
<i>Sprachverständnis (SV)</i>	-1.72	2.11	0.05	0.84	0.47	0.90	1.08	0.99
<i>Grammatik (GRA)</i>	-1.38	2.11	-0.03	0.84	0.21	0.81	0.43	0.78
Phonol. Bewusstheit (PB)	-0.97	1.58	0.02	0.56	1.54	0.74	1.96	1.00
Kontrollgruppe	<i>n</i> = 57				<i>n</i> = 56		<i>n</i> = 52	
Sprachkompetenz (SK)	-1.56	1.71	-0.02	0.79	0.37	0.71	0.55	0.87
<i>Wortschatz (WS)</i>	-2.19	1.51	-0.04	0.89	0.48	0.72	0.75	0.91
<i>Sprachverständnis (SV)</i>	-1.52	1.51	-0.02	0.83	0.39	0.82	1.13	1.29
<i>Grammatik (GRA)</i>	-1.38	2.39	0.01	0.92	0.24	0.88	0.44	0.95
Phonol. Bewusstheit (PB)	-0.89	1.57	-0.07	0.60	0.88	0.95	2.04	1.05

^a Um die Entwicklungsverläufe über die Zeit anschaulich abbilden zu können, wurden die ermittelten z-Werte der Posttestung und des Follow Up auf Basis der Prätestleistung normiert.

7.3 Kurz- und langfristige Trainingseffekte

7.3.1 Effekte auf die Phonologische Bewusstheit

Abbildung 5 zeigt die Entwicklungsverläufe der Untersuchungsgruppen in der Phonologischen Bewusstheit vom Prä- zum Post-Test sowie zum Follow Up. Um die Entwicklung über die Zeit anschaulich darzustellen wurden für alle MZP z-Werte gebildet und an den Vortestergebnissen normiert.

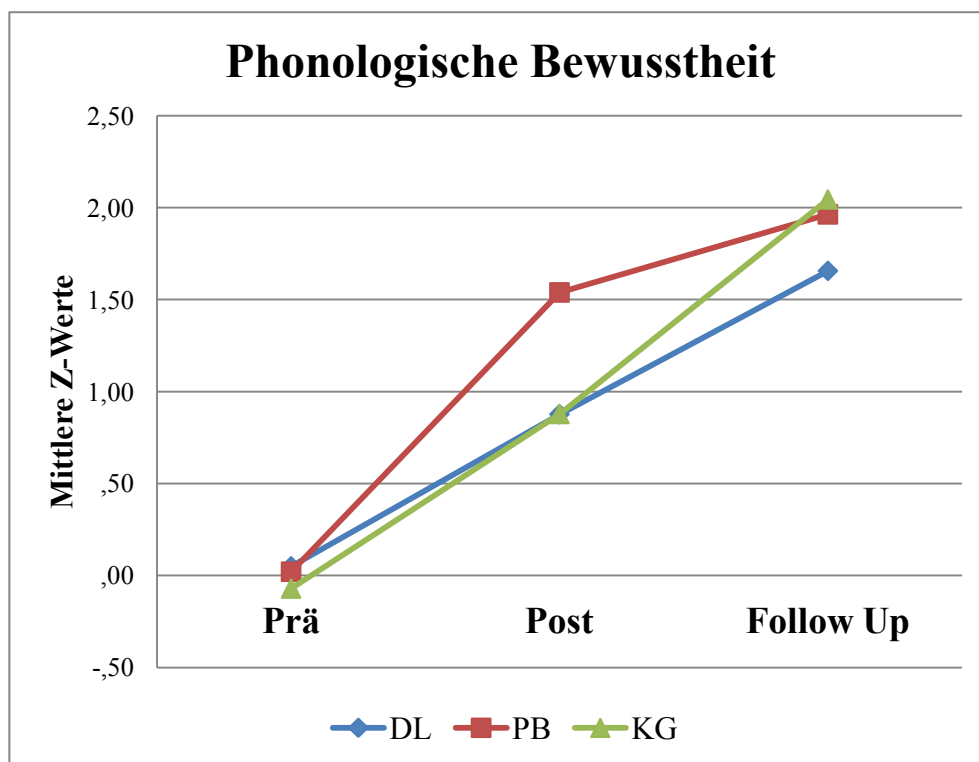


Abbildung 5 Durchschnittliche Leistungen aller Gruppen in der Phonologischen Bewusstheit im Prä-, Post- und Follow Up-Test.

Zur Analyse von kurzfristigen Trainingseffekten unmittelbar nach der Intervention wurde eine ANCOVA mit der Posttestleistung in der Phonologischen Bewusstheit als abhängige Variable sowie der Gruppenzugehörigkeit als unabhängigem Faktor berechnet. Als Kovariaten wurden Prätestleistung und Intelligenz berücksichtigt. Erwartungsgemäß konnte die Prätestleistung einen signifikanten Beitrag zur Varianzaufklärung im Posttest leisten ($F[1;160] = 68.77$; $p < .01$; $\eta^2 = .30$). Auch die Intelligenzleistung hatte signifikanten Einfluss auf die Posttestergebnisse ($F[1;160] = 16.07$; $p < .01$; $\eta^2 = .09$). Das heißt, die IQ-Leistung konnte 9% Nachtest-Varianz erklären. Außerdem ergab sich ein signifikanter Effekt der Gruppe ($F[2;160] = 15.24$; $p < .01$; $\eta^2 = .15$). Post hoc-Analysen zeigten, dass die Kinder des PB-Trainings im Nachtest signifikant bessere Leistungen in Aufgaben zur Phonologischen

Bewusstheit erzielten als Kinder der KG ($p < .01$) und Kinder der DL-Gruppe ($p < .01$). Kein signifikanter Leistungsunterschied ergab sich zwischen den Leistungen DL und KG.

Um die Größe der spezifischen Effekte des PB-Trainings empirisch einzuordnen, wurden korrigierte Effektstärken d_{korr} berechnet (Klauer & Büchel, 1993). Nach Cohen (1988) handelt es sich bei Effekten ab $d = 0.2$ um einen kleinen, ab $d = 0.5$ um einen mittleren und ab $d = 0.8$ um einen großen Effekt. Die Überlegenheit von PB gegenüber KG entsprach einem mittleren Effekt ($d_{korr} = 0.62$). Die Überlegenheit von PB gegenüber dem DL entsprach einem großen Effekt ($d_{korr} = 0.84$). Insgesamt konnten mit dem gewählten Analyseverfahren 51% Gesamtvarianz im Posttest aufgeklärt werden ($R^2 = .51$).

Im Follow Up klärte die Prätestleistung nur noch 14% der Leistungsvarianz auf ($F[1;149] = 24.82$; $p < .01$; $\eta^2 = .14$). Die graphische Darstellung der Werte ergab eine stark rechtsgipflige Verteilung, sodass von Deckeneffekten auszugehen ist. Eine Prüfung mit dem Shapiro-Wilks-Test bestätigte, dass die Follow Up-Ergebnisse in der Phonologischen Bewusstheit nicht normalverteilt sind. Es ist damit von einer eingeschränkten Varianz auszugehen. Die Varianzeinschränkung ist dafür verantwortlich, dass Prä-Test und Follow Up-Leistungen nur noch schwach miteinander korrelieren.

Die Intelligenzleistung hatte weiterhin einen signifikanten Einfluss auf die Follow Up-Leistungen ($F[1;149] = 5.77$; $p < .05$; $\eta^2 = .04$). Allerdings ergab sich kein signifikanter Gruppeneffekt. ($F[2;149] = 2.86$; $p = .06$; $\eta^2 = .04$). Das heißt, dass sich kein statistisch bedeutsamer Unterschied zwischen den Trainingsgruppen und der Kontrollgruppe feststellen ließ. Aufgrund der berichteten Deckeneffekte ist zu vermuten, dass die eingesetzten Aufgaben etwaige Leistungsunterschiede zwischen den Gruppen zur Mitte des 1. Schuljahres nicht mehr adäquat abbilden konnten. Somit ist keine gesicherte Aussage über die langfristige Stabilität der Trainingseffekte möglich.

7.3.2 Effekte auf weitere Sprachentwicklungsbereiche

7.3.2.1 Allgemeine Sprachentwicklung (AS)

Abbildung 6 zeigt die Entwicklungsverläufe der Untersuchungsgruppen in der allgemeinen Sprachkompetenz zum Prä- und Post-Test sowie zum Follow Up.

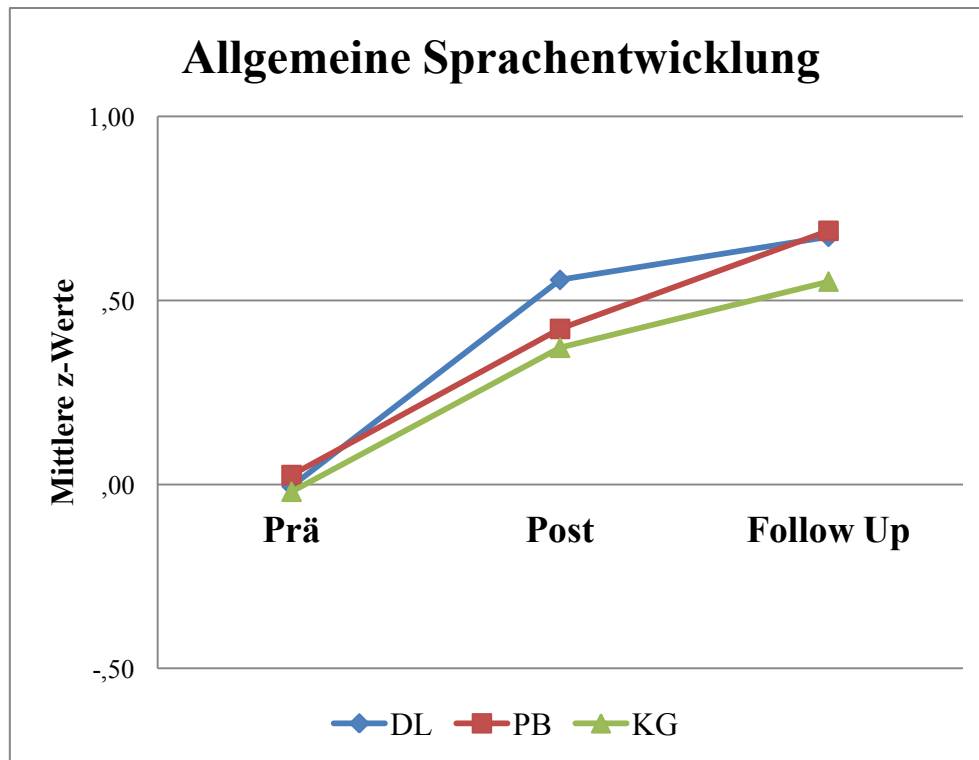


Abbildung 6 Durchschnittliche Leistungen aller Gruppen in „Allgemeine Sprachentwicklung“ im Prä-, Post- und Follow Up-Test.

In den berechneten ANCOVAs zur Ermittlung kurz- und langfristiger Leistungsunterschiede wurden die Prätestleistung und die Intelligenz als Kovariaten berücksichtigt. Im Prä Post-Vergleich konnte die Prätestleistung 66% der Varianz zum Nachtest aufklären ($F[1;160] = 306.90$; $p < .01$; $\eta^2 = .66$). Die Intelligenzleistung hatte keinen signifikanten Einfluss auf die Ergebnisse im Nachtest ($F[1;160] = 1.03$; $p = .31$; $\eta^2 = .01$). Weiter konnte ein signifikanter Effekt der Gruppe festgestellt werden ($F[2;160] = 4.00$; $p < .05$; $\eta^2 = .05$).

DL-Kinder konnten ihre allgemeine Sprachkompetenz sowohl im Vergleich zu KG ($p < .05$) als auch im Vergleich zu PB ($p < .05$) signifikant verbessern. Die korrigierten Effektstärken ergaben im Vergleich von DL vs. KG $d_{\text{kor}} = 0.20$ und DL vs. PB $d_{\text{kor}} = 0.19$. Demnach konnten kleine Effekte verzeichnet werden. Insgesamt klärte das Modell 73% der Gesamtvarianz im Posttest ($R^2 = .73$).

Zum Follow Up konnte die Prätestleistung erneut in hohem Maße zur Varianzaufklärung beitragen ($F[1;116] = 190.74$; $p < .01$; $\eta^2 = .62$).

Die Intelligenzleistung stellte im Follow Up den zweiten bedeutsamen Einflussfaktor ($F[1;116] = 4.00$; $p < .05$; $\eta^2 = .03$).

Tendenziell zeigt sich die Überlegenheit beider Experimentalgruppen gegenüber der Kontrollgruppe. Die Effektstärken von DL und PB waren im Follow Up allerdings so gering,

dass der Unterschied zur KG das Signifikanzniveau verfehlte (DL vs. KG: $d_{\text{kor}} = 0.11$, PB vs. KG: $d_{\text{kor}} = 0.12$). Insgesamt konnte das Modell 70% der Varianz der allgemeinen Sprachkompetenz im Follow Up aufklären ($R^2 = .70$). Die Befunde stützen die Hypothese, dass eine Förderung durch Dialogisches Lesen einen positiven Einfluss auf die Entwicklung der allgemeinen Sprachkompetenz hat. Allerdings ergaben sich nur kleine Effekte, die sich nicht als nachhaltig erwiesen.

Im Folgenden sollen die Entwicklungsverläufe innerhalb der Sprachentwicklungsdomänen Wortschatz, Sprachverständnis und Grammatik näher betrachtet werden.

7.3.2.2 Wortschatzentwicklung (WS)

Abbildung 7 zeigt die Entwicklungsverläufe der Untersuchungsgruppen im Wortschatz vom Prä- zum Post-Test sowie zum Follow Up.

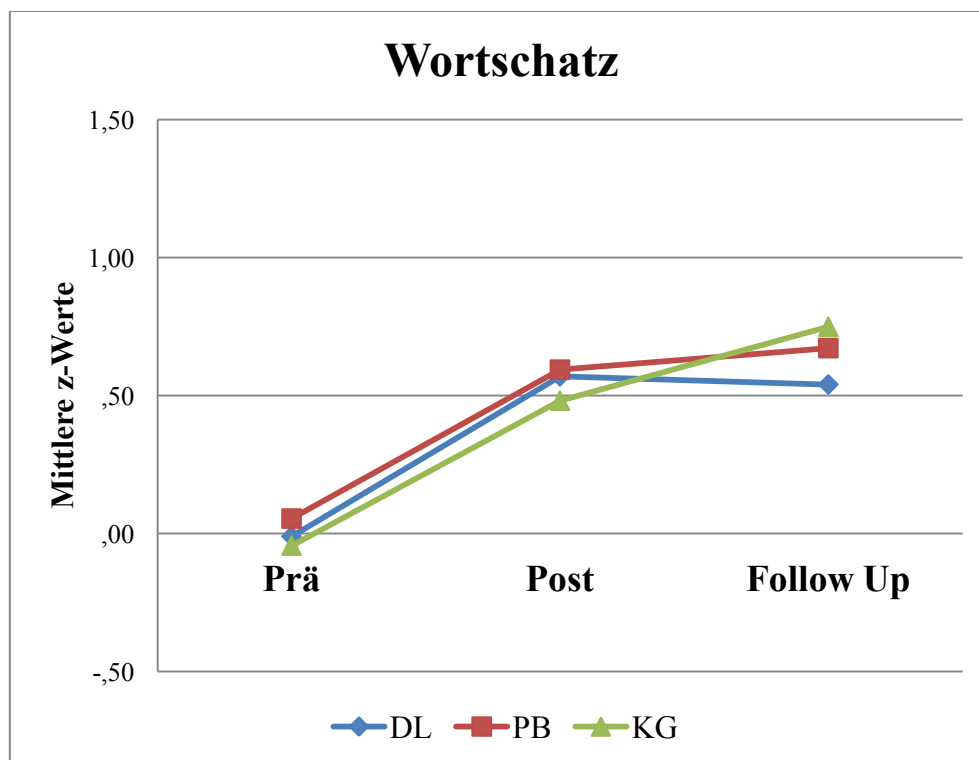


Abbildung 7 Durchschnittliche Leistungen aller Gruppen im Wortschatz im Prä-, Post- und Follow Up-Test.

Wie Abbildung 7 verdeutlicht, ergaben sich über alle Gruppen hinweg nahezu parallele Entwicklungsverläufe in der Wortschatzentwicklung. Zum Posttest konnten weder die Gruppenzugehörigkeit ($F[2;160] = 0.59$; $p = .56$; $\eta^2 = .01$) noch die Intelligenzleistung ($F[1;160] = 2.77$; $p = .10$; $\eta^2 = .02$) einen signifikanten Beitrag zur Varianzaufklärung leisten; nach Abschluss der Förderung konnten für die Wortschatzentwicklung keine statistisch signifikanten Unterschiede zwischen den Untersuchungsgruppen festgestellt werden. Die

Vortestleistung erklärte 56% der Varianz im Posttest auf ($F[1;160] = 205.37$; $p < .01$; $\eta^2 = .56$). Insgesamt konnte das Modell 66% der Varianz im Posttest erklären

Auch langfristig ließ sich kein bedeutsamer Effekt der Gruppe für die Wortschatzentwicklung feststellen ($F[2;149] = 0.96$; $p = 39$; $\eta^2 = .01$). Im Follow Up klärte die Vortestleistung erneut die meiste Varianz ($F[1;149] = 145.08$; $p < .01$; $\eta^2 = .49$). Die Intelligenzleistung hatte keinen signifikanten Einfluss auf die Langezeitentwicklung ($F[1;149] = 1.20$; $p = 27$; $\eta^2 = .01$). Das Modell klärte insgesamt 57% Varianz der Wortschatzentwicklung im Follow Up auf ($R^2 = .57$).

7.3.2.3 Sprachverständnis (SV)

Abbildung 8 zeigt die Entwicklungsverläufe der Untersuchungsgruppen im Wortschatz zum Prä- und Post-Test sowie zum Follow Up.

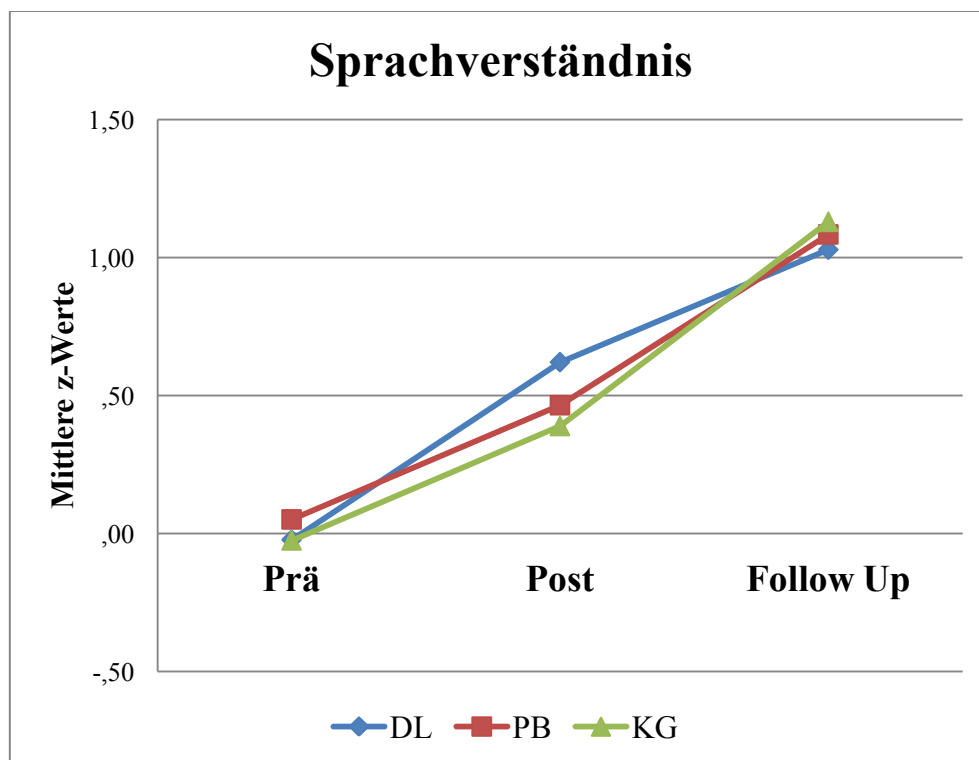


Abbildung 8 Durchschnittliche Leistungen aller Gruppen im Sprachverständnis im Prä-, Post- und Follow Up-Test.

Im Bereich Sprachverständnis konnten erneut die Leistungen im Vortest den größten Anteil der Varianz im Nachtest (42%) erklären ($F[1;160] = 116.19$; $p < .01$; $\eta^2 = .42$). Außerdem hatte die Intelligenzleistung bedeutsamen Einfluss auf das Sprachverständnis im Nachtest ($F[1;160] = 5.08$; $p < .05$; $\eta^2 = .03$).

Weiter ergab sich ein signifikanter Effekt der Trainingsgruppe ($F[2;160] = 3.35; p < .05; \eta^2 = .04$). DL-Kinder zeigten signifikant bessere Nachtestleistungen als die PB-Gruppe ($p < .05$) und ebenfalls signifikant bessere Leistungen als die Kontrollgruppe ($p < .05$). Kein signifikanter Unterschied fand sich zwischen PB und KG ($p = .88$).

Im Vergleich zur KG ergab sich für die DL-Gruppe eine geringe Effektstärke ($d_{\text{kor}} = 0.26$). Auch im Vergleich mit PB wurde ein kleiner Effekt für DL festgestellt ($d_{\text{kor}} = 0.25$). Insgesamt konnten über das eingesetzte Modell 54% der Varianz im Nachtest erklärt werden ($R^2 = .54$).

Die Leistungsunterschiede im Nachtest konnten im Follow Up nicht bestätigt werden.

Hier konnte ausschließlich die Prätestleistung einen bedeutsamen Beitrag zur Varianzaufklärung leisten ($F[1;149] = 104.14; p < .01; \eta^2 = .41$). Weder die Intelligenzleistung ($F[1;149] = 2.37; p = .13; \eta^2 = .02$) noch die Gruppenzugehörigkeit ($F[2;149] = 0.47; p = .64; \eta^2 = .01$) lieferte zu diesem Zeitpunkt einen signifikanten Beitrag zur Varianzaufklärung. Insgesamt klärte das Modell 49% der Gesamtvarianz auf.

Zusammenfassend ist festzuhalten, dass Kinder, die nach den Prinzipien des Dialogischen Lesens gefördert wurden, kurzzeitig bessere Leistungen im Sprachverständnis erzielten als die Kinder der KG und bessere Leistungen als die PB-Gruppe. Allerdings konnten nur geringe Effektstärken festgestellt werden. Langfristig erwiesen sich die Effekte als nicht stabil.

7.3.2.4 Grammatik (GRA)

Abbildung 9 zeigt die Entwicklungsverläufe der Untersuchungsgruppen in der Grammatikleistung zum Prä- und Post-Test sowie zum Follow Up.

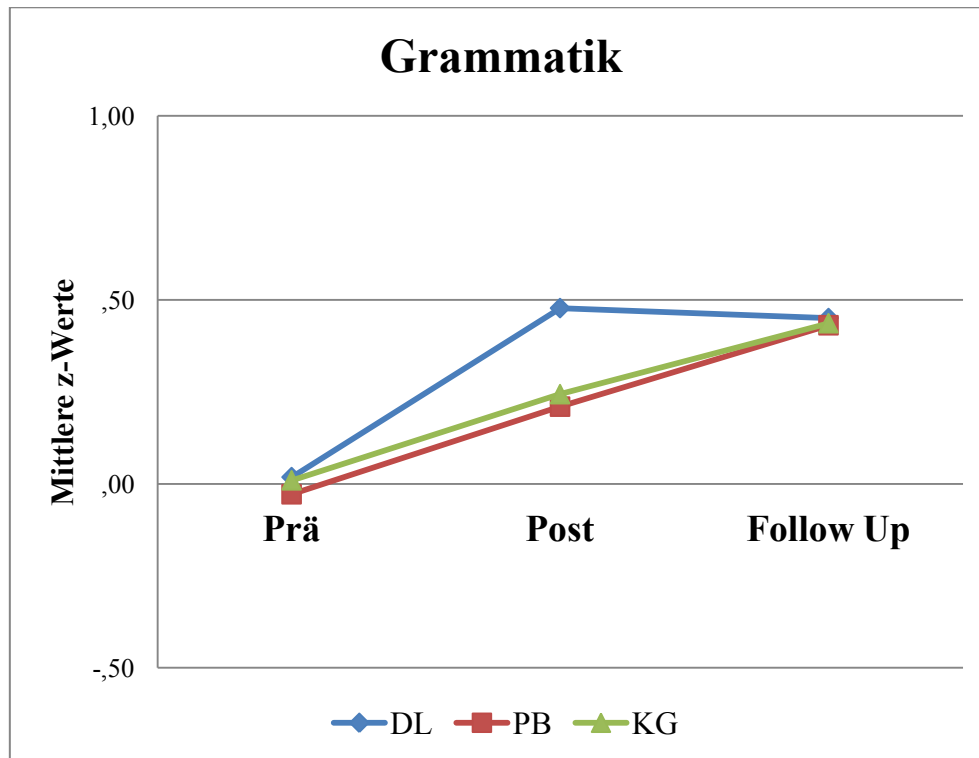


Abbildung 9 Durchschnittliche Leistungen aller Gruppen in der Grammatikleistung im Prä-, Post- und Follow Up-Test.

Im Bereich Grammatik klärte die Vortestleistung 50% der Varianz ($F[1;160] = 160.56$; $p < .01$; $\eta^2 = .50$). Zusätzlich konnten 7% der Varianz im Nachtest durch die Gruppenzugehörigkeit erklärt werden. Tendenziell war die DL-Gruppe der PB-Gruppe und der Kontrollgruppe mit jeweils kleinen Effektstärken überlegen (DL vs. PB $d_{\text{korrr}} = .21$; DL vs. KG $d_{\text{korrr}} = .20$). Allerdings verfehlte der Faktor Gruppe knapp das erforderliche Signifikanz-Niveau von $p < .05$, so dass die gefundenen Gruppenunterschiede nicht statistisch bedeutsam sind ($F[2;160] = 3.04$; $p = .05$; $\eta^2 = .04$). Betrachtet man jedoch die Ergebnisse im Gesamtbild stützen sie die Hypothese, dass Dialogisches Lesen ein breites Spektrum sprachlicher Kompetenzen fördert. Die Intelligenzleistung hatte keinen bedeutsamen Einfluss auf die Posttest-Ergebnisse ($F[1;160] = 1.10$; $p = .30$; $\eta^2 = .01$). Insgesamt konnten mit dem Modell 57% der Varianz im Nachtest aufgeklärt werden ($R^2 = .57$).

Zum Follow Up ergaben sich keine weiteren signifikanten Gruppenunterschiede ($F[1;116] = 6.17; p < .05; \eta^2 = .05$). Wie aus Abbildung 9 hervor geht, konnten die PB-Gruppe und die Kontrollgruppe ihre Grammatikkompetenzen zum Follow Up steigern, während die DL-Gruppe im Vergleich auf dem erreichten Niveau blieb.

46% der Varianz im Follow Up konnte die Vortestleistung erklären ($F[1;116] = 97.70; p < .01; \eta^2 = .46$). Außerdem zeigte sich ein signifikanter Einfluss der Intelligenzleistung ($F[1;116] = 6.17; p < .05; \eta^2 = .05$).

7.3.2.5 Erweiterte Auswertung der Satzbildungsaufgabe

Tabelle 15 fasst die deskriptiven Statistiken, der erweiterten Auswertung des Subtests *Satzbildung* (Petermann et al., 2010) zusammen. Die erweiterte Auswertung sollte die Entwicklung von morphologischen und syntaktischen Kompetenzen, sowie die Entwicklung der Artikelbildung isoliert erfassen.

Tabelle 15 Deskriptive Statistiken der durchschnittlich erreichten Rohpunkte in der erweiterten Auswertung der Satzbildungsaufgabe.

	Prä-Test					Post-Test			Follow Up		
	Min	Max	N	M	SD	N	M	SD	N	M	SD
Gesamtstichprobe											
Morphologie	0	12	156	4.88	3.8	151	5.80	3.88	140	8.75	3.35
Artikelbildung	0	12	156	4.41	3.9	151	5.10	3.96	140	6.78	3.82
Syntax	0	12	156	5.30	4.1	151	6.53	4.20	140	6.31	3.62
Dialogisches Lesen											
Morphologie	0	12	46	5.20	4.04	44	6.68	3.96	42	8.93	3.31
Artikelbildung	0	12	46	4.30	3.86	44	6.27	4.40	42	6.29	3.89
Syntax	0	12	46	5.46	4.00	44	6.61	4.02	42	5.95	3.60
Phonologische Bewusstheit											
Morphologie	0	12	54	4.67	3.55	52	6.19	3.33	48	8.35	3.34
Artikelbildung	0	12	54	4.41	3.86	52	6.27	4.40	48	6.29	3.86
Syntax	0	12	54	5.19	4.01	52	7.33	3.83	48	5.88	3.47
Kontrollgruppe											
Morphologie	0	12	56	4.84	3.97	55	4.73	4.11	50	8.98	3.43
Artikelbildung	0	12	56	4.50	3.97	55	4.44	4.05	50	7.60	3.80
Syntax	0	12	56	5.29	4.28	55	5.71	4.59	50	7.04	3.72

Zur statistischen Analyse der Ergebnisse wurden wie schon in den Abschnitten zuvor ANCOVAs berechnet. Die Ergebnisse werden in den folgenden Abschnitten dargestellt.

7.3.2.5.1 Morphologie

Im Bereich Morphologie ergab sich im Post-Test ein signifikanter Unterschied zwischen den Gruppen ($F[2;146] = 5.36$; $p < .01$; $\eta^2 = .07$). Sowohl das DL ($p < .05$) als auch PB ($p < .05$) waren der KG überlegen. Die korrigierten Effektstärken lagen für beide Gruppen im mittleren Bereich (DL: $d_{\text{kor}} = 0.39$ und PB: $d_{\text{kor}} = 0.43$). Weiter leisteten die Vortestleistung ($F[1;146] = 41.32$; $p < .01$; $\eta^2 = .22$) und die Intelligenz ($F[1;146] = 7.93$; $p < .01$; $\eta^2 = .05$) einen signifikanten Beitrag zur Varianzaufklärung im Nachtest. Insgesamt erbrachte das Modell 35% Varianzaufklärung ($R^2 = .35$).

Bei der Analyse der Follow Up-Ergebnisse blieb der zum Nachtest gefundene Gruppeneffekt nicht erhalten ($F[2;135] = .47$; $p = .63$; $\eta^2 = .01$). Das heißt, vier Monate nach der Förderung unterschieden sich die Leistungen im Bereich Morphologie zwischen den Gruppen nicht mehr signifikant. Tendenziell konnten alle Gruppen ihre Ausgangsleistung im Prätest zum Follow Up um 3 bis 4 Rohpunkte zu verbessern.

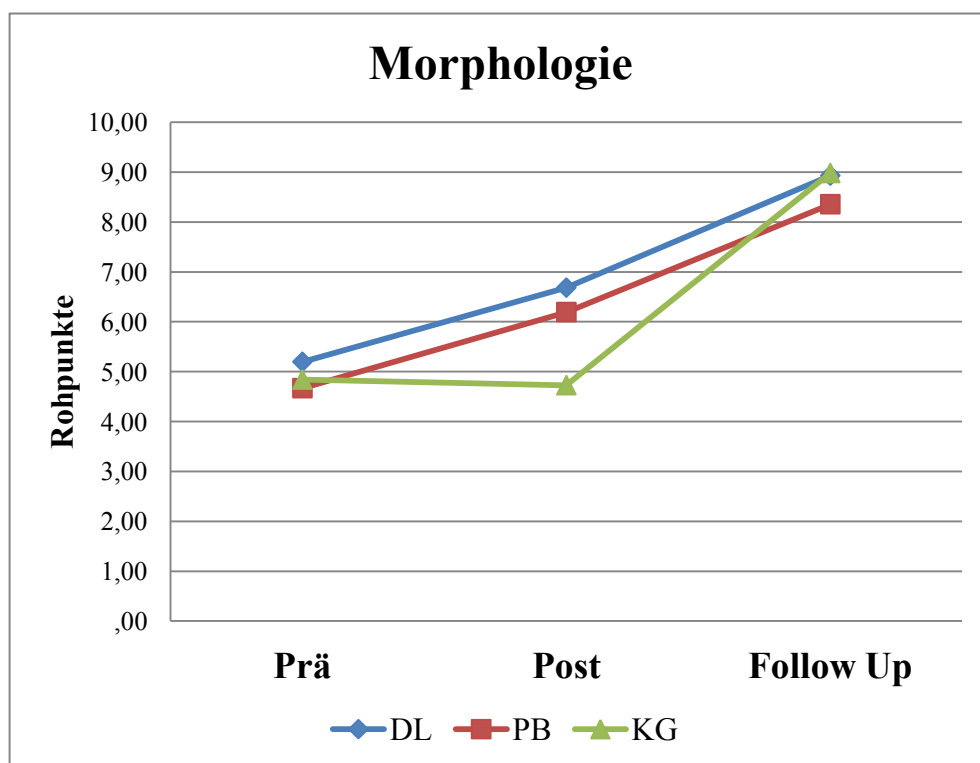


Abbildung 10 Durchschnittliche Leistungen aller Gruppen in „Morphologie“ im Prä-, Post- und Follow Up-Test.

7.3.2.5.2 Artikelbildung

In der Artikelbildung klärte die Vortestleistung in der Prä-Post Analyse 25% der Varianz auf ($F[1;146] = 49.66$; $p < .01$; $\eta^2 = .25$). Außerdem ergab sich ein signifikanter Einfluss der Intelligenzleistung ($F[1;146] = 4.86$; $p < .05$; $\eta^2 = .03$). Auch die Gruppe hatte Einfluss auf die Artikelbildung im Posttest ($F[2;146] = 5.91$; $p < .01$; $\eta^2 = .08$). Die DL war KG und PB überlegen ($p < .05$). Zwischen DL und KG ergab sich eine mittlere Effektstärke von $d_{\text{kor}} = 0.48$. Für die Entwicklung zwischen DL und PB wurde ebenfalls ein mittlerer Effekt gemessen ($d_{\text{kor}} = 0.41$). Langfristig blieb der Trainingseffekt jedoch nicht stabil. Das heißt, zum Follow Up konnten keine Unterschiede zwischen den Untersuchungsgruppen nachgewiesen werden ($F[2;135] = 2.09$; $p = .13$; $\eta^2 = .03$). Die Kovariaten Prätestleistung ($F[1;135] = 22.31$; $p < .01$; $\eta^2 = .14$) und Intelligenz ($F[1;135] = 6.81$; $p < .05$; $\eta^2 = .05$) klärten 19% der Gesamtvarianz auf. Für das im Follow Up angewandte Modell wurde $R^2 = .24$ ermittelt.

Wie aus den Leistungsentwicklungsverläufen in Abbildung 11 ersichtlich wird, ergab sich zum Follow Up eine unerwartete Leistungsverbesserung der Kontrollgruppe. Die Leistungsunterschiede zwischen den Gruppen erwiesen sich allerdings nicht als statistisch bedeutsam.

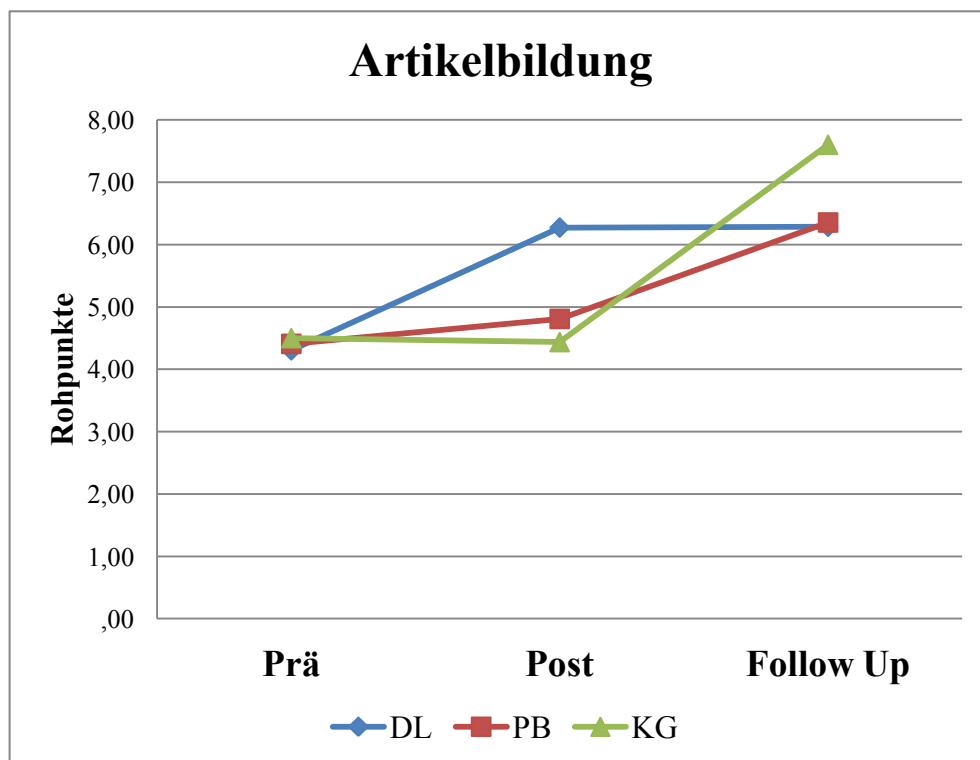


Abbildung 11 Durchschnittliche Leistungen aller Gruppen in „Artikelbildung“ im Prä-, Post- und Follow Up-Test

7.3.2.5.3 Syntax

Im Bereich Syntax lieferte die Vortestleistung mit 29% den größten Beitrag zur Varianzaufklärung ($F[1;146] = 58.30$; $p < .01$; $\eta^2 = .29$) im Nachtest. Auch die Intelligenzleistung hatte einen bedeutsamen Einfluss auf die Nachtestergebnisse ($F[1;146] = 4.40$; $p < .05$; $\eta^2 = .03$). Weiter ergab sich ein signifikanter Effekt der Gruppe ($F[2;146] = 3.26$; $p < .05$; $\eta^2 = .04$). Es zeigte sich, dass Kinder des PB-Trainings besser im Bereich der Syntax abschnitten als Kinder der Kontrollgruppe ($p < .05$). Es ergab sich eine moderate Effektstärke von $d_{\text{kor}} = 0.41$. Auch die DL-Gruppe war der Kontrollgruppe mit einer kleinen Effektstärke ($d_{\text{kor}} = 0.17$) überlegen; allerdings verfehlte der Effekt das statistische Signifikanzniveau. Die unmittelbaren Fördererfolge blieben zum Follow Up nicht stabil. Die Gruppenzugehörigkeit leistete vier Monate nach dem Nachtest keinen substantiellen Beitrag zur Varianzaufklärung in der Syntaxentwicklung ($F[2;135] = 1.81$; $p = .17$; $\eta^2 = .03$). Einzig Vortestleistung ($F[1;135] = 24.02$; $p < .01$; $\eta^2 = .15$) und Intelligenz ($F[1;135] = 4.51$; $p < .05$; $\eta^2 = .03$) konnten auch im Follow Up Varianz erklären.

Wie schon zuvor in der Artikelbildung verschlechterten sich PB und DL im Follow Up tendenziell, während sich die KG überraschender Weise sogar verbessern konnte. Das verwendete Modell konnte im Vortest 34 % der Varianz aufklären. Zum Follow Up lag R^2 bei .18.

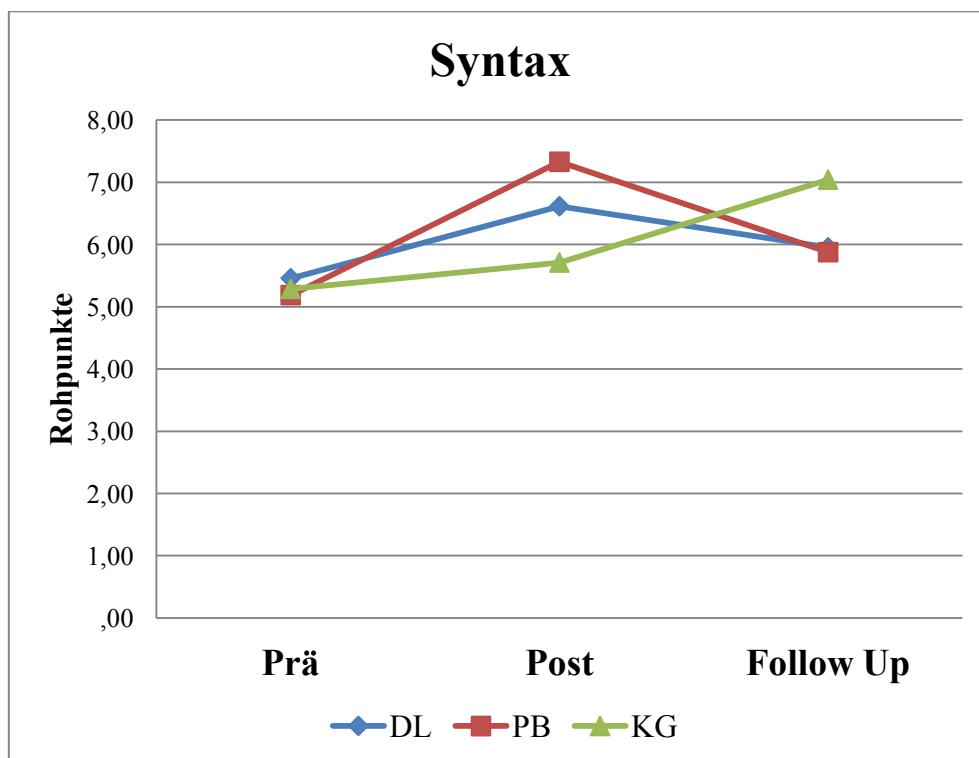


Abbildung 12 Durchschnittliche Leistungen aller Gruppen in „Syntax“ im Prä-, Post- und Follow Up-Test..

7.3.3 Zusammenfassung der kurz- und der langfristigen Trainingseffekte

Die Ergebnisse zeigten, dass sowohl eine Förderung im Sinne des Dialogischen Lesens (DL) als auch ein Training der Phonologischen Bewusstheit (PB) sprachförderliche Effekte bei Risikokindern im letzten Kindergartenjahr erzielten. Das Wirksamkeitsspektrum der beiden Interventionen fiel wie erwartet sehr unterschiedlich aus.

Kinder der PB-Gruppe erzielten die größten Leistungszuwächse in trainingsnahen Aufgaben zur Phonologischen Bewusstheit und waren hier sowohl dem Dialogischen Lesen als auch der Kontrollgruppe überlegen. Dafür waren für Kinder, denen dialogisch vorgelesen wurde, breit gestreute Effekte in mehreren Sprachentwicklungsbereichen zu verzeichnen. So zeigte sich die DL-Gruppe dem PB-Training und der Kontrollgruppe in Sprachentwicklungsbereichen jenseits der Phonologischen Bewusstheit überlegen. Differentielle Analysen ergaben, dass die Überlegenheit der DL-Gruppe vor allem auf Verbesserungen im Sprachverständnis und in der Grammatik zurückzuführen war. Allerdings konnten für die DL-Gruppe insgesamt nur kleine Trainingseffekte festgestellt werden.

Eine differentielle Auswertung des SET-Subtets *Satzbildung* (Petermann et al., 2010) ergab, dass sich Kinder der DL-Gruppe in der Artikelbildung kurzzeitig stärker verbesserten, als die Kinder des PB-Training und der Kontrollgruppe. Die Zuwächse entsprachen mittleren Effektstärken. Auch in der Morphologie ergab sich im Vergleich zur Kontrollgruppe ein mittlerer Effekt für Kinder, denen dialogisch vorgelesen wurde.

Überraschender Weise waren auch für Kinder, die am PB-Training teilnahmen, Effekte in der Grammatik festzustellen. Kinder der PB-Gruppe waren, wie die DL-Gruppe, der Kontrollgruppe in puncto Morphologie bei moderater Effektstärke überlegen. Außerdem zeigte sich das PB-Training in der Syntax gegenüber der Kontrollgruppe mit mittlerer Effektstärke im Vorteil. Zur DL-Gruppe konnte kein Vorsprung festgestellt werden.

Die beschriebenen Fördereffekte ließen sich nur direkt nach der Interventionsphase nachweisen. Zum Follow Up, das drei Monate nach der Einschulung durchgeführt wurde, konnten die Effekte in beiden Experimentalgruppen nicht bestätigt werden. Der pädagogische Wert dieser Effekte wird in Kapitel 8 zu diskutieren sein.

7.4 Einfluss der Durchführungsqualität

Um den Einfluss der Durchführungsqualität auf die sprachliche Entwicklung im Nachtest zu überprüfen, wurden schrittweise Regressionsanalysen berechnet. Abhängige Variablen waren die Nachtestleistungen in den Bereichen *Allgemeine Sprachkompetenz*, *Wortschatz*, *Sprachverständnis* und *Grammatik*. Als mögliche Einflussfaktoren wurden die Durchführungsqualität, die Vortestleistung und die Intelligenz in den Analysen berücksichtigt. Im Folgenden werden die Ergebnisse der schrittweisen Regressionsanalysen tabellarisch dargestellt.

Tabelle 16 zeigt die schrittweise Regression für die abhängige Variable *Allgemeine Sprachkompetenz*. Wie erwartet trug die Vortestleistung im ersten Schritt in besonders hohem Maße zur Varianzklärung bei. Sie erklärte 75% der Gesamtvarianz im Nachtest.

In einem zweiten Schritt wurde die Durchführungsqualität in die Regressionsanalyse einbezogen. Sie konnte zur weiteren Varianzaufklärung beitragen ($\Delta R^2 = 0.02$; $F(1;52) = 4.27$, $p < .05$). Die Hinzunahme der Durchführungsqualität klärte zusätzlich zur Vortestleistung 2% Varianz auf ($t(52) = 2.07$, $p < .05$). Der positive Regressionskoeffizient ($\beta = .14$) sagt aus, dass Kinder von Trainern, denen hohe Durchführungsqualität im Dialogischen Lesens attestiert wurde, bessere Leistungen im Nachtest zeigten als Kinder von Trainern, bei denen eine geringere Durchführungsqualität festgestellt wurde.

Auch wenn es sich dabei nur um eine sehr kleine Einflussgröße handelt, zeigt das Ergebnis, dass Risikokinder etwas bessere Leistungen in den allgemeinen Sprachkompetenzen erbrachten, wenn die Trainer die DL-Prinzipien häufig und konsequent umsetzten.

Die Intelligenzleistung hatte keinen bedeutsamen Einfluss auf die Nachtestleistung.

Tabelle 16 Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung in „Allgemeiner Sprachentwicklung“ im Dialogische Lesen (n = 55).

		Beta	T	p
Schritt 1				
	Vortestleistung	.87	12.58	< .01
Schritt 2				
	Vortestleistung	.84	12.24	< .01
	Durchführungsqualität	.14	2.07	< .05

Anmerkungen: $R^2 = .75$ für Schritt 1; $\Delta R^2 = .02$ für Schritt 2.

In drei weiteren Analysen wurde der Einfluss der Durchführungsqualität auf die Entwicklung im *Wortschatz*, *Sprachverständnis* und *Grammatik* näher untersucht. Tabelle 17 fasst die Ergebnisse der hierarchischen Regressionsanalyse für die Wortschatzentwicklung zusammen.

Tabelle 17 Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung im Wortschatz für Dialogisches Lesen (n = 55).

	Beta	T	p
Schritt 1			
Vortestleistung	.82	10.55	< .01
Schritt 2			
Vortestleistung	.78	10.41	< .01
Durchführungsqualität	.20	2.68	< .05

Anmerkungen: $R^2 = .68$ für Schritt 1; $\Delta R^2 = .04$ für Schritt 2.

Im *Wortschatz* klärte die Vortestleistung im ersten Schritt 68% der Nachtestvarianz auf. Im zweiten Schritt wurde die Durchführungsqualität in den Berechnungen berücksichtigt. Sie hatte ebenfalls einen statistisch bedeutsamen Einfluss ($F(1;52) = 7.20$, $p < .05$). Die Hinzunahme der Durchführungsqualität klärte zusätzlich 4% der Varianz im Nachtest auf ($t(52) = 2.07$, $p < .05$).

Erneut konnte ein positiver Regressionskoeffizient für die Durchführungsqualität festgestellt werden. Das heißt, je höher die Durchführungsqualität in den Gruppen eingeschätzt wurde, desto besser fielen die Nachtestergebnisse in der Wortschatzentwicklung aus.

In den zuvor berichteten Befunden zur Wortschatzentwicklung (vgl. Kap. 7.3.3) waren zwar keine Entwicklungsunterschiede zu Gunsten der DL-Gruppe festzustellen, allerdings deutet der signifikante Einfluss der Durchführungsqualität darauf hin, dass bei durchweg hoher Durchführungsqualität auch Fördereffekte im Wortschatz möglich gewesen wären.

Die Intelligenzleistung erwies sich erneut nicht als bedeutsamer Einflussfaktor.

Die weiteren Tabellen zeigen die Ergebnisse der stufenweisen Regressionsanalysen für die abhängigen Variablen *Sprachverständnis* und *Grammatik* (Tabelle 18 und 19). In beiden Fällen ließ sich kein bedeutsamer Einfluss der Durchführungsqualität nachweisen.

Für *Sprachverständnis* klärte ausschließlich die Vortestleistung statistisch bedeutsam 58% der Varianz im Nachtest auf. Im Bereich *Grammatik* konnte die Vortestleistung 65% der Varianz erklären. Die Intelligenzleistung hatte in beiden Domänen erneut keinen bedeutsamen Einfluss.

Tabelle 18 Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung im Sprachverständnis für Dialogisches Lesen (n = 55).

	Beta	T	p
Schritt 1			
Vortestleistung	.76	8.63	< .01

Anmerkungen: $R^2 = .58$ für Schritt 1

Tabelle 19 Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage des Einflusses der Durchführungsqualität auf die Nachtestleistung im Bereich Grammatik für Dialogisches Lesen (n = 55).

	Beta	T	p
Schritt 1			
Vortestleistung	.81	10.04	< .01

Anmerkungen: $R^2 = .66$ für Schritt 1

Zusammenfassend kann festgestellt werden, dass es sich bei der Durchführungsqualität um einen bedeutsamen Einflussfaktor bei der Erklärung von Fördereffekten handelt. Die vorliegenden Ergebnisse sind im Folgenden zu diskutieren.

8 Diskussion

Die durchgeführte Studie hatte zum Ziel die sprachförderliche Wirkung von Dialogischem Lesen (DL) als Kleingruppen-Förderung bei Risikokindern im letzten Kindergartenjahr zu untersuchen. Da mit Blick auf den internationalen Forschungsstand durchaus breitgestreute Effekte in unterschiedlichen Bereichen der Sprachentwicklung zu erwarten waren, wurden Leistungsveränderungen im Wortschatz, im Sprachverständnis, in der Grammatik und in der Phonologischen Bewusstheit untersucht.

Über ein Sprachscreening, an dem 660 Kindergartenkinder teilnahmen, wurden 174 Risikokinder mit Sprachauffälligkeiten ausgelesen. Sie wurden zufällig zu einer Sprachförderung im Sinne des Dialogischen Lesens, einem Training der Phonologischen Bewusstheit oder einer Kontrollgruppe, die das Sprachförderangebot ihrer Kita erhielt, zugewiesen. Aus organisatorischen Gründen konnte die randomisierte Zuweisung nicht auf Individualebene erfolgen. Förderkinder aus demselben Kindergarten wurden immer gemeinsam zu einer Untersuchungsbedingung zugewiesen. Auf diese Weise konnte die Förderung im gewohnten Umfeld der Kinder stattfinden und potentieller Drop Out durch Mehrbelastung für Eltern und Erzieher vermieden werden.

Die Ergebnisse der Untersuchung bestätigen die positiven internationalen Befunde zum Dialogischen Lesen (z. B. National Early Literacy Panel, 2008). Risikokinder, die am Dialogischen Lesen teilnahmen, zeigten im Nachtest bessere Sprachkompetenzen als Kinder, die ein Training zur Phonologischen Bewusstheit (PB) erhalten hatten und bessere Sprachkompetenzen als die Kinder der Kontrollgruppe (KG). Differenzierte Analysen zeigten, dass sich Kinder, die am Dialogischen Lesen teilnahmen, im Sprachverständnis in grammatikalischen Kompetenzen (Morphologie und Artikelbildung) bedeutsam verbesserten. Kinder, die am PB-Training teilnahmen, erzielten im Nachtest signifikant bessere Leistungen in der Phonologischen Bewusstheit. Überraschenderweise zeigten auch sie bessere Leistungen in der morphologischen Entwicklung als die KG. In der Syntax war das PB-Training dem DL und der KG überlegen. Zum Follow Up blieb keiner der beschriebenen Effekte stabil.

Im Folgenden sind die Befunde für die erhobenen Teilbereiche im Einzelnen zu diskutieren.

8.1 Phonologische Bewusstheit

Im Bereich der Phonologischen Bewusstheit zeigte die PB-Gruppe im Nachtest erwartungsgemäß die stärksten Leistungszuwächse. Sie war der Kontrollgruppe und der DL-Gruppe signifikant überlegen. Es ergaben sich moderate bis große Effektstärken zugunsten des PB-Trainings.

Die kurzfristigen Fördererfolge bestätigen die zahlreichen nationalen und internationalen Befunde zur Wirksamkeit eines PB-Trainings in der Phonologischen Bewusstheit (Bus & van Ijzendoorn, 1999; Schneider, Küspert, Roth, Visé & Marx, 1997).

Der, im Vergleich mit zu DL, deutliche Entwicklungszugewinn der PB-Gruppe in der Phonologischen Bewusstheit stützt die Hypothese, dass ein formbasiertes Sprachförderprogramm, das formalsprachliche Aspekte isoliert und explizit fördert, größere Fördereffekte in diesem speziellen Bereich erzielt, als eine inhaltsfokussierte Fördermethode (Norris & Ortega, 2000).

Anders als die Trainer der PB-Gruppe wurden die Trainer des Dialogischen Lesens nicht explizit auf die Vermittlung von Aspekten der Phonologischen Bewusstheit geschult. Dennoch legten einige internationale Befunde die Vermutung nahe, dass auch durch Dialogisches Lesen Fördereffekte in der Phonologischen Bewusstheit zu erreichen sind (Aram, 2006; Mol, Bus & de Jong, 2009; Swanson et al. 2011). Dies bestätigen die Ergebnisse der eigenen Untersuchung nicht. Im Nachtest schnitten Kinder, die im Sinne des Dialogischen Lesens gefördert wurden, in der Phonologischen Bewusstheit nicht signifikant besser ab als die Kinder der Kontrollgruppe. Die Ursache hierfür kann auf die Untersuchungsbedingungen der Kontrollgruppe zurückzuführen sein. Wie in Kapitel 6.3.3 bereits erwähnt, handelte es sich bei der KG nicht um eine Non-Treatment-Gruppe, da die am konventionellen Sprachförderangebot ihrer Kitas teilnahmen. In einer Befragung vor der Interventionsphase gaben viele Kindergärten an, Elemente aus „Hören, lauschen, lernen“ oder anderen Programm zur Förderung der Phonologischen Bewusstheit im Kindergartenalltag einzusetzen. Eine einheitliche Fördermethode konnte allerdings nicht festgestellt werden. Zur Frage, wie oft und in welchem Umfang Aufgaben zur Phonologischen Bewusstheit in den Kitas angeboten werden, machten die Kindergärten keine genauen Angaben.

Grundsätzlich ist davon auszugehen, dass Elemente aus Trainings zur Phonologischen Bewusstheit im Förderangebot der KG enthalten waren. Dies könnte erklären, warum sich im

Nachtest hier kein signifikanter Vorteil des Dialogischen Lesens im Vergleich mit der KG finden ließ.

Möglicherweise hat auch die Dauer der Intervention eine Rolle gespielt. Mol et al. (2009) fanden in ihrer Metaanalyse von Studien zur Wirksamkeit des Dialogischen Lesens bei Risikokindern im Kindergarten einen signifikanten Zusammenhang von Interventionsdauer und Effekten in der Phonologischen Bewusstheit. Studien, die Dialogisches Lesen 16 Wochen und weniger einsetzen, wiesen durchschnittlich kleinere Effekte in der Phonologischen Bewusstheit auf ($d = 0.21$) als Studien mit einer längeren Interventionsdauer ($d = 0.60$). In der eigenen Untersuchung betrug die Interventionsdauer zwölf Wochen. Effekte des Dialogischen Lesens in der Phonologische Bewusstheit sind also erst bei einer längeren Interventionsdauer zu erwarten. Dies sollte in Folgestudien Berücksichtigung finden.

Das Follow Up zur Mitte des ersten Schulbesuchsjahres konnte, anders als in den Untersuchungen von Schneider und Kollegen (z.B. Roth & Schneider, 2002; Schneider, Roth & Ennemoser, 2000), keine Langzeiteffekte für die PB-Gruppe in der Phonologischen Bewusstheit finden. Dieser Befund könnte darauf zurückzuführen sein, dass in der eigenen Studie eine modifizierte Variante von *Hören, lauschen, lernen* zum Einsatz kam. Im Original erhalten Kinder über 20 Wochen täglich 10 bis 15 Minuten Förderung (Küspert & Schneider, 2006). Da für die eigene Untersuchung ein Interventionszeitraum von 12 Wochen eingeplant war, musste die Struktur des PB-Trainings angepasst werden. Um Vergleichbarkeit zu gewährleisten hatte das Training den selben zeitlichen Umfang wie das Dialogische Lesen. Die Kinder nahmen zwei bis drei Mal wöchentlich für je 30 Minuten an der Förderung teil. Auch wenn sehr darauf geachtet wurde, dass sich der Aufbau des Programms und die Verteilung der Übungen möglichst wenig von der konventionellen Variante unterscheiden (vgl. Kap. 6.3.2.2), ist anzunehmen, dass die originale Anwendung von 10 bis 15 Minuten zu einer höheren Förderintensität führt. Die Einteilung des Lernstoffes in kleine Einheiten entspricht dem Prinzip des *verteilten Lernens* (Buchner, 2012). Vergleichende Untersuchungen zeigten, dass die selben Inhalte, wenn sie durch verteiltes Lernen erworben wurden, länger erinnert werden als Inhalte, die durch massiertes Lernen erworben wurden (Bahrick & Phelps, 1987). Durch die Modifikation des Trainings ist die langzeitliche Entwicklung der PB-Gruppe deshalb nur schwerlich mit den Befunden von Schneider und Kollegen zu vergleichen. Bei einer längeren Interventionsphase hätten voraussichtlich weniger Modifikationen am Trainingsprogramm vorgenommen werden müssen. So hätten möglicherweise auch langfristige Erfolge aufgezeigt werden können.

Im Gegensatz zu den Untersuchungen von Schneider und Kollegen wurden im Follow Up mögliche Transfereffekte auf die Lese-Rechtschreibentwicklung nicht untersucht. Dies sollte in Folgeuntersuchungen Beachtung finden.

Weiter zeigen die Ergebnisse, dass die KG und die DL-Gruppe im Follow Up ihren Rückstand zur PB-Gruppe aufholen konnten. Diese Entwicklung ist allerdings wenig überraschend, wenn man bedenkt, dass die Förderung der Phonologischen Bewusstheit mittlerweile fester Bestandteil des Deutschunterrichts in der 1. und 2. Klasse ist und sich auch in den Bildungsstandards für den Primarbereich etabliert hat (vgl. KMK, 2005). Während sich die PB-Gruppe nach den deutlichen Effekten im Nachtest zum Follow Up nur geringfügiger steigern konnte, verbesserten sich die Leistungen von DL und Kontrollgruppe zu diesem Zeitpunkt möglicherweise stärker, weil Kompetenzen der Phonologischen Bewusstheit in der Schule expliziter gelehrt und gelernt wurden als zuvor im Kindergarten. Die gefunden Deckeneffekte für Entwicklung der PB-Gruppe im Follow Up sprechen dafür, dass die Kinder dieser Gruppe Phonologische Bewusstheit in der 1. Klasse so sicher beherrschten, dass keine Weiterentwicklung mehr abgebildet werden konnte. Allerdings müssen auch die eingesetzten Tests kritisch betrachtet werden. Aufgaben wie das Reimen aus dem BISC von Jansen (2002) haben sich schon im Kindergarten als zu einfach erwiesen. Für Folgeuntersuchungen sollten die Testaufgabenauswahl in der Phonologischen Bewusstheit überarbeitet werden.

Die ausgebliebenen Follow Up-Effekte sind möglicherweise auch auf Beschulungsunterschiede innerhalb der Gruppen zurückzuführen. Die Follow Up-Analysen berücksichtigten die Daten aller geförderten Kinder unabhängig davon, ob sie nach der Förderung regulär ins erste Schuljahr eingeschult wurden oder nicht. Bezogen auf die Einschulung ergaben sich zum Follow Up unerwartete Unterschiede zwischen den Gruppen. Während aus der PB-Gruppe acht Kinder (16,3%) und aus der KG neun Kinder (17,3%) im Follow Up weiterhin ihren Kindergarten oder eine Vorklasse besuchten, waren es in der DL-Gruppe 16 Kinder (30,2%), die nicht regulär eingeschult wurden. Es ist möglich, dass die DL-Gruppe in der Phonologischen Bewusstheit zum Follow Up weniger gut abschnitt, weil in dieser Gruppe durch Zufall deutlich weniger Kinder eingeschult wurden als in den Vergleichsgruppen. Zusätzliche Analysen, bei denen die nicht-eingeschulten Kinder aus den Daten ausgeschlossen wurden, ergaben im Follow Up deutlich geringere Unterschiede zwischen den Gruppen.

8.2 Weitere Sprachentwicklung

Als zentrales Ergebnis der Untersuchung konnte festgestellt werden, dass sich Kinder, die am Dialogischen Lesen teilnahmen, in ihren Sprachkompetenzen verbesserten. Es ergaben sich kleine Effektstärken im Vergleich zur KG ($d_{\text{kor}} = 0.20$) und zum PB-Training ($d_{\text{kor}} = 0.19$). Dass Kinder des Dialogischen Lesens im Nachtest dem PB-Training in allgemeinen Sprachkompetenzen jenseits der Phonologischen Bewusstheit überlegen waren, ist ein Indiz für das sprachförderliche Potential der Methode.

Während sich für das formbasierte Training der Phonologischen Bewusstheit – wie zuvor beschrieben – spezifische Trainingseffekte feststellen ließen, zeigen die Ergebnisse in der Sprachentwicklung, dass von einer inhaltsbasierten Methode im Sinne des Dialogischen Lesens ein breiteres Effektspektrum in unterschiedlichen Sprachentwicklungsbereichen zu erwarten ist.

Die positive Wirkung von DL für die Entwicklung von Sprachkompetenzen steht in Einklang mit aktuellen Forschungsergebnissen internationaler Studien. Swanson et al. (2011) fassten in einer Metanalyse die Ergebnisse aus 23 Interventionsstudien zusammen, in denen ausschließlich Risikokinder entweder mit Dialogischem Lesen nach Whitehurst oder einer vergleichbaren Vorlesemethode gefördert wurden. Die Autoren fanden heraus, dass DL-geförderte Kinder der KG in Sprachentwicklungsmaßen mit einer durchschnittlichen Effektstärke von 0.29 überlegen waren. Die einzelnen Effektstärken hatten eine Range von -0.48 bis 1.79. Auch wenn der Leistungsunterschied zwischen DL- und Kontrollgruppe das statistische Signifikanzniveau knapp verfehlte, werten die Autoren das Ergebnis als Hinweis für das sprachförderliche Potential der Methode.

Die durchschnittliche Intensität der Interventionen bei Swanson und Kollege ist gut mit der eigenen Untersuchung vergleichbar. Die Autoren berichten, dass Lehrer und geschulte Trainer durchschnittlich 29 Sitzungen Dialogisches Lesen durchführten. Die durchschnittliche Länge einer Sitzung lag bei 20 Minuten. Im Vergleich wurden in der eigenen Studie 24 Sitzungen à 30 Minuten durchgeführt. Die leicht höheren Effektstärken bei Swanson und Kollegen könnten auf die geringfügig höhere Intensität der Interventionen zurückzuführen sein.

Durch eine höhere Förderintensität in der eigenen Untersuchung hätte vielleicht auch die tendenzielle Überlegenheit der DL-Gruppe im Follow Up statistisch bestätigt werden können. Folgestudien sollten daher einen längeren Interventionszeitraum wählen.

Auch wären möglicherweise deutlichere Fördereffekte zugunsten des Dialogischen Lesens gefunden worden, wenn die Untersuchung auch jüngere Risikokinder berücksichtigt hätte. Diese Vermutung legen die Ergebnisse einer Metaanalyse von Mol, Bus und de Jong (2009) nahe. Die Autoren fassten 31 Studien zusammen, in denen entweder Dialogisches Lesen nach Whitehurst oder eine ähnliche Methode des interaktiven Vorlesens eingesetzt wurde. Insgesamt stellen sie für Experimentalgruppen mit ausgelesenen Risikokindern einen durchschnittlichen Effekt von 0.57 fest. Bezüglich des Stichprobenalters zeigte sich, dass jüngere Preschool-Kinder durchschnittlich höhere Effektstärken erzielten ($d = 0.60$) als ältere Kindergartenkinder ($d = 0.42$). Mit Blick auf die eigenen Befunde sollten zukünftige Untersuchungen also einen Vergleich von Effekten bei jüngeren und älteren Risikokindern anstreben.

Ferner wäre es interessant die Wirksamkeit des Dialogischen Lesens auch für Nicht-Risikokinder zu evaluieren. Mol und Kollegen konnten in einer Metanalyse von Studien, in denen Eltern ihren Kindern dialogisch vorlasen, zeigen, dass Risikokinder grundsätzlich mehr von Dialogischem Lesen profitierten, als normalentwickelte Kinder (Mol, Bus, de Jong & Smeets, 2008). Lasen geschulte Trainer oder Lehrer ihren Kindern dialogisch vor, waren keine Unterschiede in der Wirksamkeit für Risiko- und Nicht-Risikogruppen zu finden. Risikokinder profitierten tendenziell sogar etwas mehr von der Förderung als Nicht-Risikokinder (Mol, Bus & de Jong, 2009). Diese Befunde wären mit weiterführenden Untersuchungen für den deutschen Sprachraum zu bestätigen.

Einschränkend ist darauf hinzuweisen, dass Metaanalysen zum Dialogischen Lesen vornehmlich Ergebnisse aus dem angloamerikanischen Raum zusammenfassen. Damit sind die Befunde nur bedingt mit der vorliegenden Arbeit vergleichbar, die erstmals die Wirkung des Dialogischen Lesens bei Risikokindern in Deutschland umfassend untersuchte. Trotz eingeschränkter Vergleichbarkeit unterscheiden sich die eigenen Befunde aber nicht grundsätzlich vom internationalen Forschungsstand. Dies spricht dafür, dass der kumulierte Gesamtscore aus Wortschatzentwicklung, Sprachverständnis und Grammatik ein angemessenes Maß zur Erfassung der Sprachentwicklung jenseits der Phonologischen Bewusstheit darstellt. Die differentiellen Effekte im Wortschatz, im Sprachverständnis und in der Grammatik werden im Folgenden ausführlich diskutiert.

8.2.1 Wortschatz

Die Ergebnisse zur Entwicklung im aktiven Wortschatz zeigen, dass alle Gruppen im Nachtest und im Follow Up mehr Wörter in einem allgemeinen Wortschatztest benennen konnten als vor der Förderung.

Überraschenderweise gab es zu keinem Zeitpunkt eine signifikante Überlegenheit von DL oder PB gegenüber der Kontrollgruppe. Dieses Ergebnis scheint zunächst den positiven Befunden von Whitehurst (1988, 1994), Mol, Bus und de Jong (2009) und Swanson et al. (2011) entgegenzustehen. Sie berichten für Dialogisches Lesen moderate bis große Fördereffekte in der Entwicklung des aktiven Wortschatzes. Allerdings beruhen die international berichteten Effekte zumeist auf Ergebnissen aus proximalen Wortschatztests. Das heißt, es wurden ausschließlich Wörter abgefragt, die häufig in den vorgelesenen Bilderbüchern erwähnt wurden.

Coyne, Simmons, Kame'enui und Stoolmiller (2004) untersuchten, ob die Effekte Dialogischen Lesens in einem proximalen Wortschatztest höher ausfielen als in einem allgemeinen und somit eher distalen Wortschatztest. Im Nachtest konnten die Autoren im proximalen Test einen signifikanten Fördereffekt der DL-Gruppe feststellen ($d = 0.74$), während im nicht-proximalen Wortschatztest kein signifikanter Leistungsunterschied zwischen den Gruppen zu finden war.

Mit Blick auf die eigenen Ergebnisse ist es wahrscheinlich, dass die sehr allgemeinen Wortschatzaufgaben des SET 5-10 die Entwicklung des Wortschatzes nicht differenziert genug abbildeten, weil keine inhaltliche Nähe zu den Vorlesegeschichten gegeben war. Möglicherweise hätten mit einem proximalen Messinstrument auch langfristige Fördereffekte durch Dialogisches Lesen festgestellt werden können.

Die Entscheidung, mit Subtests aus dem SET 5-10 distale Wortschatzaufgaben in der Studie zu verwenden, hatte methodische Gründe. Hager (2000) formuliert als Paradigma für vergleichende Evaluationsstudien, dass an das zu untersuchende Experimentalprogramm die gleichen Ziele zu stellen sind, wie an das Alternativprogramm. Durch den Einsatz eines proximalen Wortschatztests, der vor allem Wörter aus dem Dialogischen Lesen berücksichtigt, wäre die PB-Gruppe systematisch benachteiligt gewesen. Auch wäre das Förderpotential von Dialogischem Lesen in Bezug auf die Wortschatzentwicklung möglicherweise überschätzt worden (*teaching to the test*).

Der einheitliche Einsatz der Wortschatzaufgaben aus dem SET 5-10 stellte hingegen sicher, dass die Wortauswahl für alle Gruppen gleichermaßen trainingsfern war. Bedeutsame

Zuwächse hätten als stärkerer Hinweis für die Generalisierbarkeit des wortschatzförderlichen Potentials des Dialogischen Lesens gewertet werden können, als bei einem proximalen Maß. Die Subtests *Bildbenennung* und *Kategoriebildung* aus dem SET 5-10 (vgl. Kap. 6.2.3.1) wurden grundsätzlich als sehr praktikable und ökonomische Aufgabenformate beurteilt. Allerdings besteht bei der Erfassung des Wortschatz das generelle Problem, dass verfügbare Instrumente nur einen extrem kleinen Ausschnitt des kindlichen Wortschatzes prüfen. Folgeuntersuchungen sollten die Entwicklung des aktiven Wortschatzes nicht nur umfänglicher, sondern auch gezielter erheben. Idealerweise sollte die Ermittlung der Wortschatzleistung gleichermaßen proximale und distale Wortfelder berücksichtigen. Bei vergleichende Evaluationsstudien mit zwei oder mehreren Experimentalgruppen sollten im proximalen Wortschatztest Wörter berücksichtigt werden, die für alle Experimentalgruppen gleichermaßen trainingsimmanent sind. Auf diesem Weg lassen sich systematische Vor- und Nachteile zwischen den Gruppen vermeiden.

Ein weiterer Grund für die ausgebliebenen Effekte im Wortschatz könnte die Größe der Fördergruppen gewesen sein. In der eigenen Untersuchung wurde das Dialogische Lesen in Kleingruppen mit drei bis fünf Kindern durchgeführt. In ihrer Metaanalyse finden Mol, Bus und de Jong (2009) einen signifikanten Einfluss der Moderatorvariablen *Interventionsleiter* und *Gruppengröße* auf die Leistungszuwächse im aktiven Wortschatz bei Risikokindern, denen dialogisch vorgelesen wurde. Geförderte Kinder profitierten am meisten, wenn der Experimentator selbst oder eine geschulte Hilfskraft die Förderung anleitete ($d = 1.38$). Kinder in Kleingruppen, in den der Experimentator oder eine geschulte Hilfskraft förderte, profitierten verglichen mit der Einzelförderung durchschnittlich weniger im Wortschatz ($d = 0.73$). Die geringsten Zuwächse zeigten Großgruppen in denen Lehrer Dialogisches Lesen anwendeten ($d = 0.47$). Für weiterführende Untersuchungen ist zu überlegen, ob die Fördergruppen grundsätzlich verkleinert werden können. Da eine individuelle Förderung bei großen Stichproben kaum zu realisieren ist, könnte sich die Gruppengröße auf maximal drei Kinder beschränken. Gegen diesen Gedanken spricht allerdings, dass für viele Praktiker eine solche Förderung im Alltag auf Grund enger Personalschlüssel wahrscheinlich nicht umzusetzen wäre. Etwaige Fördereffekte wären damit untersuchungsspezifisch und ließen keine Verallgemeinerungen für die Praxis zu.

8.2.2 Sprachverständnis

Im Bereich des Sprachverständnisses erzielten Kinder, die am Dialogischen Lesen teilnahmen, im Nachtest signifikant bessere Leistungen als Kinder des PB-Trainings und der Kontrollgruppe. Es ergaben sich kleine bis mittlere Effektstärken (DL vs. KG: $d_{\text{kor}} = 0.26$; DL vs. PB: $d_{\text{kor}} = 0.25$). Die Ergebnisse der Follow Up-Erhebung Mitte der 1. Klasse zeigten, dass PB und KG ihren Rückstand zur DL-Gruppe soweit aufholen konnten, dass kein signifikanter Unterschied zwischen den Gruppen mehr bestand. Die kurzfristigen Fördereffekte stehen in Einklang mit den Ergebnissen des internationalen Forschungsstands.

Swanson und Kollegen (2011) berichten in ihrer Metaanalyse von Studien zum Dialogischen Lesen, dass der Bereich „Comprehension“ in 18 Risikostichproben erfasst wurde. Die Ergebnisse zeigten, dass Dialogisches Lesen den Kontrollgruppen im Sprachverständnis mit einer durchschnittlichen Effektstärke von 0.60 überlegen war. Der Range der gefundenen Effektstärken reichte von -0.52 bis 1.47.

Dass der berichtete Durchschnittseffekt im Sprachverständnis bei Swanson et al. etwas höher ausfiel als in der eigenen Untersuchung, könnte auf die unterschiedliche Operationalisierung der Sprachverständnisvariablen in den berücksichtigten Studien zurückzuführen sein. Häufig wurde das Sprachverständnis über das Nacherzählen von Geschichten (Verhallen, Bus & de Jong, 2006) oder über das Fragenbeantworten erhoben (Morrow, 1988). Die eigene Untersuchung stützte sich derweil auf die SET-Aufgaben *Handlungssequenzen* und *Fragen zum Text* (vgl. Kap. 6.2.3.2).

Während internationale Studien häufig proximale, also trainingsnahe Aufgaben für die Messung des Sprachverständnisses verwendeten, kamen in der eigenen Studie Aufgaben eines allgemeinen Sprachentwicklungstests zum Einsatz, die inhaltlich nicht auf die Förderungen abgestimmt waren. Die gefundenen Effekte für Dialogisches Lesen zeigen, dass geförderte Kinder in Aufgaben, die nicht auf die Trainingsinhalte abgestimmt waren, bessere Sprachverständnisleistungen zeigten, als die Vergleichsgruppen. Dieses Ergebnis bestärkt die Annahme, dass es mit Dialogischem Lesen möglich ist, das Sprachverständnis von Risikokindern umfassend zu fördern. Trotzdem sollten künftige Untersuchungen die Sprachverständnisleistung umfänglich erfassen und dabei auch auf trainingsnahe Aufgabenformate zurückgreifen.

Für die eigene Studie wurde in der Planungsphase zusätzlich zu den SET-Aufgaben ein Test zum Frageverstehen entwickelt. Da sich allerdings im Vortest Deckeneffekte in allen Gruppen ergaben, erwies sich der Test zur Abbildung von Fördereffekten als ungeeignet. Für weiterführende Untersuchungen sollte der Test erweitert und erneut erprobt werden.

Dass sich die gefundenen Effekte zugunsten der DL-Gruppe im Follow Up nicht als stabil erwiesen, ist möglicherweise auf die bereits angesprochene kurze Dauer der Interventionsphase zurückzuführen. Da in allen Gruppen ein deutlicher Entwicklungsfortschritt zu beobachten war, könnte die Entwicklung des Sprachverständnis auch durch Sprachanregungen im Schulunterricht verstärkt worden sein. Es ist anzunehmen, dass alle Förderkinder durch den Schulunterricht neue Spracherfahrungen sammeln konnten und möglicherweise mehr sprachlichen Input und gezielteres Sprachfeedback erhielten als noch im Kindergarten. Das heißt, eine abschließende Aussage über die Langzeitwirkung von Dialogischem Lesen kann nur schwerlich getroffen werden, da das Follow Up im Vergleich zum Nachtest unter stark veränderten Rahmenbedingungen stattfand. Zukünftige Untersuchungen sollten nach Möglichkeit zwei Follow Up-Erhebungen in ihr Studien-Design aufnehmen. Eine erste Follow Up-Erhebung sollte bereits im Kindergarten erfolgen. Eine zweite Erhebung kann in der Schulzeit erfolgen. In der eigenen Untersuchung war es nicht möglich einen früheren Zeitpunkt für das Follow Up zu wählen.

Wie bereits für die vorherigen Sprachentwicklungsbereiche formuliert, sollten Folgeuntersuchungen das sprachförderliche Potential des Dialogischen Lesens auch für jüngere Stichproben untersuchen. Die Metaanalyse von Mol, Bus und de Jong (2009) legt nahe, dass jüngere Kinder in rezeptiven Sprachentwicklungsmaßen mehr von der Intervention profitieren als ältere. Dieser Befund ist für deutschsprachige Stichproben zu bestätigen.

8.2.3 Grammatik

Im Bereich Grammatik zeigte die DL-Gruppe im Nachtest signifikant bessere Leistungen als die Kontrollgruppe und die PB-Gruppe (DL vs. KG: $d_{\text{kor}} = 0.21$ d; DL vs. PB: $d_{\text{kor}} = 0.20$). Dieses Ergebnis stützt die Annahme, dass Risikokinder vom Dialogischen Lesen in mehreren Bereich der Sprachentwicklung (Sprachverständnis und Grammatik) profitieren.

Die kleinen Effektstärken können, wie zuvor erwähnt, auf den relativ kurzen Interventionszeitraum und das hohe Durchschnittsalter der Stichprobe zurückgehen. Außerdem ist zu beachten, dass die Trainer während des Dialogischen Lesens kein formalsprachliches Wissen (z.B. Artikelbildung, Syntax, etc.) vermittelten. Die Fördereffekte in der Grammatik sind daher tatsächlich auf die Modellierungstechniken und impliziten Rückmeldungen zurückzuführen, die Trainer beim Dialogische Lesen einsetzten.

Zur Follow Up-Erhebung ergaben sich keine Unterschiede in den Gruppen. Einige Gründe hierfür (kurzer Interventionszeitraum, Alter der Stichprobe, ausgelesene Risikokinder) wurden bereits diskutiert.

Insgesamt ist festzustellen, dass die kurzfristigen Fördereffekte in der Grammatik mit Blick auf den internationalen Stand der Forschung neue Hinweise bezüglich des Wirkungsspektrums von Dialogischem Lesen liefern. Denn, bislang wurde die Entwicklung Grammatikleistung nur in wenigen Studien berücksichtigt. Hinweise finden sich bei Lonigan und Whitehurst (1998), die in ihrer Untersuchung zeigen konnten, dass Kinder, die von ihren Eltern dialogisch vorgelesen bekamen, häufiger Verben und generell mehr Worte verwendeten, als die Kinder einer Kontrollgruppe. Allerdings kamen in der Untersuchung keine Instrumente zum Einsatz, die feststellen konnten, in welchen Grammatikkompetenzen sich die Kinder genau verbesserten.

Hinweise, dass Dialogisches Lesen die morphologische Entwicklung bei deutschsprachig aufwachsenden Kindern mit Migrationshintergrund fördert, fanden Ennemoser, Kuhl und Pepouna (2013) in einer Pilotstudie. Die Autoren weisen darauf hin, dass weiterführende Belege für diesen Wirkungszusammenhang zu erbringen sind.

Auf Grund dieser vielversprechenden Befunde galt es auch in der eigenen Untersuchung einzelne Teilbereiche der grammatikalischen Entwicklung zu untersuchen und das Wirkungsspektrum des Dialogischen Lesens für Risikokinder genau abzustecken. Informationen zur morphologischen und syntaktischen Entwicklung der Stichprobe sowie für die Entwicklung der Artikelbildung ergaben sich aus einer erweiterten Auswertung der Satzbildungsaufgabe aus dem SET 5-10 (vgl. Kap. 6.2.3.3.3). Geschulte Hilfskräfte beurteilten für jedes Item Kompetenzen in Morphologie und Syntax sowie korrekter Artikelverwendung.

Hinsichtlich der Morphologieentwicklung bestätigte die eigene Untersuchung die Befunde von Ennemoser, Kuhl und Pepouna (2013). Kinder, die am Dialogischen Lesen teilnahmen, verbesserten ihre morphologische Kompetenzen im Nachtest stärker als Kinder der Kontrollgruppe. Der gefundene moderate Effekt ($d_{\text{kor}} = 0.39$) steht in Einklang mit den Befunden von Ennemoser und Kollegen, die einen leicht stärkeren, aber ebenfalls moderaten Effekt im Subtest *Morphologische Regelbildung* aus dem SET-K (Grimm, Aktaş & Frevert, 2001) feststellten ($d_{\text{kor}} = 0.41$).

Überraschenderweise war in der eigenen Untersuchung bei Morphologie im Nachtest auch ein moderater Effekt für das PB-Training festzustellen ($d_{\text{kor}} = 0.43$). Die Leistungen von DL und

PB unterschieden sich nicht bedeutsam voneinander. Möglicherweise konnten Kinder des PB-Trainings in der Morphologie profitieren, weil Übungen zur Phonologischen Bewusstheit häufige Analyse- und Syntheseleistungen erfordern und dadurch auch implizit morphologische Eigenschaften von Wörtern geübt wurden.

Befunde einer Untersuchung, die Lyster (2002) mit norwegischen Kindergartenkindern durchführte, stützen diese Annahme. Sie fand, dass durch ein PB-Training neben den erwarteten Haupteffekten auch Erfolge in morphologischen Kompetenzen zu verzeichnen waren. Indes erzielten Kinder, die ein Training zur Förderung von morphologischen Kompetenzen erhielten, ebenso Effekte in der Phonologischen Bewusstheit.

Casalis und Colé (2009) versuchten Lysters Befunde bei französischen Kindergartenkindern zu bestätigen. Anders als Lyster fanden die Autoren positiven Einfluss eines Trainings der Phonologischen Bewusstheit auf morphologische Fähigkeiten der PB-Trainingsgruppe, aber keinen Einfluss eines Trainings zur Förderung von morphologischen Kompetenzen für die Entwicklung der Phonologischen Bewusstheit. Sie vermuten, dass Segmentationsprozesse, die sich möglicherweise durch ein PB-Training verbesserten, hilfreich für die Bestimmung von morphematischen Manipulationen eines Wortes sein könnten. Allerdings hat umgekehrt die Förderung von morphologischem Regelwissen, nach Meinung der Autoren, keinen Einfluss auf phonologische Segmentationsleistungen (Casalis & Colé, 2009, S. 129).

Für deutschsprachige Stichproben ließen sich auch nach umfangreichen Recherchebemühungen keine vergleichbaren Interventionsstudien finden. Weitere Folgeuntersuchungen sind nötig, um die Befunde zu bestätigen.

Langfristig erwiesen sich die Effekte in der Morphologie nicht als stabil. In der Follow Up-Erhebung war zwischen den Experimentalgruppen und der Kontrollgruppe kein signifikanter Leistungsunterschied mehr festzustellen. Die Ergebnisse zeigten, dass Kinder der Kontrollgruppe nach der Einschulung in die 1. Klasse ihren Rückstand gegenüber der DL-Gruppe und dem PB-Training aufholten. Wie schon für die Follow Up-Ergebnisse in der Phonologischen Bewusstheit diskutiert (vgl. Kap. 8.1), kann auch für die langfristigen Ergebnisse in der Morphologieentwicklung angenommen werden, dass der Beginn der Schulzeit für die sprachliche Weiterentwicklung von Risikokindern von besonderer Bedeutung ist. Mit Blick auf die Befunde von Lyster (2002) und Casalis & Colé (2002) könnte die Leistungssteigerung der Kontrollgruppe in der Morphologie auf die verstärkte Förderung der Phonologischen Bewusstheit ab der 1. Klasse zurückzuführen sein.

Vorteile der Experimentalgruppen, die bereits im Kindergarten sprachlich gefördert wurden, könnten sich in Form von Transfereffekten in der Lese- und Rechtschreibentwicklung

niederschlagen. Um diese Annahme zu bestätigen, sollten weiterführende Studien die schulische Entwicklung der Förderkinder erfassen und mögliche transferierende Effekte von Sprachfördermaßnahmen untersuchen.

Neben Effekten in der Morphologie ergaben sich in der eigenen Untersuchung auch Effekte in der Artikelbildung. Kinder, die am Dialogischen Lesen teilnahmen, erzielten dabei im Nachtest signifikant bessere Leistungen als die Kontrollgruppe ($d_{\text{korrr}} = 0.48$) und als die PB-Trainingsgruppe ($d_{\text{korrr}} = 0.41$). Da die Artikelbildung einen besonderen grammatikalischen Aspekt der deutschen Sprache darstellt, können nur schwerlich internationale Befunde zum Vergleich herangezogen werden. Nach bisherigem Kenntnisstand liegen bislang keine weiteren Untersuchungen vor, die das förderliche Potential von Dialogischem Lesen für die Artikelbildung im Deutschen näher untersuchten.

Der positive Effekt von Dialogischem Lesen auf die Artikelbildung kann möglicherweise darauf zurückgeführt werden, dass geförderte Kinder durch das Nacherzählen der Geschichten häufiger auf die korrekte Verwendung von Artikeln achten mussten, als Kinder in der PB-Gruppe. Durch die Modellierungs- und Feedbacktechniken des Dialogischen Lesens haben Kinder in der DL-Gruppe häufiger implizite Rückmeldungen zur grammatikalischen Qualität der Aussage erhalten, als Kinder im PB-Training. Außerdem forderten die Trainer der DL-Gruppe ihre Kinder nach einer impliziten Korrektur dazu auf, den modellierten Satz zu wiederholen. Daher ist anzunehmen, dass der hochfrequente Einsatz der Modellierungs- und Feedbacktechniken beim Dialogischen Lesen für die positiven Effekte in der Artikelbildung verantwortlich ist.

Da die Entwicklung der Artikelbildung nach Meinung einiger Autoren vermutlich in engem Zusammenhang mit der Entwicklung morphologischer Kompetenzen steht (z.B. Szagun, 2006), ist mit Blick auf die bereits diskutierten Effekte festzustellen, dass es gelang Kindern durch Dialogisches Vorlesen tiefere Einsichten in die morphologische Struktur des Deutschen zu vermitteln. Da zu diesem Befund bislang keine weiteren Studienergebnisse vorliegen, sollten Folgeuntersuchungen das Effektspektrum des Dialogischen Lesens in Bezug auf die Förderung grammatikalischer Kompetenzen weiter untersuchen.

Zum Follow Up blieben die Fördereffekte in der Artikelbildung nicht stabil. Es ergaben sich keine signifikanten Unterschiede zwischen den Gruppen, wobei erneut eine deutliche Verbesserung der Kontrollgruppe gegenüber den Leistungen der DL-Gruppe und PB-Gruppe verzeichnet werden konnte. Möglicherweise wirkte sich an dieser Stelle der bereits erwähnte

hohe Anteil von nicht-ingeschulten Kindern auf die Leistungsentwicklung der DL-Gruppe aus. In Nachfolgestudien sollte der Testungszeitraum deshalb so geplant werden, dass nach Möglichkeit eine erste Follow Up-Untersuchung bereits im Kindergarten durchgeführt wird. Auch sollte in zukünftigen Studien der Interventionszeitraum erhöht werden, um mögliche Langzeiteffekte auch nach der Einschulung nachweisen zu können.

Als letzter grammatikalischer Teilbereich wurden die Entwicklungen in der Syntax analysiert. Sowohl die DL-Gruppe als auch das PB-Training erzielten im Nachtest bessere Leistungen als die KG. Dabei ergab sich für das PB-Training ein signifikanter Unterschied im Vergleich zur Kontrollgruppe ($d_{\text{kor}} = 0.41$), während der Unterschied zwischen DL und Kontrollgruppe das .05-Signifikanzniveau verfehlte ($d_{\text{kor}} = 0.16$).

Der positive Einfluss des PB-Trainings auf die syntaktischen Kompetenzen der PB-Gruppe lässt sich dadurch erklären, dass die Förderkinder in der modifizierten Variante des PB-Trainings Übungsaufgaben aus dem Bereich „Sätze und Wörter“ häufiger wiederholten als es im Original vorgesehen war. Es ist möglich, dass dadurch syntaktische Kompetenzen im PB-Training gezielter geübt wurden. Dies festigt auch die Annahme, dass ein formbasiertes Sprachförderprogramm größere Fördereffekte in den spezifisch geförderten Teilbereichen der Sprachentwicklung erzielt, als eine inhaltsbasierte Fördermethode wie Dialogisches Lesen. Nationalen und internationalen Forschungsarbeiten wiesen auf den Zusammenhang zwischen Phonologischer Bewusstheit und anderen Sprachentwicklungsbereichen hin (z.B. Grimm 1995; Scarborough, 2001). Marx, Weber & Schneider (2005a) beobachteten bei Kindern mit einer Sprachentwicklungsstörung deutlich größere Schwierigkeiten im Erwerb der Phonologischen Bewusstheit als bei normalentwickelten Kindern. Die Frage, ob ein Training der Phonologischen Bewusstheit von Risikokindern weitere Sprachkompetenzen oberhalb der phonologischen Ebene fördern kann, hat in der Forschung bislang kaum Beachtung gefunden. Die Befunde im Bereich der Syntax sind dahingehend erste Hinweise, die in Folgestudien Bestätigung finden sollten.

Langfristig blieben auch die Syntax-Effekte nicht stabil. Wie bei den zuvor diskutierten Befunden, zeigte sich auch bei den Syntax-Ergebnissen, dass die KG grundsätzlich stärkere Verbesserungen erzielt als DL und die PB-Gruppe.

Aus statistischer Sicht war für die Syntaxvariable im Follow Up sehr geringe Varianzaufklärung zu beobachten ($R^2 = .18$). Dies bedeutet, dass die eingeschlossenen Kovariaten einen Großteil der Varianz im Follow Up nicht erklärten. Möglicherweise spielen

auch externe Einflussfaktoren wie der sozio-ökonomische Status der Förderkinder für die langfristige Entwicklung von Sprachkompetenzen eine nicht zu unterschätzende Rolle. Sie sollten in weiteren Untersuchungen verstärkt kontrolliert werden.

8.3 Einfluss der Durchführungsqualität

Neben der Wirksamkeitsprüfung von Dialogischem Lesen ist die vorliegende Untersuchung der Frage nachgegangen, in welchem Maße die Fördereffekte des Dialogischen Lesens auf die Qualität der Durchführung zurückzuführen sind. Zur Erfassung der Durchführungsqualität wurde ein Beobachtungsbogen entwickelt und zur Beurteilung von einzelnen Fördersitzungen eingesetzt. Insgesamt waren neun geschulte studentische Hilfskräfte für die Durchführung des Dialogischen Lesens verantwortlich. Zwei externe Rater beurteilten die Qualität der Fördersitzungen.

Für die allgemeine Sprachentwicklung konnte festgestellt werden, dass eine hohe Durchführungsqualität während der Förderung einen positiven Einfluss auf die Entwicklung im Nachtest hatte. Insgesamt konnte die Durchführungsqualität 2% der Varianz im Nachtest zusätzlich zum Einfluss der Vortestleistung (75%) erklären.

Ein signifikanter Einfluss ergab sich auch in der Wortschatzentwicklung. Hier konnte die Durchführungsqualität mit 4% zur Varianzaufklärung beitragen, während die Vortestleistung 68% erklärte. Für die Entwicklung des Sprachverständnis und der Grammatik wurde kein bedeutsamer Einfluss der Durchführungsqualität festgestellt.

Die Ergebnisse zeigen, dass es im Zuge der Untersuchung gelungen ist, die Durchführungsqualität bei der Durchführung des Dialogischen Lesens zu erfassen. Allerdings sollten die Ergebnisse lediglich als erste Hinweise verstanden werden, da eine umfassende Validierung des Messinstruments noch nicht vorgenommen werden konnte. In Folgeuntersuchungen sollte der Beobachtungsbogen weiterentwickelt und im Besonderen die Konstruktvalidität an vergleichbaren Messinstrumenten geprüft werden. Dies war zum Zeitpunkt der Untersuchung nicht möglich, weil bisher keine vergleichbaren Beobachtungsbögen zur Erfassung der Durchführungsqualität vorlagen.

In angloamerikanischen Studien wurde die Durchführungsqualität des Dialogischen Lesens zumeist bei Eltern erfasst, die ihrem Kind in einer Einzelsituation vorlasen (vgl. Kap. 3.2.1). Die Ergebnisse zeigen, dass sich der eigene Beurteilungsbogen auch zum Einsatz eignet,

wenn Dialogisches Lesen im Rahmen einer institutionellen Kleingruppenförderung durchgeführt wird. Natürlich wäre es denkbar, den Bogen auch im Rahmen einer Einzelförderung und als Coachingmaßnahme für Eltern einzusetzen.

Insgesamt erwies sich das verwendete Globalmaß aus Ratingitems, Zählitems und die Benotung der Fördersitzung als guter Indikator zur Beurteilung der Durchführungsqualität. Bei Betrachtung der Einzelskalen zeigte sich, dass sich die Durchführungsqualität besser über die Zählitems und die Benotung der Fördersitzung erfassen ließ, als über die Ratingitems. Möglicherweise besaß die entwickelte Ratingskala nicht genügend Trennschärfe und erzeugte somit zu wenig Varianz in den Raterurteilen.

Da die vorliegende Arbeit lediglich erste Impulse für die Entwicklung eines Beurteilungsbogens zur Erfassung der Durchführungsqualität beim Dialogischen Lesen geben kann, sind weiterführende Untersuchungen unbedingt notwendig. Besonders die Frage, mit welchen Itemformaten sich die Durchführungsqualität zuverlässig erfassen lässt, sollte in zukünftigen Arbeiten intensiver diskutiert werden. Dabei wäre zu beachten, dass auch trainingsferne Faktoren, wie die Motivation von Trainern und Kindern sowie der Umgang mit Störungen und institutionelle Voraussetzungen möglicherweise einen bedeutsamen Einfluss auf die Qualität einer Förderung haben können. Diese Aspekte sollten bei einer möglichen Überarbeitung des vorliegenden Bogens stärker beachtet werden.

Auch für die Implementation des Dialogischen Lesens in die pädagogische Praxis kann die Entwicklung eines Instrumentes zur Beurteilung der Durchführungsqualität, wie hier geschehen, von besonderem Wert sein. Die Ergebnisse zeigen, dass sich die Qualität bei der Durchführung des Dialogischen Lesens trotz gut geschulter Hilfskräfte und gut kontrollierten Untersuchungsbedingungen bedeutsam unterscheiden kann. Kinder aus Gruppen, in denen Trainer die DL-Prinzipien besser umsetzten, profitierten in der allgemeinen Sprachentwicklung und im Wortschatz mehr, als Kinder aus Gruppen mit geringerer Durchführungsqualität. So hätten sich bei noch höherer Durchführungsqualität möglicherweise auch signifikante Fördereffekte in der Wortschatzentwicklung zum Nachtest feststellen lassen. In der pädagogischen Praxis sind weitaus höhere Qualitätsunterscheide bei der Durchführung des Dialogischen Lesens zu erwarten, als in den kontrollierten Bedingungen der eigenen Untersuchung. Deshalb sind für eine gute Implementation der Methode umfassende Schulungs- und Supervisionsmaßangebote für Erzieherinnen und Erzieher von Nöten. Der entwickelte Beobachtungsbogen könnte im Bereich der Supervision als hilfreiches

Instrument zum Einsatz kommen, um Erzieherinnen und Erziehern Rückmeldungen zur praktischen Umsetzung des Dialogischen Lesens zu geben. Weitere Studien sollten sich deshalb mit der Konzeption von Erzieher-Coachings befassen.

Da die sprachförderlichen Prinzipien des Dialogischen Lesens nicht ausschließlich beim Vorlesen, sondern praktisch in jeder Alltagssituation von Erziehern eingesetzt werden können, bietet der entwickelte Beobachtungsbogen auch die Möglichkeit, die Qualität von alltagsimmanenten Sprachförderprinzipien zu beurteilen. Auch dies könnte ein Vorhaben für weitere Untersuchungen sein.

Fazit und Ausblick

In der vorliegenden Arbeit wurde das förderliche Potential des Dialogischen Lesens (DL) als bedeutungsfokussierte Sprachfördermethode (*focus on meaning*) bei Risikokindern in Kindertagesstätten untersucht. Dabei wurden die Fördereffekte des DL mit den Effekten eines Trainings der Phonologischen Bewusstheit (*focus on form*) und einer Kontrollgruppe, die am konventionellen Sprachförderangebot der Kindergärten teilnahm, verglichen. Externe Rater beurteilten außerdem die Durchführungsqualität im Dialogischen Lesen (Treatmentvalidität).

Die Ergebnisse zeigen, dass Risikokinder durch eine Förderung im Sinne des Dialogischen Lesens kurzzeitig ihre sprachlichen Kompetenzen verbessern konnten. In Bezug auf das Wirksamkeitsspektrum konnten für DL geförderte Kinder breitgestreute Effekte in unterschiedlichen Sprachentwicklungsbereichen nachgewiesen werden. Sie zeigten verbesserte Leistungen im Sprachverständnis und in der Grammatik. Weiterführende Analysen der Grammatikentwicklung ergaben, dass DL-Kinder besonders in der Morphologie und in der Artikelbildung profitieren konnten.

Die Ergebnisse der Untersuchung stehen in Einklang mit internationalen Forschungsbefunden (National Early Literacy Panel, 2008). Sie stützen die Hypothese, dass Risikokinder durch den Einsatz einer bedeutungsfokussierten Sprachfördermethode in unterschiedlichen Sprachentwicklungsbereichen profitieren konnten. Erzielt wurden zumeist kleine Effektstärken.

Kinder, die an einem formfokussierten Training der Phonologischen Bewusstheit (PB) teilgenommen hatten, waren der DL-Gruppe im Bereich der Phonologischen Bewusstheit überlegen. Gemessen wurde eine moderate Effektstärke. Neben diesem trainingsspezifischen Effekt ergaben sich für die PB-Gruppe keine breitgestreuten Fördereffekte in der Sprachentwicklung. Allerdings zeigte sich die PB-Gruppe kurzzeitig in ihren morpho-syntaktischen Fähigkeiten verbessert.

Die Ergebnisse der PB-Gruppe bestätigen die Annahme, dass mit einer formfokussierenden Sprachfördermethode relativ große Fördereffekte in trainingsnahen Entwicklungsbereichen erzielt werden können, während von einer bedeutungsfokussierten Maßnahme wie dem

Dialogischen Lesen ein eher breites Wirkungsspektrum mit kleineren Effekten in verschiedenen Bereichen der Sprachentwicklung zu erwarten ist.

In weiteren Analysen wurde die Bedeutung der Durchführungsqualität für die Wirksamkeit von Dialogischem Lesen untersucht. Es ergab sich, dass DL geförderte Kinder größere Leistungszuwächse in der allgemeinen Sprachentwicklung und im Wortschatz erzielten, wenn die Durchführungsqualität in den DL-Gruppen von externen Ratern als hoch bewertet wurde.

Die Ergebnisse bestätigen, dass die Durchführungsqualität eine bedeutsame Einflussgröße für die Wirksamkeit des Dialogischen Lesens darstellt. Um in zukünftigen Untersuchungen größere Fördereffekte erzielen zu können, sollten Coachings zum Dialogischen Lesen verbessert werden. Durch ein gezieltes Supervisionsangebot während der Interventionsphase bei dem den Trainern gezielte Rückmeldungen zur Umsetzung der Methode gegeben werden, könnte die Durchführungsqualität besser kontrolliert werden. Auch sollte der eingesetzte Beobachtungsbogen zur Beurteilung der Durchführungsqualität weiter überarbeitet und um Items ergänzt werden, die den Einfluss der äußeren Rahmenbedingungen unter denen eine Förderung stattfindet, noch besser erfassen.

Da sich in der vorliegenden Untersuchung sowohl ein bedeutungsfokussierter Sprachförderansatz als auch ein formfokussiertes Training als wirksame Interventionsmaßnahmen erwiesen haben, sollten weitere Studien die Wirksamkeit einer kombinierten Sprachfördermaßnahme evaluieren. Im Rahmen des Dialogischen Lesens könnten zum Beispiel gezielte Übungen zur Förderung der Phonologischen Bewusstheit oder des grammatikalischen Regelwissens eingebaut werden. Internationalen Studien berichten bereits positive Effekte für die kombinierte Anwendung des Dialogischen Lesens mit gleichzeitiger Förderung von schriftsprachlichen Vorläuferkompetenzen (Aram, 2006; Mol et al., 2009; Swanson et al., 2011). Für deutschsprachige Stichproben stehen vergleichbare Untersuchungen bislang noch aus.

Literaturverzeichnis

- Aram, D. (2006). Early literacy interventions: The relative roles of storybook reading, alphabetic activities, and their combination. *Reading & Writing*, 19 (5), 489-515.
- Arnold, D. H., Lonigan, C. J., Whitehurst, G. J. & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology*, 86 (2), 235-243.
- Baden-Württemberg-Stiftung (Hrsg.). (2011). *"Sag mal was" - Sprachförderung für Vorschulkinder: Zur Evaluation des Programms der Baden-Württemberg Stiftung. Sprachförderung im Spannungsfeld zwischen Wissenschaft und Praxis* (Schriftenreihe der Baden-Württemberg Stiftung, Bd. 57, 1. Aufl.). Tübingen: Francke, A.
- Bahrack, H. P. & Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (2), 344-349.
- Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Hrsg.), *Joint attention. Its origins and role in development* (S. 131–158). Hillsdale, N.J: Lawrence Erlbaum Associates.
- Barrett, M. (2001). An introduction to the nature of language and to the central themes and issues in the study of language development. In M. Barrett (Hrsg.), *The development of language* (Studies in developmental psychology, S. 1–24). Hove: Psychology Press.
- Barrett, M. (Hrsg.). (2001). *The development of language* (Studies in developmental psychology). Hove: Psychology Press.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P. S., Reznick, J. S. et al. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, 21 (1), 85-123.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W. et al. (Hrsg.). (2001). *Pisa 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (Rev. Nachdr. der Erstausg.). Opladen: Leske + Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.). (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verl. für Sozialwiss.
- Berkel, C., Mauricio, A. M., Schoenfelder, E. & Sandler, I. N. (2011). Putting the Pieces Together: An Integrated Model of Program Implementation. *Prev Sci*, 12 (1), 23-33.

- Bickman, L. & Athay, M. (2009). The Worst of all Possible Program Evaluation Outcomes. In A. R. Stiffman (Hrsg.), *The Field Research Survival Guide* (S. 174–204). Oxford University Press.
- Bickman, L., Riemer, M., Brown, J. L., Jones, S. M., Flay, B. R., Li, K.-K. et al. (2009). Approaches to Measuring Implementation Fidelity in School-Based Program Evaluations. *Journal of Research in Character Education*, 7 (2), 75-101.
- Bijeljic-Babic, R., Bertoncini, J. & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, 29 (4), 711-721.
- Bisiacchi, P. S. & Kaufmann, L. (2008). *Neuropsychologisches Screening für 5-11-jährige: BVN/NPS 5-11*. Gardolo: Ed. Erickson.
- Blewitt, P., Rump, K. M., Shealy, S. E. & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, 101 (2), 294-304.
- Bortz, J. & Schuster, C. (2010). Statistik für Human- und Sozialwissenschaftler. *Statistik für Human- und Sozialwissenschaftler*.
- Brain, W. R. (1967). *Speech disorders: Aphasia, apraxia and agnosia* (2. Aufl.). London: Butterworths.
- Brandl, M. (2010). *Die kindliche Sprachentwicklung*, Autonome Provinz Bozen-Südtirol Abteilung 24 - Familie und Sozialwesen Familienservicestelle. Verfügbar unter www.provinz.bz.it/sozialwesen/download/sprache.pdf.
- Brannon, D. & Dauksas, L. (2012a). Studying the Impact of Dialogic Training on Mothers with Lower Education Levels. *National Teacher Education Journal*, 5 (4), 9-19.
- Brannon, D. & Dauksas, L. (2012b). Studying the Effect Dialogic Reading has on Family Members' Verbal Interactions During Shared Reading. *SRATE Journal*, 21 (2), 9-20.
- Bruner, J. S. (1983). *Child's talk: Learning to use language* (1. Aufl.). New York NY u.a: Norton.
- Brunner, M. & Schöler, H. (2002). *HASE - Heidelberger Auditives Screening in der Einschulungsuntersuchung*. Wertingen: Westra Elektroakustik.
- Buchner, A. (2012). Funktionen und Modelle des Gedächtnisses. In H.-O. Karnath & P. Thier (Hrsg.), *Kognitive Neurowissenschaften* (Springer-Lehrbuch, S. 541-551). Springer Berlin Heidelberg.

- Bullheller, S. & Häcker, H. (2002). *Raven's Coloured Progressive Matrices (CPM): Deutsche Version*.
- Bus, A. G. & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91 (3), 403-414.
- Buschmann, A. (2009). *Heidelberger Elterntraining zur frühen Sprachförderung: Trainermanual* (1. Aufl.). München: Urban & Fischer.
- Buschmann, A. & Jooss, B. (2011). Effektivität eines sprachbasierten Interaktionstrainings für pädagogisches Fachpersonal. *Verhaltenstherapie & psychosoziale Praxis*, 43 (2), 303-312.
- Carpenter, M., Nagell, K. & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr Soc Res Child Dev*, 63 (4), i-vi, 1-143.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, New York: Cambridge University Press.
- Casalis, S. & Cole, P. (2009). On the relationship between morphological and phonological awareness: Effects of training in kindergarten and in first-grade reading. *First Language*, 29 (1), 113-142.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use* (Convergence). New York: Praeger.
- Chomsky, N. (1988). *Sprache und Geist: Mit einem Anhang: Linguistik und Politik* (Suhrkamp Taschenbuch Wissenschaft, Bd. 19, 4. Aufl.). Frankfurt am Main: Suhrkamp.
- Christophe, A., Mehler, J. & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, 2 (3), 385-394.
- Clahsen, H. (1982). *Spracherwerb in der Kindheit*, Univ. Tübingen, Wuppertal
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Auflage). New York, N.Y: Psychology Press.
- Coyne, M. D., Simmons, D. C., Kame'enui, E. J. & Stoolmiller, M. (2004). Teaching Vocabulary During Shared Storybook Readings: An Examination of Differential Effects. *Exceptionality*, 12 (3), 145-162.
- Crystal, D. (2010). *The Cambridge encyclopedia of language* (3. Aufl.). Cambridge: Cambridge Univ. Press.

- Cutler, A. (1999). Prosodische Struktur und Worterkennung. In A. D. Friederici (Hrsg.), *Sprachrezeption* (Enzyklopädie der Psychologie Theorie und Forschung Sprache, Bd. 2, S. 49–84). Göttingen: Hogrefe Verl. für Psychologie.
- Dane, A. V. & Schneider, B. H. (1998). Program integrity in primary and early secondary Prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18 (1), 23-45.
- Dannenbauer, F. M. (2002). Grammatik. In S. Baumgartner & I. Füssenich (Hrsg.), *Sprachtherapie mit Kindern. Grundlagen und Verfahren* (UTB für Wissenschaft Uni-Taschenbücher, Bd. 8188, 5. Aufl., S. 256–319). München: Reinhardt.
- Darsow, A., Paetsch, J., Stanat, P. & Felbrich, A. (2012). Ansätze der Zweitsprachförderung: Eine Systematisierung. *Unterrichtswissenschaft*, 40 (1), 64-82.
- DeCasper, A. J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C. & Maugeais, R. (1994). Fetal Reactions to Recurrent Maternal Speech. *Infant Behavior and Development*, 17 (2), 159-164.
- DeKeyser, R. M. (1995). Learning Second Language Grammar Rules. *Stud. Sec. Lang. Acq.*, 17 (03), 379.
- Deutsch, W. & El Mogharbel, C. (2007). Thema Sprachentwicklung: Ein einführender Rundblick. In H. Schöler & A. Welling (Hrsg.), *Sonderpädagogik der Sprache* (Handbuch Sonderpädagogik, Bd. 1, S. 5–18). Göttingen: Hogrefe.
- Deutsch, W. & El Mogharbel, C. (2007). Von der Stimme zur Sprache: Die Ontogenese von Phonetik, Phonologie und Prosodie. In H. Schöler & A. Welling (Hrsg.), *Sonderpädagogik der Sprache* (Handbuch Sonderpädagogik, Bd. 1, S. 19–28). Göttingen: Hogrefe.
- Dilling, H. & Schulte-Markwort, E. (Hrsg.). (2014). *Internationale Klassifikation psychischer Störungen: ICD-10 Kapitel V (F) klinisch-diagnostische Leitlinien* (9. Aufl., unter Berücks. der Änderungen entsprechend ICD-10-GM 2014). Bern: Huber.
- Dimitrov, D. M. & Rumrill, P. D. (2003). Pretest-posttest designs and measurement of change. *Work*, 20 (2), 159-165.
- Dromi, E. (2001). Early lexical development. In M. Barrett (Hrsg.), *The development of language* (Studies in developmental psychology, S. 99–131). Hove: Psychology Press.

- Durlak, J. A. & DuPre, E. P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *Am J Community Psychol*, 41 (3-4), 327-350.
- Dusenbury, L., Brannigan, R., Falco, M. & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res*, 18 (2), 237-256.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. (1971). Speech perception in infants. *Science*, 171 (3968), 303-306.
- El Mogharbel, C. & Deutsch, W. (2007). Pargmatik: Sprachentwicklung im Kontext sozialen Handelns. In H. Schöler & A. Welling (Hrsg.), *Sonderpädagogik der Sprache* (Handbuch Sonderpädagogik, Bd. 1, S. 57–67). Göttingen: Hogrefe.
- Ellis, N. C. (Hrsg.). (1994). *Implicit and explicit learning of languages*. London: Academic Press.
- Engels, Köller & Koopmans. (2011). *Zweiter Integrationsindikatorenbericht: Erstellt für die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration*, Bundesregierung. Zugriff am 11.08.2015. Verfügbar unter http://www.bundesregierung.de/Content/DE/Publikation/IB/2012-01-12-zweiter-indikatorenbericht.pdf?__blob=publicationFile.
- Ennemoser, M., Kuhl, J. & Pepouna, S. (2013). Evaluation des Dialogischen Lesens zur Sprachförderung bei Kindern mit Migrationshintergrund*. *Zeitschrift für Pädagogische Psychologie*, 27 (4), 229-239.
- Ennemoser, M., Marx, P., Weber, J. & Schneider, W. (2012). Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44 (2), 53-67.
- Ennemoser, M., Pepouna, S. & Hartung, N. (2012). Kulturfaire und prognostisch valide Erfassung des phonologischen Arbeitsgedächtnisses bei Kindern mit Migrationshintergrund. *Unterrichtswissenschaft*, 40 (1), 26-46.
- Esser, H. (2011). Was ist praktischer als eine gute Theorie? Anmerkungen zur wissenschaftlichen Begründung und Evaluation von praktischen Maßnahmen (nicht nur) im Bereich der Modellversuche zur Sprachförderung. In Baden-Württemberg-Stiftung (Hrsg.), *"Sag mal was" - Sprachförderung für Vorschulkinder. Zur Evaluation des*

- Programms der Baden-Württemberg Stiftung. Sprachförderung im Spannungsfeld zwischen Wissenschaft und Praxis* (Schriftenreihe der Baden-Württemberg Stiftung, Bd. 57, 1. Aufl., S. 49–60). Tübingen: Francke, A.
- Ferguson, C. A. (1964). Baby Talk in Six Languages. *American Anthropologist*, 66 (6), 103-114.
- Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)* (3. Aufl.). Los Angeles: Sage Publ.
- Frey, A., Duhm, E., Althaus, D., Heinz, P. & Mengelkamp, C. (2008). *BBK 3-6: Der Beobachtungsbogen für 3-6jährige Kinder*. Göttingen: Hogrefe.
- Frey, A., Mengelkamp, C. & Riebel, C. (2010). Ein Vergleich von theoriegeleiteter und faktorenanalytischer Skalenbildung beim Beobachtungsbogen für Kinder (BBK 3-6). *Empirische Pädagogik*, 24 (3).
- Friedrich, M. & Friederici, A. D. (2005). Phonotactic Knowledge and Lexical–Semantic Processing in One-year-olds: Brain Responses to Words and Nonsense Words in Picture Contexts. *Journal of Cognitive Neuroscience*, 17 (11), 1785-1802.
- Ganger, J. & Brent, M. R. (2004). Reexamining the Vocabulary Spurt. *Developmental Psychology*, 40 (4), 621-632.
- Gasteiger-Klicper, B., Knapp, W. & Kucharz, D. (2011). Die wissenschaftliche Begleitforschung durch die Pädagogische Hochschule Weingarten. In Baden-Württemberg-Stiftung (Hrsg.), *"Sag mal was" - Sprachförderung für Vorschulkinder. Zur Evaluation des Programms der Baden-Württemberg Stiftung. Sprachförderung im Spannungsfeld zwischen Wissenschaft und Praxis* (Schriftenreihe der Baden-Württemberg Stiftung, Bd. 57, 1. Aufl., S. 94–101). Tübingen: Francke, A.
- Gathercole, S. E. & Baddeley, A. D. (2004). *Working memory and language* (Essays in cognitive psychology, Reprint. in paperback). Hove: Erlbaum.
- Glass, G. V., Peckham, P. D. & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research*, 42 (3), 237-288.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1 (1), 3-55.

- Goldfield, B. A. & Reznick, J. S. (1990). Early lexical acquisition: rate, content, and the vocabulary spurt. *Journal of Child Language*, 17 (1), 171-1983.
- Gollwitzer, M. & Jäger, R. S. (2007). *Evaluation: Workbook* (1. Aufl.). Weinheim u.a: Beltz PVU.
- Grimm, H. (1995). Gestörter Sprachlernprozeß: Ursachen und schulische Folgen. In W. Niemeyer & B. Behrens (Hrsg.), *Kommunikation und Lese-Rechtschreibschwäche. Sprachaneignung, Lesen, Schreiben, Rechtschreiben ; Beiträge der internationalen Bremer Arbeitstagung des Wissenschaftlichen Instituts für Schulpraxis, Bremen, Bremen, 17. - 20. September 1991. Mit Beitr. von Birgit Behrens* . (S. S. 53–70). Bochum: Winkler.
- Grimm, H. (Hrsg.). (2000). *Sprachentwicklung* (Sprache, Bd. 3). Göttingen: Hogrefe Verl. für Psychologie.
- Grimm, H. (2003). *Störungen der Sprachentwicklung: Grundlagen - Ursachen - Diagnose - Intervention - Prävention* (2. Aufl.). Göttingen: Hogrefe Verl. für Psychologie.
- Grimm, H., Aktas, M. & Kießig, U. (2003). *Sprachscreening für das Vorschulalter (SSV)*. Göttingen: Hogrefe.
- Grimm, H., Aktaş, M. & Frevert, S. (2001). *SETK 3-5: Sprachentwicklungstest für drei- bis fünfjährige Kinder ; Diagnose von Sprachverarbeitungsfähigkeiten und auditiven Gedächtnisleistungen ; Manual*. Göttingen: Hogrefe.
- Grimm, H. & Doil, H. (2000). *ELFRA: Elternfragebögen für die Früherkennung von Risikokindern ; Handanweisung*. Göttingen: Hogrefe.
- Grimm, H. & Schöler, H. (1991). *Heidelberger Sprachentwicklungstest (HSET)* (Westermann-Test, 2., verb). Göttingen: Verl. für Psychologie Hogrefe.
- Grimm, H. & Weinert, S. (2002). Sprachentwicklung. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (Lehrbuch, 5. Aufl., S. 517–550). Weinheim u.a: Beltz PVU.
- Hager, W. (2000). Planung von Untersuchungen zur Prüfung von Wirksamkeits- und Wirksamkeitsunterschiedshypothesen. In W. Hager, J.-L. Patry & H. Brenzing (Hrsg.), *Handbuch, Evaluation psychologischer Interventionsmassnahmen. Standards und Kriterien* (S. 202–239). Hans Huber Verlag: Bern.
- Hamilton, B. L. (1977). An Empirical Investigation of the Effects of Heterogeneous Regression Slopes in Analysis of Covariance. *Educational and Psychological Measurement*, 37 (3), 701-712.

- Hasselhorn, M., Seidler-Brandler, U. & Körner, K. (2000). Ist das "Nachsprechen von Kunstwörtern" für die Entwicklungsdiagnostik des phonologischen Arbeitsgedächtnisses geeignet? In M. Hasselhorn, W. Schneider & H. Marx (Hrsg.), *Diagnostik von Leserechtschreibschwierigkeiten* (S. 119–134). Göttingen: Hogrefe.
- Hasselhorn, M. & Werner, I. (2000). Zur Bedeutung des phonologischen Arbeitsgedächtnisses für die Sprachentwicklung. In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 363–378). Göttingen: Hogrefe Verl. für Psychologie.
- Häuser, D. & Jülich, B.-R. (2006). *Handlung und Sprache. Das Sprachförderprogramm*. Berlin: NIF.
- Häuser, D., Kasielke, E. & Scheidereiter, U. (1994). *KISTE: Kindersprachtest für das Vorschulalter* (Deutsche Schultests). Weinheim [u.a.]: Beltz.
- Hennon, E., Hirsh-Pasek, K. & Golinkoff, R. M. (2000). Die besondere Reise vom Fötus zum spracherwerbenden Kind. In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 41–103). Göttingen: Hogrefe Verl. für Psychologie.
- Hess, M., Möller, D. & Probst, P. (2008). Durchführung und Evaluation eines Elterntrainings bei Sprachentwicklungsverzögerung. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 57 (3), 197-215.
- Hessisches Kultusministerium. (2011). *Bildungsstandards und Inhaltsfelder – Das neue Kerncurriculum für Hessen: Primarstufe Deutsch* (Hessisches Kultusministerium, Hrsg.). Wiesbaden
- Hirsh-Pasek, K. & Golinkoff, R. M. (1993). Skeletal supports for grammatical learning: What the infant brings to the language learning task. In C. Rovee-Collier & L. P. Lipsitt (Hrsg.), *Advances in infancy research* (Vol. 8). Norwood NJ: Ablex Publ.
- Hirsh-Pasek, K. & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, Mass: MIT Press.
- Hoff-Ginsberg, E. (2000). Soziale Umwelt und Sprachlernen. In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 463–494). Göttingen: Hogrefe Verl. für Psychologie.
- Höhle, B. (2004). Sprachwahrnehmung und Spracherwerb im ersten Lebensjahr. *Sprache-Stimme-Gehör*, 28 (1), 2-7.

- Höhle, B. (2005). Der Einstieg in die Grammatik: Spracherwerb während des ersten Lebensjahres. *Forum Logopädie*, 19 (6), 16-21.
- Höhle, B. & Weissenborn, J. (2000). Lauter Laute? Lautsegmente und Silben in der Sprachperzeption und im Spracherwerb. In R. Thieroff, M. Tamrat, N. Fuhrhop & O. Teuber (Hrsg.), *Deutsche Grammatik in Theorie und Praxis* (S. 1–11). Berlin, Boston: DE GRUYTER.
- Huebner, C. E. (2000). Promoting Toddlers' Language Development Through Community-Based Intervention. *Journal of Applied Developmental Psychology*, 21 (5), 513-535.
- Jampert, K., Best, P., Guadatiello, A., Holler, D. & Zehnbaue, A. (2007). *Schlüsselkompetenz Sprache: Sprachliche Bildung und Förderung im Kindergarten ; Konzepte, Projekte und Maßnahmen ; Ergebnisse der bundesweiten Recherche im DJI-Projekt "Schlüsselkompetenz Sprache" zu Maßnahmen und Aktivitäten im Bereich der sprachlichen Bildung und Sprachförderung in Tageseinrichtungen für Kinder* (2. Aufl.). Weimar [u.a.]: Netz-Verl.
- Janczyk, M., Schöler, H. & Grabowski, J. (2004). Arbeitsgedächtnis und Aufmerksamkeit bei Vorschulkindern mit gestörter und unauffälliger Sprachentwicklung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 36 (4), 200-206.
- Janosch. (2007). *Post für den Tiger: Die Geschichte, wie der kleine Bär und der kleine Tiger die Briefpost, die Luftpost und das Telefon erfinden ; [30 Jahre Tigerente]* [Neuausg. mit neuem Einband]. Weinheim u.a: Beltz & Gelberg.
- Jansen, H., Mannhpaut, G., Marx, H. & Skowronek, H. (2002). *Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten: BISC ; Manual* (2. Aufl.). Göttingen [u.a.]: Hogrefe, Verl. für Psychologie.
- Johnson, E. K., Jusczyk, P. W., Cutler, A. & Norris, D. (2003). Lexical viability constraints on speech segmentation by infants. *Cognitive Psychology*, 46 (1), 65-97.
- Jungmann, T. (2007). Frühförderung unter entwicklungspsychologischer Perspektive. In H. Schöler & A. Welling (Hrsg.), *Sonderpädagogik der Sprache* (Handbuch Sonderpädagogik, Bd. 1, S. 722–741). Göttingen: Hogrefe.
- Jusczyk, P. W. (2000). *The discovery of spoken language* (A Bradford book, 1. Aufl.). Cambridge, Mass: MIT Press.

- Jusczyk, P. W. (2001). Learning language: What infants know about it, and what we don't know about that. In E. Dupoux (Hrsg.), *Language, brain, and cognitive development* (S. 363–377). Cambridge, Mass: MIT Press.
- Jusczyk, P. W., Houston, D. & Goodman, M. (1998). Speech perception during the first year. In A. Slater (Hrsg.), *Perceptual development. Visual, auditory, and speech perception in infancy* (Studies in developmental psychology, S. 357–388). Hove: Psychology Press.
- Justice, L. M. & Pullen, P. C. (2003). Promising Interventions for Promoting Emergent Literacy Skills: Three Evidence-Based Approaches. *Topics in Early Childhood Special Education, 23* (3), 99-113.
- Justice, L. M. & Ezell, H. K. (2002). Use of Storybook Reading to Increase Print Awareness in At-Risk Children. *Am J Speech Lang Pathol, 11* (1), 17.
- Kaltenbacher, E. (2008). *Deutsch für den Schulstart. Ein Förderprogramm für Vorschüler und Schulanfänger mit Deutsch als Erst- und Zweitsprache*. Heidelberg: Seminar für Deutsch als Fremdsprache.
- Kany, W. & Schöler, H. (2010). *Fokus: Sprachdiagnostik: Leitfaden zur Sprachstandsbestimmung im Kindergarten* (Sprache & Literacy, 2. Aufl.). Berlin: Scriptor.
- Karmiloff-Smith, A. (1999). Taking Development Seriously. *Human Development, 42* (6), 325-327.
- Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B. et al. (1998). Statistical Practices of Educational Researchers: An Analysis of their ANOVA, MANOVA, and ANCOVA Analyses. *Review of Educational Research, 68* (3), 350-386.
- Klauer, K. J. & Büchel, F. P. (1993). *Kognitives Training*. Göttingen, Seattle: Hogrefe.
- Kleeck, A., Stahl, S. & Bauer, E. (Hrsg.). (2003). *On Reading Books to Children: Parents and Teachers*. Mahwah, NJ: Erlbaum.
- Klieme, E. (2010). *PISA 2009: Bilanz nach einem Jahrzehnt*. Münster: Waxmann.
- Kuczaj, S. A. (2001). The world of words: Thoughts on the development of a lexicon. In M. Barrett (Hrsg.), *The development of language* (Studies in developmental psychology, S. 133–190). Hove: Psychology Press.
- Kuger, S., Sechtig, J. & Anders, Y. (2012). Kompensatorische (Sprach-)Förderung. *Frühe Bildung, 1* (4), 181-193.

- Kuhl, P. K., Williams, K. A. L. F., Stevens, K. N. & Lindblom, B. (1992). Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age. *Science*, 255 (31), 606-608.
- Kushnir, C. C. & Blake, J. (1996). The nature of the cognitive deficit in specific language impairment. *First Language*, 16 (46), 21-40.
- Küspert, P. & Schneider, W. (2006). *Hören, lauschen, lernen* (5., überarb.). Göttingen: Vandenhoeck und Ruprecht.
- Lienert, G. A., Raatz, U. & Lienert-Raatz. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Beltz Psychologie-Verl.-Union.
- Lisker, A. (2011). *Additive Maßnahmen zur vorschulischen Sprachförderung in den Bundesländern: Expertise im Auftrag des Deutschen Jugendinstituts*. München: Verlag Deutsches Jugendinstitut.
- Lonigan, C. J. & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*, 13 (2), 263-290.
- Lundberg, I., Frost, J. & Petersen, O.-P. (1988). Effects of an Extensive Program for Stimulating Phonological Awareness in Preschool Children. *Reading Research Quarterly*, 23 (3), 263-284.
- Lyster, S.-A. (2002). The effects of morphological versus phonological awareness training in kindergarten on reading development. *Reading and Writing*, 15 (3-4), 261-294.
- Mandel, D. R., Jusczyk, P. W. & Kemler-Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, 53 (2), 155-180.
- Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins, and domain specificity. In M. R. Gunnar & M. P. Maratsos (Hrsg.), *Modularity and constraints in language and cognition* (The Minnesota symposia on child psychology, Bd. 25, S. 59–101). Hillsdale, N.J.: L. Erlbaum Associates.
- Marulis, L. M. & Neuman, S. B. (2010). The Effects of Vocabulary Intervention on Young Children's Word Learning: A Meta-Analysis. *Review of Educational Research*, 80 (3), 300-335.

- Marx, P., Weber, J. & Schneider, W. (2005a). Phonologische Bewusstheit und ihre Förderung bei Kindern mit Störungen der Sprachentwicklung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37 (2), 80-90.
- Marx, P., Weber, J. & Schneider, W. (2005b). Langfristige Auswirkungen einer Förderung der phonologischen Bewusstheit bei Kindern mit Defiziten in der Sprachentwicklung. *Die Sprachheilarbeit*, 50 (6), 280-285.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29 (2), 143-178.
- Menyuk, P. (2000). Wichtige Aspekte der lexikalischen und semantischen Entwicklung. In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 171–192). Göttingen: Hogrefe Verl. für Psychologie.
- Metz, D., Fröhlich, L. P. & Petermann, F. (2010). *Schulbasierte Förderung der phonologischen Bewusstheit und sprachlicher Kompetenzen: Das Lobo-Schulprogramm*. Göttingen [u.a.]: Hogrefe.
- Meyer, S., Jungheim, M. & Ptok, M. (2011). Kindgerichtete Sprache: Für den Spracherwerb wirklich nützlich? *HNO*, 59 (11), 1129-1134.
- Miller, G. A. & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *J Abnorm Psychol*, 110 (1), 40-48.
- Mol, S. E., Bus, A. G. & Jong, M. T. de. (2009). Interactive Book Reading in Early Education: A Tool to Stimulate Print Knowledge as Well as Oral Language. *Review of Educational Research*, 79 (2), 979-1007.
- Mol, S. E. & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137 (2), 267-296.
- Mol, S. E., Bus, A. G., Jong, M. T. de & Smeets, D. J. H. (2008). Added Value of Dialogic Parent–Child Book Readings: A Meta-Analysis. *Early Education and Development*, 19 (1), 7-26.
- Morrow, L. M. (1988). Young Children's Responses to One-to-One Story Readings in School Settings. *Reading Research Quarterly*, 23 (1), 89-107.
- Mowbray, C. T., Holter, M., Teague, G. B. & Bybee, D. (2003). Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation*, 24 (3), 315-340.

- National Early Literacy Panel (U.S.). (2008). *Developing early literacy: Report of the National Early Literacy Panel*. [Washington, D.C.]: National Institute for Literacy.
- Nazzi, T., Bertoncini, J. & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24 (3), 756-766.
- Neuman, S. B. & Carta, J. J. (2011). Advancing the measurement of quality for early childhood programs that support early language and literacy development. In M. J. Zaslow, I. Martinez-Beck, K. Tout & T. Halle (Hrsg.), *Quality measurement in early childhood settings* (S. 51–76). Baltimore, Md: Paul H. Brookes Pub. Co.
- Neuman, S. B. & Dickinson, D. K. (Hrsg.). (2001). *Handbook of early literacy research*. New York: Guilford Press.
- Norris, J. M. & Ortega, L. (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *Language Learning*, 50 (3), 417-528.
- Nußbeck, S. (2007). *Sprache: Entwicklung, Störungen und Intervention* (Module angewandter Psychologie, 1. Aufl.). Stuttgart: Kohlhammer.
- Oller, D. K. & Eilers, R. E. (1988). The Role of Audition in Infant Babbling. *Child Development*, 59 (2), 441-449.
- Owens, R. E. (2005). *Language development: An introduction* (6. Aufl.). Boston Mass. u.a: Pearson/Allyn and Bacon.
- Papoušek, M. (2012). Kommunikation und Sprachentwicklung im ersten Lebensjahr. In M. Cierpka (Hrsg.), *Frühe Kindheit 0 – 3. Beratung und Psychotherapie für Eltern mit Säuglingen und Kleinkindern* (S. 69–80). Berlin Heidelberg: Springer-Verlag Berlin Heidelberg.
- Penner, Z. (2000). Phonologische Entwicklung: Eine Übersicht. In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 105–139). Göttingen: Hogrefe Verl. für Psychologie.
- Penner, Z. (2003). *Programm sprachliche Frühförderung von fremdsprachigen Kindern im Kindergarten. Programmhandbuch.*: Kon-Lab.
- Petermann, F., Metz, D. & Fröhlich, L. P. (2010). *SET 5-10: Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren*. Göttingen: Hogrefe.

- Pinker, S. (2000). *The language instinct: How the mind creates language*. (22. Aufl.). New York, NY: HarperPerennial.
- Plaut, D. C. & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Hrsg.), *The emergence of language* (Carnegie Mellon symposia on cognition, S. 381–416). Mahwah NJ. u.a: Lawrence Erlbaum.
- Plunkett, K. & Schafer, G. (2001). Early speech perception and word learning. In M. Barrett (Hrsg.), *The development of language* (Studies in developmental psychology, S. 51–71). Hove: Psychology Press.
- Pressley, M. & Harris, K. (1994). Increasing the quality of educational intervention research. *Educ Psychol Rev*, 6 (3), 191-208.
- Redder, A. S. K., Hasselhorn, M., Forschner, S., Fickermann, D., Ehlich, K., Becker-Mrotzek, M. et al. (2011). *Bilanz und Konzeptualisierung von strukturierter Forschung zu "Sprachdiagnostik und Sprachförderung"* (ZUSE-Berichte, Bd. 2). Hamburg: ZUSE.
- Ricken, G. & Wechsler, D. (2007). *Hannover-Wechsler-Intelligenztest für Kinder im Vorschulalter - III: HAWIVA-III ; Übersetzung und Adaption der WPPSI-III von David Wechsler*. Bern: Huber [u.a.].
- Ritterfeld, U. (2000). Welchen und wieviel Input braucht das Kind? In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 403–432). Göttingen: Hogrefe Verl. für Psychologie.
- Ritterfeld, U. (2005). Interventionsprinzipien bei Spracherwerbsstörungen. *Heilpädagogik online* (3), 4-29.
- Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, 88 (2), 245-258.
- Roos, J., Polotzek, S. & Schöler, H. (2010). *Evaluationsstudie zur Sprachförderung von Vorschulkindern: Wissenschaftliche Begleitung der Sprachfördermaßnahmen im Programm „Sag’ mal was – Sprachförderung für Vorschulkinder*, Pädagogische Hochschule Heidelberg. Zugriff am 05.09.14. Verfügbar unter http://www.sagmalwas-bw.de/media/WiBe%201/pdf/EVAS_Abschlussbericht_Januar2010.pdf.

- Roos, J. & Schöler, H. (2007). Sprachentwicklungsdiagnostik mittels standardisierter Tests. In H. Schöler & A. Welling (Hrsg.), *Sonderpädagogik der Sprache* (Handbuch Sonderpädagogik, Bd. 1, S. 531–550). Göttingen: Hogrefe.
- Roth, E. & Schneider, W. (2002). Langzeiteffekte einer Förderung der phonologischen Bewusstheit und der Buchstabenkenntnis auf den Schriftspracherwerb. *Zeitschrift für Pädagogische Psychologie*, 16 (2), 99-107.
- Rothweiler, M. & Kauschke, C. (2007). Lexikalischer Erwerb. In H. Schöler & A. Welling (Hrsg.), *Sonderpädagogik der Sprache* (Handbuch Sonderpädagogik, Bd. 1, S. 42–57). Göttingen: Hogrefe.
- Rückert, E. M., Kunze, S., Schillert, M. & Schulte-Körne, G. (2010). Prävention von Leserechtschreibschwierigkeiten. *Kindheit und Entwicklung*, 19 (2), 82-89.
- Rückert, E., Kunze, S. & Schulte-Körne, G. (2010). *Lass uns lesen!: Ein Eltern-Kind-Training zur Vorbereitung auf das Lesen- und Schreibenlernen*. Bochum: Winkler.
- Sachse, S., Budde, N., Rinker, T. & Groth, K. (2012). Evaluation einer Sprachfördermaßnahme für Vorschulkinder. *Frühe Bildung*, 1 (4), 194-201.
- Scarborough, H. S. (2001). A discussion of evidence, theory, and practice connecting early language and literacy to later reading disabilities. In S. B. Neuman & D. K. Dickinson (Hrsg.), *Handbook of early literacy research* (S. 97–110). New York: Guilford Press.
- Schlesiger, C. (2009). *Sprachtherapeutische Frühintervention für Late Talkers: Eine randomisierte und kontrollierte Studie zur Effektivität eines direkten und kindzentrierten Konzeptes* (Beiträge zur Gesundheits- und Therapiewissenschaft, Bd. 4). Idstein: Schulz-Kirchner Verlag.
- Schneider, W., Küspert, P., Roth, E., Visé, M. & Marx, H. (1997). Short- and long-term effects of training phonological awareness in kindergarten: Evidence from two German studies. *Journal of Experimental Child Psychology* (66), 311-340.
- Schneider, W. & Näslund, J. (1993). The impact of early metalinguistic competencies and memory capacities on reading and spelling in elementary school: Results of the Munich Longitudinal Study on the Genesis of individual competencies (LOGIC). *European Journal of Psychology of Education* (8), 273-288.

- Schneider, W., Visé, M., Reimers, P. & Blässer, B. (1994). Auswirkungen eines Trainings der sprachlichen Bewußtheit auf den Schriftspracherwerb in der Schule. *Zeitschrift für Pädagogische Psychologie*, 8 (3/4), 177-188.
- Schneider, W. (2004). Frühe Entwicklung von Lesekompetenz: Zur Relevanz vorschulischer Sprachkompetenzen. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz* (S. 13–36). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schneider, W., Roth, E. & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology*, 92 (2), 284-295.
- Schneider, W., Roth, E. & Küspert, P. (1999). Frühe Prävention von Lese-Rechtschreibproblemen. *Kindheit und Entwicklung*, 8 (3), 147-152.
- Siegmüller, J. & Kauschke, C. (Hrsg.). (2011). *Leitfaden Sprache, Sprechen, Stimme, Schlucken* (2. Aufl.). München: Urban & Fischer in Elsevier.
- Simon, S. & Sachse, S. (2013). Anregung der Sprachentwicklung durch ein Interaktionstraining für Erzieherinnen. *Diskurs Kindheits- und Jugendforschung* (4), 379-397.
- Skowronek, H. & Marx, H. (1989). Die Bielefelder Längsschnittstudie zur Früherkennung von Risiken der Lese-Rechtschreibschwäche: Theoretischer Hintergrund und erste Befunde. *Heilpädagogische Forschung*, 15, 38-49.
- Slavin, R. E. (2008). Evidence-based Reform in Education: what will it take? *eerj*, 7 (1), 124.
- Stanat, P. (2006). Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 189–219). Wiesbaden: VS Verl. für Sozialwiss.
- Stevens, J. (2007). *Intermediate statistics: A modern approach* (3. Aufl.). New York: Lawrence Erlbaum Associates.
- Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C. et al. (2011). A Synthesis of Read-Aloud Interventions on Early Reading Outcomes Among Preschool Through Third Graders at Risk for Reading Difficulties. *Journal of Learning Disabilities*, 44 (3), 258-275.

- Szagun, G. (2006). *Sprachentwicklung beim Kind* (Vollst. überarb. Neuausg.). Weinheim, Basel: Beltz.
- Tomasello, M. (2000). Preceiving intentions and learning words in the second year of life. In M. Bowermann & S. Levinson (Hrsg.), *Language acquisition and conceptual development* (S. 132–158). Hove: Psychology Press.
- Tomasello, M. (2001). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11 (1-2), 61-82.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, Mass: Harvard University Press.
- Tracy, R. (2000). Sprache und Sprachentwicklung: Was wird erworben? In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 3–39). Göttingen: Hogrefe Verl. für Psychologie.
- Tracy, R. (2003). *Sprachliche Frühförderung - Konzeptuelle Grundlagen eines Programms zur Förderung von Deutsch als Zweitsprache im Vorschulalter*. Mannheim: Universität Mannheim, Forschungs- und Kontaktstelle Mehrsprachigkeit.
- Tracy, R. (2008). *Wie Kinder Sprachen lernen und wie wir sie dabei unterstützen können* (2. Aufl.). Tübingen: Francke.
- Ulich, M. & Mayr, T. (2004). *SISMIK: Sprachverhalten und Interesse an Sprache bei Migrantenkindern in Kindertageseinrichtungen ; Begleitheft zum Beobachtungsbogen SISMIK*. Freiburg im Breisgau: Herder.
- Verhallen, M. J. A. J., Bus, A. G. & Jong, M. T. de. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology*, 98 (2), 410-419.
- Warren, S. F., Yoder, P. J., Gazdag, G. E., Kim, K. & Jones, H. A. (1993). Facilitating Prelinguistic Communication Skills in Young Children With Developmental Delay. *J Speech Hear Res*, 36 (1), 83.
- Weber, A. & Bacher, S. (2011). Ein Programm für Vorschulkinder: "Sag mal 'was". In Baden-Württemberg-Stiftung (Hrsg.), *"Sag mal was" - Sprachförderung für Vorschulkinder. Zur Evaluation des Programms der Baden-Württemberg Stiftung. Sprachförderung im Spannungsfeld zwischen Wissenschaft und Praxis* (Schriftenreihe der Baden-Württemberg Stiftung, Bd. 57, 1. Aufl., S. 14–31). Tübingen: Francke, A.

- Weinert, S. (2000). Beziehungen zwischen Sprach- und Denkentwicklung. In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 311–362). Göttingen: Hogrefe Verl. für Psychologie.
- Weinert, S. (2006). Sprachentwicklung. In W. Schneider & B. Sodian (Hrsg.), *Kognitive Entwicklung* (S. 609–719). Göttingen: Hogrefe.
- Weinert, S. & Lockl, K. (2008). Sprachförderung. In F. Petermann & W. Schneider (Hrsg.), *Angewandte Entwicklungspsychologie* (Entwicklungspsychologie, Bd. 7, Bd. 7, S. 91–134). Göttingen: Hogrefe Verl. für Psychologie.
- Weissenborn, J. (2000). Der Erwerb von Morphologie und Syntax. In H. Grimm (Hrsg.), *Sprachentwicklung* (Sprache, Bd. 3, S. 141–169). Göttingen: Hogrefe Verl. für Psychologie.
- Whitehurst, G. J., Falco, F. L. 1., Lonigan, C. J. 1., Fischel, J. E. 1., DeBaryshe, B. D. 1., Valdez-Menchaca, M. C. 1. et al. (1988). Accelerating Language Development Through Picture Book Reading. *Developmental Psychology*, 24 (4), 552-559.
- Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L. & et al. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, 30 (5), 679-689.
- Whitehurst, G. J., Fischel, J. E., Lonigan, C. J., Valdez-Menchaca, M. C. & et, a. (1991). Treatment of early expressive language delay: If, when, and how. *Topics in Language Disorders*, 11 (4), 55-68.
- Whitehurst, G. J., Zevenbergen, A. A., Crone, D. A., Schultz, M. D., Velting, O. N. & Fischel, J. E. (1999). Outcomes of an emergent literacy intervention from Head Start through second grade. *Journal of Educational Psychology*, 91 (2), 261-272.
- Wildt, A. R. & Ahtola, O. (1978). *Analysis of covariance* (Quantitative applications in the social sciences, ser. no. 07-012). Beverly Hills, Calif: SAGE Publications.
- Wirtz, M. (2004). Bestimmung der Güte von Beurteilereinschätzungen mittels der Intraklassenkorrelation und Verbesserung von Beurteilereinschätzungen: Determining the quality of rater judgments using intraclass correlation and possibilities of improving rater judgments. *Die Rehabilitation*, 43 (6), 384-389.
- Wirtz, M. A. (2013). Lexikon der Psychologie: Dorsch - 2014/2015. *Dorsch Lexikon der Psychologie*.

- Wirtz, M. A. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe Verl. für Psychologie.
- Wolf, K. M., Stanat, P. & Wendt, W. (2011). *EkoS - Evaluation der kompensatorischen Sprachförderung: Abschlussbericht*. Zugriff am 19.08.15. Verfügbar unter <http://www.isq-bb.de/uploads/media/ekos-bericht-3-110216.pdf>.
- Wu, C. F. J. (1986). Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis. *Ann. Statist.*, 14 (4), 1261-1295.
- Wygotskij, L. S. (1978). *Mind in society*. Cambridge: Harvard University Press.
- Zevenbergen, A. A. & Whitehurst, G. J. (2003). Dialogic reading. A shared picture book intervention for preschoolers. In A. Kleeck, S. Stahl & E. Bauer (Hrsg.), *On Reading Books to Children: Parents and Teachers* (S. 302–320). Mahwah, NJ: Erlbaum.
- Zollinger, B. (1997). *Die Entdeckung der Sprache* (Beiträge zur Heil- und Sonderpädagogik, Bd. 16, 3. Aufl.). Bern ; Wien u.a: Haupt.

Anhang

Tabelle 20 Korrelationen zwischen den Subtests zur Erfassung sprachlicher Kompetenzen zu allen Messzeitpunkten.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Vortest Bildbenennung	(1)	,723"	,660"	,498"	,606"	,698"	,841"	,652"	,658"	,435"	,559"	,611"	,806"	,642"	,589"	,466"	,537"	,603"
Vortest Kategoriebildung	(2)		,591"	,511"	,574"	,573"	,675"	,649"	,569"	,510"	,544"	,484"	,618"	,557"	,540"	,540"	,490"	,440"
Vortest Handlungssequenzen	(3)	,660"		,591"	,511"	,567"	,588"	,509"	,615"	,453"	,452"	,514"	,547"	,480"	,606"	,462"	,495"	,443"
Vortest Textverständnis	(4)	,498"	,511"		,457"	,458"	,429"	,467"	,447"	,595"	,357"	,272"	,441"	,441"	,501"	,524"	,376"	,219"
Vortest Satzbildung	(5)	,606"	,574"	,511"		,496"	,612"	,575"	,539"	,446"	,706"	,478"	,586"	,514"	,569"	,484"	,586"	,446"
Vortest Singular-Plural-Bildung	(6)	,698"	,573"	,567"	,458"	,496"	1	,638"	,553"	,450"	,423"	,643"	,663"	,517"	,493"	,417"	,422"	,634"
Nachtest Bildbenennung	(7)	,841"	,675"	,588"	,429"	,612"		,748"	,697"	,511"	,611"	,659"	,825"	,621"	,551"	,473"	,510"	,578"
Nachtest Kategoriebildung	(8)	,652"	,649"	,509"	,467"	,575"	,553"	1	,578"	,566"	,632"	,574"	,636"	,593"	,513"	,467"	,508"	,481"
Nachtest Handlungssequenzen	(9)	,658"	,569"	,615"	,447"	,539"	,555"	,697"	1	,498"	,548"	,505"	,626"	,528"	,627"	,493"	,502"	,432"
Nachtest Textverständnis	(10)	,435"	,510"	,453"	,595"	,446"	,450"	,511"	,498"	1	,519"	,355"	,478"	,430"	,455"	,587"	,457"	,375"
Nachtest Satzbildung	(11)	,559"	,544"	,452"	,357"	,706"	,423"	,611"	,548"	,519"	1	,549"	,557"	,466"	,579"	,499"	,587"	,456"
Nachtest Singular-Plural-Bildung	(12)	,611"	,484"	,514"	,272"	,478"	,643"	,659"	,574"	,355"	,549"	1	,600"	,428"	,463"	,336"	,407"	,649"
Follow Up Bildbenennung	(13)	,806"	,618"	,547"	,441"	,586"	,663"	,825"	,626"	,478"	,557"	,600"	1	,747"	,641"	,516"	,601"	,651"
Follow Up Kategoriebildung	(14)	,642"	,557"	,480"	,441"	,514"	,517"	,621"	,528"	,430"	,466"	,428"	,747"	1	,606"	,546"	,621"	,538"
Follow Up Handlungssequenzen	(15)	,589"	,540"	,606"	,501"	,569"	,493"	,551"	,627"	,455"	,579"	,463"	,641"	,606"	1	,551"	,610"	,487"
Follow Up Textverständnis	(16)	,466"	,540"	,462"	,524"	,484"	,417"	,473"	,493"	,587"	,499"	,336"	,516"	,546"	,551"	1	,471"	,351"
Follow Up Satzbildung	(17)	,537"	,490"	,495"	,376"	,586"	,422"	,510"	,508"	,457"	,587"	,407"	,601"	,621"	,610"	,471"	1	,448"
Follow Up Singular-Plural-Bildung	(18)	,603"	,440"	,443"	,219"	,446"	,634"	,578"	,432"	,375"	,456"	,649"	,651"	,538"	,487"	,351"	,448"	1

**. Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

*. Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 174

N = 165

N = 154

N = 150

N = 121

N = 118

Tabelle 21 Korrelationen der Subtests zur Erfassung der Phonologischen Bewusstheit zu allen Messzeitpunkten.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Vortest Phonemsynthese I (ohne Abbildung)	1	,360**	,293**	0,059	,514**	,329**	,358**	,166*	,345**	,253**	,333**	,170*
Vortest Anlautidentifikation	,360**	1	0,147	,196**	,270**	,297**	,288**	,240**	,219**	,234**	,221**	0,1
Vortest Phonemsynthese II (mit Abbildung)	,293**	0,147	1	0,146	,156*	0,126	,381**	0,096	0,151	0,154	0,103	,175*
Vortest Reimen	0,059	,196**	0,146	1	,159*	,217**	,220**	,514**	,234**	,310**	,263**	,398**
Nachtest Phonemsynthese I (ohne Abbildung)	,514**	,270**	,156*	,159*	1	,403**	,526**	,278**	,475**	,324**	,493**	,338**
Nachtest Anlautidentifikation	,329**	,297**	0,126	,217**	,403**	1	,390**	,342**	,242**	,341**	,393**	,203*
Nachtest Phonemsynthese II (ohne Abbildung)	,358**	,288**	,381**	,220**	,526**	,390**	1	,216**	,305**	,223**	,431**	,228**
Nachtest Reimen	,166*	,240**	0,096	,514**	,278**	,342**	,216**	1	,269**	,379**	,275**	,517**
Follow Up Phonemsynthese I (ohne Abbildung)	,345**	,219**	0,151	,234**	,475**	,242**	,305**	,269**	1	,612**	,652**	,373**
Follow Up Anlautidentifikation	,253**	,234**	0,154	,310**	,324**	,341**	,223**	,379**	,612**	1	,601**	,468**
Follow Up Phonemsynthese II (mit Abbildung)	,333**	,221**	0,103	,263**	,493**	,393**	,431**	,275**	,652**	,601**	1	,326**
Follow Up Reimen	,170*	0,1	,175*	,398**	,338**	,203*	,228**	,517**	,373**	,468**	,326**	1

** . Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

* . Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 174

N = 165

N = 154

N = 150

Tabelle 22 Korrelationen zwischen den abhängigen Variablen zu allen Messzeitpunkten.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Vorlest Allgemeine Sprachentwicklung	1	,921**	,880**	,901**	,401**	,848**	,803**	,759**	,719**	,486**	,832**	,767**	,733**	,722	,368**
Vorlest Wortschatz	,921**	1	,713**	,763**	,379**	,798**	,806**	,673**	,670**	,453**	,765**	,750**	,643**	,658	,318**
Vorlest Sprachverständnis	,880**	,713**	1	,675**	,347**	,690**	,621**	,711**	,528**	,422	,693**	,604**	,694**	,547**	,349**
Vorlest Grammatik	,901**	,763**	,675**	1	,357**	,798**	,736**	,666**	,741**	,437**	,761**	,704**	,637**	,717**	,323**
Vorlest Phonologische Bewusstheit	,401**	,379**	,347**	,357**	1	,486**	,449**	,404**	,451**	,594**	,455**	,379**	,345**	,462**	,443**
Nachtest Allgemeine Sprachentwicklung	,848**	,798**	,690**	,798**	,486**	1	,917**	,875**	,895**	,532**	,825**	,744**	,710**	,722	,374**
Nachtest Wortschatz	,803**	,806**	,621**	,736**	,449**	,917**	1	,724**	,750**	,514**	,762**	,757**	,600**	,653	,315**
Nachtest Sprachverständnis	,759**	,673**	,711**	,666**	,404**	,875**	,724**	1	,638**	,460**	,746**	,632**	,708**	,596**	,354**
Nachtest Grammatik	,719**	,670**	,528**	,741**	,451**	,895**	,750**	,638**	1	,456**	,720**	,611**	,598**	,692**	,333**
Nachtest Phonologische Bewusstheit	,486**	,453**	,422**	,437**	,594**	,532**	,514**	,460**	,456**	1	,505**	,419**	,420**	,492**	,596**
Follow Up Allgemeine Sprachentwicklung	,832**	,765**	,693**	,761**	,455**	,825**	,762**	,746**	,720**	,505**	1	,893**	,894**	,881**	,454**
Follow Up Wortschatz	,767**	,750**	,604**	,704**	,379**	,744**	,757**	,632**	,611**	,419**	,893**	1	,686**	,759**	,391**
Follow Up Sprachverständnis	,733**	,643**	,694**	,637**	,345**	,710**	,600**	,708**	,598**	,420**	,894**	,686**	1	,634**	,434**
Follow Up Grammatik	,722**	,658**	,547**	,717**	,462**	,722**	,653**	,596**	,692**	,492**	,881**	,759**	,634**	1	,438**
Follow Up Phonologische Bewusstheit	,368**	,318**	,349**	,323**	,443**	,374**	,315**	,354**	,333**	,596**	,454**	,391**	,434**	,438**	1

** . Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

* . Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 174

N = 165

N = 154

N = 150

N = 121

N = 100

Tabelle 23 Korrelationen zwischen den Skalen der erweiterten Auswertung der Satzbildungsaufgabe über alle Messzeitpunkte.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Vortest Morphologie bei Satzbildung	1	,895**	,876**	,517**	,503**	,536**	,296**	,364**	,382**
Vortest Syntax bei Satzbildung		1	,856**	,545**	,572**	,540**	,362**	,431**	,432**
Vortest Artikelbildung bei Satzbildung			1	,471**	,482**	,530**	,298**	,399**	,424**
Nachtest Morphologie bei Satzbildung				1	,905**	,890**	,463**	,477**	,529**
Nachtest Syntax bei Satzbildung					,905**	,849**	,480**	,512**	,538**
Nachtest Artikelbildung bei Satzbildung					,890**	1	,403**	,508**	,525**
Follow Up Morphologie bei Satzbildung							1	,754**	,760**
Follow Up Syntax bei Satzbildung								1	,926**
Follow Up Artikelbildung bei Satzbildung									1

**, Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

*, Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 156

N = 153

N = 151

N = 140

Tabelle 24 Korrelation zwischen den Raterurteilen in den Items zur Erfassung der Durchführungsqualität (Fortsetzung auf S.161).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
R1 Zählen: Geschlossene Fragen	(1) 1	.866**	.401**	-0.035	.701**	.729**	.642**	.840**	.677**	.283*	.023	.334**	.664**	.512**	.839**
R1 Zählen: Offene Fragen	(2) .866**	1	.487**	-0.245	.565**	.519**	.590**	.725**	.836**	.448**	-0.079	-0.001	.524**	.388**	.875**
R1 Zählen: Erweiterung der Aussage	(3) .401**	.487**	1	-0.002	.0202	.037	-0.001	.474**	.0176	.904**	-0.062	-0.168	-0.089	-0.111	.421**
R1 Zählen: korrektive Wiederholung	(4) -0.035	-0.245	-0.002	1	-0.227	.0116	-0.254	.098	-.264*	-0.1	.849**	.06	-0.126	.001	-0.047
R1 Zählen: Nachfragen	(5) .701**	.565**	.0202	-0.227	1	.611**	.257*	.401**	.461**	.0139	-0.205	.597**	.659**	.268*	.596**
R1 Zählen: Alltagsbezug	(6) .729**	.519**	.037	.0116	.611**	1	.277*	.451**	.511**	-0.021	.048	.521**	.926**	.0137	.517**
R1 Zählen: Sätze vervollständigen	(7) .642**	.590**	-0.001	-0.254	.257*	.277*	1	.567**	.644**	-0.068	-0.194	.021	.320*	.870**	.456**
R1 Rating: Geschlossene Fragen	(8) .840**	.725**	.474**	.098	.401**	.451**	.567**	1	.434**	.282*	.0187	.0236	.347**	.445**	.772**
R1 Rating: Offene Fragen	(9) .677**	.836**	.0176	-.264*	.461**	.511**	.644**	.434**	1	.263*	-0.09	-0.008	.596**	.534**	.725**
R1 Rating: Erweiterung der Aussage	(10) .283*	.448**	.904**	-0.1	.0139	-0.021	-0.068	.282*	.263*	1	-0.189	-0.251	-0.062	-0.131	.466**
R1 Rating: korrektive Wiederholung	(11) 0.023	-0.079	-0.062	.849**	-0.205	.0248	-0.194	.0187	-0.09	-0.189	1	.006	.0059	-0.048	.003
R1 Rating: Nachfragen	(12) .334**	-0.001	-0.168	.006	.597**	.521**	.021	.0236	-0.008	-0.251	.006	1	.549**	.0153	.0205
R1 Rating: Alltagsbezug	(13) .664**	.524**	-0.089	-0.126	.659**	.926**	.320*	.347**	.596**	-0.062	.0059	.549**	1	.018	.523**
R1 Rating: Sätze vervollständigen	(14) .512**	.388**	-0.111	.001	.268*	.0137	.870**	.445**	.534**	-0.131	-0.048	.0153	.018	1	.418**
R1 Gesamtnote	(15) .839**	.875**	.421**	-0.047	.596**	.517**	.456**	.772**	.725**	.466**	.003	.0205	.523**	.418**	1
R2 Zählen: Geschlossene Fragen	(16) .855**	.881**	.611**	-.021	.434**	.426**	.530**	.850**	.660**	.548**	.030	-0.010	.388**	.385**	.872**
R2 Zählen: Offene Fragen	(17) .874**	.806**	.008	-.183	.631**	.666**	.805**	.634**	.830**	-.012	-.055	.274*	.712**	.712**	.753**
R2 Zählen: Erweiterung der Aussage	(18) .129	-.197	-.194	.423**	-.071	.014	.373**	.245	-.197	-.398**	.290*	.129	-.078	.463**	-.145
R2 Zählen: korrektive Wiederholung	(19) .737**	.690**	.358**	.107	.576**	.301*	.334**	.705**	.479**	.308*	.036	.203	.285*	.408**	.822**
R2 Zählen: Nachfragen	(20) .327*	.483**	.672**	.110	.130	-.122	-.031	.473**	.119	.556**	.003	-.295*	-.214	-.113	.429**
R2 Zählen: Alltagsbezug	(21) .854**	.663**	.263*	-.125	.424**	.636**	.639**	.815**	.500**	.162	-.123	.348**	.563**	.433**	.675**
R2 Zählen: Sätze vervollständigen	(22) .625**	.554**	-.041	-.227	.244	.230	.990**	.587**	.586**	-.120	-.186	.044	.271*	.888**	.445**
R2 Rating: Geschlossene Fragen	(23) .605**	.661**	.371**	.136	.290*	.433**	.356**	.549**	.574**	.312*	.396**	-.212	.424**	.261*	.563**
R2 Rating: Offene Fragen	(24) .727**	.520**	-.154	-.047	.648**	.809**	.549**	.413**	.649**	-.160	.066	.557**	.886**	.537**	.526**
R2 Rating: Erweiterung der Aussage	(25) .056	-.266*	-.269*	.608**	-.173	.236	.112	.202	-.221	-.412**	.477**	.384**	.102	.190	-.108
R2 Rating: korrektive Wiederholung	(26) .629**	.487**	.223	.159	.591**	.390**	.118	.614**	.260*	.153	.140	.398**	.331*	.239	.669**
R2 Rating: Nachfragen	(27) .189	.343**	.572**	.198	.274*	.090	-.343**	.194	.096	.469**	.155	-.028	.028	-.377**	.268*
R2 Rating: Alltagsbezug	(28) .774**	.577**	.003	-.051	.595**	.851**	.511**	.599**	.618**	.036	.038	.597**	.882**	.420**	.702**
R2 Rating: Sätze vervollständigen	(29) .548**	.481**	-.298*	-.176	.372**	.293*	.889**	.438**	.624**	-.362**	-.065	.165	.364**	.888**	.408**
R2 Gesamtnote	(30) .922**	.815**	.260*	.043	.619**	.815**	.576**	.763**	.735**	.144	.220	.314*	.789**	.430**	.756**

**, Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

*, Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 59

	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
R1 Zählen: Geschlossene Fragen	(1) ,855**	,874**	,129	,737**	,327*	,854**	,625**	,605**	,727**	,056	,629**	,189	,774**	,548**	,922**
R1 Zählen: Offene Fragen	(2) ,881**	,806**	-,197	,690**	,483**	,663**	,554**	,661**	,520**	-,266*	,487**	,343**	,577**	,481**	,815**
R1 Zählen: Erweiterung der Aussage	(3) ,611**	,008	-,194	,358**	,672**	,263*	-,041	,371**	-,154	-,269*	,223	,572**	,003	-,298*	,260*
R1 Zählen: korrektive Wiederholung	(4) -,021	-,183	,423**	,107	,110	-,125	-,227	,136	-,047	,608**	,159	,198	-,051	-,176	,043
R1 Zählen: Nachfragen	(5) ,434**	,631**	-,071	,576**	,130	,424**	,244	,290*	,648**	-,173	,591**	,274*	,595**	,372**	,619**
R1 Zählen: Alltagsbezug	(6) ,426**	,666**	,014	,301*	-,122	,636**	,230	,433**	,809**	,236	,390**	,090	,851**	,293*	,815**
R1 Zählen: Sätze vervollständigen	(7) ,530**	,805**	,373**	,334**	-,031	,639**	,990**	,356**	,549**	,112	,118	-,343**	,511**	,889**	,576**
R1 Rating: Geschlossene Fragen	(8) ,850**	,634**	,245	,705**	,473**	,815**	,587**	,549**	,413**	,202	,614**	,194	,599**	,438**	,763**
R1 Rating: Offene Fragen	(9) ,660**	,830**	-,197	,479**	,119	,500**	,586**	,574**	,649**	-,221	,260*	,096	,618**	,624**	,735**
R1 Rating: Erweiterung der Aussage	(10) ,548**	-,012	-,398**	,308*	,556**	,162	-,120	,312*	-,160	-,412**	,153	,469**	,036	-,362**	,144
R1 Rating: korrektive Wiederholung	(11) ,030	-,055	,290*	,036	,003	-,123	-,186	,396**	,066	,477**	,140	,155	,038	-,065	,220
R1 Rating: Nachfragen	(12) -,010	,274*	,129	,203	-,295*	,044	-,212	,557**	,384**	,398**	-,028	,597**	,165	,314*	,789**
R1 Rating: Alltagsbezug	(13) ,388**	,712**	-,078	,285*	-,214	,563**	,271*	,424**	,886**	,102	,331*	,028	,882**	,364**	,789**
R1 Rating: Sätze vervollständigen	(14) ,385**	,712**	,463**	,408**	-,113	,433**	,888**	,261*	,537**	,190	,239	-,377**	,420**	,888**	,430**
R1 Gesamtnote	(15) ,872**	,753**	-,145	,822**	,429**	,675**	,445**	,563**	,526**	-,108	,669**	,268*	,702**	,408**	,756**
R2 Zählen: Geschlossene Fragen	(16) 1	,678**	,082	,815**	,614**	,742**	,508**	,688**	,430**	-,024	,531**	,338**	,554**	,340**	,798**
R2 Zählen: Offene Fragen	(17) ,678**	1	,0116	,615**	,042	,736**	,790**	,551**	,847**	0	,493**	-,085	,785**	,825**	,859**
R2 Zählen: Erweiterung der Aussage	(18) 0,082	0,116	1	,0102	-,0141	0,197	,390**	0,034	0,144	,757**	-,0108	-,367**	0,08	,346**	0,139
R2 Zählen: korrektive Wiederholung	(19) ,815**	,615**	0,102	1	,641**	,570**	,373**	,496**	,430**	-,033	,812**	,401**	,435**	,363**	,645**
R2 Zählen: Nachfragen	(20) ,614**	0,042	-,0141	,641**	1	0,196	0,002	,366**	-,0221	-,0249	,481**	,788**	-,0156	-,0159	0,231
R2 Zählen: Alltagsbezug	(21) ,742**	,736**	0,197	,570**	0,196	1	,638**	,330*	,607**	0,231	,494**	-,071	,784**	,468**	,762**
R2 Zählen: Sätze vervollständigen	(22) ,508**	,790**	,390**	,373**	0,002	,638**	1	,332*	,531**	0,115	0,195	-,339**	,480**	,906**	,547**
R2 Rating: Geschlossene Fragen	(23) ,688**	,551**	0,034	,496**	,366**	,330*	,332*	1	,482**	-,0197	,421**	,300*	,343**	,307*	,745**
R2 Rating: Offene Fragen	(24) ,430**	,847**	0,144	,430**	-,0221	,607**	,531**	,482**	1	0,144	,452**	-,0101	,840**	,629**	,823**
R2 Rating: Erweiterung der Aussage	(25) -,024	0	,757**	-,0333	-,0249	0,231	0,115	-,0197	0,144	1	-,0154	-,025	,283*	0,096	0,126
R2 Rating: korrektive Wiederholung	(26) ,531**	,493**	-,0108	,812**	,461**	,494**	0,195	,421**	,452**	-,0154	1	,362**	,420**	,259*	,546**
R2 Rating: Nachfragen	(27) ,338**	-,085	-,367**	,401**	,788**	-,071	-,339**	,300*	-,0101	-,025	,362**	1	-,118	-,356**	0,218
R2 Rating: Alltagsbezug	(28) ,554**	,785**	0,08	,435**	-,0156	,784**	,480**	,343**	,840**	,283*	,420**	-,0118	1	,485**	,800**
R2 Rating: Sätze vervollständigen	(29) ,340**	,825**	,346**	,363**	-,0159	,468**	,906**	,307*	,629**	0,096	,259*	-,356**	,485**	1	,537**
R2 Gesamtnote	(30) ,798**	,859**	0,139	,645**	0,231	,762**	,547**	,745**	,823**	0,126	,546**	0,218	,800**	,537**	1

** : Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

* : Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 59

Tabelle 25 Korrelationen zwischen den Mittelwerten beider Rater und den Gesamtscores zur Durchführungsqualität (Fortsetzung auf S. 163).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
MW Zählen: Geschlossene Fragen	1	,894**	,566**	0,0488	,790**	,780**	,594**	,813**	,679**
MW Zählen: Offene Fragen		1	,284*	-0,152	,743**	,755**	,726**	,729**	,846**
MW Zählen: Erweiterung der Aussage			1	-0,076	,427**	0,1138	-0,02	,484**	-0,088
MW Zählen: korrektive Wiederholung				1	-0,015	0,0676	0,093	0,1438	-0,057
MW Zählen: Nachfragen					1	,589**	,346**	,559**	,625**
MW Zählen: Alltagsbezug						1	,496**	,557**	,769**
MW Zählen: Sätze vervollständigen							1	,457**	,616**
MW Rating: Geschlossene Fragen								1	,587**
MW Rating: Offene Fragen									1
MW Rating: Erweiterung der Aussage									0
MW Rating: korrektive Wiederholung									0
MW Rating: Nachfragen									
MW Rating: Alltagsbezug									
MW Rating: Sätze vervollständigen									
MW Gesamtnote									
Durchführungsqualität (nur Zählitems)									
Durchführungsqualität (nur Ratings)									
Durchführungsqualität (Zählen + Rating + Note)									

**, Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

*, Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 59

	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
MW Zählen: Geschlossene Fragen	(1)	,411**	0,0221	,559**	,638**	,467**	,947**	,972**	,818**
MW Zählen: Offene Fragen	(2)	0,1967	-0,125	,487**	,714**	,670**	,924**	,884**	,865**
MW Zählen: Erweiterung der Aussage	(3)	,894**	-0,206	,261*	-0,122	-0,206	,402**	,478**	0,094
MW Zählen: korrektive Wiederholung	(4)	-0,225	,869**	0,0528	-0,039	0,1783	-0,015	0,2009	0,1943
MW Zählen: Nachfragen	(5)	,383**	-0,115	,806**	,565**	,414**	,816**	,807**	,719**
MW Zählen: Alltagsbezug	(6)	0,0474	0,2058	,549**	,895**	,393**	,805**	,791**	,831**
MW Zählen: Sätze vervollständigen	(7)	-,270*	-0,016	0,1474	,425**	,918**	,534**	,673**	,651**
MW Rating: Geschlossene Fragen	(8)	,367**	0,1198	,435**	,469**	,379**	,804**	,779**	,732**
MW Rating: Offene Fragen	(9)	-0,033	0,022	,469**	,869**	,663**	,793**	,705**	,914**
MW Rating: Erweiterung der Aussage	(10)	1	-0,252	0,2336	-0,04	-,388**	,347**	,299*	0,0595
MW Rating: korrektive Wiederholung	(11)	-0,252	1	0,0366	0,1681	0,062	0,0577	0,1323	,259*
MW Rating: Nachfragen	(12)	0,2336	0,0366	1	,496**	,267*	,639**	,595**	,614**
MW Rating: Alltagsbezug	(13)	-0,04	0,1681	,496**	1	,415**	,772**	,640**	,855**
MW Rating: Sätze vervollständigen	(14)	-,388**	0,062	,267*	,415**	1	,493**	,590**	,683**
MW Gesamtnote	(15)	,347**	0,0577	,639**	,772**	,493**	1	,918**	,904**
Durchführungsqualität (nur Zählitems)	(16)	,299*	0,1323	,595**	,640**	,590**	,918**	1	,868**
Durchführungsqualität (nur Ratings)	(17)	0,0595	,259*	,614**	,855**	,683**	,904**	,868**	1
Durchführungsqualität (Zählen + Rating + Note)	(18)	0,2357	0,1572	,640**	,796**	,609**	,977**	,953**	,963**
1									

**, Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

*, Korrelation ist bei Niveau 0,05 signifikant (zweiseitig).

N = 59

Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe, alle Ausführungen, die anderen Schriften wörtlich oder sinngemäß entnommen wurden, kenntlich gemacht sind und die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat. Ich stimme zu, dass die vorliegende Arbeit mit einer Anti-Plagiatssoftware überprüft werden darf.

Marburg, 13.10.2015

Nils Hartung