Effektive Sprachförderung konsequent zu Ende gedacht!

Von der Wirksamkeit einer Sprachfördermaßnahme im Alltag bis zur kontinuierlichen Überwachung des Entwicklungsverlaufs mithilfe einer neu entwickelten Lernverlaufsdiagnostik

Kumulative Dissertation
zur Erlangung des Grades der Philosophie (Dr. phil.)
der Pädagogische Hochschule Ludwigsburg

vorgelegt von Melanie Besca

aus Rottweil

Ludwigsburg 2025

Erstgutachter: Prof. Dr. Marco Ennemoser

Zweitgutachterin: Prof. Dr. Margit Berg

Drittgutachterin: Prof. Dr. Steffi Sachse

Datum des Abschlusses der mündlichen Prüfung: 08.04.2025

Inhaltsverzeichnis

1 Wissenschaftliches Tätigkeitsfeld und Forschungsarbeiten im Rahmen der Promotion 1

1.1 Promotionsrelevante Publikationen 1

1.2 Weitere Publikationen, Poster und Vorträge 2

2 Einleitung und Forschungsfragen 4

3 Wirksamkeit des Dialogischen Lesens im Kita-Alltag 9

3.1 Ergebnisse einer Studie zum Vergleich der Wirksamkeit des Dialogischen Lesens in kontrollierten und realitätsnahen Studienbedingungen (Publikation D) 9

3.2 Implikationen und offene Fragen zu Publikation D 11

4 Lernverlaufsdiagnostik zur Evaluation von individuellen Veränderungen 13

4.1 Begriffsklärung Lernverlaufsdiagnostik 13

4.2 Verschiedene Funktionen von Lernverlaufsdiagnostik 14

4.3 Einsatz von Lernverlaufsdiagnostik im binnendifferenzierten Unterricht (Publikation A) ... 15

4.4 Implikationen und offene Fragen 17

5 Entwicklung einer Lernverlaufsdiagnostik sprachlicher Fähigkeiten (LVD Sprache) 19

5.1 Das Sätze Nachsprechen als Globalmaß sprachlicher Fähigkeiten 19

5.2 Entwicklung der Sätze für die LVD Sprache 20

6 Erprobung der LVD Sprache im Quer- und Längsschnitt 24

6.1 Rekrutierung und Stichproben 24

6.2 Vorgehen bei der Erprobung einer Lernverlaufsdiagnostik 25

6.3 Erhebungsinstrumente 27

6.4 Erste Erprobungsergebnisse der LVD Sprache (Publikation B) 29

6.5 Ergebnisse der Pilotierung am SBBZ Sprache (Publikation C) 31

6.6 Ergebnisse der Pilotierung im Kindergarten (Publikation E) 32

6.7 Erste Ergebnisse aus Einzelfallstudien 35

7 Diskussion 38

8 Fazit und Ausblick 45

9 Literaturverzeichnis 49

10 Anhang 57

10.1 Danksagung 57

10.2 Erklärung zu den Publikationen 58

10.3 Eidesstaatliche Erklärung 59

1 Wissenschaftliches Tätigkeitsfeld und Forschungsarbeiten im Rahmen der Promotion

Die Promotion fand im Förderschwerpunkt Sprache (Psychologie/ Diagnostik), Fakultät für Teilhabewissenschaften der Pädagogischen Hochschule Ludwigsburg statt. Der inhaltliche Schwerpunkt des Forschungsgebietes umfasste die evidenzbasierte Sprachförderung in Kindestagesstätten und die Lernverlaufsdiagnostik sprachlicher Fähigkeiten.

Die promotionsrelevanten Publikationen entstanden im Rahmen des BMBF-geförderten Projekts „Dialogisches Lesen als Tool für eine wirksame Praxis der Sprachförderung in Kindertagesstätten (DiaLes)“, Förderkennzeichen 01GJ1703A sowie im eigenverantwortlich durchgeführten Projekt zur Entwicklung und Erprobung einer Lernverlaufsdiagnostik sprachlicher Fähigkeiten. Darüber hinaus wurden innerhalb der Forschungstätigkeit sowie der Tätigkeit als akademische Mitarbeiterin im Förderschwerpunkt Sprache noch weitere Poster, Publikationen und Vorträge bei Tagungen verfasst und präsentiert.

1.1 Promotionsrelevante Publikationen

Die vorliegende kumulative Dissertation basiert auf den folgenden Veröffentlichungen, die in der nachfolgenden Arbeit mit den Buchstaben A-E gekennzeichnet werden. Die fünf Publikationen wurden in nationalen und internationalen Zeitschriften und einem Sammelband veröffentlicht, angenommen und eingereicht. Vier von fünf Publikationen wurden als Erstautorin geschrieben. Die Beiträge C und D haben das Peer-Review-Verfahren durchlaufen, der Beitrag E ist ebenfalls in dieser Kategorie eingereicht. In der Erklärung zu den Publikationen im Anhang wird der Eigenanteil bzw. die Rolle der Mitautor*innen zu jeder Publikation dargestellt. Eine Kopie der Publikationen befindet sich in dieser Reihenfolge im Anhang.

A. Ennemoser, M. & Besca, M. (2021). Präventive Mehrebenenförderung. *Pädagogik*, 73(4), 44–47.

B. Besca, M. & Ennemoser, M. (2022). Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten und zur Evaluation sprachlicher Fördermaßnahmen? In M. Spreer, M. Wahl & H. Beek (Hrsg.), *Sprachheilpädagogik aktuell: Bd. 4, Sprachentwicklung im Dialog: Digitalität - Kommunikation - Partizipation* (S. 205–212). Schulz-Kirchner Verlag.

C. Besca, M., Krauss, T. & Ennemoser, M. (2023). Pilotierung einer Lernverlaufsdiagnostik bei Kindern mit Sprachentwicklungsstörung am SBBZ Sprache. *Forschung Sprache* (2), 16–26. https://www.forschung-sprache.eu/fileadmin/user_upload/Dateien/Heftausgaben/2023-2/Forschung_Sprache_2_2023_Besca_Krauss_Ennemoser2.pdf

D. Besca, M., Reiner, K. & Ennemoser, M. (2025): Evaluation des Dialogischen Lesens im Kita-Alltag. Ein Vergleich der Wirksamkeit in kontrollierten und realitätsnahen Studienbedingungen. *Logos 33 (1)*.

E. Besca, M. & Ennemoser, M. (eingereicht): Assessing Oral Language Progress in Children Aged 4–6. Evaluation of the Psychometric Properties.

1.2 Weitere Publikationen, Poster und Vorträge

Publikationen in Sammelbänden und Fachzeitschriften

- Christl, J., Verl, E., Besca, M., Krajewski, K. & Ennemoser, M. (2024): Entwicklungsorientierte Diagnostik mathematischer Basiskompetenzen in Förderschulen Lernen, nicht-gymnasialer und gymnasialer Sekundarstufe von der 5. bis 9. Klasse – Eine Längsschnittstudie (EnDiMath). In: Beck, K.; Ferdigg, R.A; Katzenbach, D.; Kett-Hauser, J.; Laux, S.; Michael Urban, M. (Hrsg.). Fachbezogene Diagnostik in der inklusiven Bildung (S. 271-286). Münster.
- Besca, M. (2023). Sprachheilpädagogische Unterstützung bei Kindern mit selektivem Mutismus. *Sprachförderung und Sprachtherapie* (1), 8–13.
- Besca, M. & Ennemoser, M. (in Druck). Sprache fördern. In G. W. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen: Förderung, Training und Therapie in der Praxis*.
- Ennemoser, M. & Besca, M. (in Druck). Response to Intervention als schulisches Förderkonzept. In G. W. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen: Förderung, Training und Therapie in der Praxis*.

Vorträge bei Tagungen und Kongressen

- Besca, M.; Ennemoser, M. (2024, 15. November): Erfassung von Sprachentwicklungsverläufen mithilfe der LVD Sprache. [Konferenzbeitrag] 13. Interdisziplinäre Tagung über Sprachentwicklungsstörungen (ISES) der GISKID, Halle an der Saale.
- Besca, M. & Ennemoser, M. (2024, 27. September). Längsschnittergebnisse der Erprobung einer Lernverlaufsdiagnostik sprachlicher Fähigkeiten bei 4- bis 6-jährigen Kindern. [Konferenzbeitrag] 35. Bundeskongress der Deutschen Gesellschaft für Sprachheilpädagogik (dgs), Heidelberg.
- Besca, M. & Ennemoser, M. (2024, 8. Juni). Pilotierung einer LVD Sprache – Analyse der Gütekriterien. [Konferenzbeitrag] Frühjahrestagung der Arbeitsgruppe Empirische Sonderpädagogische Forschung (AESF), Zürich.
- Besca, M. & Ennemoser, M. (2023, 28. Februar). Entwicklung und Erprobung einer Lernverlaufsdiagnostik Sprache. [Konferenzbeitrag] Gesellschaft für Empirische Bildungsforschung (10. Tagung der GEBF), Essen.
- Ennemoser, M. & Besca, M. (2023, 2. März). Die Wirksamkeit adaptiver Sprachförderung nach den Prinzipien des Dialogischen Lesens. [Konferenzbeitrag] Gesellschaft für Empirische Bildungsforschung (10. Tagung der GEBF), Essen.
- Besca, M. & Ennemoser, M. (2022, 9. März). Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten? [Konferenzbeitrag]. 9. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), digital.
- Besca, M. & Ennemoser, M. (2022, 19. Mai). Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten und zur Evaluation sprachlicher Fördermaßnahmen? [Konferenzbeitrag]. 37. Arbeitstagung der Fachgruppe "Psychologie der Kommunikation und ihrer Störungen", Rauischholzhausen.
- Besca, M. & Ennemoser, M. (2022, 22. September). Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten und zur Evaluation sprachlicher Fördermaßnahmen? [Konferenzbeitrag]. 34. Bundeskongress der Deutschen Gesellschaft für Sprachheilpädagogik (dgs), Berlin.

- Besca, M., Ennemoser, M. & Reiner, K. (2020, 20. November). Sprachliche Fähigkeiten von Kindern einfach und effektiv mit dem Dialogischen Lesen (DL) im Alltag fördern: aktuelle Forschungsergebnisse [Konferenzbeitrag]. 11. Interdisziplinäre Tagung über Sprachentwicklungsstörungen (ISES) der GISKID, digital.
- Besca, M., Ennemoser, M., Reiner, K. & Kuhn, D. (2019, 9. September). Kann die Wirksamkeit des Dialogischen Lesens durch die Kombination mit einem Grammatiktraining optimiert werden? [Konferenzbeitrag]. Tagung der Fachgruppen „Pädagogische Psychologie“ und „Entwicklungspsychologie“ (paEpsy), Leipzig.
- Kuhn, D., Reiner, K., Besca, M. & Ennemoser, M. (2018, 15. September). Dialogisches Lesen - Varianten des dialogischen Lesens zur Sprachförderung in Kindertagesstätten [Konferenzbeitrag]. 51. Kongress der Deutschen Gesellschaft für Psychologie (DGPs), Frankfurt am Main.

Poster

- Besca, M. & Ennemoser, M. (2022, 25. November). Pilotierung einer Lernverlaufsdiagnostik sprachlicher Fähigkeiten am SBBZ Sprache [Posterpräsentation]. 12. Interdisziplinäre Tagung über Sprachentwicklungsstörungen (ISES). GISKID, Marburg.
- Besca, M. & Ennemoser, M. (2020, 19. November). Sätze Nachsprechen als Lernverlaufsdiagnostik sprachlicher Fähigkeiten? [Posterpräsentation]. Nachwuchsschool der 11. Interdisziplinäre Tagung über Sprachentwicklungsstörungen (ISES). GISKID, digital.
- Besca, M. & Ennemoser, M. (2020, 11. Februar). Eignet sich das Nachsprechen von Sätzen als Möglichkeit zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten? [Posterpräsentation]. Forschungstag, Ludwigsburg.
- Besca, M., Reiner, K., Kuhn, D. & Ennemoser, M. (2018, 7. Februar). Dialogisches Lesen als Tool für eine wirksame Praxis der Sprachförderung in Kindertagesstätten (DiaLes) [Posterpräsentation]. Forschungstag, Ludwigsburg.

2 Einleitung und Forschungsfragen

Die Relevanz sprachlicher Fähigkeiten für den individuellen Bildungserfolg (Glück & Spreer, 2015) sowie die Bildungssprache im Grundschulalter (Wimmer et al., 2022) sind unbestritten. "Oral language abilities enable children to learn to read, and they predict future academic achievement and life outcomes." (Adlof & Hogan, 2019, S. 210). Insbesondere für Kinder mit Zuwanderungshintergrund zeigt sich ein deutlicher Nachteil hinsichtlich ihrer sprachlichen Voraussetzungen in Bezug auf schulische Leistungen (Kempert et al., 2016). Auch Kinder mit einer Sprachentwicklungsstörung sind in ihrer schulischen Laufbahn häufig beeinträchtigt und erreichen mitunter eine weniger qualifizierte berufliche Beschäftigung oder Position (Conti-Ramsden et al., 2018). Der Sprachförderung im Elementarbereich kommt daher eine immer größere Bedeutung zu. Hier werden verschiedene Präventionsmaßnahmen unterschieden. Eine alltagsintegrierte sprachliche Bildung ist für alle Kinder wichtig, um Risiken für die sprachliche Entwicklung zu minimieren (primäre Prävention). Kinder mit Sprachförderbedarf und Risiko für Bildungsmisserfolge benötigen dagegen zusätzliche pädagogische Förderung (sekundäre Prävention). Kinder mit Sprachentwicklungsstörungen bedürfen einer therapeutischen Intervention (tertiäre Prävention) (Titz & Hasselhorn, 2017).

Für den Bereich der sekundären Prävention gibt es mittlerweile eine Vielzahl an Sprachfördermaßnahmen (Überblick in Egert & Hopf, 2016), darunter auch evidenzbasierte Programme wie „Mit Kindern im Gespräch“ und das Heidelberger Interaktionstraining. Jedoch ist die Befundlage oft inkonsistent (Egert & Hopf, 2016). Bisherige Versuche sowohl im vorschulischen als auch im schulischen Bereich haben nur begrenzt zu Verbesserungen geführt. Die Ergebnisse von Evaluations- und Begleitstudien zu Sprachförderkonzepten im Kita-Bereich fallen überwiegend enttäuschend aus oder weisen keine bzw. nur geringe Effekte auf die Sprachkompetenz von Kindern auf (Überblick z.B. Schneider, 2018 und Egert et al., 2020). Auch langfristig zeigen die Sprachfördermaßnahmen noch wenig Wirkung, wie die aktuellen PISA-Ergebnisse von 2022 zeigen. Hier wurden nicht mehr die sprachlichen Fähigkeiten direkt überprüft, jedoch die Leseleistungen, die in einem direkten Zusammenhang mit den vorschulischen sprachlichen Fähigkeiten stehen (z.B. Ennemoser et al., 2012). Die PISA-Ergebnisse zeigen, dass sich die Leseleistungen im Vergleich zu den Jahren 2018 und 2015 signifikant verschlechtert haben und ein signifikanter Anstieg des Anteils besonders leseschwacher Schüler*innen zu verzeichnen ist (Lewalter et al., 2023).

Aus diesem Grund wurden aktuell in Deutschland zwei umfangreiche Bildungsmaßnahmen verabschiedet bzw. geplant. Das bislang größte und langfristigste Bildungsprogramm in der Geschichte der Bundesrepublik Deutschland, das Startchancenprogramm, zielt darauf ab, Schulen in sozial benachteiligten Regionen sowie solche, in denen ein hoher Anteil an Kindern mit Migrationshintergrund lernt, in besonderem Maße zu fördern (Bundesministerium für Bildung und Forschung, 2024). In Baden-

Württemberg ist darüber hinaus das Paket "Sprachfit" zur Sprachförderung von Kindern geplant. Das Programm beinhaltet unter anderem eine verpflichtende Sprachförderung von vier Stunden pro Woche in Gruppen. Diese soll im Jahr vor der Einschulung für Kita-Kinder stattfinden, bei denen im Rahmen ihrer Einschulungsuntersuchung erhebliche Sprachdefizite festgestellt. Zudem umfasst es sogenannte Juniorklassen in Grundschulen für Kinder, die auch nach Absolvierung der Sprachförderung in den Kita-Gruppen noch nicht über ausreichende Deutschkenntnisse verfügen. Des Weiteren ist vorgesehen, dass in den Klassenstufen 1 und 2 eine Sprachförderung als regulärer Bestandteil des Lehrplanes implementiert wird (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2024).

Für die verpflichtende Sprachförderung von vier Stunden pro Woche in Gruppen im Jahr bietet sich das Dialogische Lesen als evidenzbasierte Maßnahme der Sprachförderung für den Kita-Bereich an. Die Wirksamkeit des Dialogischen Lesens konnte in zahlreichen Studien nachgewiesen werden (z.B. Ennemoser et al., 2015; Pillinger & Vardy, 2022; Towson et al., 2017.) Allerdings kann aus Wirksamkeitsnachweisen, die im Rahmen wissenschaftlicher Untersuchungen unter gut kontrollierten Bedingungen gewonnen wurden, keine Garantie abgeleitet werden, dass die jeweilige Sprachfördermaßnahme auch im Kita-Alltag wirksam ist. Die Durchführungsbedingungen im Alltag sind im Gegensatz zu einer gut kontrollierten Bedingung in wissenschaftlichen Studien starken Schwankungen unterworfen (Treatment-Validität; Ennemoser, 2006). Zudem wird die Förderung in den meisten Fällen nicht von pädagogischen Fachkräften, sondern von Projektmitarbeiter*innen durchgeführt (Marulis & Neuman, 2010; Mol et al., 2009). Für das Dialogische Lesen liegen bereits umfangreiche Wirksamkeitsnachweise vor, allerdings besteht noch Forschungsbedarf bei der Frage nach der Wirksamkeit bei der Durchführung im regulären Kita-Alltag. Das beinhaltet auch, dass die Förderung von pädagogischen Fachkräften selbst durchgeführt wird. Die Frage nach der Wirksamkeit im Kita-Alltag kommt häufig im Rahmen von Untersuchungen zur Evidenzbasierung von Sprachfördermaßnahmen zu kurz (Rauch & Hartig, 2018) und soll daher im Rahmen dieser Arbeit in Kapitel 3 beantwortet werden.

> Forschungsfrage 1: Ist das Dialogische Lesen als Sprachfördermaßnahme auch unter den weniger gut kontrollierbaren Alltagsbedingungen wirksam?

Die Auswahl einer evidenzbasierten Sprachfördermaßnahme, auch wenn Nachweise für ihre Wirksamkeit im Kita-Alltag vorliegen, reicht jedoch nicht aus, um sicherzustellen, dass alle Kinder auch ausreichend von dieser Maßnahme profitieren. Evidenzbasierung ist ein wesentliches Kriterium bei der Auswahl einer Sprachfördermaßnahme, aber keine Garantie dafür, dass jedes Kind im Verlauf der Förderung automatisch in ausreichendem Maße von der gewählten Maßnahme profitiert. Die Auswahl einer Sprachfördermaßnahme vor dem Hintergrund aktueller wissenschaftlicher Erkenntnisse, die auch als externe Evidenz bezeichnet wird (Y. Blumenthal & Mahlau, 2015), kann daher als notwendige Voraussetzung, aber nicht als hinreichende Maßnahme zur Qualitätssicherung angesehen werden.

Um den Erfolg einer Sprachfördermaßnahme überprüfen zu können, sollte eine Evaluation nach Abschluss der Maßnahme selbstverständlich sein. Y. Blumenthal und Mahlau (2015) stellen dies in ihrem Beitrag sehr anschaulich dar, so dass deutlich wird, dass zu einem evidenzbasierten Entscheidungsprozess eine abschließende Evaluation gehört, in der überprüft wird, ob die gewählte Maßnahme für das einzelne Kind, die Gruppe oder die Klasse wirksam war.

Die Evaluation einer Fördermaßnahme am Ende der Maßnahme, z.B. nach einem halben oder ganzen Jahr, birgt jedoch die Gefahr, dass wertvolle Zeit verloren geht, wenn zu diesem Zeitpunkt festgestellt wird, dass einzelne Kinder nicht ausreichend von der Förderung profitiert haben. Idealerweise sollte die Entwicklung bereits prozessbegleitend evaluiert werden. Um eine Sprachfördermaßnahme bereits prozessbegleitend zu evaluieren, eignen sich Verfahren der Lernverlaufsdiagnostik. Dabei handelt es sich um kurze, schnell durchführbare und ökonomisch auswertbare Tests, die aus mehreren Paralleltests bestehen und in kurzen Zeitabständen (z.B. wöchentlich, monatlich) eingesetzt werden können. Auf diese Weise kann eine Lernentwicklung sensitiv erfasst oder eine Fördermaßnahme evaluiert werden (Deno, 2003). Werden Entwicklungsveränderungen bereits prozessbegleitend zu einer Fördermaßnahme evaluiert, kann frühzeitig festgestellt werden, in welchem Ausmaß Kinder von einer Förderung profitieren. In diesem Kontext werden die Begriffe "Responder" und "Nonresponder" verwendet (vgl. RTI-Konzept, Deno, 2003). Bei Kindern, die nicht ausreichend von einer Fördermaßnahme profitieren, kann dann frühzeitig reagiert werden, indem die Maßnahmen modifiziert oder ggf. eine noch differenziertere und umfangreichere Diagnostik eingeleitet wird.

Die prozessbegleitende Evaluation einer Sprachfördermaßnahme mit Hilfe einer Lernverlaufsdiagnostik stellt eine sinnvolle Ergänzung zu einer abschließenden Evaluation dar. Für den Bereich der mündlichen sprachlichen Fähigkeiten liegen derzeit jedoch keine Verfahren der Lernverlaufsdiagnostik vor. Hier besteht sowohl für den vorschulischen als auch für den schulischen Bereich eine Forschungslücke (Besca & Ennemoser, 2022; Mahlau et al., 2014). Bisherige Verfahren zur Lernverlaufsdiagnostik liegen vor allem für Schulkinder im Bereich der Schriftsprache vor (Überblick z.B. in Winkes & Till, 2023 oder Gebhardt et al., 2021). Aufgabe der Forschung ist daher, in Zukunft die Entwicklung von Verfahren der Lernverlaufsdiagnostik für mündliche Sprachkompetenzen voranzutreiben (Adlof & Hogan, 2019) und „weiterhin an zuverlässigen und praktikablen Instrumenten [zu arbeiten], spezifisch auch für die mündliche Sprache und das therapeutische (Einzel-)Setting“ (Winkes & Till, 2023, S. 254). Auch Ehlich et al. (2008) weisen darauf hin, dass die einmalige Erfassung des Sprachstandes mit einem Messinstrument nicht ausreicht, um eine sprachliche Qualifikationsverbesserung zu erreichen. Es wird als notwendig erachtet, „vorhandene Verfahren zur Erfassung des Sprachstandes und deren Veränderung mit Blick auf gezielte Förderungen und deren Evaluation“ weiterzuentwickeln (Ehlich et al., 2008, S. 12). Für Kinder mit Migrationshintergrund stellen Gogolin und Kolleg*innen in ihrem Gutachten

2003 fest, dass neben der Anschlussfähigkeit von Verfahren zur Sprachstandsfeststellung und Sprachförderung auch die Sprachentwicklung längsschnittlich evaluiert werden sollte, um die gewählte Sprachfördermaßnahme zu evaluieren (Gogolin et al., 2003).

Da es bisher keine Lernverlaufsdiagnostik im Bereich der mündlichen Sprache gibt, muss diese erst entwickelt und erprobt werden. Für die Konstruktion einer Lernverlaufsdiagnostik kann ein robuster Indikator herangezogen werden, der eine Kompetenz als Gesamtmaß repräsentiert (Gebhardt et al., 2021). Das Nachsprechen von Sätzen scheint ein vielversprechendes Globalmaß sprachlicher Fähigkeiten zu sein. „Sentence repetition appears to be a valuable tool for language assessment because it draws upon a wide range of language processing skills." (Klem et al., 2015, S. 146). Aus der dargestellten Forschungslücke ergibt sich der zweite Forschungsschwerpunkt dieser Arbeit.

> Forschungsfrage 2: Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten?

Zur Beantwortung dieser Frage wird in Kapitel 4 die Lernverlaufsdiagnostik als Methode zur systematischen Erfassung und Evaluation von Entwicklungsveränderungen vorgestellt. In Kapitel 5 wird die die Entwicklung einer Lernverlaufsdiagnostik sprachlicher Fähigkeiten (LVD Sprache) vorgestellt, gefolgt von der Erprobung der LVD Sprache im Quer- und Längsschnitt anhand der Überprüfung der klassischen Gütekriterien, der Änderungssensitivität sowie der Eignung für verschiedene Zielgruppen in Kapitel 6. Im Kapitel 7 werden beide Forschungsfragen abschließend beantwortet und diskutiert, bevor in Kapitel 8 ein Ausblick auf das weitere Vorgehen und mögliche Forschungsschwerpunkte gegeben wird.

Übergeordnetes Ziel der vorliegenden Arbeit ist es, die Wirksamkeit von Sprachförderung konsequent zu Ende zu denken, indem zum einen die Wirksamkeit einer Sprachfördermaßnahme am Beispiel des Dialogischen Lesens im Kita-Alltag überprüft wird und zum anderen eine Lernverlaufsdiagnostik sprachlicher Fähigkeiten als neue Möglichkeit, die Wirksamkeit von Sprachförderung im Einzelfall zu evaluieren, entwickelt und erprobt wird.

Die folgende Abbildung (Abb. 1) soll den idealtypischen Ablauf bei der Auswahl und Evaluation einer Sprachfördermaßnahme darstellen, um die Forschungsschwerpunkte und den Aufbau dieser Arbeit zu verdeutlichen. Wird bei Kindern ein Sprachförderbedarf festgestellt, sollte eine evidenzbasierte Sprachfördermaßnahme zur Durchführung der Förderung ausgewählt werden (Punkt A in der Abbildung). Idealerweise sollte eine Sprachfördermaßnahme ausgewählt werden, für die Wirksamkeitsnachweise im Alltag vorliegen (Punkt B in der Abbildung). Für das Dialogische Lesen soll im Rahmen dieser Arbeit die Wirksamkeit im Kita-Alltag überprüft werden (Kapitel 3). Um die Wirksamkeit der Sprachfördermaßnahme nicht erst nach deren Abschluss, sondern bereits prozessbegleitend zu evaluieren, kann ein Verfahren der Lernverlaufsdiagnostik eingesetzt werden

Da es für den Bereich der mündlichen sprachlichen Fähigkeiten noch keine Verfahren der Lernverlaufsdiagnostik gibt, muss zunächst ein Verfahren der LVD entwickelt und erprobt werden (Kapitel 4-6). Mit Hilfe einer Lernverlaufsdiagnostik kann dann der Sprachentwicklungsverlauf evaluiert werden (ab Punkt B), um möglichst frühzeitig die Kinder zu identifizieren, die nicht ausreichend von der Sprachfördermaßnahme profitieren, um die Maßnahme unmittelbar modifizieren oder ggf. weitere diagnostische Schritte einleiten zu können (Punkt E). Für die Kinder, die ausreichend von der Fördermaßnahme profitieren, wir die Förderung fortgesetzt (Punkt D).

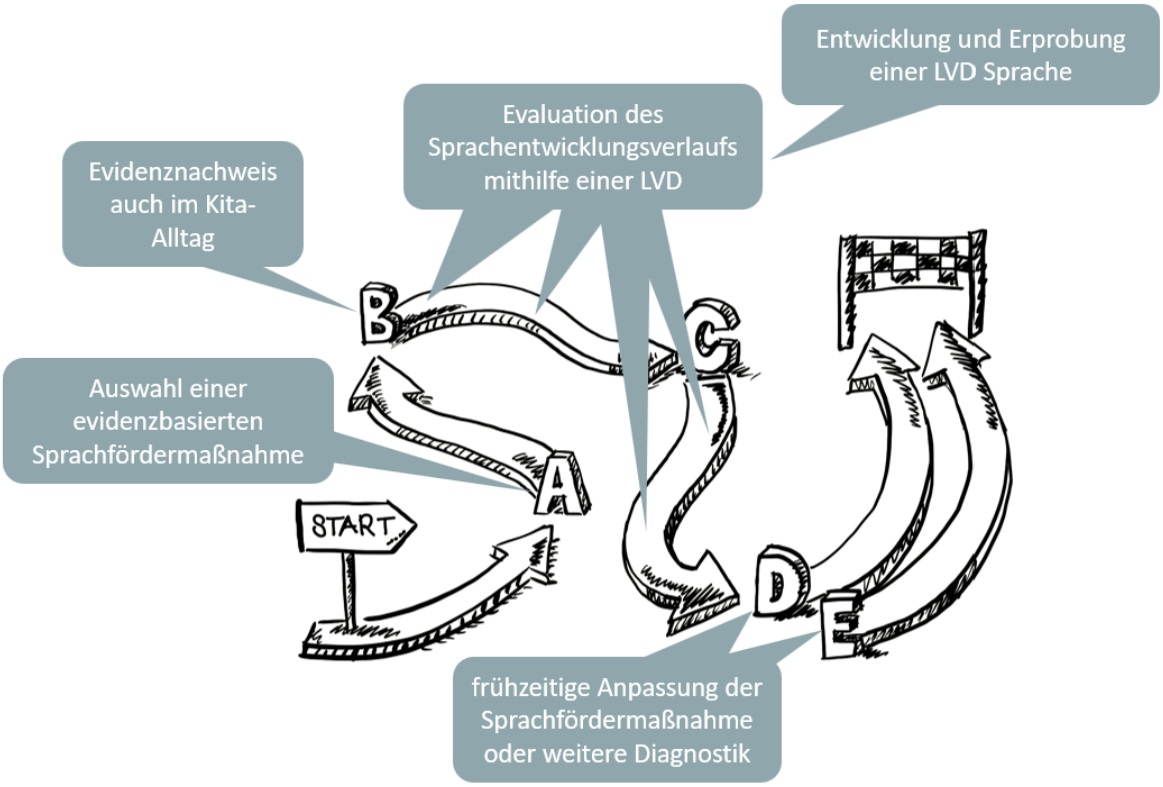

Abbildung 1 Vorgehen bei der Auswahl und Evaluation einer Sprachfördermaßnahme

3 Wirksamkeit des Dialogischen Lesens im Kita-Alltag

3.1 Ergebnisse einer Studie zum Vergleich der Wirksamkeit des Dialogischen Lesens in kontrollierten und realitätsnahen Studienbedingungen (Publikation D)

Beim Dialogischen Lesen handelt es sich um eine Sprachfördermethode, für die mittlerweile sowohl auf internationaler als auch auf nationaler Ebene umfangreiche Wirksamkeitsnachweise vorliegen. In der im Folgenden vorgestellten Studie konnte zum einen die Wirksamkeit des Dialogischen Lesens unter realitätsnahen Studienbedingungen, d.h. im regulären Kita-Alltag, im Vergleich zu einer Kontrollgruppe nachgewiesen werden. Zum anderen konnte durch den Vergleich mit den Ergebnissen einer Studie in einem kontrollierten Setting gezeigt werden, dass pädagogische Fachkräfte im Kita-Alltag ähnliche Fördereffekte erzielen können wie Fachkräfte in einem kontrollierten Setting. Damit konnten die bisherigen Erkenntnisse zum Dialogischen Lesen erweitert werden. Darüber hinaus konnte gezeigt werden, dass es möglich ist, die Methode des Dialogischen Lesens zeitökonomisch mit einer Fortbildung von 1 ½ Tagen zu vermitteln. Diese Ergebnisse liefern somit einen wichtigen Beleg für die Implementierbarkeit des Dialogischen Lesens in den Kita-Alltag.

> Besca, M., Reiner, K. & Ennemoser, M. (2025): Evaluation des Dialogischen Lesens im Kita-Alltag. Ein Vergleich der Wirksamkeit in kontrollierten und realitätsnahen Studienbedingungen. *Logos 33 (1)*.

Theoretischer Hintergrund

Im Zentrum des Dialogischen Lesens steht der Einsatz von Sprachförderstrategien, die auf der natürlichen Eltern-Kind-Kommunikation basieren und erstmals von Whitehurst et al. (1988) systematisch untersucht wurden. Diese Strategien werden eingesetzt, um die sprachlichen Fähigkeiten der Kinder durch gezielte Fragen, das Modellieren von Äußerungen sowie durch Verstärkung und Lob zu fördern (Ennemoser et al., 2013). Beim Dialogischen Lesen wird das Setting der gemeinsamen Bilderbuchbetrachtung genutzt, um diese Sprachförderstrategien gezielt einzusetzen. Sowohl auf internationaler (Fitton et al., 2018; Flack et al., 2018; Mol et al., 2009; Pillinger & Vardy, 2022; Swanson et al., 2011; Towson et al., 2017) als auch auf nationaler Ebene (Ennemoser & Hartung, 2017; Ennemoser et al., 2013; Ennemoser et al., 2015) liegen umfangreiche Wirksamkeitsstudien und Metaanalysen zum Dialogischen Lesen vor. Diese berichten überwiegend positive Effekte auf eine Bandbreite an Sprach- und Literacyfähigkeiten im Kita-Alter (Neuman & Kaefer, 2018; Pillinger & Vardy, 2022) oder für unterschiedliche Zielgruppen, z.B. mehrsprachige Kinder (Ennemoser et al., 2013). Allerdings gibt es auch Studien, die weniger positiv ausfallen (Noble et al., 2019) oder die berichten, dass es den Fachkräften nicht gelingt, ihre Sprache an den Sprachentwicklungsstand des Kindes anzupassen oder Förderstrategien adaptiv einzusetzen (Geyer & Müller, 2021).

Meta-Analyse zeigen, dass sich die Durchführungsbedingungen (Gesamtdauer der Förderung, Häufigkeit pro Woche, Dauer pro Sitzung und Materialauswahl) in den Studien zum Dialogischen Lesen deutlich unterscheiden, dies aber keinen oder nur einen geringen Einfluss auf die Wirksamkeit hat (z.B. Marulis & Neuman, 2010; Mol et al., 2009). Es ist daher davon auszugehen, dass das Dialogische Lesen auch unter weniger kontrollierten Bedingungen im Kita-Alltag wirksam sein könnte. Die Befundlage zum Einfluss der Förderkraft auf die Sprachfördereffekte ist nicht eindeutig (Flack et al., 2018; Marulis & Neuman, 2010; Mol et al., 2009).

Zahlreiche Studien zeigen, dass die Qualifizierung der Fachkräfte eine entscheidende Rolle für den Erfolg von Sprachfördermaßnahmen spielt (Böse et al., 2023; Buschmann & Jooss, 2011; Kammermeyer et al., 2019; Towson et al., 2017; Voltmer et al., 2021). Bekannt ist auch, dass Fortbildungen, die auf dem Ansatz des situierten Lernens basieren (z.B. Stark & Klauer, 2019) und praxisnahe Reflexionen einbeziehen (Egert et al., 2017), besonders effektiv sind.

Vor dem Hintergrund bestehender Forschungslücken zur Qualifikation der durchführenden Fachkräfte sowie zum Dialogischen Lesen im Kita-Alltag möchte die vorliegende Studie folgende Hypothesen beantworten: (1) Es gibt keine signifikanten Unterschiede in den sprachlichen Fortschritten zwischen Kindern, die von pädagogischen Fachkräften und solchen, die von wissenschaftlichen Mitarbeiter*innen gefördert werden. (2) Kinder, die von pädagogischen Fachkräften im Kita-Alltag mit Dialogischem Lesen gefördert wurden, machen größere Fortschritte als Kinder einer Kontrollgruppe. (3) Es gibt keine signifikanten Unterschiede in den sprachlichen Fortschritten zwischen Kindern, die im Kita-Alltag gefördert wurden, und solchen, die unter kontrollierten Bedingungen gefördert wurden.

Methode

Die vorliegende Studie analysiert Daten aus zwei Studien zum Dialogischen Lesen, um dessen Wirksamkeit bei Vorschulkindern mit Sprachförderbedarf in Kleingruppen zu untersuchen. Beide Studien folgten einem Prä-Post-Interventions-Design mit Experimental- und Kontrollgruppe und beinhalteten Sprachscreenings zur Identifikation des Sprachförderbedarfs. In der ersten Studie wurden 374 Kinder aus 76 Kindertagesstätten in Mittelhessen, in der Folgestudie 201 Kinder aus 46 Kindertagesstätten in Baden-Württemberg untersucht. Die Förderung erfolgte durch pädagogische Fachkräfte oder wissenschaftliche Mitarbeiter*innen, die hierfür eine 1 ½-tägige Fortbildung erhielten. Im Prä- und Posttest wurden die sprachlichen Fähigkeiten mit vier Subtests (Bildbenennung, Handlungssequenz, Satz- und Pluralbildung) des SET 5-10 (Petermann, 2018) überprüft. Die Daten wurden deskriptiv und inferenzstatistisch mit Varianzanalysen ausgewertet, um die Wirksamkeit der Intervention zu beurteilen. Die Besonderheit der gesamten Untersuchung bestand darin, dass die Förderung der Kinder aus Studie 1 von pädagogischen Fachkräften oder wissenschaftlichen Mitarbeiter*innen in einem eng kontrollierten Studiendesign durchgeführt wurde, während die Förderung der Folgestudie im regulären Kita-

Alltag von pädagogischen Fachkräften durchgeführt wurde. Die Fortbildung wurde für die Folgestudie auf der Grundlage aktueller wissenschaftlicher Erkenntnisse und der Ergebnisse der ersten Studie aktualisiert.

Ergebnisse und Diskussion

Für die Daten der Studie 1 zeigten sich keine Unterschiede zwischen Kindern, die von pädagogischen Fachkräften, und solchen, die von wissenschaftlichen Mitarbeiter*innen gefördert wurden. Allerdings unterschieden sich die von pädagogischen Fachkräften geförderten Kinder nicht von der Kontrollgruppe. Ein Follow-up-Test zu Studie 1 zeigte jedoch langfristig stabilere und signifikant größere Lernzuwächse bei den von pädagogischen Fachkräften geförderten Kindern (Dickel-Lehnigk et al., 2022).

Mit den Daten der Folgestudien konnte gezeigt werden, dass die Kinder, die von den pädagogischen Fachkräften im Kita-Alltag mit dem Dialogischen Lesen gefördert wurden, signifikant größere sprachliche Kompetenzzuwächse aufwiesen als die Kinder der Kontrollgruppe. Fördereffekte zeigten sich im Wortschatz und im Sprachverständnis, nicht aber in der Grammatik. Besonders ermutigend sind die Fördereffekte auf das Sprachverständnis als komplexe sprachliche Fähigkeit (Pillinger & Vardy, 2022). Die Anzahl der Fördersitzungen hatte keinen Einfluss auf die Fördereffekte, was als Hinweis darauf gewertet werden kann, dass die pädagogischen Fachkräfte die Sprachförderstrategien vermutlich auch in ihrem sonstigen Alltag anwenden.

Durch den Vergleich der Daten aus Studie 1 und der Folgestudie konnte gezeigt werden, dass die Fördererfolge im Kita-Alltag und unter eng kontrollierten Bedingungen vergleichbar sind. Dies stellt eine wichtige Erkenntnis für eine mögliche Implementierung des Dialogischen Lesens in den Kita-Alltag dar. Darüber hinaus legen die Ergebnisse nahe, dass pädagogische Fachkräfte auch im Kita-Alltag mit relativ geringem Fortbildungsaufwand signifikante Sprachfördererfolge erzielen können. Zukünftige Forschung sollte sich auf die langfristigen Effekte und die Übertragbarkeit der Sprachförderstrategien auf andere Alltagssituationen konzentrieren.

3.2 Implikationen und offene Fragen zu Publikation D

Als Aspekte zukünftiger Forschung wurden in diesem Beitrag zum einen Follow-up-Untersuchungen als zentral erachtet, um die Fördereffekte auch langfristig zu sichern. Zum anderen sollten die Sprachförderstrategien auch in anderen Alltagsformaten untersucht werden, um herauszufinden, ob die Sprachförderstrategien auch auf andere Formate im Kita-Alltag übertragbar sind. Darüber hinaus sollte die noch offene Frage beantwortet werden, wie es gelingen kann, frühzeitig Kinder zu identifizieren, die trotz der evidenzbasierten Fördermethode nicht ausreichend von der Förderung profitieren.

In weiteren Analysen zur Studie 1 zum Dialogischen Lesen konnte festgestellt werden, dass nicht alle Kinder gleichermaßen von der Förderung mit dem Dialogischen Lesen profitieren (Dickel-Lehnigk et al., 2022). Betrachtet man die geförderte Gruppe genauer, so zeigt sich, dass von den Kindern mit Sprachförderbedarf, deren Ergebnisse im Prätest zwischen 0,5 und 1,5 Standardabweichungen unter dem Mittelwert liegen, nach einem halben Jahr Förderung 70% der Kinder im Posttest und 83% der Kinder im Follow-up keinen Sprachförderbedarf mehr aufweisen. Das bedeutet aber auch, dass bei 17% auch nach der Förderung noch Sprachförderbedarf besteht (Dickel-Lehnigk et al., 2022). Liegt bei den Kindern Sprachförderbedarf im Sinne einer Sprachentwicklungsstörung vor, d.h. mehr als 1,5 Standardabweichungen unter dem Mittelwert, so weisen im Posttest noch 19% und im Follow-up 12% der Kinder Ergebnisse von mehr als 1,5 Standardabweichungen unter dem Mittelwert auf (Dickel-Lehnigk et al., 2022). Bei diesen Kindern mit Sprachförderbedarf, deren Ergebnisse zwischen 0,5 und 1,5 Standardabweichungen bzw. mehr als 1,5 Standardabweichungen unter dem Mittelwert liegen, spricht man von sogenannten Nonrespondern, d.h. diese Kinder konnten noch nicht ausreichend von der Fördermaßnahme profitieren. Hier wären weiterführende diagnostische Maßnahmen notwendig, um festzustellen, ob die Maßnahme modifiziert oder intensiviert werden muss oder ob ein anderer Förder- oder Therapiebedarf besteht.

Im Rahmen wissenschaftlicher Forschungsprojekte werden Eltern und pädagogische Fachkräfte in der Regel über die Ergebnisse von Post- oder Follow-up-Tests informiert, so dass hier die Möglichkeit besteht, zu erkennen, welche Kinder nicht ausreichend von der Förderung profitiert haben und ob weiterhin Förderbedarf besteht, d.h. ob es sich um sogenannte Nonresponder handelt. Findet Sprachförderung im regulären Kita-Alltag statt, wird der Erfolg der Maßnahme in der Regel nicht systematisch für einzelne Kinder dokumentiert oder überprüft. Dadurch geht wertvolle Zeit für einzelne Kinder verloren, in der eine Maßnahme verändert, angepasst oder eine erneute Überprüfung eingeleitet werden könnte. Findet eine erneute Überprüfung des Sprachentwicklungsstandes statt, so erfolgt diese in der Regel erst nach einem halben oder ganzen Jahr. Auch hier geht wertvolle Zeit verloren, in der die Kinder bereits von einer veränderten Förderung oder einer spezifischen Therapie profitieren könnten.

Um die Wirksamkeit einer Sprachfördermaßnahme nicht erst nach einem halben oder ganzen Jahr zu überprüfen, bedarf es diagnostischer Verfahren, die auch parallel zur Sprachförderung eingesetzt werden können (vgl. Adlof & Hogan, 2019; Ehlich et al., 2008; Gogolin et al., 2003; Winkes & Till, 2023). Hierfür eignen sich Verfahren der Lernverlaufsdiagnostik, wie sie im folgenden Kapitel beschrieben werden.

4 Lernverlaufsdiagnostik zur Evaluation von individuellen Veränderungen

4.1 Begriffsklärung Lernverlaufsdiagnostik

Übergeordnetes Ziel jedes diagnostischen Handelns im professionellen Kontext sind „die Prävention […], die Indikation eines Therapie- /Unterstützungsbedarfes und die Grundlage der Interventionsplanung […] sowie die Verlaufskontrolle und Evaluation“ (Spreer, 2018, S. 10). Es existiert eine Vielzahl an Bezeichnungen, um das diagnostische Handeln zu beschreiben. Nach Breitenbach (2020) können die Begrifflichkeiten zwei großen Gruppen zugeordnet werden, die sich an den zentralen diagnostischen Strategien orientieren, einerseits der Selektionsstrategie und andererseits der Modifikationsstrategie. In der statusorientierten Selektionsdiagnostik geht es vorrangig um das Zuweisen zu Lerngruppen, einem Lernort oder unterschiedlichen Bildungswegen. Die Modifikationsstrategie hingegen meint die Analyse von Lernprozessen (Ingenkamp und Lissmann 2008 in Breitenbach, 2020). Der Selektionsstrategie lassen sich die Begriffe der Statusdiagnostik, summativen Diagnostik und Platzierungsdiagnostik zuordnen, der Modifikationsstrategie die Begriffe der Prozessdiagnostik, formativen Diagnostik, Förderdiagnostik sowie der schulischen und inklusiven Diagnostik.

In der psychologischen Diagnostik dominieren die Begriffe der Modifikation und Selektion. Wenn Veränderungsprozesse abgebildet werden spricht man von einer Modifikation oder auch Prozessdiagnostik. Werden von einem aktuellen Stand ausgehend Entscheidungen getroffen oder Prognosen gestellt, wird der Begriff Selektion verwendet. In der pädagogischen Diagnostik werden vorrangig die Begriffe Förder- und Platzierungsdiagnostik verwendet. Die Förderdiagnostik bezieht sich auf einzelne Personen und hat das Ziel, die Kompetenzen und individuellen Lern- und Lebensbedingungen eines Kindes oder Jugendlichen über einen längeren Zeitraum hinweg zu erfassen, um seinen individuellen Lern- und Entwicklungsprozess zu beschreiben. Die Platzierungsdiagnostik hingegen bezieht sich meist auf Institutionen und beurteilt die Kompetenzen eines Kindes oder Jugendlichen anhand der zur Verfügung stehenden Lernorte und Bildungswege und deren Anforderungen (Breitenbach, 2020).

Aus der Evaluationsforschung stammen die Begriffe der formativen und summativen Diagnostik. Diese gehen auf Scriven (1967) zurück und wurden durch Bloom (1969) erweitert (in Klauer, 2014). Wie bei der Prozess- oder Förderdiagnostik geht es bei der formativen Diagnostik um die kontinuierliche Überprüfung eines Veränderungsprozesses. In der summativen Diagnostik wird hingegen der Lern- und Entwicklungsprozess abschließend beurteilt und evaluiert. Nach Bloom (1969) ist die formative Diagnostik nicht nur eine Dokumentation des Veränderungsprozesses, sondern auch eine Förderung und

Verbesserung des Lernprozesses sowie eine Rückmeldung der Ergebnisse an Lernenden und Lehrenden.

Im Alltag findet sich keine harte Grenze zwischen den beiden diagnostischen Strategien. Meist gehen diese ineinander über und werden als Mischstrategie oder Strategiekombination bezeichnet (Breitenbach, 2020). So führt eine Statusdiagnostik möglicherweise zu einem Wechsel an einen anderen Lernort, an dem aufgrund von Förder- und Prozessdiagnostik der Lernprozess evaluiert und optimiert wird oder die Ergebnisse eines Förderdiagnostik oder formativen Diagnostik führen zu einer Platzierungsdiagnostik.

Eine Mischform stellt auch die Lernverlaufs- oder Lernprozessdiagnostik dar. Durch mehrfach wiederholte Messungen des Lernstandes in kurzen Zeitintervallen kann, vergleichbar der Statusdiagnostik, der jeweils aktuelle Lernstand dokumentiert und anschließend der Lernprozess dargestellt und evaluiert werden. An die Lernverlaufsdiagnostik sollte sich eine Förderdiagnostik anschließen, wenn festgestellt wurde, dass ein Kind nicht in erwartetem Maß vom Unterricht oder einer Förderung profitiert hat. Nur so können mögliche Bedingungen und Ursachen des mangelnden Fortschritts genauer analysiert und die Förderung anschließend verändert oder optimiert werden.

Im deutschen Sprachraum hat sich der von Strathmann und Klauer (2010) eingeführte Begriff der Lernverlaufsdiagnostik durchgesetzt. Der von Deno (2003) geprägte Begriff des Curriculum-Based Measurement (CBM) kann hingegen durch seine Übersetzung als curriculumbasierte Messungen zu falschen Assoziationen führen. Bei der CBM geht es nicht um standardisierte Schulleistungstests, sondern um Tests, die optimalerweise von der Lehrkraft zum jeweils aktuellen Thema entwickelt wurden, um hier den Lernfortschritt zu überprüfen.

4.2 Verschiedene Funktionen von Lernverlaufsdiagnostik

Die kontinuierliche Lernverlaufsdiagnostik hat ihren Ursprung im Response-to-Intervention-Ansatz (RTI). Dabei handelt es sich um ein schulisches Rahmenkonzept, das seit den 1960er Jahren in den USA zur Prävention von Lernstörungen und schulischen Schwierigkeiten sowie zur integrierten Beschulung von Kindern mit sonderpädagogischem Förderbedarf entwickelt wurde (Deno, 2003). In den USA ist dieser Ansatz seit einigen Jahren auch gesetzlich verankert. Die Funktion der Lernverlaufsdiagnostik ist dabei die systematische Erfassung des Lern- und Entwicklungsstandes sowie des Lernverlaufs, um die Lehr- und Förderkräfte dabei zu unterstützen, Kinder mit Lernproblemen oder -schwierigkeiten frühzeitig zu erkennen und angemessen darauf zu reagieren (Deno, 2003).

Die verschiedenen Funktionen der Lernverlaufsdiagnostik sollen an dieser Stelle näher erläutert werden. Als Ausgangspunkt einer Bildungsmaßnahme übernimmt die Lernverlaufsdiagnostik eine Art Screening-Funktion, in der Förderbedarfe möglichst frühzeitig und präzise erfasst werden (Voß et al.,

2017). Im regulären Schulunterricht kann anschließend mit Tests der Lernverlaufsdiagnostik der Lern- und Entwicklungsverlauf dokumentiert werden, um zu evaluieren, ob die Schüler*innen vom Unterricht profitieren (z.B. S. Blumenthal, 2022; Breitenbach, 2020). Ein aktuelles Beispiel ist das Rügener Inklusionsmodell, in dem über mehrere Jahre u.a. der Einsatz von Tests zur Lernverlaufsdiagnostik sowie die daraus abgeleiteten Konsequenzen evaluiert wurden (Hartke et al., 2022). Weitere Beispiele sind die Internetplattformen quop und levumi, die sowohl Tests zur LVD als auch konkrete Fördervorschläge anbieten (S. Blumenthal et al., 2022; Mühling et al., 2017; Souvignier, 2018).

Die Ergebnisse einer Lernverlaufsdiagnostik können als Grundlage für datenbasierte Förderentscheidungen nicht nur im schulischen Kontext, sondern auch zur Evaluation der Wirksamkeit einer Förderung herangezogen werden (Deno, 2003). Von besonderem Interesse sind hier die sogenannten Nonresponder, d.h. Kinder, die nicht im gewünschten Maße von einer Förderung profitieren (D. Fuchs et al., 2004). Werden Nonresponder rechtzeitig erkannt, kann eine Fördermaßnahme angepasst oder modifiziert werden. In diesem Sinne hat die Lernverlaufsdiagnostik das Potenzial, einen wichtigen Beitrag zur Prävention von Lern- und Entwicklungsstörungen zu leisten. Neben der konkreten Anwendung in Schule und Therapie kann die Lernverlaufsdiagnostik auch in der Forschung im Rahmen von multiplen Einzelfallstudien eingesetzt werden (z.B. Grosche, 2011).

Durch aktuelle Forschungsbemühungen, z.B. im Rahmen des Rügener Inklusionsmodell oder der Internetplattformen quop und levumi gewinnt die Lernverlaufsdiagnostik auch in Deutschland zunehmend an Bedeutung. Vor allem im Zuge der Bemühungen um Inklusion und dem Bedarf, Bildungsbenachteiligungen insbesondere bei Schüler*innen aus sozial schwächeren Familien oder Familien mit Migrationshintergrund auszugleichen. Im folgenden Kapitel wird diskutiert, ob die Verfahren der Lernverlaufsdiagnostik auch unabhängig vom RTI-Konzept im binnendifferenzierten Unterricht eingesetzt werden können.

4.3 Einsatz von Lernverlaufsdiagnostik im binnendifferenzierten Unterricht (Publikation A)

Die hier zusammengefassten Inhalte stammen aus nachstehender Veröffentlichung:

> Ennemoser, M. & Besca, M. (2021). Präventive Mehrebenenförderung. *Pädagogik*, 73(4), 44–47.

Der RTI-Ansatz ist ein schulisches Rahmenkonzept, dessen Wirksamkeit umfassend empirisch belegt ist (Burns et al., 2005). Es besteht aus drei aufeinander aufbauenden Förderebenen, die auf die individuellen Lernbedürfnisse der Schüler*innen abgestimmt sind. Im Mittelpunkt des Ansatzes steht die präventive Förderung, wobei der mangelnde Lernerfolg nicht als Schwäche der Schüler*innen interpretiert wird, sondern als unzureichende Anpassung des Unterrichts an die individuellen

Voraussetzungen der Lernenden. Ein Kernelement ist dabei die kontinuierliche Lernverlaufsdiagnostik, die es ermöglicht, ungünstige Entwicklungen frühzeitig zu erkennen und durch abgestufte Maßnahmen zu intervenieren (Deno, 2003).

Die erste Ebene des RTI-Ansatzes ist der reguläre Unterricht, in dem binnendifferenzierende Maßnahmen eingesetzt werden. Ziel ist es, die individuellen Lernvoraussetzungen bestmöglich zu berücksichtigen und dadurch zu verhindern, dass leistungsschwächere Schüler*innen zurückfallen. Trotz qualitativ hochwertigen Unterrichts profitieren nicht alle Schüler*innen gleichermaßen davon. Deshalb wird auf der zweiten Ebene eine gezielte Förderung in Kleingruppen durchgeführt. Hier steht nicht die maximale Individualisierung im Vordergrund, sondern die Anwendung empirisch erprobter Förderprogramme bei Lernschwierigkeiten (Grünke, 2006), die strukturiert und fokussiert auf das Aufholen von Basiswissen und Kompetenzen abzielen.

Reichen diese Maßnahmen nicht aus, um den gewünschten Lernerfolg zu erzielen, kommt die dritte Stufe zum Einsatz. Diese richtet sich an Schüler*innen, die selbst von einer intensiven Förderung in Kleingruppen nicht profitieren. Hier sind individualisierte, oft lerntherapeutische oder sonderpädagogische Maßnahmen erforderlich, die über die Möglichkeiten des regulären Unterrichts hinausgehen. Die Ursachen für anhaltende Schwierigkeiten werden oft in sozial-emotionalen oder kognitiven Defiziten gesucht, die in einer noch differenzierteren Diagnostik analysiert werden müssen.

Das gestufte Vorgehen des RTI-Ansatzes kann somit auch als eine systematische Abfolge von Differenzierungsstrategien betrachtet werden, wobei der Klassenverband als Grundstruktur erhalten bleibt. Die drei Ebenen des RTI-Modells stellen eine fortlaufende Anpassung der Fördermaßnahmen dar, die sich an den Lernfortschritten der Schüler*innen orientieren. Wichtig ist dabei ein evidenzbasierter Unterricht, in dem empirisch erprobte Methoden flexibel eingesetzt werden, ohne den Unterricht in starre Programme zu zwängen.

Bis zur zweiten Ebene lassen sich viele Elemente des RTI-Ansatzes ohne größere schulstrukturelle Veränderungen umsetzen. So können Lernverlaufsdiagnostik und gezielte Förderprogramme insbesondere in den Bereichen Lesen, Rechtschreiben und Mathematik, mit relativ geringem Aufwand in den Unterricht integriert werden (Überblick in Hasselhorn et al., 2014). Die dritte Ebene, die stark individualisierte Fördermaßnahmen erfordert, bleibt jedoch weitgehend auf sonderpädagogische und therapeutische Angebote außerhalb des regulären Unterrichts beschränkt.

Zusammenfassend lässt sich sagen, dass das schrittweise Vorgehen des RTI-Ansatzes und der Einsatz von Lernverlaufsdiagnostik wertvolle Impulse für die Differenzierung im Unterricht liefern.

4.4 Implikationen und offene Fragen

Ziel des Einsatzes von Lernverlaufsdiagnostik ist es, Schüler*innen, die nicht ausreichend von Förder- oder Unterrichtsmaßnahmen profitieren, möglichst frühzeitig zu identifizieren, um die Maßnahmen dann entsprechend modifizieren zu können (Deno, 2003). Neben dem Einsatz von Verfahren zur Lernverlaufsdiagnostik als zentrales Element im Rahmen des RIT-Ansatzes oder als Methode im binnendifferenzierten Unterricht erscheinen Verfahren zur Lernverlaufsdiagnostik daher auch zur prozessbegleitenden Evaluation von Sprachfördermaßnahmen geeignet.

Auch im vorschulischen Bereich wäre eine systematische Abfolge von Differenzierungsstrategien denkbar. Alle Kinder erhalten in der Kita sprachliche Bildungsangebote im Sinne einer primären Prävention, um sprachliche Entwicklungsrisiken zu minimieren. Mit Hilfe eines regelmäßig eingesetzten Screeningverfahrens könnte dann festgestellt werden, bei welchen Kindern ein erhöhter Förderbedarf vorliegt. Diese Kinder sollten dann ein zusätzliches Förderangebot (sekundäre Prävention) auf der zweiten Ebene (vgl. RTI-Ansatz) erhalten. Dieses Angebot würde als zusätzliche Förderung in der jeweiligen Kita stattfinden. Zum Einsatz sollten evidenzbasierte Sprachfördermaßnahmen kommen, die von pädagogischen Fachkräften oder qualifizierten Sprachförderkräften durchgeführt werden. Parallel zur Fördermaßnahme sollte dann systematisch evaluiert werden, in welchem Ausmaß die Kinder von der Förderung profitieren. Wird hierbei festgestellt, dass Kinder nicht ausreichend profitieren, sollte eine stärker individualisierte Förderung, z.B. in Form einer Sprachtherapie erfolgen oder eine differenzierte Diagnostik, um herauszufinden, ob noch weitere Förderbedarfe vorliegen. Dies entspricht der dritten Förderebene oder wird auch als tertiäre Prävention bezeichnet.

Gerade im Bereich der Sprachförderung wäre es wichtig, die Wirksamkeit der Maßnahmen auch prozessbegleitend zu evaluieren. Obwohl in allen Bundesländern vorschulische Screeningverfahren durchgeführt werden und zahlreiche und zum Teil evidenzbasierte Förderkonzepte vorliegen, wird die Wirksamkeit der Maßnahmen selten im Einzelfall überprüft. Dies wäre jedoch besonders relevant, da die Wirksamkeit evaluierter Fördermaßnahmen nicht auf jeden Einzelfall übertragbar ist (Hasselhorn, 2017 in Rauch & Hartig, 2018). Nur wenn die Wirksamkeit einer Förderung auch im Einzelfall evaluiert wird, kann ein Kind entsprechend seinem Förder- und Entwicklungsbedarfes angemessen in seiner sprachlichen Entwicklung unterstützt werden. Nicht alle Kinder im Kita-Alter benötigen eine gezielte Sprachfördermaßnahme, so dass es sinnvoll wäre, nur einzelne Kinder zu fördern, bei denen ein Sprachförderbedarf durch ein Sprachscreening festgestellt wurde. Begleitend sollte dann der Entwicklungsverlauf überwacht werden, um zu evaluieren, ob die ausgewählte Sprachfördermaßnahme bei diesem Kind tatsächlich wirksam ist oder ob sie ggf. angepasst werden muss. Wenn diese individuellen Fördermaßnahmen dann systematisch erfasst und ausgewertet werden, kann dies darüber hinaus zu

einem alternativen Evidenznachweis führen, in Form von kumulativen Einzelfallstudien (Kohler et al., 2021) oder einer systematischen Replikation von Einzelfällen (Jain & Spieß, 2012; Kulawiak & Wilbert).

Verfahren zur Lernverlaufsdiagnostik liegen bisher fast ausschließlich für den schulischen Kontext in den Bereichen Lesen, Rechtschreiben und Mathematik vor (Übersicht z.B. in Gebhardt et al., 2021), wobei hier wiederum Verfahren zum Lesen dominieren (Winkes & Till, 2023). Für den Bereich der mündlichen Sprache besteht allerdings noch eine Forschungslücke (Besca & Ennemoser, 2022; Mahlau et al., 2014). Im Rahmen des Rügener Inklusionsmodells wurden zwar bereits Verfahren zur Lernverlaufsdiagnostik für Morphologie, Syntax und Phonetik entwickelt und eingesetzt, diese sind jedoch noch nicht publiziert (Mahlau et al., 2014). Diese Forschungslücke besteht auch im englischsprachigen Raum, da auch hier nur wenige Verfahren im Bereich der gesprochenen Sprache existieren. Ein Beispiel ist der Wortschatztest auf easycmb.com. Adlof und Hogan (2019) fordern daher zu Recht, dass Forschungsmittel gezielt eingesetzt werden sollten, um u.a. Verfahren zu entwickeln, die zur prozessbegleitenden Evaluation der mündlichen Sprache von Kindern im Schulalter eingesetzt werden können. Im deutschsprachigen Raum gibt es derzeit lediglich das hier im Folgenden vorgestellte Forschungsprojekt zur Entwicklung und Erprobung einer Lernverlaufsdiagnostik Sprache (LVD Sprache) zur Erfassung mündliche Sprachkompetenzen. Außerdem gibt es ein neues Forschungsprojekt zur Entwicklung einer Lernverlaufsdiagnostik zur Messung von Sprachverständnisleistungen in Unterricht und Therapie (Althelmig & Jungjohann, 2024).

5 Entwicklung einer Lernverlaufsdiagnostik sprachlicher Fähigkeiten (LVD Sprache)

Verfahren zur Lernverlaufsdiagnostik bestehen aus mehreren kurzen, schnell durchführbaren und ökonomisch auswertbaren Tests, die regelmäßig, z.B. in wöchentlichen oder monatlichen Abständen durchgeführt werden. Dies erfordert eine Vielzahl von Paralleltests mit vergleichbarem Schwierigkeitsgrad, die aber ein und dieselbe Kompetenz messen. Um Paralleltests für eine Lernverlaufsdiagnostik zu entwickeln, können diese auf zwei verschiedene Arten generiert werden. Beim Curriculum Sampling werden aus einem großen Aufgabenpool Aufgaben ausgewählt, die stellvertretend für bestimmte Inhalte z.B. eines Schulfachs stehen (Breitenbach, 2020). Alternativ kann ein robuster Indikator verwendet werden, der eine bestimmte Fähigkeit als Globalmaß repräsentiert (Gebhardt et al., 2021). Ein typisches Beispiel für ein Globalmaß ist die Anzahl richtig gelesener Wörter pro Minute als Indikator für die Lesekompetenz (Lernverlaufsdiagnostik Lesen LDL, Walter, 2009). Im Bereich der sprachlichen Fähigkeiten scheint das Nachsprechen von Sätzen ein vielversprechendes Globalmaß zu sein, aus dem eine LVD konstruiert werden könnte.

5.1 Das Sätze Nachsprechen als Globalmaß sprachlicher Fähigkeiten

Das Sätze Nachsprechen scheint aus mehreren Gründen ein vielversprechendes Globalmaß bzw. ein robuster Indikator zur Erfassung sprachlicher Fähigkeiten zu sein. Beim Sätze Nachsprechen werden dem Kind grammatikalisch korrekte, aber semantisch meist sinnlose Sätze vorgesprochen, die es anschließend nachsprechen soll. Gezählt werden nur die Wörter, die vollständig korrekt nachgesprochen werden. Schon Schöler stellte bereits 1999 fest, dass eine „korrekte Reproduktion bzw. Rekonstruktion […] eine grammatische und semantische Verarbeitung des Satzes [erfordert]“ (Schöler, 1999, S. 65). Allerdings scheint die Nachsprechleistung auch von der Kapazität des phonologischen Arbeitsgedächtnisses abhängig zu sein (Klicpera 1982 in Goldammer et al., 2011). Goldammer et al. (2010) stellten in einer Studie fest, dass die Varianz der Leistungen im Sätze Nachsprechen durch die Leistungen im Wortschatz und im phonologischen Arbeitsgedächtnis aufgeklärt wurden und die grammatischen Kompetenzen keinen weiteren Beitrag leisten konnten. Im Gegensatz dazu stellt Polišenská et al. (2015) in ihrer Studie fest, dass die Fähigkeit des Sätze Nachsprechens stärker von der Vertrautheit mit der Morphosyntax und der lexikalischen Phonologie abhängt als von der Semantik. Zusammenfassend lässt sich festhalten, dass das Sätze Nachsprechen wesentlich auf Leistungen im Bereich des Wortschatzes, der Grammatik und des phonologischen Arbeitsgedächtnisses basiert und somit zentrale Aspekte sprachlicher Kompetenzen umfasst.

> “Sentence repetition is best seen as a reflection of an underlying language ability factor rather than as a measure of a separate construct with a specific role in language processing. Sentence

> repetition appears to be a valuable tool for language assessment because it draws upon a wide range of language processing skills." (Klem et al., 2015, S. 146)

Das Sätze Nachsprechen wird bereits in vielen statusdiagnostischen Verfahren eingesetzt, um z.B. morphologische und syntaktische Fähigkeiten zu überprüfen (P-ITPA, Esser & Wyschkon, 2010). Das Sätze Nachsprechen eignet sich auch zur Identifikation von Risikokinder für Schriftspracherwerbsstörungen (HASE, Schöler & Brunner, 2008). Goldammer et al. (2010) fanden heraus, dass die Satzgedächtnisleistung im Vorschulalter der stärkste Prädiktor für die spätere Lese- und Rechtschreibleistung in der Grundschule ist. Darüber hinaus ist das Nachsprechen von Sätzen eine etablierte Methode zur Feststellung einer Sprachentwicklungsstörung (SETK 3-5, Grimm, 2015). Schöler et al. konnten bereits 1997 zeigen, dass eine Sprachentwicklungsstörung ausschließlich über die Nachsprechleistung diagnostizieren werden kann (Schöler et al., 1997). Das Nachsprechen von Sätzen wird sogar als „Königsweg für die Diagnostik und Differentialdiagnostik bei Sprachentwicklungsstörungen" bezeichnet (Schöler & Scheib, 2004, S. 40). Dieser Befund konnte inzwischen mehrfach repliziert werden (Abed Ibrahim et al., 2018; Leclercq et al., 2014; Theodorou et al., 2017). Darüber hinaus konnte gezeigt werden, dass bilinguale Kinder mit einer Sprachentwicklungsstörung, im Gegensatz zu monolingualen Kindern durch Aufgaben des Sätze Nachsprechens nicht benachteiligt werden (Abed Ibrahim et al., 2018; Tuller et al., 2018; Ziethe et al., 2013), insbesondere wenn der Einfluss des Wortschatzumfangs berücksichtigt wird (Komeili/ Marshall 2013).

Das Nachsprechen von Sätzen scheint nicht nur ein vielversprechender Indikator sprachlicher Fähigkeiten zu sein, sondern auch ein Aufgabenformat, das einem natürlichen Lernprozess sehr nahekommt. "Imitating someone else's behaviour, including language, is an innate ability of humans and other species, and repeating sounds, words, and sentences is often part of uninstructed and instructed teaching and learning processes" (Marinis, Armon-Lotem 2015, S.97).

Auch in der praktischen Anwendung erscheint das Sätze Nachsprechen gut geeignet. Das Aufgabenformat des Vor- und Nachsprechens ist schnell und unkompliziert anwendbar und eine ökonomische Auswertung über die Summe der korrekt nachgesprochenen Wörter ist möglich. Theoretisch sind unendlich viele Paralleltests denkbar, so dass eine wiederholte Durchführung möglich ist.

5.2 Entwicklung der Sätze für die LVD Sprache

Für die Konstruktion der Paralleltests zum Sätze Nachsprechen wurden der Wortschatzpool des Child-Lex-Korpus (Schroeder et al., 2015), sowie aktuelle Befunde zur Grammatikentwicklung von sprachunauffälligen Kindern im Alter von vier bis neun Jahren herangezogen (GED 4-9, Ulrich, 2017). Basierend auf diesen Erkenntnissen zur Grammatikentwicklung wurden Satzschablonen entwickelt, die wesentliche morphologische und syntaktische Strukturen enthalten, die Kinder im Alter zwischen vier

und neun Jahren erwerben. Diese Satzschablonen wurden dann mit Hilfe eines eigens dafür entwickelten Excel-Programms systematisch, aber zufällig mit Wörtern aus dem Wortschatzpool des ChildLex-Corpus gefüllt.

Um die in diesem Altersbereich erworbenen syntaktischen Strukturen möglichst repräsentativ abzubilden, enthalten die Sätze Haupt- und Nebensätze (temporale, kausale, finale Nebensätze sowie eingeschobene Relativsätze) und zudem Strukturen, die als Meilensteine im Erwerb der deutschen Wortstellung gelten, wie Topikalisierung und Satzklammer (Hamann & Abed Ibrahim, 2017). Darüber hinaus wurden einige der komplexen Strukturen integriert, von denen bekannt ist, dass sie für Kinder mit Sprachentwicklungsstörung sprachübergreifend schwierig sind, darunter Objektfragen, Subjekt- und Objektrelativsätze, finite Komplementsätze und Passive (Marinis & Armon-Lotem, 2015). Auf Fragesätze, die besonders für Kinder mit Sprachentwicklungsstörung sprachübergreifend als schwierig gelten und die 2. Person Singular als Indikator für eine nahezu vollständig erworbene Subjekt-Verb-Kongruenz wurde verzichtet. Diese Strukturen können ohne zusätzliche Erklärung oder einen sinnvollen Handlungskontext als direkte Fragen an das Kind verstanden werden, die dann vom Kind nicht wiederholt, sondern beantwortet werden.

Das Nachsprechen von Sätzen gilt in vielen Testverfahren als aussagekräftiges Maß zur Erfassung grammatischer Fähigkeiten (Grimm, 2015; Polišenská et al., 2015). Auch bei der Entwicklung des Deutschen Tests zum Sätze Nachsprechen (LITMUS-SRT von Hamann & Abed Ibrahim, 2017) wurde ein stärkerer Fokus auf die grammatischen Fähigkeiten gelegt, indem der Wortschatz bewusst einfach gehalten wurde, mit der Begründung, dass ein komplexer Wortschatz bei mehrsprachig aufwachsenden Kindern häufig zu Schwierigkeiten führen kann (Marinis & Armon-Lotem, 2015). Wenn das Nachsprechen von Sätzen zur Diagnose einer Sprachentwicklungsstörung verwendet wird, weisen Komeili und Marshall (2013) darauf hin, dass sich die diagnostische Genauigkeit erhöht, wenn der Einfluss des Wortschatzumfangs berücksichtigt wird. Wird das Sätze Nachsprechen hingegen als Prädiktor für Lese- und Rechtschreibleistungen in der Grundschule herangezogen, so zeigt sich, dass „das Nachsprechen von Sätzen sowohl Anforderungen an das phonologische Arbeitsgedächtnis als auch an das semantische Wissen (Wortschatz) im Langzeitgedächtnis stellt“ (Goldammer et al., 2011, S. 2). Dabei ist zu berücksichtigen, dass die verschiedenen Tests zum Sätze Nachsprechen unterschiedlich konstruiert sind. Mit dem hier vorgestellten, neu entwickelten Verfahren der LVD Sprache soll überprüft werden, ob mit dem Konstrukt des Nachsprechens von Sätzen sprachliche Fähigkeiten in Bezug auf Grammatik und Wortschatz über die Zeit erfasst werden können. Daher wurde bewusst ein komplexerer Wortschatz gewählt.

Die Wörter, mit denen die Satzschablonen der LVD Sprache gefüllt wurden, stammen aus dem Wortschatzpool des ChildLex-Korpus. Dieser Wortschatzpool repräsentiert die schriftsprachliche Umwelt,

mit der Kinder im Alter von sechs bis acht Jahren konfrontiert werden und die für die spätere Lese- und Schreibentwicklung entscheidend ist (Schroeder et al., 2015). Er setzt sich aus Vorlese-, Selbstlese- und Schulbüchern zusammen und repräsentiert somit auch den schulischen Wortschatz. Dieser Wortschatz wurde gezielt ausgewählt, um zu erfassen, welche Schwierigkeiten Kinder mit Sprachförderbedarf im Hinblick auf die schulischen Anforderungen noch haben.

Der Wortschatz von Kindern differenziert sich etwa ab dem dritten Lebensjahr in Bezug auf Ober- und Unterbegriffe, Flexionen, zusammengesetzte Wörter, Derivationen, Pronomina, Artikel, Präpositionen, Konjunktionen und Relativpronomen aus (Bockmann et al., 2020). In den Sätzen der LVD Sprache wurden daher folgende Elemente berücksichtigt: einfache und zusammengesetzte Nomen; attributive, prädikative und adverbiale Adjektive; finite modale und auxiliare Verben sowie infinite Vollverben; Partizip Perfekt-Verbformen; bestimmte und unbestimmte Artikel; Adverbien; Konjunktionen und Pronomen (Personal-, Relativ-, Possessiv-, Demonstrativ- und Indefinitpronomen).

Insgesamt wurden 11 Paralleltests entwickelt, die jeweils aus 16 Sätzen mit zunehmender grammatischer Komplexität und Satzlänge bestehen. Die ersten vier Sätze sind semantisch sinnvoll, die Sätze 5-16 sind semantisch sinnlos, aber grammatisch korrekt. Im Folgenden soll kurz beschrieben werden, welche Bedeutung die Satzlänge für die Nachsprechleistung hat und warum die LVD Sprache semantisch sinnlose und sinnvolle Sätze enthält.

Aus der Forschung ist bekannt, dass durch das Nachsprechen von Sätzen implizites Wissen erfasst werden kann, insbesondere, wenn die Sätze lang genug sind (z.B. Ellis, 2005 und Erlam, 2006 in Marinis & Armon-Lotem, 2015). Dadurch wird ein einfaches „Kopieren“ des Satzes verhindert. Stattdessen wird auf das vorhandene Sprachsystem zurückgegriffen, um den Satz zu wiederholen. Dabei wird der gehörte Satz analysiert und auf Basis des vorhandenen grammatischen und semantischen Wissens rekonstruiert (Goldammer et al., 2011; Marinis & Armon-Lotem, 2015). Im Umkehrschluss bedeutet dies, dass Kinder nur solche Sätze nachsprechen können, deren spezifische Strukturen sie bereits erworben haben. In den Testverfahren ESGRAF-R und MuSE-Pro zur Erfassung grammatischer Fähigkeiten findet sich ein ähnliches Vorgehen beim Nachsprechen von Nebensätzen. Dem Kind wird ein Haupt- und Nebensatz vorgegeben, den es wiederholen soll. Hierbei wird davon ausgegangen, dass das Kind den Nebensatz mit Verbendstellung nur dann korrekt bilden kann, wenn es die grammatische Struktur erworben hat (Berg, 2015; Motsch & Rietz, 2019). Sind die Sätze beim Sätze Nachsprechen hingegen zu kurz, kann das Kind den Satz einfach „Nachplappern“ und den Prozess der Analyse und Rekonstruktion umgehen. Hiermit wird dann nur die Arbeitsgedächtniskapazität des Kindes erfasst (Marinis & Armon-Lotem, 2015). Die Sätze der LVD Sprache wurden mit unterschiedlicher Länge entwickelt, um den unterschiedlichen Fähigkeiten der Kinder in Bezug auf Alter, Arbeitsgedächtniskapazität und sprachliche Fähigkeiten gerecht zu werden. Damit soll sowohl das

erwartete Anfangsniveau als auch das angestrebte Lernzielniveau sprachlicher Fähigkeiten abgedeckt werden (Gebhardt et al., 2021; Ulrich, 2017).

Beim Nachsprechen von Sätzen werden gezielt morphologisch-syntaktische Strukturformen verwendet, um bereits erworbenes Grammatikwissen zu überprüfen. Um zu verhindern, dass der Satz mit Hilfe semantischer Informationen besser analysiert und reproduziert werden kann, werden semantisch sinnlose Sätze verwendet (Grimm, 2015). Insbesondere Kinder mit einer Sprachentwicklungsstörung sind auf den Bedeutungskontext angewiesen und haben nachhaltige Schwierigkeiten, den syntaktischen Kontext zu nutzen. Bei diesen Kindern kann die Nachsprechleistung völlig zusammenbrechen, wenn sie die semantischen Hinweise nicht nutzen können (Grimm, 2015). Um das gesamte Spektrum sprachlicher Fähigkeiten von unauffällig bis sprachentwicklungsgestört abzudecken, wurden sowohl semantisch sinnvolle als auch semantisch sinnlose Sätze mit einer zunehmender Satzlänge konstruiert.

In einem letzten Schritt wurden die Sätze manuell überarbeitet, um sicherzustellen, dass keine Wörter aus dem gleichen semantischen Feld in gleichen Satz auftauchen oder sich keine Wörter in einem Paralleltest wiederholen. Bei der Überarbeitung der neu konstruierten Sätze wurde auch auf eine gleichmäßige Verteilung von geschlechterspezifischen Begriffen und Themen geachtet, um eine mögliche Diskriminierung in diesem Bereich zu vermeiden.

Für die Anwendung der LVD Sprache existieren genaue Anweisungen zur Durchführung und Auswertung. Die Sätze werden dem Kind ohne Betonung mit verdecktem Mundbild vorgesprochen. Die Auswertung erfolgt über die Anzahl der vollständig richtig nachgesprochenen Wörter. Pro Satz sind abhängig von der Länge 5-12 Wörter, insgesamt maximal 134-137 Wörter pro Paralleltest möglich.

6 Erprobung der LVD Sprache im Quer- und Längsschnitt

6.1 Rekrutierung und Stichproben

Die Erprobung der LVD Sprache erfolgte in mehreren Schritten und setzt sich aus unterschiedlichen Stichproben zusammen. Die Rekrutierung der Kinder erfolgte über die direkte Ansprache von Kita- und Schulleitungen bzw. Lehrkräfte in Einrichtungen im ländlichem oder städtischem Raum Baden -Württembergs (Landkreise Ludwigsburg, Stuttgart, Calw, Rottweil, Heilbronn, Rhein-Neckar-Kreis, Rems-Murr-Kreis, Main-Tauber-Kreis; Bodenseekreis). Die Zustimmung zur Teilnahme wurde von den Eltern schriftlich über eine Einverständniserklärung eingeholt, die gleichzeitig auch Fragen zu den Sprachen, zur Sprachentwicklung und zu bisherigen Sprachfördermaßnahmen der Kinder enthielt. Die Paralleltests wurden entweder von Studierenden im Rahmen ihrer Bachelor- oder Masterarbeiten, von studentischen Hilfskräften oder von der Doktorandin selbst durchgeführt. Einschlusskriterium für die Stichproben waren ein- und mehrsprachig aufwachsende Kinder ab vier Jahren in der Kita sowie Kinder der 1. bis 3. Klasse in der Grundschule bzw. mit einer Sprachentwicklungsstörung am Sonderpädagogischen Bildungs- und Beratungszentraum mit Förderschwerpunkt Sprache (SBBZ Sprache). Da es sich beim Sätze Nachsprechen um ein abstraktes Aufgabenformat handelt, scheint eine zuverlässige Durchführung dieses Formates erst ab vier Jahren möglich zu sein (Grimm, 2015). Kinder mit massiven Aussprachstörungen wurden ausgeschlossen, da eine Auswertung nicht zuverlässig möglich gewesen wäre.

Die Paralleltests der LVD Sprache wurden in mehreren Phasen sowohl im Querschnitt als auch im Längsschnitt erprobt. Im Querschnitt wurden alle Paralleltests pro Kind möglichst innerhalb von zwei Wochen durchgeführt. Im Längsschnitt wurde jeweils ein Paralleltest alle vier Wochen im Verlauf eines halben Jahres durchgeführt. Zusätzlich wurden die Paralleltests begleitend zu Einzelinterventionen erprobt.

Querschnitt

Die erste Version der Paralleltests wurde im Juli 2020 mit neun Schüler*innen einer 3. Klasse Grundschule erprobt. Die Kinder waren zwischen 8;9 und 9;5 Jahre alt. Die zweite Version der Paralleltests wurde im Oktober 2020 mit 24 Kindern aus zwei Kitas im Alter zwischen 4;0 und 6;2 Jahren erprobt. Die dritte Version der Paralleltests wurde im Zeitraum zwischen Juni und Dezember 2021 mit 86 Kindern aus sechs Kitas im Alter zwischen 4;0 und 6;8 Jahren sowie mit 19 Kindern der 1. und 2. Klasse eines SBBZ Sprache erprobt (Alter zwischen 7;1 und 9;0 Jahren).

Längsschnitt

Die vierte und aktuelle Version der Paralleltests wurde 2022 mit 88 Kindern aus fünf Kitas über einen Zeitraum von einem halben Jahr sowie mit 19 Kindern der 2. und 3. Klasse eines SBBZ Sprache

durchgeführt. Anschließend wurden die Paralleltests 2023 mit 99 Kindern der 1. bis 3. Klasse in fünf Grundschulen erprobt. Zusätzlich wurde die LVD Sprache mit 44 Kita-Kindern als Kontrollgruppe durchgeführt. Dabei wurden die Prä- und Posttests sowie jeweils ein Paralleltest zu Beginn und am Ende des halben Jahres eingesetzt.

6.2 Vorgehen bei der Erprobung einer Lernverlaufsdiagnostik

Wie bei einem Testverfahren zur Statusdiagnostik müssen auch bei einem Testverfahren zur Lernverlaufsdiagnostik die psychometrischen Gütekriterien überprüft werden, zu denen die Hauptgütekriterien Objektivität, Reliabilität und Validität gehören. Darüber hinaus sind bei einem Verfahren der Lernverlaufsdiagnostik die Homogenität des Gesamttests, die Eindimensionalität des Verfahrens und die Änderungssensitivität zu prüfen (Gebhardt et al., 2021). Zur Überprüfung der Gütekriterien der LVD Sprache werden Daten aus Querschnitts- und Längsschnittuntersuchungen herangezogen. Das konkrete Vorgehen wird im Folgenden dargestellt. Die Ergebnisse werden in den Kapiteln 6.4. bis 6.6. dargestellt.

Die Objektivität gliedert sich in Durchführungs-, Auswertungs-, und Interpretationsobjektivität. Ein Verfahren kann dann als objektiv bezeichnet werden, wenn Durchführung, Auswertung und Interpretation unabhängig von der durchführenden Person sind (Pospeschill, 2022). Für die LVD Sprache liegen Vorgaben zu einer standardisierte Testdurchführung sowie eindeutige Vorgaben für die Auswertung vor. Vorgaben zur Interpretation liegen noch nicht vor, da die LVD Sprache noch in der Erprobungsphase ist.

In einem ersten Schritt wurde die LVD Sprache an einer großen Stichprobe im Querschnitt zu einem Messzeitpunkt überprüft (Klauer, 2006; Wilbert & Linnemann, 2011). Hierbei wurden Trennschärfe, Itemschwierigkeit, Homogenität und interne Konsistenz anhand der Klassischen Testtheorie überprüft. Eine empfohlene Itemanalyse auf Basis der Item-Response-Theorie scheint aufgrund der Konstruktion der Sätze mit unterschiedlichen Satzlängen an dieser Stelle keine zusätzlichen Erkenntnisse zu liefern. Eine dichotome Auswertung auf Satzebene (z.B. 1,0; Satz richtig vs. falsch) sowie wie eine polytome Erweiterung (z.B. 2,1,0; richtig/ teilweise richtig/ falsch) würde zu einem erheblichen Informationsverlust führen. Der Summenwert würde sich von maximal 134-137 auf 16 Punkte reduzieren und eine sensitive Erfassung von Unterschieden oder Veränderungen wäre nicht mehr möglich. Zudem wäre dies in der Auswertung nicht praktikabel.

Die Eindimensionalität des Verfahrens wurde mittels explorativer und konfirmatorischer Faktorenanalyse überprüft (z.B. Förster et al., 2023). Damit soll nachgewiesen werden, dass alle Sätze eines Paralleltests der LVD Sprache nur auf einen gemeinsamen Faktor zurückzuführen sind (Hattie, 1985) und keine weiteren Faktoren extrahiert werden können.

Zur Beurteilung der Paralleltest-Reliabilität wurden die Längsschnittdaten herangezogen und die jeweils benachbarten Messzeitpunkte miteinander korreliert (Walter, 2011). Die Validität der LVD Sprache wurde anhand von Korrelationen mit konvergenten und divergenten Konstrukten bestimmt (Pospeschill, 2022). Die Unterschiede zwischen den Korrelationen werden mit Hilfe von Fisher's-Z beschrieben.

Ein zentraler Schritt bei der Evaluation einer Lernverlaufsdiagnostik ist der Nachweis der Änderungssensitivität, d.h. ob mit einem Verfahren überhaupt Veränderungen im individuellen Lernverlauf gemessen werden können (L. S. Fuchs, 2004; Gebhardt et al., 2021). Es muss gewährleistet sein, dass eine Lernverlaufsdiagnostik Veränderungen der Fähigkeit über die Zeit sensitiv erfasst und insbesondere bei Kindern mit Entwicklungsverzögerungen oder -störungen auch kleine Lernfortschritte erkannt werden. Zuerst wurde die Paralleltest-Reliabilität anhand einer Korrelation zwischen den einzelnen Messzeitpunkten berechnet und anschließend die Mittelwerte aller Messzeitpunkte mithilfe inferenzstatistischer Verfahren auf signifikante Unterschiede geprüft (Souvignier et al., 2014; Walter, 2011). Voraussetzung sind gleiche Zeitabstände zwischen den Messungen (Gebhardt et al., 2021), sowie mindestens sechs Messungen, um den mittleren Lernzuwachs schätzen zu können (Christ et al., 2013 in Gebhardt et al., 2021). Wenn mit einem Verfahren Änderungen erfasst werden können, muss zudem ausgeschlossen werden, dass diese nicht auf die Vertrautheit mit dem Aufgabenformat, den sog. Übungseffekt zurückzuführen sind (Wilbert & Linnemann, 2011). Dazu wurde eine Varianzanalyse mit Messwiederholung zum ersten und letzten Messzeitpunkt für die Experimental- und Kontrollgruppe gerechnet.

Mit Hilfe von Regressionsanalysen können Lernverläufe z.B. getrennt nach Perzentilen berechnet werden, um einerseits das individuelle Ausgangsniveau und andererseits ggf. unterschiedliche Zuwachsraten zu berücksichtigen (Jungjohann et al., 2021). Da die Wortschatz- und Grammatikentwicklung im Kindesalter eine große Variabilität aufweist, werden die Entwicklungsverläufe getrennt nach Quartilen bzw. nach Anzahl der Sprachen betrachtet. Eine große Variabilität im produktiven Wortschatz zeigt sich mit einem Umfang von 300-500 Wörter im Alter von drei Jahren bis zu 4000-8000 Wörter am Ende der 2. Klasse (Kauschke, 2023). Erst ab etwa 12 Jahren verlangsamt sich der Zuwachs deutlich. Auch die Grammatikentwicklung scheint nicht wie bisher angenommen mit etwa vier Jahren weitgehend abgeschlossen zu sein, sondern dauert bis weit in die Grundschulzeit an, wie aktuelle Auswertungen der Studie GED 4-9 zeigen (Ulrich, 2017). Auch für Kinder mit Deutsch als Zweitsprache zeigen sich große Unterschiede in der Entwicklung, die u.a. von den Inputbedingungen oder der Kontaktdauer mit dem Deutschen abhängen (Chilla, 2020).

6.3 Erhebungsinstrumente

Im Rahmen der Erprobung im Quer- und Längsschnitt wurden einzelne Untertests aus den Testverfahren zur Überprüfung der konvergenten und divergenten Validität der neu entwickelten LVD Sprache eingesetzt. Im Folgenden werden die Verfahren vorgestellt, auf die in den Testergebnissen Bezug genommen wird.

Konvergente Validität

Um den Wortschatz der Kinder zu erfassen, wurde der Untertest Bildbenennung aus dem SET 5-10 (Petermann, 2018) ausgewählt. Der SET 5-10 ist ein standardisierter Test zur umfassenden Sprachstandsfeststellung für Kinder im Alter von 5 bis 10 Jahren. Er eignet sich unter anderem für Kinder mit Sprachentwicklungsverzögerungen und -störungen sowie für Kinder mit Migrationshintergrund. Obwohl der SET 5-10 erst ab fünf Jahren konzipiert ist, wurde er auch für die Gruppe der vierjährigen Kinder eingesetzt, da sich in früheren Studien gezeigt hat, dass er in dieser Altersgruppe gut differenziert. Der Untertest Bildbenennung besteht aus 40 Bildern, die Begriffe und Handlungen darstellen, die das Kind benennen soll. Richtige Lösungen werden jeweils mit einem Punkt bewertet. Zulässige Alternativlösungen finden sich auf dem Protokollbogen. Aus den summierten Rohwerten können mit Hilfe des Manuals Prozentränge und T-Werte für Kinder im Alter von 5 bis 10 Jahren berechnet werden (Tabellen in Jahresschritten).

Aus dem P-ITPA (Esser & Wyschkon, 2010), einem standardisierten Test zur Erfassung sprachlicher und schriftsprachlicher Auffälligkeiten, wurde der Untertest Grammatik verwendet. Der P-ITPA eignet sich für Kinder im Alter von vier Jahren bis zum Ende der fünften Klasse und kann sowohl bei ein- und mehrsprachigen Kindern zur Erfassung der sprachlichen Fähigkeiten im Deutschen eingesetzt werden. Er bietet Normwerte in Halbjahresschritten an. Der Untertest Grammatik umfasst aus 57 Items, mit denen das grammatische Wissen zu Pluralbildung, Komparation, Bildung von Perfekt und Präteritum sowie zur korrekten Verwendung von Akkusativ, Dativ und Genitiv zu erfasst wird. Dem Kind wird ein unvollständiger Satz oder Satzanfang vorgelesen, den es richtig beenden soll. Zusätzlich werden Bildkarten gezeigt, um den Inhalt zu verdeutlichen.

Weder der SET 5-10 noch der P-ITPA bieten separate Normdaten im Hinblick auf mehrsprachige Kinder an oder gehen auf das Thema der Mehrsprachigkeit ein. Dass es nahezu unmöglich ist, Normwerte für die sehr heterogene Gruppe der mehrsprachig aufwachsenden Kinder zur Verfügung zu stellen, hat die Neunormierung der PDSS gezeigt, die in ihrer neuen Version nun sogar Orientierungswerte für mehrsprachige Kinder anbietet (Kauschke et al., 2023). In der Praxis hat sich jedoch gezeigt, dass beide Testverfahren geeignet erscheinen, um Kompetenzen in der deutschen Sprache zu erfassen (z.B. Lüke, 2011).

Aus dem SET-K 3-5 (Grimm, 2015) wurde der statusdiagnostische Untertest zum Satzgedächtnis verwendet. Dabei handelt es sich um einen standardisierten Test für 4- bis 5-Jährige zur Erfassung der rezeptiven und produktiven Sprachverarbeitungsfähigkeiten sowie der auditiven Gedächtnisleistungen. Er kann sowohl bei einsprachigen Kindern als auch bei Kindern mit Deutsch als Zweitsprache eingesetzt werden. Der Untertest Satzgedächtnis erfasst bereits erworbenes grammatisches Wissen und eignet sich für Kinder von 4,0 bis 5,11 Jahren, hier liegen Normwerte in Halbjahresschritten vor. Er besteht aus 15 Items bzw. Sätzen, die dem Kind einmal vorgesprochen werden. Anschließend soll das Kind den Satz nachsprechen. Jedes vollständig richtig nachgesprochene Wort wird mit einem Punkt bewertet, so dass sich ein Gesamtwert von maximal 119 Rohpunkten ergibt.

Die Mottier-Silben sind ein informeller Test zur Erfassung des phonologischen Arbeitsgedächtnisses, der ursprünglich als Ergänzung zum Züricher Leseprobe konzipiert wurde (Mottier, 1951). Aktuell liegen verschiedene Normierungsstichproben vor, z.B. für 4- bis 6-jährige Kinder von Risse und Kiese-Himmel (2009), für 5- bis 17-jährige Kinder von Wild und Fleck (2013) und für 3- bis 4-jährige Kinder von Frey und Kolleg*innen (2024). Die Mottier-Silben bestehen aus 30 Kunstwörtern (zwei- bis sechssilbig), die dem Kind standardisiert im Sekundentakt pro Silbe ohne Betonung mit verdecktem Mundbild vorgesprochen werden. Wird das Kunstwort richtig nachgesprochen, wird dies mit einem Punkt bewertet.

Der MSVK ist ein standardisiertes Verfahren zur Untersuchung des Sprachverständnisses bei Kindern im Vorschulalter und frühen Schulalter (Lohaus & Elben, 2000). Es liegen geschlechtsspezifische Prozentränge und T-Werte für Kindergartenkinder ab 5 Jahren und Kinder der ersten Klasse vor. Aus dem MSVK wurden die beiden Untertests Satzverständnis und Instruktionsverständnis verwendet. Beim Satzverständnis wird dem Kind ein Satz vorgelesen und es muss aus einer Auswahl von drei Bildern das richtige ankreuzen, wobei maximal 18 Punkte erreicht werden können. Beim Instruktionsverständnis wird eine aus mehreren Informationen bestehende Instruktion gegeben, die das Kind im Testheft ausführen muss, wobei jede vollständig richtige Antwort einen Punkt ergibt und maximal acht Punkte erreicht werden können.

Die Informationen zu den Sprachen der Kinder, zur Einschätzung des Sprachentwicklungsstandes und des Sprachförderbedarfs aus Sicht der Eltern sowie zum Kontakt mit Deutsch als Zweitsprache wurden per Fragebogen erhoben.

Divergente Validität

Der WPPSI ist ein standardisiertes Testverfahren zur Erfassung allgemeiner und spezifischer kognitiver Fähigkeiten von Kindergarten- und Vorschulkindern. Für die vorliegende Studie wurden die Untertests Bildkonzepte und Matrizen zur differenzierteren Erfassung der fluiden Intelligenz aus dem WPPSI-III für das Alter von 2;6 bis 7;7 Jahren ausgewählt (Petermann & Daseking, 2018). Im Untertest

Bildkonzepte werden dem Kind Karten mit zwei oder drei Bildreihen vorgelegt. Die Aufgabe besteht darin, eine Gruppe von Bildern mit einem gemeinsamen Merkmal oder Konzept zusammenzustellen, indem aus jeder Reihe ein passendes Bild ausgewählt wird. Jede richtige Lösung wird mit einem Punkt bewertet, wobei maximal 28 Punkte erreicht werden können. Im Untertest Matrizen wird dem Kind ein unvollständiges Muster gezeigt, und es soll aus fünf Antwortmöglichkeiten den fehlenden Teil des Musters auswählen. Jede richtige Lösung wird mit einem Punkt bewertet, maximal können 35 Punkte erreicht werden.

Der MBK ist eine Testbatterie, die mathematische Kompetenzen standardisiert und lehrplanunabhängig erfasst. Aus dem MBK 0 für Kindergartenkinder im Alter von 3;6 bis 7 Jahren wurden die Untertests Zahlenfolge und Ziffernkenntnis ausgewählt (Krajewski, 2018). Aus dem MBK 2-4 für Kinder im Grundschulalter wurden die Untertests Zehn mehr und Basisrechnen ausgewählt (Krajewski & Ennemoser, in Vorb.). Im Untertest Zahlenfolge sollen die Kinder vorwärts und rückwärts zählen und die Nachbarzahlen vor und nach einer Zahl im Zahlenraum bis 20 benennen. Im Untertest Ziffernkenntnis soll das Kind verschiedene Ziffern im Zahlenraum bis 20 benennen. Die Aufgabe im Untertest Zehn mehr besteht darin, zu einer vorgegebenen Zahl immer die Zahl aufzuschreiben, die genau zehn mehr ist. Die Bearbeitungszeit ist auf eine Minute begrenzt, es können maximal 16 Punkte erreicht werden. Beim Basisrechnen soll das Kind innerhalb von 2 Minuten möglichst viele Aufgaben (Addition, Subtraktion, Multiplikation, Division) im Zahlenraum bis 100 lösen, maximal sind 40 Punkte möglich.

Ein in der Praxis häufig verwendeter Indikator zur Erfassung des Bildungs- und sozioökonomischen Status ist die sogenannte Bücherfrage, d.h. die Frage nach dem Buchbestand in der Familie (Heppt et al., 2022). Dabei werden Eltern oder Kinder in der Regel gebeten, auf einer fünfstufigen Skala einzuschätzen, wie viele Bücher zu Hause vorhanden sind. Da es bisher keine Erfahrungen mit dem Einsatz bei Kindergartenkinder vorliegen, wurde zur Veranschaulichung der Skala (1 = keine oder nur sehr wenige Bücher bis 5 = genug Bücher, um drei oder mehr Bücherregale zu füllen) eigenes Bildmaterial entwickelt.

6.4 Erste Erprobungsergebnisse der LVD Sprache (Publikation B)

In diesem Beitrag werden erste Ergebnisse der zu diesem Zeitpunkt vorliegenden Quer- und Längsschnittdaten vorgestellt, um erste Erkenntnisse über die Eignung des Nachsprechens von Sätzen als Möglichkeit der Lernverlaufsdiagnostik sprachlicher Fähigkeiten zu gewinnen. Die Daten wurden hinsichtlich der Vergleichbarkeit der neu entwickelten Paralleltests der LVD Sprache sowie der Itemschwierigkeit, Reliabilität und Validität untersucht.

> Besca, M. & Ennemoser, M. (2022). Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten und zur Evaluation sprachlicher Fördermaßnahmen? In M.

Spreer, M. Wahl & H. Beek (Hrsg.), *Sprachheilpädagogik aktuell: Bd. 4, Sprachentwicklung im Dialog: Digitalität - Kommunikation - Partizipation* (S. 205–212). Schulz-Kirchner Verlag.

Methode (Stichproben, Testverfahren und Testdesign)

Für die Analysen wurde die Querschnittsstichprobe aus Kita und SBBZ Sprache im Zeitraum von Juni bis Dezember 2021 verwendet. In die Auswertung konnten die Daten von 86 Kindern aus sechs Kitas im Alter zwischen 4,0 und 6,8 Jahren sowie von 19 Kindern der 1. und 2. Klasse eines SBBZ Sprache (Alter zwischen 7,1 und 9,0 Jahren) einbezogen werden. Darüber hinaus konnte ein Teil der bereits ausgewerteten Daten des Prätests aus der Längsschnittstudie in den Kitas für erste Analysen zur Validität herangezogen werden.

Zur Überprüfung der konvergenten und divergenten Validität wurden der Untertest Sätze Nachsprechen aus dem SETK 3-5 (Grimm, 2015) bzw. P-IPTPA (Esser & Wyschkon, 2010), der Untertest Bildbenennung aus dem SET 5-10 (Petermann, 2018), der Untertest Grammatik aus dem P-IPTA (Esser & Wyschkon, 2010) und die Mottier-Silben (Mottier, 1951) als Arbeitsgedächtnismaß im Prä- oder Posttest eingesetzt. Aus dem WPPSI-III (Petermann & Daseking, 2018) wurden die Untertests Bildkonzepte und Matrizen als Indikator für die Begabung verwendet. Der sozioökonomische Status wurde mit Hilfe der Bücherfrage (Heppt et al., 2022) und die mathematischen Fähigkeiten anhand des MBK 0 bzw. MBK 2-4 (Krajewski, 2018; Krajewski & Ennemoser, in Vorbereitung) erhoben.

Ergebnisse und Diskussion

Mit den ersten Ergebnissen der Erprobung in Kita und SBBZ Sprache konnte gezeigt werden, dass die Paralleltests der LVD Sprache einen vergleichbaren Schwierigkeitsgrad aufweisen und hinsichtlich der Reliabilität sehr gute interne Konsistenzen zwischen .92 und .94 sowie sehr gute Trennschärfen von .46 bis .86 berichtet werden können. Bei der Itemschwierigkeit der einzelnen Sätze ist eine ansteigende Schwierigkeit zu erkennen, die aufgrund der zunehmenden grammatischen Komplexität und Satzlänge zu erwarten war. Die Korrelationen mit konvergenten Maßen zeigten eine hohe Übereinstimmung mit Testergebnissen, die dasselbe Konstrukt messen und keine Übereinstimmung mit divergenten Maßen. Die explorative Faktorenanalyse lieferte neben den Ergebnissen zur Validität einen weiteren Hinweis für das Sätze Nachsprechen als Globalmaß sprachlicher Fähigkeiten.

Die vorliegenden Testergebnisse geben erste Anhaltspunkte, dass das Nachsprechen von Sätzen eine vielversprechende Methode zur Erfassung sprachlicher Fähigkeiten darstellt. Die neu entwickelte LVD Sprache kann die klassischen Gütekriterien Itemschwierigkeit, Reliabilität und Validität erfüllen. Ergebnisse zur Erprobung im Längsschnitt lagen zu diesem Zeitpunkt noch nicht vor.

6.5 Ergebnisse der Pilotierung am SBBZ Sprache (Publikation C)

In der vorliegenden Studie wurde die neu entwickelte LVD Sprache bei Kindern mit Sprachentwicklungsstörungen (SES), die an einem Sonderpädagogischen Bildungs- und Beratungszentrum mit dem Förderschwerpunkt Sprache (SBBZ Sprache) in Baden-Württemberg beschult werden, erprobt. Das Hauptziel der Studie war die Überprüfung der Validität und Reliabilität der LVD Sprache für diese Zielgruppe.

> Besca, M., Krauss, T. & Ennemoser, M. (2023). Pilotierung einer Lernverlaufsdiagnostik bei Kindern mit Sprachentwicklungsstörung am SBBZ Sprache. *Forschung Sprache* (2), 16–26.

Methode (Stichproben, Testverfahren und Testdesign)

Die Daten von drei Teilstichproben konnten in die Analysen einbezogen werden. Stichprobe 1 bestand aus 19 Kindern der 1. und 2. Klasse eines SBBZ Sprache im Alter von 7;1 bis 9;0 Jahren (*M* = 7;8 Jahre). Mit diesen Kindern wurden alle Paralleltests innerhalb von drei Wochen durchgeführt. Stichprobe 2 bestand aus 19 Kindern der 2. und 3. Klasse eines SBBZ Sprache mit einem Durchschnittsalter von 8;8 Jahren. Mit dieser Gruppe wurden die LVD Sprache im Längsschnitt erprobt und weitere Validitätsmaße im Prä- und Posttest erhoben. Bei Stichprobe 3 handelt es sich um 31 Kinder aus der 1. Klasse an SBBZs Sprache, die im Durchschnitt 7;8 Jahre alt waren. Mit dieser Gruppe wurde sowohl ein Paralleltest als auch ein Sprachverständnistest als Validitätsmaß durchgeführt. In allen Stichproben waren ein- und mehrsprachig aufwachsende Kinder vertreten.

Zur Überprüfung der konvergenten Validität wurden folgende Testverfahren eingesetzt: der Subtest Bildbenennung (Wortschatz) aus dem SET 5-10 (Petermann, 2018), die Subtests Grammatik und Sätze Nachsprechen aus dem P-ITPA (Esser & Wyschkon, 2010), der Marburger Sprachverständnistest MSVK (Lohaus & Elben, 2000) und die Mottier-Silben (Mottier, 1951) zur Erfassung der Arbeitsgedächtnisleistung. Die divergente Validität wurde mit den Subtests Zehn mehr und Basisrechnen aus dem Test der mathematischen Basiskompetenzen MBK 2-4 (Krajewski & Ennemoser, in Vorb.) überprüft.

Ergebnisse und Diskussion

Die Ergebnisse zeigen, dass die verschiedenen Paralleltests vergleichbar schwierig sind und sich daher für eine Messung im Längsschnitt eignen. Die interne Konsistenz der Tests war sehr hoch (.90 bis .94), was auf eine zuverlässige Messung hinweist. Die Korrelationen der LVD Sprache mit bestehenden Sprachkompetenztests waren durchweg hoch signifikant, bei divergenten Maßen zeigte sich kein Zusammenhang.

Darüber hinaus konnte festgestellt werden, dass das Verfahren sensitiv auf Veränderungen der Nachsprechleistung reagiert. Bereits nach vier Monaten konnten signifikante Leistungszuwächse

festgestellt werden, was auf die Änderungssensitivität des Verfahrens hinweist. Allerdings wurde auch die Möglichkeit von Übungseffekten diskutiert, weshalb in zukünftigen Studien eine Kontrollgruppe einbezogen werden soll. Die explorative Analyse der individuellen Lernverläufe ergab unterschiedliche Entwicklungsmuster: Während einige Kinder kontinuierliche Fortschritte zeigten, stagnierten oder verschlechterten sich die Leistungen anderer. Dieses Ergebnis findet sich auch in anderen Studien (z.B. Jungjohann et al., 2021) und ist an einem SBBZ Sprache bei Kindern mit einer diagnostizierten Sprachentwicklungsstörung zu erwarten.

Mit den Ergebnissen der Untersuchung der Kinder am SBBZ Sprache konnten die Erkenntnisse der ersten Untersuchung in der Kita hinsichtlich Reliabilität und Validität erfolgreich repliziert werden. Da Ergebnisse von Kita-Kindern ohne sprachliche Beeinträchtigung nicht automatisch auf Schulkinder mit Sprachentwicklungsstörung übertragen werden können, ist dies nicht selbstverständlich. Darüber hinaus konnte ein erster Nachweis für die Änderungssensitivität der LVD Sprache erbracht werden. Die Analyse individueller Lernverläufe gibt erste Hinweise auf die praktische Relevanz. Der Einsatz der LVD Sprache könnte Lehrkräften helfen, individuelle Schwankungen in der Nachsprechleistung frühzeitig zu erkennen und gezielt darauf zu reagieren.

6.6 Ergebnisse der Pilotierung im Kindergarten (Publikation E)

Mit den vollständigen Daten der Erprobung im Längsschnitt sollen die psychometrischen Eigenschaften der neu entwickelten LVD Sprache zur Messung des sprachlichen Entwicklungsverlaufs für Kinder im Alter von vier bis sechs Jahren umfassend überprüft werden.

> Besca, M. & Ennemoser, M. (eingereicht): Assessing Oral Language Progress in Children Aged 4–6. Evaluation of the Psychometric Properties.

Methode (Stichproben, Testverfahren und Testdesign)

Diese Studie verwendet die Daten von zwei Stichproben: einer Stichprobe im Querschnitt von 87 Kindern im Alter von vier bis sechs Jahren aus sechs Kindertagesstätten in Baden-Württemberg und einer Stichprobe im Längsschnitt von 127 Kita-Kindern zwischen vier und sechs Jahren, aus sieben Kindertagesstätten in ländlichen oder städtischen Gebieten in Baden-Württemberg. Im Querschnitt wurden alle Paralleltests der LVD Sprache mit jedem Kind innerhalb von zwei Wochen durchgeführt. Die Längsschnittstudie umfasste eine Experimentalgruppe (88 Kinder), mit der alle vier Wochen ein Paralleltest durchgeführt wurde und eine Kontrollgruppe (39 Kinder), die nur zu Beginn und am Ende getestet wurde, um einen möglichen Übungseffekt durch wiederholtes Messen auszuschließen.

Zur Überprüfung der Validität wurden der Untertest Bildbenennung (Wortschatz) aus dem SET 5-10 (Petermann, 2018), der Untertest Grammatik aus der P-ITPA (Esser & Wyschkon, 2010), der

statusdiagnostische Test zum Sätze Nachsprechen aus dem SET-K 3-5 (Grimm, 2015) sowie die Mottier-Silben (phonologisches Arbeitsgedächtnis) (Mottier, 1951) eingesetzt. Für die divergente Validität wurden die Untertests Zahlenfolge und Ziffernkenntnis aus dem Test der mathematischen Basiskompetenzen MBK 0 (Krajewski, 2018) verwendet. Der soziökonomische Status wurde über die Bücherfrage (Heppt et al., 2022) bei den Kindern erhoben. Zusätzlich wurde die Begabung anhand von der Untertests Matrizen und Bildkonzepte aus dem WPPSI (Petermann & Daseking, 2018) erfasst. Die Informationen zu den Sprachen der Kinder, zur Sprachentwicklung, zum Kontakt mit DaZ und zum Sprachförderbedarf wurden per Fragebogen erhoben.

Ergebnisse

Die Ergebnisse der Studie geben einen umfassenden Einblick in die psychometrischen Eigenschaften und die Anwendung der LVD Sprache. Die Überprüfung der Homogenität anhand der Querschnittsdaten ergab, dass die Itemschwierigkeiten der Paralleltests der LVD Sprache im mittleren Schwierigkeitsbereich sehr nahe beieinander liegen. Die LVD Sprache weist eine hohe interne Konsistenz (Cronbachs Alpha = .99) und ausgezeichnete Trennschärfe mit Werten zwischen .93 und .96 auf. Der mittlere Inter-Item-Korrelationswert als Homogenitätsindex liegt bei .91. Die Eindimensionalität der LVD Sprache wurde mit Hilfe einer konfirmatorischen Faktorenanalyse bestätigt (CFI = 1.00, RSMEA = .00), 92% der Varianz werden durch den Faktor Sprachkompetenz erklärt.

Die Paralleltest-Reliabilität zeigt sehr hohe Korrelationen zwischen den benachbarten Messzeitpunkten (r = .97, p<.001). Die konvergente Validität der LVD Sprache ist hoch, insbesondere im Vergleich zu einem Wortschatztest, einem Grammatiktest und einem statusdiagnostischen Test zum Sätze Nachsprechen. Die Werte liegen zwischen .81 bis .94, was einem großen Effekt entspricht. Die Korrelationen mit den mathematischen Basiskompetenzen sind ebenfalls hoch (r = .56). Die Korrelationen mit anderen divergenten Maßen sind eher moderat (Arbeitsgedächtnis, r =.30/.34; sozioökonomischer Status, r =.44) oder gering (Begabung, r =.27). Die Korrelationen zwischen der LVD Sprache und den konvergenten Maßen (Wortschatz, Grammatik, statusdiagnostische Satzwiederholungsaufgabe) sind signifikant höher als die Korrelationen zwischen der LVD Sprache und den divergenten Maßen (Begabung, mathematische Basiskompetenzen).

Die deskriptive Beschreibung der Rohwerte über ein halbes Jahr zeigt einen kontinuierlichen Anstieg der sprachlichen Fähigkeiten, insbesondere in den unteren Quartilen. Dies kann als mittlerer Effekt interpretiert werden (d_{Cohen} = .63 und .73). Eine Varianzanalyse mit Messwiederholung mit den Gesamtdaten ergab, dass sich die Messzeitpunkte signifikant unterscheiden $F(4.70,408.91) = 9.00$, $p < .001$, mit einer mittleren Effektstärke von $\eta^2 = .09$. Nachfolgende Analysen mit T-Tests ergaben, dass ab dem 5. Messzeitpunkt signifikante Unterschiede zum 1. Messzeitpunkt vorliegen. Werden nur die unteren drei Quartile in die Analyse einbezogen, zeigen sich bereits ab dem 3. Messzeitpunkt signifikante

Unterschiede. Der sogenannte Übungseffekt konnte durch den Vergleich der Experimental- vs. Kontrollgruppe mit Hilfe einer Varianzanalyse ausgeschlossen werden.

Die Analyse der unterschiedlichen Entwicklungsverläufe anhand von Trendlinien macht deutlich, dass mehrsprachige Kinder und Kinder in den unteren Quartilen von einem niedrigeres Ausgangsniveau starten, aber deutliche Zuwachsraten von 1,5 bis 1,7 Wörtern pro Messzeitpunkt aufweisen. Kinder in den beiden oberen Quartilen starten von einem hohen Ausgangsniveau und zeigen nur geringe Zuwachsraten. Ein- und mehrsprachig aufwachsende Kinder unterscheiden sich im Anfangsniveau, zeigen aber vergleichbare Zuwachsraten von 0,8 bis 1,1 Wörtern pro Messzeitpunkt.

Diskussion

In der vorliegenden Studie konnte gezeigt werden, dass die LVD Sprache ein geeignetes Instrument zur Erfassung der Sprachentwicklung von vier- bis sechsjährigen Kindern ist, insbesondere für Kinder mit einem niedrigen sprachlichen Ausgangsniveau. Die Studie zeigt, dass das Instrument eine eindimensionale Struktur sowie eine hohe Homogenität und Reliabilität aufweist. Die hohen Korrelationen mit konvergenten Tests zum Wortschatz, zur Grammatik und zum Sätze Nachsprechen belegen, dass die LVD Sprache zur Erfassung sprachlicher Fähigkeiten geeignet zu sein scheint. Die Ergebnisse der Längsschnittstudie sprechen für eine sensitive Erfassung der sprachlichen Entwicklung.

Die moderate Korrelation der LVD Sprache mit dem Arbeitsgedächtnismaß kann dadurch erklärt werden, dass der Einfluss des phonologischen Arbeitsgedächtnisses auf das Wortschatzlernen ab dem Vorschulalter abnimmt und stattdessen das zuvor erworbene Wortschatzwissen eine größere Rolle spielt (Gathercole et al., 1997).

Die Längsschnittdaten zeigen, dass Kinder mit einem niedrigen sprachlichen Ausgangsniveau deutliche Fortschritte zeigen, während bei Kindern in höheren Leistungsbereichen eine Stagnation zu beobachten ist. Stagnationen oder negative Verläufe werden auch in anderen Studien beschrieben (L. S. Fuchs et al., 1993), allerdings könnte diese Stagnation durch die Testkonstruktion bedingt sein. Möglicherweise sind anspruchsvollere Sätze notwendig, um die sprachliche Entwicklung in diesen Gruppen besser zu erfassen. Für Kinder mit geringeren sprachlichen Fähigkeiten scheint die LVD Sprache jedoch ein geeignetes Maß für die sprachliche Entwicklung zu sein.

Mehrsprachige Kinder zeigen im Vergleich zu einsprachigen Kindern insgesamt niedrigere sprachliche Fähigkeiten, aber vergleichbare Wachstumsraten. Dieses Ergebnis stimmt mit anderen Studien überein, die zeigen, dass es mitunter Jahre dauern kann, bis mehrsprachige Kinder das gleiche Sprachniveau wie einsprachige Kinder erreichen (Collier & Thomas, 2017). Dies kann als Hinweis interpretiert werden, dass es möglich es ist, mit der LVD Sprache die unterschiedlichen Fähigkeiten ein- und mehrsprachiger Kinder abzubilden.

Limitationen und Ausblick

Bei der Interpretation der Reliabilitätswerte ist zu beachten, dass sich diese nur auf die getestete Stichprobe, die eingesetzten Testverfahren und diesen Messzeitraum beziehen. Mit den vorliegenden Ergebnissen ist es jedoch gelungen, die Ergebnisse zu replizieren, die bereits für Kinder mit einer Sprachentwicklungsstörung berichtet wurden (Besca et al., 2023). Weitere Erkenntnisse werden sich aus der Analyse der Daten von rund 100 Grundschulkindern ergeben, die derzeit ausgewertet werden.

Während der Entwicklung der LVD Sprache wurden keine Unterschiede auf die Nachsprechleistung zwischen einer mündlich vorgesprochenen und digitalen Präsentation der Sätze gefunden. Jedoch haben aktuelle Studien zu den Mottier-Silben gezeigt, dass die Art der Durchführung (analog vs. digital, langsam vs. schnell) einen signifikanten Einfluss auf die Nachsprechleistung haben kann (Heeg et al., 2024). Daher soll zukünftig eine digitale Version der LVD Sprache entwickelt werden. Zudem erleichtert und verbessert eine digitale Version die Auswertung.

Die Studie lieferte erste Hinweise zur Änderungssensitivität anhand der Analyse der Mittelwertsunterschiede. Eine weitere Möglichkeit zur Untersuchung der Änderungssensitivität ist ein Kontrollgruppendesign, wobei eine Gruppe eine Förderung erhält, während die Kontrollgruppe nicht gefördert wird (Breitenbach, 2020). Anschließend werden die Fortschritte verglichen, um zu analysieren, ob auch kurzfristige und eher kleine Veränderungen durch eine Intervention erfasst werden können.

Fazit

Die LVD Sprache scheint geeignet, die Sprachentwicklung von Kindern von vier bis sechs Jahren reliabel, valide und änderungssensitiv zu erfassen, insbesondere von Kindern mit geringen sprachlichen Fähigkeiten.

6.7 Erste Ergebnisse aus Einzelfallstudien

Die Paralleltests der LVD Sprache wurden im Rahmen von Abschlussarbeiten parallel zu Interventionsmaßnahmen erprobt. Exemplarisch werden die Ergebnisse einer Intervention im Bereich der Grammatik vorgestellt (Scheeff, 2022). Dazu wurden zwei Kinder aus der Stichprobe im Längsschnitt ausgewählt. Sie erhielten eine sechswöchige sprachtherapeutische Intervention zur Verbendstellung im Nebensatz (Kontextoptimierung nach Motsch, 2017). Zusätzlich zu den in der Studie eingesetzten Testverfahren wurde mit den Kindern der Untertest zur Verbstellung aus dem ESGRAF-R (Motsch & Rietz, 2019) im Prä-, Post- und Follow-Up durchgeführt. Die Forschungsfrage war hier, ob sich die Veränderungen, die durch eine Intervention erreicht werden können, auch in der Entwicklungskurve der LVD Sprache zeigen.

Ergebnisse Grammatiktests aus ESGRAF-R und P-ITPA

Die Auswertung der statusdiagnostischen Testverfahren des Prä-, Post- und Follow-Up-Tests macht deutlich, dass beide Kinder von der Intervention sowohl in Bezug auf das konkrete Therapieziel als auch in Bezug auf ihre allgemeinen grammatischen Fähigkeiten profitiert haben. In der Verbendstellung erreichte Kind 1 im Prätest eine Korrektheit von 73,3%, im Posttest und Follow-up eine 100%-Korrektheit. Kind 2 erreichte im Posttest eine Korrektheit von 46,7%, im Posttest 77,8% und im Follow-up-Test76,9%. Im Untertest Grammatik des P-ITPA verbesserte sich Kind 1 von einem T-Wert von 22 auf 32, Kind 2 von einem T-Wert von 33 auf 37.

Deskriptive Beschreibung der Ergebnisse der LVD Sprache

Die Ergebnisse der LVD Sprache wurden als Liniendiagramm dargestellt (vgl. Abbildung 2) und die Entwicklungsverläufe in Anlehnung an Strathmann und Klauer (2010) visuell inspiziert und deskriptiv beschrieben. Der Interventionszeitraum ist farblich markiert. Das sehr niedrige Ausgangsniveau der LVD Sprache der beiden Kinder (Kind 1 Rohwert 46, Kind 2 Rohwert 30) kann den beiden unteren Leistungsquartilen zugeordnet werden (vgl. Besca & Ennemoser, eingereicht). Bei Kind 1 kann beobachtet werden, dass sich bereits gegen Ende der Intervention zum 5. MZP bzw. nach der Intervention zum 6. Messzeitpunkt eine Veränderung der Lernkurve zeigt. Bei Kind 2 ist erst nach der Intervention zum 6. und 7. MZP eine Veränderung in der Lernkurve zu beobachten.

Analyse der Ergebnisse der LVD Sprache anhand von Trendlinien

Um zu prüfen, ob sich die beobachtete Veränderung der Lernkurve vor und nach der Intervention auf die Intervention zuführen lässt und nicht die allgemeine Entwicklung widerspiegelt, werden in einem nächsten Schritt die Trendlinien der beiden Kinder mit den Daten der Gesamtstichprobe verglichen. Trendlinie können zur Interpretation der Entwicklungsverläufe genutzt werden (Jungjohann et al., 2021), wenn noch keine Normdaten vorliegen (Förster et al., 2017). In den Studien zur Erprobung am SBBZ Sprache (Kap. 6.5) und zur Erprobung im Kindergarten (Kap. 6.6) wurden bereits Trendlinien zur Interpretation der Entwicklungsverläufe genutzt.

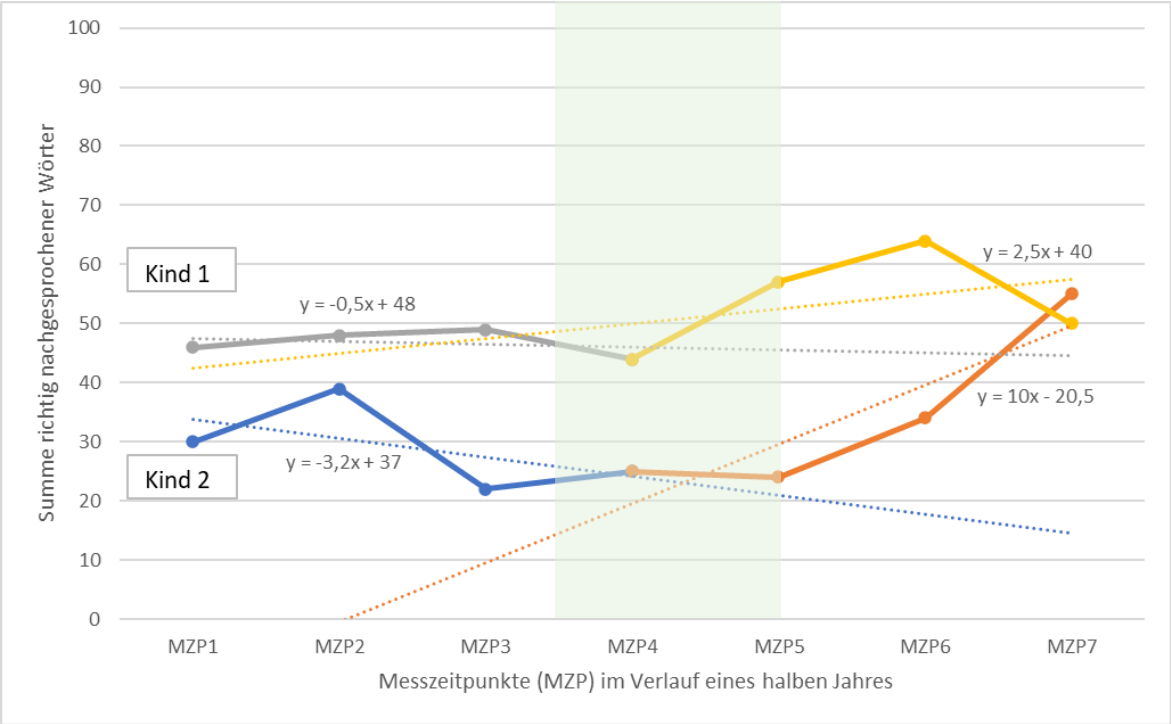

Abbildung 2 Entwicklungsverlauf anhand der LVD Sprache im Verlauf eines halben Jahres

Es wurden separate Trendlinien für Baseline und vor/nach der Intervention berechnet, um Veränderung zwischen Baseline und Intervention bzw. nach der Intervention zu beschreiben. Der vierte Messzeitpunkt wurde ebenfalls in die Baseline einbezogen, da zwei Wochen nach Interventionsbeginn noch keine Veränderungen zu erwarten waren. Für Kind 1 zeigt die Trendlinie der ersten vier Messzeitpunkte einen stagnierenden Verlauf (Steigung -0,5 Wörter pro Messzeitpunkt), für Kind 2 einen absteigenden Verlauf (Steigung -3,2 Wörter pro Messzeitpunkt). Die Trendlinien der Messzeitpunkte während und nach der Intervention zeigen nun für beide Kinder einen deutlich ansteigenden Verlauf mit 2,5 Wörtern pro Messzeitpunkt für Kind 1 und mit 10 Wörtern pro Messzeitpunkt für Kind 2 im Vergleich zur eigenen Baseline. Im Vergleich zu den Lernverläufen der unteren beiden Quartile der Gesamtstichprobe (niedriges Ausgansniveau und Zuwachsraten von 1,5 bis 1,7 Wörtern pro Messzeitpunkt) zeigen die Ergebnisse der beiden geförderten Kinder einen Zuwachs, der deutlich über dem erwarteten Anstieg liegt. Dieses Ergebnis deckt sich mit den festgestellten Verbesserungen der beiden Kinder im Post- und Follow-up -Test der statusdiagnostischen Grammatiktests im Vergleich zum Prätest.

In Bezug auf die Fragestellung kann das Ergebnis dieser Einzelfallstudie als erster Hinweis gedeutet werden, dass die Fortschritte, die durch eine Intervention erzielt wurden, änderungssensitiv erfasst mit Hilfe der LVD Sprache sichtbar gemacht werden konnten. Weitere Interventionsstudien im Einzel- oder Gruppendesign sollten sich zur Untersuchung der Änderungssensitivität anschließen.

7 Diskussion

Übergeordnetes Ziel der vorliegenden Arbeit ist es, die Wirksamkeit von Sprachförderung konsequent zu Ende zu denken, indem zum einen die Wirksamkeit einer Sprachfördermaßnahme am Beispiel des Dialogischen Lesens im Kita-Alltag überprüft wird und zum anderen eine Lernverlaufsdiagnostik sprachlicher Fähigkeiten als neue Möglichkeit der Evaluation von Sprachförderung im Einzelfall entwickelt und erprobt wird.

> Forschungsfrage 1: Ist das Dialogische Lesen als Sprachfördermaßnahme auch unter den weniger gut kontrollierbaren Alltagsbedingungen wirksam?

Die Wirksamkeit des Dialogischen Lesens im Kita-Alltag wurde in dieser Arbeit durch den Vergleich zweier Studien zum Dialogischen Lesen (Studie 1 und Folgestudie) untersucht. In Studie 1 zeigten sich im Posttest keine signifikanten Unterschiede im Sprachtest der Kinder, die von pädagogischen Fachkräften oder wissenschaftlichen Mitarbeiter*innen gefördert wurden, d.h. es gelingt den pädagogischen Fachkräften ähnliche Fördereffekte zu erzielen wie den wissenschaftliche Mitarbeiter*innen. Zu einem ähnlichen Ergebnis kamen Flack und Kolleg*innen (2018) in ihrer Meta-Analyse, allerdings nur bezogen auf das Wortverständnis. Andere Meta-Analysen berichten jedoch, dass die wissenschaftlichen Mitarbeiter*innen größere Fördereffekte erzielten als die pädagogischen Fachkräfte (Marulis & Neuman, 2010; Mol et al., 2009). Ein deskriptiver Unterschied zugunsten der wissenschaftlichen Mitarbeiter*innen findet sich auch in der Studie 1. Die Follow-Up-Ergebnisse in Studie 1 weisen jedoch darauf hin, dass die von pädagogischen Fachkräften geförderten Kinder langfristig stabilere und signifikant größere Lernzuwächse erreichen als die von wissenschaftlichen Mitarbeiter*innen geförderten Kinder (Dickel-Lehnigk et al., 2022). Die Analyse zeigte auch, dass sich die Ergebnisse der von pädagogischen Fachkräften geförderten Kinder nicht signifikant von denen der Kontrollgruppe unterschieden. Als mögliche Erklärung wurde das nachträglich angewandte einheitliche Auswahlkriterium und die daraus resultierende kleinere Stichprobe diskutiert, die zu einer geringeren Teststärke geführt haben könnte. Auch hier deuten die Follow-Up-Ergebnisse darauf hin, dass die pädagogischen Fachkräfte langfristig die größeren Fördererfolge erzielen. Die langfristige Absicherung von Fördererfolge durch Follow-Up-Untersuchungen sollte bei allen zukünftigen Forschungen berücksichtigt werden.

In der Folgestudie wurden die Daten von 201 Kindern dahingehend analysiert, ob die pädagogischen Fachkräfte bei der Durchführung des Dialogischen Lesens im Kita-Alltag größere Fördereffekte erzielen als eine nicht geförderte Kontrollgruppe. Die Analysen ergaben, dass die Kinder, die von den pädagogischen Fachkräften im Kita-Alltag mit dem Dialogischen Lesen gefördert wurden, signifikant größere sprachliche Kompetenzzuwächse aufwiesen als die Kinder der Kontrollgruppe. Fördereffekte zeigten sich in den Bereichen Wortschatz und Sprachverständnis, jedoch nicht in der Grammatik. Die

Fördereffekte auf das Sprachverständnis als komplexe sprachliche Fähigkeit sind besonders vielversprechend, da sie bisher nur in wenigen Studien untersucht wurden (Pillinger & Vardy, 2022). Als möglicher Grund für die ausbleibenden Ergebnisse im Bereich Grammatik wurde der verwendete Test diskutiert, der vermutlich die Fortschritte insbesondere bei mehrsprachigen Kindern nicht sensitiv genug erfassen konnte (Kaltenbacher, 2011). Ein direkter Vergleich der beiden Studien war jedoch nur möglich, da die gleichen Testverfahren wie in Studie 1 verwendet wurden. In zukünftigen Studien sollten Testverfahren gewählt werden, wie z.B. die neunormierte PDSS, die Orientierungswerte für mehrsprachige Kinder anbieten (Kauschke et al., 2023). Eine weitere Erkenntnis bezieht sich auf die die Anzahl der Fördersitzungen, hier zeigte sich eine große Varianz von min. 3 bis max. 27 Fördersitzungen, mit denen das Dialogische Lesen im Alltag umgesetzt wurde. Auf die Fördereffekte hatte das jedoch keinen Einfluss. Dies kann als Hinweis darauf gewertet werden, dass die pädagogischen Fachkräfte die Sprachförderstrategien vermutlich auch in anderen Alltagssituationen anwenden. Ob und in welche anderen Alltagssituationen die Sprachförderstrategien übertragen werden, sollte in zukünftigen Forschungsvorhaben genauer untersucht werden.

Ein Vergleich aller Gruppen aus Studie 1 und der Folgestudie wurde durchgeführt, um die Frage zu beantworten, ob es den pädagogischen Fachkräften gelingt im Kita-Alltag ähnliche Fördereffekte zu erzielen wie pädagogischen Fachkräften unter eng kontrollierten Bedingungen. Ein besonderer Vorteil der gesamten Studie waren die identischen Einschlusskriterien und Testverfahren, um die Gruppen überhaupt miteinander vergleichen zu können. Die Analysen ergaben, dass die Fördererfolge im Kita-Alltag und unter eng kontrollierten Bedingungen vergleichbar sind. Dies stellt eine wichtige Erkenntnis für die Durchführung und eine mögliche Implementierung des Dialogischen Lesens im Kita-Alltag dar, insbesondere vor dem Hintergrund einer Fortbildung der Fachkräfte und der Durchführung der Förderung unter alltagstauglichen Bedingungen. Eine weitere wichtige Erkenntnis dieser Studie stellt der Aspekt der Fortbildung und Supervision mit überschaubarem Aufwand dar, da wirksame Fortbildungen i.d.R. zeitintensiv und langfristig angelegt sind (Lipowsky, 2010). Das Dialogische Lesen scheint jedoch eine Methode zu sein, die mit einer eher kurzen Fortbildung von 1 ½ Tagen und einer Supervision an einem Nachmittag vermittelt werden kann. Frühere Studien von Whitehurst und Kolleg*innen (1988) berichten ähnliche Erkenntnisse bei der Fortbildung von Eltern.

Insgesamt liefern die Ergebnisse einen wichtigen Beitrag zur Frage nach der Wirksamkeit von Sprachfördermaßnahmen im Kita-Alltag, die im Zuge der Bemühungen um Evidenzbasierung häufig vernachlässigt wird (Rauch & Hartig, 2018). In dieser Studie konnte die Wirksamkeit des Dialogischen Lesens im Kita-Alltag im Vergleich zu einer Kontrollgruppe nachgewiesen werden. Im direkten Vergleich mit den Ergebnissen einer Studie, in der die Durchführungsbedingungen aus forschungsmethodischen Gründen deutlich enger kontrolliert wurden, zeigte sich, dass pädagogische

Fachkräfte im Kita-Alltag ähnliche Fördereffekte erzielen können wie Fachkräfte in einem eng kontrollierten Setting. Damit können bestehende Erkenntnisse zur Wirksamkeit des Dialogischen Lesens bzw. des Einsatzes von Sprachförderstrategien im Kita-Alltag erweitert werden (Böse et al., 2023; Voltmer et al., 2021)

Auch wenn bei der Auswahl einer evidenzbasierten Maßnahme ein Nachweis der Wirksamkeit im Alltag vorliegt, kann dies nur der erste Schritt eines evidenzbasierten Vorgehens sein. Er garantiert noch nicht, dass alle Kinder ausreichend profitieren. Weitere Analysen zu Studie 1 haben gezeigt, dass auch nach einer Förderung mit dem Dialogischen Lesen mit signifikanten Fördereffekten immer noch 12-19% der Kinder in den Post- und Follow-Up-Tests einen Sprachförderbedarf mit Ergebnissen zwischen 0,5 und 1,5 Standardabweichungen bzw. mehr als 1,5 Standardabweichungen unter dem Mittelwert aufweisen (Dickel-Lehnigk et al., 2022). Bei dieser Gruppe von Kindern spricht man von sogenannten Nonrespondern, d.h. diese Kinder haben nicht ausreichend von der Förderung profitiert. Im Rahmen von wissenschaftlichen Untersuchungen werden Eltern und Fachkräfte meist über die Ergebnisse informiert, so dass die Wahrscheinlichkeit hoch ist, dass diese Kinder anschließend noch eine weitere oder andere Förderung erhalten. Findet die Sprachförderung im Rahmen des regulären Kita-Alltags statt, so wird der Erfolg der Maßnahme in der Regel nicht systematisch für einzelne Kinder evaluiert. Eine Evaluation nach Abschluss einer Intervention gehört jedoch eigentlich selbstverständlich zu einem evidenzbasierten Vorgehen dazu (Y. Blumenthal & Mahlau, 2015). Findet eine Evaluation jedoch erst nach Abschluss einer Intervention statt, kann wertvolle Zeit verloren gehen, in der eine Maßnahme verändert, angepasst oder eine weitere Überprüfung eingeleitet werden könnte, wenn erst dann festgestellt wird, dass Kinder nicht ausreichend von einer Förderung profitieren. Idealerweise sollten die Fördermaßnahmen bereits prozessbegleitend evaluiert werden. Hierfür eignen sich Verfahren der Lernverlaufsdiagostik, mit denen Nonresponder frühzeitig identifiziert werden können (Deno, 2003).

Bisher werden Verfahren der Lernverlaufsdiagnostik vor allem im schulischen Kontext eingesetzt. In Kapitel 4.3 wurde ausführlich dargestellt, dass sich die Lernverlaufsdiagnostik auch für den binnendifferenzierten Unterricht eignet und hier als systematische Abfolge von Differenzierungsstrategien losgelöst vom schulischen Rahmenkonzept einsetzt werden kann. Eine systematische Abfolge von Differenzierungsstrategien wäre auch im vorschulischen Bereich denkbar: Alle Kinder erhalten sprachliche Bildungsangebote im Sinne einer primären Prävention, um sprachliche Entwicklungsrisiken zu minimieren. Mit Hilfe eines regelmäßig eingesetzten Screeningverfahrens kann dann festgestellt werden, bei welchen Kindern ein erhöhter Förderbedarf besteht. Diese erhalten dann auf der zweiten (Differenzierungs-)Ebene (vgl. RTI-Ansatz) ein zusätzliches Förderangebot (sekundäre Prävention). Diese findet z.B. im Rahmen einer zusätzlichen Kleingruppenförderung in der Kita statt. Eine zusätzliche Kleingruppenförderung ist auch Bestandteil der aktuell geplanten Bildungsmaßnahme Sprachfit in

Baden-Württemberg. Würde parallel zur Fördermaßnahme evaluiert werden, inwieweit die Kinder von der Förderung profitieren, könnte frühzeitig darauf reagiert werden.

Für den Bereich der mündlichen sprachlichen Fähigkeiten und für die Zielgruppe der Kita-Kinder besteht jedoch noch eine Forschungslücke (Besca & Ennemoser, 2022; Mahlau et al., 2014). Gerade für Kinder mit Sprachförderbedarf wäre eine systematische Evaluation von Fördermaßnahmen wichtig, da ein nicht unerheblicher Anteil von Kindern nicht ausreichend von Fördermaßnahmen profitiert, auch wenn eine evidenzbasierte Maßnahme eingesetzt wurde (vgl. Kap. 3.2). Auch die ernüchternden Ergebnisse von Begleit- und Evaluationsstudien machen deutlich, dass hier noch dringender Verbesserungsbedarf beim Einsatz von Sprachfördermaßnahmen besteht. Diese Forschungslücke ist sowohl national als auch international erkannt worden, denn verschiedene Autoren fordern, dass zukünftig die Entwicklung von LVD-Verfahren für mündliche Sprachkompetenzen vorangetrieben (Adlof & Hogan, 2019) und zuverlässige und praktikable Instrumente entwickelt werden, die spezifisch auch für die mündliche Sprache und das therapeutische (Einzel-)Setting geeignet sind (Winkes & Till, 2023).

Für die Entwicklung eines Verfahrens zur LVD kann ein robuster Indikator verwendet werden, der eine bestimmte Fähigkeit als Globalmaß repräsentiert (Gebhardt et al., 2021). Das Sätze Nachsprechen scheint aus mehreren Gründen ein vielversprechendes Globalmaß bzw. ein robuster Indikator zur Erfassung sprachlicher Fähigkeiten zu sein (Klem et al., 2015). Daher wurde im zweiten Teil dieser Arbeit die folgende Frage untersucht:

> Forschungsfrage 2: Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten?

Diese Frage ist in mehreren Schritten zu beantworten. Um überhaupt prüfen zu können, ob sich das Sätze Nachsprechen für eine Lernverlaufsdiagnostik sprachlicher Fähigkeiten eignet, musste eine Lernverlaufsdiagnostik entwickelt und konstruiert werden. Diese wurde anschließend quer- und längsschnittlich anhand der klassischen Gütekriterien sowie zusätzlich auf Änderungssensitivität, Testhomogenität und Eindimensionalität geprüft (Gebhardt et al., 2021). In der Arbeit werden die Ergebnisse der Erprobung bei ein- und mehrsprachigen Kita-Kindern im Alter von vier bis sechs Jahren (n = 127) sowie bei Kindern mit einer Sprachentwicklungsstörung am SBBZ Sprache (n=19) dargestellt (ausführlich in Kap. 6.5 – 6.6).

Die Ergebnisse der Itemanalyse zeigen eine überzeugende Homogenität der Paralleltests der LVD Sprache als Gesamtskala hin. Die Itemschwierigkeiten der Paralleltests liegen im mittleren Schwierigkeitsbereich sehr nahe beieinander. Ebenso zeigen sich eine hohe interne Konsistenz (Cronbachs Alpha = .99 für die Kita-Stichrobe insgesamt, Cronbachs Alpha = .90 bis .94 pro Paralleltest am SBBZ Sprache) und ausgezeichnete Trennschärfen mit Werten zwischen .93 und .96 für die Kita-Stichprobe und .25 bis .82 für die Stichprobe am SBBZ Sprache. Der mittlere Inter-Item-Korrelationswert

liegt als Homogenitätsindex bei .91. Daraus lässt sich ableiten, dass die Paralleltests einen vergleichbaren Schwierigkeitsgrad aufweisen und somit für den Einsatz im Verlauf geeignet sind. Die vereinzelten Trennschärfen von unter .3 am SBBZ Sprache lassen sich durch die kleine Stichprobe und die sehr unterschiedlichen sprachlichen Fähigkeiten der Kinder mit Sprachentwicklungsstörung erklären. Hervorzuheben ist, dass aufgrund der unterschiedlichen Stichproben die eigenen Daten repliziert werden konnten. Eine Replikation der Ergebnisse von Kita-Kindern ohne sprachliche Beeinträchtigung mit den Ergebnissen von Schulkindern mit Sprachentwicklungsstörung ist nicht selbstverständlich.

Ein Nachweis der Eindimensionalität der LVD Sprache konnte durch die explorative und konfirmatorische Faktorenanalyse erbracht werden. Die Faktorladungen liegen zwischen .95 und .97 in einem sehr hohen Bereich, 92% der Varianz konnten durch den Faktor Sprachkompetenz erklärt werden. Dies kann als zentraler Hinweis darauf gewertet werden, dass die LVD Sprache als Globalmaß sprachlicher Fähigkeiten interpretiert werden kann. Die Ergebnisse der Paralleltest-Reliabilität zeigen, dass die LVD Sprache zuverlässig und stabil mit einem geringen Messfehler misst. Die Korrelationen zwischen den benachbarten Messzeitpunkten liegen mit $r = .97$, $p<.001$ in einem sehr hohen Bereich.

Die Ergebnisse zur Validität ergaben für alle Teilstichproben hoch signifikante Korrelationen der LVD Sprache mit bestehenden Sprachkompetenztests (Kita: .81 bis .94, SBBZ Sprache .67 bis .89). Korrelationen mit divergenten Maßen ergaben keinen oder geringe Zusammenhänge. Im Vergleich der Korrelationswerte sind die Korrelationen zwischen der LVD Sprache und den konvergenten Maßen (Wortschatz, Grammatik, statusdiagnostische Satzwiederholungsaufgabe) signifikant höher als die Korrelationen zwischen der LVD Sprache und den divergenten Maßen (Begabung, mathematische Basiskompetenzen). Die moderate Korrelation mit dem Arbeitsgedächtnismaß lässt sich durch den abnehmenden Einfluss des phonologischen Arbeitsgedächtnisses auf den Wortschatzerwerb ab dem Vorschulalter und die zunehmende Bedeutung des zuvor erworbenen Wortschatzwissens erklären (Gathercole et al., 1997). Die Validitätsbefunde decken sich mit den Ergebnissen bisheriger Studien zum Sätze Nachsprechen, in denen das Sätze Nachsprechen als Spiegelbild eines zugrundeliegenden Sprachfähigkeitsfaktors beschrieben wird, der sich auf eine breite Palette an Sprachverarbeitungsfähigkeiten bezieht (Klem et al., 2015). Zusammenhänge mit Fähigkeiten im Wortschatz, in der Grammatik und im phonologischen Arbeitsgedächtnis stimmen ebenfalls mit Ergebnissen aus bisherigen Studien überein (Goldammer et al., 2011; Polišenská et al., 2015). Aus den Ergebnissen der verschiedenen Stichproben lässt sich zudem ableiten, dass die LVD Sprache sowohl für sprachunauffällige Kita-Kinder als auch für Schulkinder mit Sprachentwicklungsstörung als Maß der sprachlichen Fähigkeiten geeignet ist. Die Daten von sprachunauffälligen Schulkindern stehen noch aus.

Da es sich bei Verfahren der Lernverlaufsdiagnostik um Veränderungsmessungen handelt, muss außerdem ein Nachweis der Änderungssensitivität erbracht werden, d.h. es muss sichergestellt werden,

dass Veränderungen von Fähigkeiten über die Zeit sensitiv erfasst werden. Erste Hinweise und Nachweise der Änderungssensitivität konnten durch die deskriptive Beschreibung der Daten, die inferenzstatistische Analyse sowie die Interpretation von Trendlinien und einer Einzelfallintervention erbracht werden. Die deskriptive Beschreibung der Daten zeigt für beide Teilstichproben einen kontinuierlichen Anstieg der sprachlichen Fähigkeiten, insbesondere in den unteren Quartilen mit d_{Cohen} = .63 und .73, was als mittlerer Effekt interpretiert werden kann. Die inferenzstatistische Analyse zeigt signifikante Unterschiede zwischen den Messzeitpunkten mit einer mittlerer Effektstärke für die Kita-Stichprobe mit η^2 = .09, für die SBBZ Sprache-Stichprobe mit η^2 = .12. Eine daran anschließende Analyse mit T-Tests ergab, dass ab dem 5. MZP Unterschiede zum 1. MZP bestehen (Kita-Stichprobe). Analysen mit den unteren drei Quartilen ergaben signifikante Unterschiede ab dem 3. MZP. Dieses Ergebnis deutet darauf hin, dass die LVD Sprache besonders gut bei Kindern mit noch geringeren sprachlichen Fähigkeiten misst. Die Analyse der Entwicklungsverläufe anhand von Trendlinien zeigt, dass bei Kindern mit einem niedrigen sprachlichen Ausgangsniveau ein deutlicher Zuwachs zu verzeichnen ist, während bei Kindern in höheren Leistungsbereichen eine Stagnation zu beobachten ist. Diese Beobachtung wurden in beiden Stichproben gemacht. Auch in anderen Studien werden Stagnationen oder negative Verläufe beschrieben (L. S. Fuchs et al., 1993), bei der LVD Sprache könnte es aber durch die Testkonstruktion bedingt sein. Um zu überprüfen, ob die sprachliche Entwicklung in höheren Leistungsbereichen erfasst werden kann, müssten sprachlich noch anspruchsvollere Sätze entwickelt werden.

Bezüglich der ein- und mehrsprachig aufwachsenden Kinder zeigt sich ein signifikanter Unterschied im Anfangsniveau. Dieses Ergebnis deckt sich mit Studien zu mehrsprachigen Kindern, die berichten, dass es mitunter Jahre dauern kann, bis mehrsprachige Kinder das gleiche Sprachniveau erreichen wie einsprachige Kinder (Collier & Thomas, 2017). Die Verläufe von ein- und mehrsprachig aufwachsenden Kindern sind jedoch vergleichbar, so dass hier nicht automatisch davon ausgegangen werden kann, dass bei mehrsprachig aufwachsenden Kindern ein Förderdarf vorliegt. Dazu müssten im Einzelfall Ausgangsniveau und Verlauf interpretiert werden.

Die Interpretation von zwei Einzelfallstudien in der Kita konnte erste Erkenntnisse liefern, dass sich Veränderungen, die durch eine grammatische Intervention erreicht wurden, auch in der Entwicklungskurve der LVD Sprache abbilden lassen. Die Interpretation von zwei Entwicklungskurven am SBBZ Sprache deutet auch hier darauf hin, dass Ergebnisse einer Interventionsmaßnahmen (in diesem Fall die Beschulung an einem SBBZ Sprache) abgebildet und Konsequenzen abgeleitet werden könnten. Im Nachhinein konnte festgestellt werden, dass bei dem Kind mit einem stagnierenden Verlauf mit einer Wiederholung der Klasse reagiert wurde, das Kind mit überdurchschnittlichen Leistungen im Vergleich zur SBBZ Sprache Stichprobe wurde in eine allgemeine Grundschule umgeschult.

Insgesamt deuten alle Ergebnisse der Überprüfung der Gütekriterien darauf hin, dass die LVD Sprache zur systematischen Erfassung der sprachlichen Entwicklung von Kita-Kindern und Schulkindern mit Sprachentwicklungsstörung geeignet ist. Somit kann die zweite Forschungsfrage zu einem Großteil beantwortet werden, d.h. das Sätze Nachsprachen erscheint zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten geeignet. Die Ergebnisse zur Änderungssensitivität geben erste Hinweise, es sollte jedoch weiter überprüft werden, um herauszufinden, ob mit der LVD Sprache auch kurzfristige Veränderungen im Einzelfall abgebildet werden können. Im folgenden Abschnitt werden die Ergebnisse zu den beiden Forschungsfragen zusammengefasst und Implikationen für die Erprobung der LVD Sprache sowie weitere Impulse für die Evaluation von Sprachfördermaßnahmen wie dem Dialogischen Lesen diskutiert.

8 Fazit und Ausblick

Übergeordnetes Ziel der vorliegenden Arbeit war es, die Wirksamkeit von Sprachförderung konsequent zu Ende zu denken, indem zum einen die Wirksamkeit einer Sprachfördermaßnahme am Beispiel des Dialogischen Lesens im Kita-Alltag überprüft wurde (Forschungsfrage 1). Zum anderen eine Lernverlaufsdiagnostik sprachlicher Fähigkeiten, als neue Möglichkeit, die Wirksamkeit von Sprachförderung im Einzelfall zu evaluieren entwickelt und erprobt wurde (Forschungsfrage 2).

Hinsichtlich der ersten Forschungsfrage liefert die Arbeit neue Erkenntnisse zur Wirksamkeit des Dialogischen Lesens im Kita-Alltag im Vergleich zu einer Durchführung unter eng kontrollierten Bedingungen. Durch den Vergleich zweier Studien konnte gezeigt werden, dass pädagogische Fachkräfte auch im Kita-Alltag mit relativ geringem Fortbildungsaufwand signifikante Sprachfördererfolge erzielen können. Da die Befundlage zum Einfluss der Förderkraft bislang nicht eindeutig ist (Flack et al., 2018; Marulis & Neuman, 2010; Mol et al., 2009) konnte hier ein wertvoller Beitrag geleistet werden. Diese Erkenntnis ist auch vor dem Hintergrund bedeutsam, dass in Baden-Württemberg eine Bildungsmaßnahme verabschiedet werden soll, die vorsieht, dass Kinder mit Sprachförderbedarf regelmäßig in Kleingruppen von Fachkräften in den Kitas sprachlich gefördert werden sollen. Die Erkenntnis zu Fortbildungsinhalten und -aufwand stellen einen weiteren wichtigen Impuls für die Qualifizierung zukünftig benötigten Fachkräften im Rahmen der geplanten Bildungsmaßnahme dar. Um dem steigenden Bedarf an qualifizierten Fachkräften für die Sprachförderung gerecht zu werden, wäre eine Qualifikation nach dem Schneeballprinzip denkbar. Dazu müsste untersucht werden, ob es den Fachkräften, die als Multiplikatoren fungieren, gelingt, andere Fachkräfte so fortzubilden, dass diese ebenfalls Fördereffekte bei Kindern mit Sprachförderbedarf erzielen.

Besonders ermutigend sind die Fördereffekte auf das Sprachverständnis als komplexe sprachliche Fähigkeit (Pillinger & Vardy, 2022). Um Veränderungen auch in der Grammatik zu erfassen, sollte zukünftig ggf. die Auswahl der Testverfahren modifiziert werden.

In der Folgestudien zeigte sich eine große Varianz in der Anzahl der durchgeführten Fördersitzungen. Diese Varianz lässt vermuten, dass die Fachkräfte die Sprachförderstrategien auch im Kita-Alltag anwenden. Ob diese Vermutung tatsächlich zutrifft, könnte ein Schwerpunkt künftiger Forschung sein. Dazu müssten Alltagssituationen daraufhin analysiert werden, ob die pädagogischen Fachkräfte die Sprachförderstrategien auch in diesen Situationen anwenden.

Nach wie vor ist die Datenlage zu langfristigen Fördererfolgen von Sprachfördermaßnahmen eher gering (Pillinger & Vardy, 2022; Towson et al., 2017); die meisten Studien berichten nur über Prä-Post-Ergebnisse (Fitton et al., 2018; Flack et al., 2018). Daher sollten in Zukunft auch langfristige Fördererfolge untersucht werden. Dabei sollte auch untersucht werden, ob das veränderte

Kommunikationsverhalten der Fachkräfte, d.h. die Anwendung der Sprachförderstrategien, nach den Follow-Up-Untersuchungen bestehen bleibt oder ob sie in alte Kommunikationsmuster zurückfallen. Eine prozessbegleitende Diagnostik in Form einer Lernverlaufsdiagnostik könnte ggf. auch als Impuls genutzt werden, das eigene sprachförderliche Handeln immer wieder zu reflektieren, da eine prozessbegleitende Evaluation den Blick auf Lernergebnisse durch Unterricht und Förderung schärft (Breitenbach 2020).

In Bezug auf die zweite Forschungsfrage konnte die vorliegende Arbeit durch die Entwicklung und Erprobung einer LVD Sprache wichtige Erkenntnisse zur Eignung des Sätze Nachsprechen zur Lernverlaufsdiagnostik mündlicher sprachlicher Fähigkeiten liefern. Die Entwicklung der LVD Sprache basiert theoretisch auf aktuellen Erkenntnissen zur Entwicklung sprachlicher Fähigkeiten in den Bereichen Wortschatz und Grammatik sowohl bei Kindern mit unauffälligem Spracherwerb als auch bei Kindern mit Sprachförder- oder Sprachtherapiebedarf. Die Ergebnisse der Erprobung im Quer- und Längsschnitt liefern relevante Erkenntnisse zur psychometrischen Güte des Verfahrens sowie zu spezifischen Kriterien einer Lernverlaufsdiagnostik: der Homogenität, Eindimensionalität und Änderungssensitivität. Erste Erkenntnisse aus Einzelfallstudien liefern zusätzliche Hinweise auf die Änderungssensitivität und praktische Relevanz der LVD Sprache. Die Prüfung der Gütekriterien war der erste und zentrale Schritt, um die zweite Forschungsfrage grundlegend zu beantworten. Es konnte gezeigt werden, dass sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten eignet. Für die weitere Erprobung der LVD Sprache sollten folgende Schritte in Betracht gezogen werden, die nun kurz skizziert werden.

Die LVD Sprache wurde bereits mit einer weiteren Stichprobe von ca. 100 Grundschulkindern erprobt. Die noch ausstehende Auswertung könnte die vorhandenen Daten replizieren und die Eignung für sprachlich unauffällige Grundschulkinder überprüfen oder im Widerspruch zu den bisher gefundenen Ergebnissen stehen.

Für weitere Untersuchungen ist eine digitale Version der LVD Sprache geplant. Ziel ist es auch, eine Auswertungssoftware zur Verfügung zu stellen, um die Auswertung weiter zu vereinfachen und zu objektivieren. Bei einer Weiterentwicklung der LVD Sprache soll auch eine Version für ältere Kinder bzw. Kinder mit sehr guten sprachlichen Fähigkeiten entwickelt werden. Der bisher beobachtete Deckeneffekt konnte testbedingt durch die Konstruktion der Sätze erklärt werden. Um auch die sprachliche Entwicklung von Kindern aus den oberen Leistungsquartilen zu erfassen, wäre zu prüfen, ob noch anspruchsvollere Sätze hinsichtlich Grammatik und Wortschatz dazu geeignet wären.

Weitere Schritte zur Überprüfung der Änderungssensitivität stellen die Evaluation einer Sprachfördermaßnahme bzw. einer Interventionsstudie mit Hilfe der LVD Sprache dar. Dies kann im Rahmen eines Studiendesigns mit Interventions- und Kontrollgruppe, aber auch durch kumulierte Einzelfallstudien

erfolgen. Für die Auswertung der Einzelfallstudien sollten neben Methoden der visuellen Inspektion auch statistische Effektstärken und weiterführende Analysemodelle genutzt werden (Wilbert & Grünke, 2015).

Die Änderungssensitivität wurde bisher mit inferenzstatistischen Methoden analysiert. Darüber hinaus bieten sich Strukturgleichungsmodelle an, wie zum Beispiel das kombinierte Wachstums- und Veränderungsmodell von Förster und Kolleg*innen (2023). Hierbei beschreibt das latente Wachstumsmodell die Veränderungen, die mit einem Instrument der Lernverlaufsdiagnostik erfasst wurden, während das latente Veränderungsmodell den Unterschied zwischen Prä- und Posttest der statusdiagnostischen Testverfahren beschreibt. In einer gemeinsamen Modellrechnung wird dann die Kovarianz zwischen den beiden Modellen (Veränderung vs. Steigung) geschätzt, um die Änderungssensitivität des Verfahrens nachzuweisen.

Langfristiges Ziel der LVD Sprache ist es, durch eine Intervention erreichte Veränderungen zu evaluieren. Die Interpretation der unterschiedlichen Lernverläufe wurde für die Gesamtstichprobe getrennt nach Quartilen anhand von Ausgangsniveau und Zuwachsraten vorgenommen (Jungjohann et al., 2021). Hier konnten erste Orientierungswerte für einen stagnierenden oder ansteigenden Verlauf abgelesen werden. Diese konnten auch für eine erste Interpretation der Einzelfallstudien genutzt werden. Um zuverlässig individuelle Zuwachsraten zu beurteilen und um feststellen zu können, ob das Kind ausreichend von einer Interventionsmaßnahmen profitiert, sind darüber hinaus querschnittliche Normwerte nötig (Förster et al., 2017). Dies ist zudem notwendig, um realistische Erwartungen formulieren zu können.

Verfahren zur Lernverlaufsdiagnostik und zum Sätze Nachsprechen werden bereits als Screening eingesetzt (Grimm, 2015; Voß et al., 2017). In einem nächsten Schritt sollte geprüft werden, ob sich die LVD Sprache ebenfalls als Screening eignet. Wenn mit Hilfe der LVD Sprache ein Entwicklungsrisiko erkannt wird, können frühzeitig Interventionen eingeleitet werden, die wiederum das Entstehen von Lern- und Verhaltensstörungen verhindern können (Breitenbach, 2020; Huber & Grosche, 2012). Ein besonderer Vorteil der LVD Sprache wäre dann, dass sie in einer anschließenden Fördermaßnahme prozessbegleitend eingesetzt werden kann. Zu prüfen wäre auch, ob ggf. ein wiederholter Einsatz der LVD Sprache besser geeignet ist, um einen Förderbedarf festzustellen, denn für den Grundschulbereich konnten Hoffmann und Böhme (2017) zeigen, dass erst durch die mehrfache Erfassung der sprachlichen Kompetenzen eine korrekte Identifikation eines Förderbedarf erreicht werden konnte. Auch Adlof und Hogan (2019) fordern für den schulischen Bereich eine regelmäßige Messung der Sprachkompetenzen, um Kinder mit Sprachentwicklungsstörungen besser zu identifizieren.

Bisherige Befunde zu Verfahren zum Sätze Nachsprechen konnten zeigen, dass mehrsprachige Kinder gegenüber einsprachigen Kindern nicht benachteiligt sind (Hamann & Abed Ibrahim, 2017). Allerdings

ist der deutsche LITMUS-SRT etwas anders konstruiert und kann alternativ ausgewertet werden. Daher können die Ergebnisse nicht automatisch auf die LVD Sprache übertragen werden und sollte daher für die LVD Sprache differenziert überprüft werden.

Zentrales Anliegen der Arbeit ist es, Sprachförderung konsequent zu Ende zu denken, d.h. von der Auswahl einer - auch im Alltag - evidenzbasierten Sprachfördermaßnahme wie dem Dialogischen Lesen bis hin zur prozessbegleitenden Evaluation der Fördermaßnahmen mit Hilfe einer Lernverlaufsdiagnostik sprachlicher Fähigkeiten. Auf diese Weise könnten langfristig die Wirksamkeit von Sprachfördermaßnahmen und die sprachlichen Fähigkeiten von Kindern mit Förderbedarf verbessert werden, um deren Nachteile in Bezug auf schulischen Leistungen auszugleichen. Dieses Vorgehen wäre auch eine Möglichkeit, das „Wait-to-Fail“-Prinzip zu überwinden, um Kindern frühzeitig die notwendige Förderung und Therapie zukommen zu lassen bzw. die Maßnahmen unmittelbar an den Bedarf anpassen zu können.

9 Literaturverzeichnis

Abed Ibrahim, L., Hamann, C. & Öwerdieck, D. (2018). Identifying Specific Language Impairment (SLI) across Different Bilingual Populations: German Sentence Repetition Task (SRT). In A. B. Bertolini & M. J. Kaplan (Hrsg.), *Proceedings of the 42nd Annual Boston University Conference on Language Development* (S. 1–14). Cascadilla Press.

Adlof, S. M. & Hogan, T. P. (2019). If we don't look, we won't see: Measuring language development to inform literacy instruction. *Policy insights from the behavioral and brain sciences*, *6*(2), 210–217. https://doi.org/10.1177/2372732219839075

Allen, L., O'Conell, A. & Kiermer, V. (2019). How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learned Publishing*(32), 71–74. https://doi.org/10.1002/leap.1210

Althelmig, M. & Jungjohann, J. (2024, 8. Juni). *Sprachverständnis via Lernverlaufsdiagnostik in Unterricht und Therapie messen.* Frühjahrestagung der Arbeitsgruppe Empirische Sonderpädagogische Forschung (AESF), Zürich.

Berg, M. (2015). *MuSE-Pro – Überprüfung grammatischer Fähigkeiten bei 5- bis 8-jährigen Kindern: Materialbox mit Manual*. Reinhardt.

Besca, M. & Ennemoser, M. (2022). Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten und zur Evaluation sprachlicher Fördermaßnahmen? In M. Spreer, M. Wahl & H. Beek (Hrsg.), *Sprachheilpädagogik aktuell: Bd. 4, Sprachentwicklung im Dialog: Digitalität - Kommunikation - Partizipation* (S. 205–212). Schulz-Kirchner Verlag.

Besca, M., Krauss, T. & Ennemoser, M. (2023). Pilotierung einer Lernverlaufsdiagnostik bei Kindern mit Sprachentwicklungsstörung am SBBZ Sprache. *Forschung Sprache*(2), 16–26. https://www.forschung-sprache.eu/fileadmin/user_upload/Dateien/Heftausgaben/2023-2/Forschung_Sprache_2_2023_Besca_Krauss_Ennemoser2.pdf

Blumenthal, S. (2022). Lernverlaufsdiagnostik. In M. Gebhardt, D. Scheer & M. Schurig (Hrsg.), *Handbuch der sonderpädagogischen Diagnostik. Grundlagen und Konzepte der Statusdiagnostik, Prozessdiagnostik und Förderplanung. Version 1.0* (S. 633–648). Universität Regensburg.

Blumenthal, S., Gebhardt, M., Förster, N. & Souvignier, E. (2022). Internetplattformen zur Diagnostik von Lernverläufen von Schülerinnen und Schülern in Deutschland – Ein Vergleich der Plattformen Lernlinie, Levumi und quop. *Zeitschrift für Heilpädagogik*, *73*, 153–167.

Blumenthal, Y. & Mahlau, K. (2015). Effektiv fördern – Wie wähle ich aus? Ein Plädoyer für die Evidenzbasierte Praxis in der schulischen Sonderpädagogik. *Zeitschrift für Heilpädagogik*, *66*, 408–421.

Bockmann, A.-K., Sachse, S. & Buschmann, A. (2020). Sprachentwicklung im Überblick. In S. Sachse, A.-K. Bockmann & A. Buschmann (Hrsg.), *Lehrbuch. Sprachentwicklung. Entwicklung - Diagnostik - Förderung im Kleinkind- und Vorschulalter: [Language development. Development - diagnostics - intervention for toddlers and pre-school children]* (S. 3–44). Springer.

Böse, J., Busch, J., Leyendecker, B. & Scherger, A.-L. (2023). Supporting second language acquisition of bilingual preschool children through professionalization of caregivers in specialized preschool programs. *Frontiers in psychology*, *14*, 1149447. https://doi.org/10.3389/fpsyg.2023.1149447

Breitenbach, E. (2020). *Diagnostik: Eine Einführung*. *Lehrbuch: Band 5*. Springer VS. https://doi.org/10.1007/978-3-658-25150-5

Bundesministerium für Bildung und Forschung. (2024). *Vereinbarung zwischen Bund und Ländern zur Umsetzung des Startchancen-Programms für die Jahre 2024 bis 2034*. https://www.bmbf.de/SharedDocs/Downloads/de/2024/blv-startchancen.pdf?__blob=publicationFile&v=4

Burns, M. K., Appleton, J. J. & Stehouwer, J. D. (2005). Meta-Analytic Review of Responsiveness-To- Intervention Research: Examining Field-Based and Research-Implemented Models. *Journal of Psychoeducational Assessment*, *23*(4), 381–394. https://doi.org/10.1177/073428290502300406

Buschmann, A. & Jooss, B. (2011). Alltagsintegrierte Sprachförderung in der Kinderkrippe: Effektivität eines sprachbasierten Interaktionstrainings für pädagogisches Fachpersonal. *Verhaltenstherapie & psychosoziale Praxis*, *43*(2), 303–312.

Chilla, S. (2020). Mehrsprachige Entwicklung. In S. Sachse, A.-K. Bockmann & A. Buschmann (Hrsg.), *Lehrbuch. Sprachentwicklung: Entwicklung – Diagnostik – Förderung im Kleinkind- und Vorschulalter* (S. 109–130). Springer.

Collier, V. P. & Thomas, W. P. (2017). Validating the Power of Bilingual Schooling: Thirty-Two Years of Large-Scale, Longitudinal Research. *Annual Review of Applied Linguistics*, *37*, 203–217. https://doi.org/10.1017/S0267190517000034

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N. & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International journal of language & communication disorders*, *53*(2), 237–255. https://doi.org/10.1111/1460-6984.12338

Deno, S. L. (2003). Developments in Curriculum-Based Measurement. *The Journal of Special Education*, *37*(3), 184–192. https://doi.org/10.1177/00224669030370030801

Dickel-Lehnigk, M., Hohmann, E. & Ennemoser, M. (2022, 11. März). *Sprachförderung mit dem Sprachförderkonzept Dialogisches Lesen in Kindertagesstätten: Wirksamkeit und Sicherung der Durchführungsqualität durch Analyse der Strategieanwendung von Förderkräften*. Konferenzbeitrag. 9. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Bamberg/ Online.

Egert, F., Eckhardt, A. G. & Fukkink, R. G. (2017). Zentrale Wirkmechanismen von Weiterbildungen zur Qualitätssteigerung in Kindertageseinrichtungen: Ein narratives Review. *Frühe Bildung*, *6*(2), 55–66. https://doi.org/10.1026/2191-9186/a000309

Egert, F., Galuschka, K., Groth, K., Hasselhorn, M. & Sachse, S. (2020). Evidenzbasierung vorschulischer sprachlicher Bildung und Förderung: Was man darunter versteht und bisher darüber weiß. In K. Blatter, K. Groth & M. Hasselhorn (Hrsg.), *Evidenzbasierte Überprüfung von Sprachförderkonzepten im Elementarbereich* (S. 3–27). Springer Fachmedien.

Egert, F. & Hopf, M. (2016). Zur Wirksamkeit von Sprachförderung in Kindertageseinrichtungen in Deutschland. *Kindheit und Entwicklung*, *25*(3), 153–163. https://doi.org/10.1026/0942-5403/a000199

Ehlich, K., Bredel, U. & Reich, H. H. (2008). Referenzrahmen zur altersspezifischen Sprachaneignung: Bildungsforschung. 29/I. *Bundesministerium für Bildung und Forschung (BMBF)*.

Ennemoser, M. (2006). Evaluations- und Implementationsforschung. In N. Groeben & B. Hurrelmann (Hrsg.), *Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik: ein Weiterbildungsprogramm* (S. 513–528). Juventa.

Ennemoser, M. & Hartung, N. (2017). Wirksamkeit verschiedener Sprachfördermaßnahmen bei Risikokindern im Vorschulalter. *Unterrichtswissenschaft*, *45*(3), 198–219.

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. DOI 10.1024/1010-0652/a000109

Ennemoser, M., Lehnigk, M., Hohmann, E. & Pepouna, S. (2015). Wirksamkeit eines Coachings für pädagogische Fachkräfte zur Optimierung der Förderpotenziale des Dialogischen Lesens. In A. Redder, J. Naumann & R. Tracey (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung - Ergebnisse* (S. 137–153). Waxmann.

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. https://doi.org/10.1026/0049-8637/a000057

Esser, G. & Wyschkon, A. (2010). *P-ITPA: German Version von the Illinois Test for Psycholinguistic Abilites, 3rd version*. Hogrefe.

Fitton, L., McIlraith, A. L. & Wood, C. L. (2018). Shared Book Reading Interventions With English Learners: A Meta-Analysis. *Review of Educational Research*, *88*(5), 712–751. https://doi.org/10.3102/0034654318790909

Flack, Z. M., Field, A. P. & Horst, J. S. (2018). The effects of shared storybook reading on word learning: A meta-analysis. *Developmental Psychology*, *54*(7), 1334–1346. https://doi.org/10.1037/dev0000512

Förster, N., Erichsen, M. & Forthmann, B. (2023). Measuring Reading Progress in Second Grade. *European Journal of Psychological Assessment*, *39*(1), 5–13. https://doi.org/10.1027/1015-5759/a000688

Förster, N., Kuhn, J.-T. & Souvignier, E. (2017). Normierung von Verfahren zur Lernverlaufsdiagnostik. *Empirische Sonderpädagogik*, 116–122.

Frey, N., Heeg, K., Eikerling, M. R., Bloder. Theresa Sophie, Starke, A., Günther, T. & Lüke, C. (2024). Normdaten des Mottier-Tests für ein- und mehrsprachige Kinder im Alter von drei bis vier Jahren. *Logos*, *32*(3), 164–173. https://doi.org/10.7345/prolog-2403164

Fuchs, D., Fuchs, L. S. & Compton, D. L. (2004). Identifying Reading Disabilities by Responsiveness-to-Instruction: Specifying Measures and Criteria. *Learning Disability Quarterly*, *27*(4), 216–227. https://doi.org/10.2307/1593674

Fuchs, L. S. (2004). The Past, Present, and Future of Curriculum-Based Measurement Research. *School Psychology Review*, *33*(2), 188–192. https://doi.org/10.1080/02796015.2004.12086241

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L. & Germann, G. (1993). Formative Evaluation of Academic Progress: How Much Growth Can We Expect? *School Psychology Review*, *22*(1), 27–48. https://doi.org/10.1080/02796015.1993.12085636

Gathercole, S. E., Hitch, G. J., Service, E. & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, *33*(6), 966–979. https://doi.org/10.1037//0012-1649.33.6.966

Gebhardt, M., Jungjohann, J. & Schurig, M. (2021). *Lernverlaufsdiagnostik im förderorientierten Unterricht. Testkonstruktionen, Instrumente, Praxis: [Learning progression assessment in support-orientated teaching. Test constructions, instruments, practice]*. Reinhardt.

Geyer, S. & Müller, A. (2021). Wie adaptiv ist Sprachförderung für Kinder mit Deutsch als Zweitsprache? *Zeitschrift für Angewandte Linguistik*, *2021*(74), 31–60. https://doi.org/10.1515/zfal-2021-2052

Glück, C. & Spreer, M [M.] (2015). Zur Bildungsrelevanz semantisch-lexikalischer Störungen. *Sprache · Stimme · Gehör*, *39*(02), 81–85. https://doi.org/10.1055/s-0035-1549914

Gogolin, I., Neumann, U. & Roth, H.-J. (2003). *Förderung von Kindern und Jugendlichen mit Migrationshintergrund*. *Materialien zur Bildungsplanung und zur Forschungsförderung: Bd. 107*. BLK. https://doi.org/10.25656/01:335

Goldammer, A. von, Mähler, C., Bockmann, A.-K. & Hasselhorn, M. (2010). Vorhersage früher Schriftsprachleistungen aus vorschulischen Kompetenzen der Sprache und der phonologischen Informationsverarbeitung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *42*(1), 48–56. https://doi.org/10.1026/0049-8637/a000005

Goldammer, A. von, Mähler, C. & Hasselhorn, M. (2011). Determinanten von Satzgedächtnis-Leistungen bei deutsch- und mehrsprachigen Vorschulkindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *43*(1), 1–15. https://doi.org/10.1026/0049-8637/a000028

Grimm, H. (2015). *SETK 3-5: Sprachentwicklungstest für drei- bis fünfjährige Kinder [Language development test for three to five-year-old children]* (3., überarbeitete und neu normierte Auflage). Hogrefe.

Grosche, M. (2011). Effekte einer direkt-instruktiven Förderung der Lesegenauigkeit. *Empirische Sonderpädagogik*, *3.* https://doi.org/10.25656/01:9321

Grünke, M. (2006). Zur Effektivität von Fördermethoden bei Kindern und Jugendlichen mit Lernstörungen. *Kindheit und Entwicklung*, *15*(4), 239–254. https://doi.org/10.1026/0942-5403.15.4.239

Hamann, C. & Abed Ibrahim, L. (2017). Methods for Identifying Specific Language Impairment in Bilingual Populations in Germany. *Frontiers in Communication*, *2*, Artikel 16. https://doi.org/10.3389/fcomm.2017.00016

Hartke, B., Blumenthal, S., Blumenthal, Y. & Mahlau, K. (2022). Zur Wirksamkeit des Rügener Inklusionsmodells (RIM) – der Präventiven und Integrativen Schule auf Rügen (PISaR) nach neun Schulbesuchsjahren. *Zeitschrift für Heilpädagogik*, *73*, 352–367.

Hasselhorn, M., Schneider, W. & Trautwein, U. (Hrsg.). (2014). *Lernverlaufsdiagnostik* (Bd. 12). Hogrefe.

Hattie, J. (1985). Methodology Review: Assessing Unidimensionality of Tests and Items. *Applied Psychological Measurement*, *9*(2), 139–164. https://doi.org/10.1177/014662168500900204

Heeg, K., Eikerling, M. R., Bloder. Theresa Sophie, Frey, N., Starke, A. & Lüke, C. (2024). Einflussfaktoren auf die Nachsprechleistung von Pseudowörtern: Von der Relevanz einer einheitlichen Durchführung des Mottier-Tests [Factors influencing the performance in nonword repetition tasks – the relevance of a standardized administration of the Mottier-Test]. *Logos*, *32*(2), 94–102.

Heppt, B., Olczyk, M. & Volodina, A. (2022). Number of books at home as an indicator of socioeconomic status: Examining its extensions and their incremental validity for academic achievement. *Social Psychology of Education*, *25*(4), 903–928. https://doi.org/10.1007/s11218-022-09704-8

Hoffmann, L. & Böhme, K. (2017). Wird sprachlicher Förderbedarf in der Grundschule sicher erkannt? Zur Klassifikationsgüte von diagnostischen Entscheidungen. *Zeitschrift für Pädagogische Psychologie*, *31*(2), 137–147. https://doi.org/10.1024/1010-0652/a000203

Huber, C. & Grosche, M. (2012). Das response-to-intervention-Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderppädagogik. *Zeitschrift für Heilpädagogik*, *8*, 312–322.

Jain, A. & Spieß, R. (2012). Versuchspläne der experimentellen Einzelfallforschung. *Empirische Sonderpädagogik*, *4.* https://doi.org/10.25656/01:9300 (Empirische Sonderpädagogik 4 (2012) 3/4, S. 211-245).

Jungjohann, J., Schurig, M. & Gebhardt, M. (2021). Pilotierung von Leseflüssigkeits- und Leseverständnistests zur Entwicklung von Instrumenten der Lernverlaufsdiagnostik. Ergebnisse einer Längsschnittstudie in der 3ten und 4ten Jahrgangsstufe: [Piloting Reading Fluency and Reading Comprehension Tests for the Development of Instruments for Curriculum-Based Measurement. Results of a Longitudinal Study in 3rd and 4th Grade]. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete.* Vorab-Onlinepublikation. https://doi.org/10.2378/vhn2021.art12d

Kaltenbacher, E. (2011). Zur Problematik der Evaluation von Sprachfördermaßnahmen. In N. Hahn & T. D. Roelcke (Hrsg.), *Materialien Deutsch als Fremdsprache: Bd. 85, Grenzen überwinden mit Deutsch: 37. Jahrestagung des Fachverbandes Deutsch als Fremdsprache an der Pädagogischen Hochschule Freiburg/Br. 2010* (S. 163–177). Universitätsverlag.

Kammermeyer, G., Leber, A., Metz, A., Roux, S., Biskup-Ackermann, B. & Fondel, E. (2019). Empirische Arbeit: Langfristige Wirkungen des Fortbildungsansatzes „Mit Kindern im Gespräch“ zur Sprachförderung in Kindertagesstätten. *Psychologie in Erziehung und Unterricht*, *66*(4), 285–302. https://doi.org/10.2378/peu2019.art20d

Kauschke, C. (2023). Der typische Spracherwerb im Deutschen - Verläufe und Erklärungsansätze. In A. Mayer & T. Ulrich (Hrsg.), *Studienbuch Sprachheilpädagogik* (1. Auflage, S. 135–166). W. Kohlhammer GmbH.

Kauschke, C., Dörfler, T., Sachse, S. & Siegmüller, J. (2023). *Patholinguistische Diagnostik bei Sprachentwicklungsstörungen: + PDSS online* (3. Auflage). Elsevier Urban & Fischer.

Kempert, S., Edele, A., Rauch, D., Wolf, K. M., Paetsch, J., Darsow, A., Maluch, J. & Stanat, P. (2016). Die Rolle der Sprache für zuwanderungsbezogene Ungleichheiten im Bildungserfolg. In C.

Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 157–241). Springer VS.

Klauer, K. J. (2006). Erfassung des Lernfortschritts durch curriculumbasierte Messung. *Heilpädagogische Forschung*(1), 16–26.

Klauer, K. J. (2014). Formative Leistungsdiagnostik: Historischer Hintergrund und Weiterentwicklung zur Lernverlaufsdiagnostik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik* (S. 1–17). Hogrefe.

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S.-A. H., Gustafsson, J.-E. & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental science*, *18*(1), 146–154. https://doi.org/10.1111/desc.12202

Kohler, J., Starke, A. & Kohmäscher, A. (2021). Einzelfallorientierung in der Evidenzbasierung der Sprachtherapie. *Forschung Sprache*(2), 80–99.

Komeili, M. & Marshall, C. R. (2013). Sentence repetition as a measure of morphosyntax in monolingual and bilingual children. *Clinical linguistics & phonetics*, *27*(2), 152–162. https://doi.org/10.3109/02699206.2012.751625

Krajewski, K. (2018). *MBK 0. Test mathematischer Basiskompetenzen im Kindergartenalter: [MBK 0, Test of basic mathematical competences]*. Hogrefe.

Kulawiak, P. R. & Wilbert, J. Komplementäre Studiendesigns zur Evidenzbasierung in der Bildungswissenschaft, 17–31. https://doi.org/10.1007/978-3-662-55377-0_2

Leclercq, A.-L., Quémart, P., Magis, D. & Maillart, C. (2014). The sentence repetition task: A powerful diagnostic tool for French children with specific language impairment. *Research in developmental disabilities*, *35*(12), 3423–3430. https://doi.org/10.1016/j.ridd.2014.08.026

Lewalter, D., Diedrich, J., Goldhammer, F., Köller, O. & Reiss, K. (Hrsg.). (2023). *PISA 2022: Analyse der Bildungsergebnisse in Deutschland*. Waxmann. https://elibrary.utb.de/doi/book/10.31244/9783830998488

Lipowsky, F. (2010). Lernen im Beruf: Empirische Befunde zur Wirksamkeit von Lehrerfortbildung. In F. H. Müller, A. Eichenberger, M. Lüders & J. Mayr (Hrsg.), *Lehrerinnen und Lehrer lernen: Konzepte und Befunden zur Lehrerfortbildung* (S. 51–70). Waxmann.

Lohaus, A. & Elben, C. E. (2000). *Marburger Sprachverständnistest für Kinder (MSVK)*. Hogrefe.

Lüke, C. (2011). Sprachdiagnostik bei mehrsprachigen Schulkindern: Anwendbarkeit des P-ITPA und des SET 5-10 zur Erfassung der deutschsprachigen Kompetenzen. *Logos. Die Fachzeitschrift für Logopädie und Sprachtherapie*, *19*(3), 164-172.

Mahlau, K., Blumenthal, Y., Diehl, Kirsten, Schöning, Anne, Sikora, S., Voß, S. & Hartke, B. (2014). Das Rügener Inklusionsmodell (RIM) – RTI in der Praxis: [The Rügen Inclusion Model (RIM) - RTI in practice]. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik* (S. 101–126). Hogrefe.

Marinis, T. & Armon-Lotem, S. (2015). Sentence Repetition. In S. Armon-Lotem, J. de Jong & N. Meir (Hrsg.), *Communication Disorders Across Languages. Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment* (S. 95–122). Multilingual Matters. https://doi.org/10.21832/9781783093137-007

Marulis, L. M. & Neuman, S. B. (2010). The Effects of Vocabulary Intervention on Young Children's Word Learning. *Review of Educational Research*, *80*(3), 300–335. https://doi.org/10.3102/0034654310377087

Ministerium für Kultus, Jugend und Sport Baden-Württemberg. (2024, 12. September). *Sprachförderung „SprachFit" – Auf den Anfang kommt es an*. https://km.baden-wuerttemberg.de/de/schule/grundschule/sprachfit

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. https://doi.org/10.3102/0034654309332561

Motsch, H.-J. (2017). *Kontextoptimierung: Evidenzbasierte Intervention bei grammatischen Störungen in Therapie und Unterricht: mit CD-ROM mit 370 Übungen* (4., völlig überarbeitete Auflage, Online-Ausgabe). Ernst Reinhardt Verlag. https://elibrary.utb.de/doi/book/10.2378/9783497604081

Motsch, H.-J. & Rietz, C. (2019). *ESGRAF 4-8* (2., aktualisierte Auflage). Ernst Reinhardt Verlag.

Mottier, G. (1951). Über Untersuchungen der Sprache lesegestörter Kinder. *Folia Phoniatrica et Logopaedica*, *3*(3), 170–177. https://doi.org/10.1159/000262507

Mühling, A., Gebhardt, M. & Diehl, K. (2017). Formative Diagnostik durch die Onlineplattform LEVUMI. *Informatik-Spektrum*, *40*(6), 556–561. https://doi.org/10.1007/s00287-017-1069-7

Neuman, S. B. & Kaefer, T. (2018). Developing low-income children's vocabulary and content knowledge through a shared book reading program. *Contemporary Educational Psychology*, *52*, 15–24. https://doi.org/10.1016/j.cedpsych.2017.12.001

Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F. & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, *28*, 100290. https://doi.org/10.1016/j.edurev.2019.100290

Petermann, F. (2018). *SET 5-10, Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren: [SET 5-10, Status diagnostic test of language skills for children aged between 5 and 10 years]* (3., aktualisierte und teilweise neu normierte Auflage). Hogrefe.

Petermann, F. & Daseking, M. (Hrsg.). (2018). *Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition: [German Version according to D. Wechsler].*

Pillinger, C. & Vardy, E. J. (2022). The story so far: A systematic review of the dialogic reading literature. *Journal of Research in Reading*, *45*(4), 533–548. https://doi.org/10.1111/1467-9817.12407

Polišenská, K., Chiat, S. & Roy, P. (2015). Sentence repetition: what does the task measure? *International journal of language & communication disorders*, *50*(1), 106–118. https://doi.org/10.1111/1460-6984.12126

Pospeschill, M. (2022). *Testtheorie, Testkonstruktion, Testevaluation* (2., aktualisierte Auflage). *UTB: Bd. 3431*. Reinhardt.

Rauch, D. & Hartig, J. (2018). Warum sollten Sprachförderkonzepte überprüft werden? Evidenzbasierung als Prinzip der Qualitätssicherung. In C. Titz, S. Weber, A. Ropeter, S. Geyer & M. Hasselhorn (Hrsg.), *Bildung durch Sprache und Schrift: Bd. 2. Konzepte zur Sprach- und Schriftsprachförderung umsetzen und überprüfen* (S. 65–77). Kohlhammer.

Risse, T. & Kiese-Himmel, C. (2009). Der Mottier-Test: Teststatistische Überprüfung an 4- bis 6-jährigen Kindern [The Mottier test: psychometric evaluation using scores from children age 4-6 years]. *HNO*, *57*(5), 523–528. https://doi.org/10.1007/s00106-008-1708-z

Scheeff, N. (2022). *Lernverlaufsdiagnostik von sprachlich geförderten Kindern einer Kindertagesstätte anhand des neu entwickelten Sprachtests „Sätze-Nachsprechen"* [Unveröffentlichte Masterarbeit]. Pädagogische Hochschule, Ludwigsburg.

Schneider, W. (2018). Nützen Sprachförderprogramme im Kindergarten, und wenn ja, unter welcher Bedingung? *Zeitschrift für Pädagogische Psychologie*, *32*(1-2), 53–74. https://doi.org/10.1024/1010-0652/a000213

Schöler, H. (1999). *IDIS - Inventar diagnostischer Informationen bei Sprachentwicklungsauffälligkeiten*. *Edition S*. Ed. S.

Schöler, H. & Brunner, M. (2008). *HASE: Heidelberger Auditives Screening in der Einschulungsuntersuchung* (2., überarbeitete und erweiterte Auflage). Westra.

Schöler, H., Fromm, W., Schakib-Ekbatan, K. & Spohn, B. (1997). *Nachsprechen: Sein Stellenwert bei der Diagnostik von Sprachentwicklungsstörungen (Arbeitsberichte aus dem Forschungsprojekt "Differentialdiagnostik", Nr. 2)*. Pädagogische Hochschule Heidelberg.

Schöler, H. & Scheib, K. (2004). Desiderate und Thesen zur Diagnostik bei Sprachentwicklungsstörungen. *Sprache · Stimme · Gehör*, *28*(1), 37–41. https://doi.org/10.1055/s-2004-815482

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A. & Kliegl, R. (2015). childLex: a lexical database of German read by children. *Behavior Research Methods*, *47*(4), 1085–1094. https://doi.org/10.3758/s13428-014-0528-1

Souvignier, E. (2018). Computerbasierte Lernverlaufsdiagnostik. *Lernen und Lernstörungen*, *7*(4), 219–223. https://doi.org/10.1024/2235-0977/a000240

Souvignier, E., Förster, N. & Salaschek, M. (2014). quop: Ein Ansatz internetbasierter Lernverlaufsdiagnostik mit Testkonzepten für Lesen und Mathematik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik* (S. 239–256). Hogrefe.

Spreer, M [Markus] (Hrsg.). (2018). *Diagnostik von Sprach- und Kommunikationsstörungen im Kindesalter: [Diagnosics of speech and communication disorders in childhood]*. Reinhardt. https://elibrary.utb.de/doi/book/10.36198/9783838549460 https://doi.org/10.36198/9783838549460

Stark, R. & Klauer, K. J. (2019). Situiertes Lernen. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5. überarbeitete und erweiterte, S. 763–770). Beltz.

Strathmann, A. M. & Klauer, K. J. (2010). Lernverlaufsdiagnostik: Ein Ansatz zur längerfristigen Lernfortschrittsmessung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *42*(2), 111–122. https://doi.org/10.1026/0049-8637/a000011

Swanson, E., Wanzek, J., Petscher, Y., Vaughn, S., Heckert, J., Cavanaugh, C., Kraft, G. & Tackett, K. (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. https://doi.org/10.1177/0022219410378444

Theodorou, E., Kambanaros, M. & Grohmann, K. K. (2017). Sentence Repetition as a Tool for Screening Morphosyntactic Abilities of Bilectal Children with SLI. *Frontiers in psychology*, *8*, 2104. https://doi.org/10.3389/fpsyg.2017.02104

Titz, C. & Hasselhorn, M. (2017). Sprachförderliche Maßnahmen im Elementarbereich: Ein erfolgsversprechender Weg zur Prävention von Bildungsmisserfolg. In M. Becker-Mrotzek & H.-J. Roth (Hrsg.), *Sprachliche Bildung-Grundlagen und Handlungsfelder* (S. 287–297). Waxmann.

Towson, J. A., Fettig, A., Fleury, V. P. & Abarca, D. L. (2017). Dialogic Reading in Early Childhood Settings: A Summary of the Evidence Base. *Topics in Early Childhood Special Education*, *37*(3), 132–146. https://doi.org/10.1177/0271121417724875

Tuller, L., Hamann, C., Chilla, S., Ferré, S., Morin, E., Prevost, P., Dos Santos, C., Abed Ibrahim, L. & Zebib, R. (2018). Identifying language impairment in bilingual children in France and in Germany. *International journal of language & communication disorders*, *53*(4), 888–904. https://doi.org/10.1111/1460-6984.12397

Ulrich, T. (2017). *Grammatikerwerb und grammatische Störungen im Kindesalter. Ergebnisse des Forschungsprojekts GED 4-9 und ihre Implikationen für sprachdiagnostische und -therapeutische Methoden: [Grammar learning and grammatical disorders in childhood. Results of the research project GED 4-9]* [Habilitationsschrift]. Universität zu Köln. https://kups.ub.uni-koeln.de/9011/

Voltmer, K., Hormann, O., Pietsch, M., Maehler, C. & Salisch, M. von (2021). Teaching the Teachers About Language Support Strategies: Effects on Young Children's Language Development. *Frontiers in psychology*, *12*, 660750. https://doi.org/10.3389/fpsyg.2021.660750

Voß, S., Sikora, S. & Hartke, B. (2017). Lernverlaufsdiagnostik als zentrales Element der Prävention von Rechenschwierigkeiten. In A. Fritz, S. Schmidt & G. Ricken (Hrsg.), *Handbuch Rechenschwäche: Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie* (3., vollständig überarbeitete Aufl., S. 339–355). Beltz.

Walter, J. (2009). *Lernfortschrittsdiagnostik Lesen (LDL): Ein curriculumbasiertes Verfahren*. Hogrefe.

Walter, J. (2011). Die Entwicklung eines auch computerbasiert einsetzbaren Instruments zur formativen Messung von Lesekompetenz. *Heilpädagogische Forschung*, *37*(3), 106–126.

Whitehurst, G. J., Falco, F. L., Lonigan C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C. & Caulfield, M. (1988). Accelerating Language Development Through Picture Book Reading. *Developmental Psychology*, *24*(4), 552–559.

Wilbert, J. & Grünke, M. (2015). Kontrollierte Einzellfallforschung: Kap. 12. In K. Koch & S. Ellinger (Hrsg.), *Empirische Forschungsmethoden in der Heil- und Sonderpädagogik: Eine Einführung* (S. 100–105). Hogrefe.

Wilbert, J. & Linnemann, M. (2011). Kriterien zur Analyse eines Tests zur Lernverlaufsdiagnostik. *Empirische Sonderpädagogik*, *3*, 225–242. https://doi.org/10.25656/01:9325

Wild, N. & Fleck, C. (2013). Neunormierung des Mottier-Tests für 5- bis 17-jährige Kinder mit Deutsch als Erst- oder als Zweitsprache: [New standard values of the 'Mottier'-Test for children between 5–17 years of age with both German as a first or second language]. *Praxis Sprache*(3), 152–157.

Wimmer, E., Scherger, A.-L. & Ritterfeld, U. (2022). Bildungssprache im Grundschulalter: Relevanz von vorschulischem Wortschatz, Grammatik und Arbeitsgedächtnis für die weitere bildungssprachliche Entwicklung bei Kindern der Dortmunder Längsschnittstudie. *forum:logopädie*, *36*(5), 8–13.

Winkes, J. & Till, C. (2023). Dynamic Assessment und Lernverlaufsdiagnostik. Perspektiven in der sprachtherapeutischen Diagnostik. *Logos. Die Fachzeitschrift für Logopädie und Sprachtherapie*(4), 244–255. https://doi.org/10.7345/prolog-2304244

Ziethe, A., Eysholdt, U. & Doellinger, M. (2013). Sentence repetition and digit span: potential markers of bilingual children with suspected SLI? *Logopedics, phoniatrics, vocology*, *38*(1), 1–10. https://doi.org/10.3109/14015439.2012.664652

10 Anhang

10.1 Danksagung

Mein herzlicher Dank gilt allen, die mich auf dem Weg zu meiner Promotion unterstützt und begleitet haben.

Prof. Dr. Marco Ennemoser danke ich für seine durchgängige Unterstützung, seinen fachlich hochbereichernden Input und sein Vertrauen in meine Arbeit von Anfang an.

Prof. Dr. Margit Berg danke ich für die motivierende und interessierte Zweitbetreuung und den fachlichen Austausch.

Mein weiterer Dank gilt Dr. Matthias Schlagmüller für seine persönliche und fachliche Unterstützung, auf die ich mich jederzeit verlassen konnte. Meinen Kolleginnen im Förderschwerpunkt Sprache, im Projekt DiaLes und EnDiMath danke ich für den interessierten und unterstützenden Austausch zu unterschiedlichen Phasen meines Vorhabens. Ich danke der GISKID-Community und den Kolleginnen des GISKID-Stammtisches für den fachlichen Austausch und die persönliche Unterstützung.

Herzlich bedanken möchte ich mich auch bei allen Kindern, Eltern und pädagogischen Fachkräften, die an den verschiedenen Projekten teilgenommen haben. Ein großes Dankeschön geht an alle Studierenden, die als Hilfskräfte oder für ihre Abschlussarbeiten in den Projekten mitgearbeitet haben.

Ich danke meinem Mann Anghel, der mich in jeder Phase der Promotion unterstützt und mir an vielen Tagen und Wochenende den Rücken freigehalten hat. Meinen Eltern danke ich von Herzen, dass sie immer an mich geglaubt und mich unterstützt haben. Ich danke meiner Familie und meinen Freundinnen Katja und Sandra, die mir ein wichtiger Rückhalt und Ausgleich waren. Ich bin so froh euch alle haben!

10.2 Erklärung zu den Publikationen

Im Folgenden wird der Eigenanteil sowie der Beitrag der Mitautor*innen an den promotionsrelevanten Publikationen in Anlehnung an das Contributor Roles Taxonomy (CRediT, Allen et al., 2019) skizziert.

I. Ennemoser, M. & Besca, M. (2021). Präventive Mehrebenenförderung. *Pädagogik*, 73(4), 44–47.

MB hat die Vorlage verfasst und die Literatur dafür recherchiert und aufbereitet. ME hat den Originalentwurf geschrieben und final überarbeitet. MB und ME haben an dem Artikel mitgewirkt und die endgültige Fassung genehmigt.

II. Besca, M. & Ennemoser, M. (2022). Eignet sich das Sätze Nachsprechen zur Lernverlaufsdiagnostik sprachlicher Fähigkeiten und zur Evaluation sprachlicher Fördermaßnahmen? In M. Spreer, M. Wahl & H. Beek (Hrsg.), *Sprachheilpädagogik aktuell: Bd. 4, Sprachentwicklung im Dialog: Digitalität - Kommunikation - Partizipation* (S. 205–212). Schulz-Kirchner Verlag.

MB hat das Manuskript geschrieben und überarbeitet, die Datenerhebung geplant, geleitet bzw. selbst Daten erhoben und die Analysen durchgeführt. ME hat das übergeordnete Forschungsziel formuliert und die finanzielle Unterstützung für die Umsetzung der Projektidee gewährleistet. MB und ME haben das Manuskript Korrektur gelesen. MB und ME haben an dem Artikel mitgewirkt und die endgültige Fassung genehmigt.

III. Besca, M., Krauss, T. & Ennemoser, M. (2023). Pilotierung einer Lernverlaufsdiagnostik bei Kindern mit Sprachentwicklungsstörung am SBBZ Sprache. *Forschung Sprache* (2), 16–26. https://www.forschung-sprache.eu/fileadmin/user_upload/Dateien/Heftausgaben/2023-2/Forschung_Sprache_2_2023_Besca_Krauss_Ennemoser2.pdf

MB und TK haben den Erstentwurf gemeinsam geschrieben (TK: 1.2, 3.2, 3.3, 4.2, 4.3, 5.2, 5.3, MB: Zusammenfassung/ Abstract, 1.1, 1.3., 1.4., 2, 3.1, 4.1, 5.1, 5,4). MB hat die finale Version des Manuskripts geschrieben. TK hat die Daten der Stichprobe 2 (Längsschnittdaten am SBBZ Sprache) erhoben, ausgewertet und analysiert. MB hat die Datenerhebung der Stichprobe 1+3 geplant, geleitet und die Analysen durchgeführt. ME hat das übergeordnete Forschungsziel formuliert und die finanzielle Unterstützung für die Umsetzung der Projektidee gewährleistet. MB, TK und ME haben das Manuskript Korrektur gelesen. Alle Autor*innen haben an dem Artikel mitgewirkt und die endgültige Fassung genehmigt.

IV. Besca, M., Reiner, K. & Ennemoser, M. (2025): Evaluation des Dialogischen Lesens im Kita-Alltag. Ein Vergleich der Wirksamkeit in kontrollierten und realitätsnahen Studienbedingungen. *Logos 33 (1)*.

MB hat das Manuskript geschrieben und überarbeitet. KR und MB haben die Datenerhebung geplant, geleitet und teilweise Daten erhoben. MB hat die Auswertung der Daten geleitet und die Daten analysiert. ME hat das übergeordnete Forschungsziel formuliert und die Drittmittel eingeworben. MB und ME haben das Manuskript Korrektur gelesen. MB und ME haben an dem Artikel mitgewirkt und die endgültige Fassung genehmigt.

V. Besca, M. & Ennemoser, M. (eingereicht): Assessing Oral Language Progress in Children Aged 4–6. Evaluation of the Psychometric Properties.

MB hat das Manuskript geschrieben und überarbeitet, die Datenerhebung geplant, geleitet bzw. selbst Daten erhoben und die Analysen durchgeführt. ME hat das übergeordnete Forschungsziel formuliert und die finanzielle Unterstützung für die Umsetzung der Projektidee gewährleistet. MB und ME haben das Manuskript Korrektur gelesen. MB und ME haben an dem Artikel mitgewirkt und die endgültige Fassung genehmigt.

10.3 Eidesstaatliche Erklärung

Hiermit versichere ich, dass die vorgelegte Dissertation selbstständig angefertigt wurde, wörtliche wie inhaltliche Entlehnungen kenntlich gemacht wurden und außer den in der Dissertation genannten keine weiteren Hilfsmittel Verwendung fanden. Ebenfalls erkläre ich, dass die Dissertation weder im Ganzen noch in Teilen Gegenstand eines Promotionsverfahrens an einer anderen Hochschule oder Gegenstand einer anderen akademischen Prüfung oder einer Staatsprüfung war.

Ludwigsburg, 28.11.2024 Melanie Besca (im Original unterzeichnet)

Ort, Datum Unterschrift