CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Die Tübinger Theater-Arbeitsgemeinschaften: Ergebnisse aus zwei Jahren DaZ-Grammatikförderung für Grundschulkinder;The Tübingen theatre working groups
Authors	Vogel, Sophie Charlotte
Publication date	31/12/2023
Original Citation	Vogel, S. C. (2023) 'Die Tübinger Theater-Arbeitsgemeinschaften: Ergebnisse aus zwei Jahren DaZ-Grammatikförderung für Grundschulkinder', Scenario: A Journal for Performative Teaching, Learning, Research, 17(2), pp. 44-73. doi: https://doi.org/10.33178/scenario.17.2.3
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.33178/scenario.17.2.3
Rights	© 2023, the Author(s). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. - https://creativecommons.org/licenses/by-nc/4.0/
Download date	2026-11-14 15:41:10
Item downloaded from	https://hdl.handle.net/10468/15461



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Volume/Jahrgang 17, Issue/Ausgabe 2(2023)
https://doi.org/10.33178/scenario.17.2.3

Die Tübinger Theater-Arbeitsgemeinschaften

Ergebnisse aus zwei Jahren DaZ-Grammatikförderung für Grundschulkinder

Sophie Charlotte Vogel

Theaterferiencamps zur Zweitsprachenförderung für sprachförderbedürftige Grundschulkinder gab es ab der Jahrtausendwende in Deutschland mehrere. Einige davon wurden evaluiert: zum Beispiel die Camps in Duisburg-Essen, Bremen, Hamburg und Tübingen (Boorsma/Moraitis, 2007; Stanat et al., 2005; Kinze, 2012; Batzel, Bohl, Bryant, 2013). Gezeigt werden konnte in den Evaluationen, dass diese Camps in den fokussierten Fördergebieten sprachförderlich wirken. Ebenfalls einheitlich wird eine Nachhaltigkeitsproblematik festgestellt (Oelschläger, 2008, S. 114-15.; Rösch, 2007, S. 288; Bryant, 2012, S. 28). Das heißt, die gefundenen Effekte verblassen über die nachfolgenden Monate hinweg (Oelschläger, 2008, S. 114 f.; Rösch, 2007, S. 288; Bryant, 2012, S. 28). In Tübingen und andernorts etablierte man als Nachhaltigkeitsmaßnahme deshalb im Anschluss Theater-Arbeitsgemeinschaften (AGs), um die in der Feriensprachförderung gewonnen, positiven Effekte beizubehalten und auszubauen (Bryant & Rummel, 2015). Nun liegen zu diesem Tübinger Nachhaltigkeitsprojekt Ergebnisse vor. Zwei Jahre lang konnten 82 Kinder am Projekt unter Einsatz derselben sogenannten dramagrammatischen Methode wie im von ihnen besuchten Theatercamp teilnehmen, in drei sprachhomogenen Interventions- und drei Kontrollgruppen (Sprachanfänger:innen, Sprachfortgeschrittene, mündlich Versierte). Die Beschreibungen des Vorgehens in insgesamt 192 AG-Doppelstunden ergeben zusammen mit Fragebogendaten sowie 490 standardisierten Sprachstandsdiagnostiken in der Grammatik, Textproduktion und Lesefähigkeit ein umfassendes, auch deskriptiv und inferenzstatistisch ausgewertetes Datenmaterial, das in der das Projekt evaluierenden Dissertationsschrift aufbereitet wurde. In diesem Artikel sollen die Ergebnisse der Anfänger:innen skizziert werden, die in allen getesteten Bereichen von der sogenannten dramagrammatischen Sprachförderung profitierten. Die dramagrammatische Methode eignet sich demnach zur Sprachförderung besonders für Sprachanfänger:innen, wobei sich auch die höheren Sprachniveaus – durch die Diagnostiken bestätigt – in vielen Bereichen verbessern konnten (Vogel, in Vorb.).

1 Empirie und Praxis – Appell und Wirklichkeit

Das Spektrum performativer Einsatzmöglichkeiten für unterschiedliche inner- und außerschulische Lernkontexte wächst und ist gefragter denn je, wie die große Publikationszahl in den letzten Jahren rund um die Begriffe ästhetisches und performatives Lernen bestätigt (Bryant & Zepter, 2022, S. 11). Das ist erfreulich für all diejenigen in Wissenschaft und Praxis,

die sich seit geraumer Zeit der Aufgabe angenommen haben, performatives Lehren und Lernen für eine breite Masse zugänglich zu machen. Dem in den letzten drei Jahrzehnten wiederholten Appell nach empirischer Forschung im Bereich des performativen Lehrens und Lernens kann dank dem wachsenden Interesse und damit einhergehender Bereitschaft, in die künstlerische Orientierung von Lehre und Forschung zu investieren, immer mehr nachgegangen werden. Zudem steigt die Zahl an Veröffentlichungen empirischer Forschungsergebnisse aus der Fremdsprachendidaktik und angrenzenden Disziplinen (Delius, 2020, S. 16). Daneben stehen schließlich die Praxisbeiträge von Lehrkräften, die von positiven Veränderungen durch den Einsatz performativer Lehr- und Lernhandlungen berichten (Dragović, 2019b, S. 36; Domkowsky & Walter, 2012, S. 104). Als Forschungsfelder, die empirische Daten generieren, stechen vor allem folgende Bereiche heraus: Solche mit Fokus rund um die Fertigkeit *Sprechen* (Bora, 2017; Hensel, 2020; Delius, 2020; Piazzoli, 2011; Sirisrimangkorn & Suwanthep, 2013) sowie solche, die sich im pädagogischen, psychisch-kognitiven Bereich ansiedeln und Fragen zu beispielsweise Einstellungen, Motivation, Empathie, Vertrauen und Reflexionsfähigkeiten von Lehrenden und Lernenden untersuchen (Sirisrimangkorn & Suwanthep, 2013; Klempin 2019; Weyer, 2019) oder solche, die diese Gebiete miteinander verknüpfen (Piazzoli, 2011; Wirag & Surkamp, 2022). Ungeachtet des jeweiligen Forschungsfokus umfasst jedoch die Projektlaufzeit zumeist nur wenige Wochen oder Monate, sodass die Aussagekraft von Ergebnissen begrenzt und generelle Schlüsse über die Wirksamkeit performativer Arbeitsweisen kaum möglich erscheinen. Die große Ausnahme ist noch immer die inzwischen über zehn Jahre alte Forschungsarbeit des DICE-Konsortiums mit 12 teilnehmenden europäischen Ländern und beinahe 5000 Jugendlichen, in welcher Aussagen über Entwicklungen in den sogenannten Schlüsselkompetenzen getroffen wurden, die theaterspielende im Gegensatz zu nicht-theaterspielenden Mitschüler:innen erreichen können (Küppers, 2011). Sehr selten wurde darüber hinaus bisher die performative Grammatikvermittlung über einen längeren Zeitraum in den Blick genommen. Ausnahmen sind beispielsweise die Dissertationen von Dragović (2019a), die über ein Jahr an serbischen Primarschulen forschte (Dragović, 2019a, S. 1), und Schappert (2020), deren Intervention fünf Monate in Integrationskursen umfasste (Schappert, 2020, S. 162). Allerdings hebt letztere selbst in ihrer Schrift zur Dramagrammatik in Alphabetisierungskursen die ungünstigen Bedingungen ihrer Forschungssituation hervor. Auf Aspekte wie die kleinen Fallzahlen durch die begrenzte Kurskapazität und die typische Fluktuation in Hinblick auf die Bleibeproblematik in Gruppen mit Geflüchteten hatte sie keinen Einfluss (Schappert, 2020, S. 151). Gerade der Bereich der Grammatikvermittlung durch performative Techniken eignet sich jedoch für Wirksamkeitsuntersuchungen, denn hier lässt sich eine mögliche Veränderung, nämlich anhand des messbaren Zuwachses sprachlicher Kompetenzen durch das Nutzen von Prä-Post-Testungen, bestenfalls mittels standardisierter Sprachdiagnostiken, verhältnismäßig einfach

feststellen (Walter, 2013, S. 1). So fasst auch Bernstein in seinem Bericht über die Podiumsdiskussion zum performativen Lehren und Lernen auf den „Sommerlichen FaDaF-Literaturtagen 2021“ zusammen, dass

> weiterhin größer angelegte Studien interessant [wären,] mit Versuchs- und Kontrollgruppen zu Effekten performativen Lernens auf den Spracherwerb, auch hinsichtlich Prüfungssituationen im Bereich des Testens und Prüfens, auch hinsichtlich langfristiger Auswirkungen auf den Studienerfolg oder generell auf den Bereich des lebenslangen Lernens. (Bernstein, 2021, S. 122)

Die Grundlagenforschung für performative Lernsettings steckt noch immer in ihren Kinderschuhen. Dieser Umstand bietet allerdings weniger einen Anlass zur Kritik, sondern schafft vielmehr einen echten Anstoß für weitere Forschungsvorhaben.

Vor diesem Hintergrund soll es im vorliegenden Artikel um die Evaluation der sprachlichen Entwicklungen zweier Sprachanfänger:innengruppen der den Tübinger Theatercamps nachfolgenden Theater-AGs gehen. Wie auch in den beiden höheren Sprachniveaus (Sprachfortgeschrittene, mündlich Versierte) wurde eine der zwei Gruppen in einer dramagrammatischen Interventionsmaßnahme in Gestalt einer Theater-AG gefördert, die andere fungierte als Kontrollgruppe mit anderer Nachmittags-AG ohne sprachförderlichen Background. Die Maßnahme sollte Aufschluss darüber geben, inwiefern sich die für die Tübinger Theatercamps (2011 - 2017) eingesetzte dramagrammatische Methode auch zur langfristigen Sprachförderung von Grundschulkindern eignete, das heißt auch, ob sich durch das Theatercamp messbare sprachliche Erfolge durch das Beibehalten dieser Fördermethode in AGs längerfristig manifestieren ließen. Die Interventionsmaßnahme war über zwei Jahre angelegt und fand im Nachmittagsband des Ganztagsschulbetriebes (Klasse 1-4) statt. Es nahmen 82 Kinder mit Förderbedarf im Deutschen sowie oben genannten unterschiedlichen Sprachständen in jeweils drei Interventions- und drei Kontrollgruppen teil. Nach Ausschlüssen aus der Intervention wegen beispielsweise im Studiendesign festgelegter Parameter (z.B. fehlende Sprachförderbedürftigkeit, Vorhandensein nachgewiesener Sprachentwicklungsstörung der Kinder) gehörten noch 25 Schüler:innen dem Anfänger:innenniveau, 22 der fortgeschrittenen und 24 der mündlich versierten Gruppe an. Die AGs fanden an drei Grundschulen in Tübingen statt und konnten von Kindern der Schulen selbst sowie von Schüler:innen der Umgebung besucht werden.

2 Vom Theatercamp zum Nachhaltigkeitskonzept

Das Tübinger Theatercamp hatte zwei Zielgruppen: Zum einen war es gedacht als Lehr- und Lernraum für Studierende des Studiengangs *Deutsch als Zweitsprache: Sprachdiagnostik und Sprachförderung*, die jeweils zu dritt eine der sechs Kindergruppen im Theatercamp förderten

und begleiteten. Zum anderen sollten die 7- bis 12-jährigen Schüler:innen kindgerecht in den Baden-Württembergischen zwei Pfingstferienwochen ihrem Sprachstand entsprechend gefördert werden (Bryant, 2012). Da sich diese Kinder in ihrem Alltag selten in sprachlich ähnlichen *peer groups* aufhielten, hatten sie oftmals wenig Antrieb, sich in einer größeren Gruppe zu präsentieren und zu äußern (Bryant & Rummel, 2015). Die Sicherheit einer Gemeinschaft, die auf dem gleichen Stand ist, und die Anwesenheit Erwachsener, die sich Zeit für ihre Gedanken nehmen und ihr Formulieren verständnisvoll unterstützen, wurde als ein Grundpfeiler der Förderung festgelegt (Bryant & Zepter, 2022, S. 399ff.). Über Prätests konnte eine Einteilung in sprachhomogene Gruppen vorgenommen werden, die darüber hinaus absicherte, dass die Kinder sprachlich in der Förderung dort ansetzen konnten, wo sie sich gerade befanden, und somit einen entwicklungsproximal abgestimmten Input bekamen (Bryant, 2012). Das gemeinsame Ziel aller Gruppen und pädagogischer Kräfte war es, eine große, öffentliche, regionalmedial beachtete Theateraufführung am letzten Theatercamptag zu zeigen.[1] Die Ferientage vor der Bühnenpräsentation gliederten sich in eine den Vormittag umfassende dramagrammatische, dem Sprachstand entsprechende Förderung anhand eines Fünfphasenmodells (siehe Abschnitt 4) und eine implizite Sprachförderung beim Kulissenbau, Kostümerstellen, Tanz und Gesang am Nachmittag (siehe Abschnitt 4).

Eine erste Evaluation zu Konzept und Methode des Theatercamps von 2011 erschien 2013 (Batzel, Bohl & Bryant, 2013). Darin konnten positive Effekte auf unterschiedliche Bereiche (z.B. auf die geförderten grammatischen Gebiete, wie beispielsweise der Umgang mit dem Nebensatzspektrum, auf die mündliche Mitarbeit in der Schule nach dem Camp, auf eine positive Einstellung gegenüber dem Bühnenauftritt etc.) aufgezeigt werden (Batzel, Bohl & Bryant, 2013, S. 105-115). Jedoch ließen sich die gemessenen sprachlichen Erfolge nicht dauerhaft beibehalten. Nach drei Theatercamps in den Jahren 2011, 2013 und 2015 folgten deshalb ab 2015 die Theater-AGs als Nachhaltigkeitsprojekt mit demselben dramagrammatischen Vorgehen und aus dem Camp bekannten Personal. Aufbau und Ergebnisse aus dem Anfänger:innenniveau werden folgend skizziert.

3 Beschreibung und Merkmale der dramagrammatischen Intervention

Auf die Erkenntnisse aus der Evaluation des Tübinger Theatercamps 2011 konnte für die Anschlussförderung in AGs aufgebaut werden. Es wurde mit derselben dramagrammatischen Methode wie im Theatercamp und weiterhin mit sprachhomogenen Gruppen gearbeitet. Die Methode umfasst fünf Phasen impliziter und expliziter Theater-Spracharbeit sowie

[1] Das pädagogische Personal umfasste nicht nur die Sprachförderkräfte der Universität, sondern ebenfalls Freizeitpädagog:innen, Sozialpädagog:innen der Stadt Tübingen und Theaterschaffende des Landestheaters Württemberg-Hohenzollern Tübingen Reutlingen.

metasprachliche Reflexion. Für die AGs meldeten sich vornehmlich Kinder aus dem vorangegangenen Theatercamp an. Noch verbliebene Plätze konnten jedoch von Mitschüler:innen freiwillig eingenommen werden, die die Klassenlehrer:innen ebenfalls als sprachförderbedürftig einschätzten. Jede der drei Interventionsgruppen mit je unterschiedlichem Sprachniveau (Sprachanfänger:innen, Sprachfortgeschrittene und mündlich Versierte) war an einer anderen Grundschule angesiedelt, welche jeweils auch eine Kontrollgruppe stellte. Kinder, deren passende Sprachniveaugruppe an einer anderen Schule als der eigenen gefördert wurde, erhielten einen Transfer dorthin. Die Zuordnung zum Sprachniveau erfolgte anhand der aus dem Theatercamp vorliegenden sprachdiagnostischen Daten. Für Drittklässler:innen bestand die Möglichkeit, noch zwei Jahre nach dem Camp weiter mit der dramagrammatischen Methode gefördert zu werden. Für Viertklässler:innen galt dies für noch ein Jahr vor dem Wechsel auf die Sekundarschule. Die folgenden Merkmale waren im Vorfeld als relevant für die Formung der Theater-AGs festgelegt worden:[2]

- Die Kinder werden ihrem Sprachstand entsprechend weiterhin in homogenen Kleingruppen dramagrammatisch gefördert.
- Die dramagrammatische Sprachförderung muss sich in den Schulalltag einbinden lassen.
- Sie muss sich methodisch und personell am Tübinger Theatercamp orientieren.
- Ein Förderzeitraum von zwei Jahren muss gewährleistet sein.
- Die Förderung muss hinsichtlich der oftmals umfangreichen Förderbedarfe von Kindern mit DaZ mehr Förderinhalte als nur sprachliche bereithalten, sondern sich auch mit Inhalten aus dem primarstuflichen Sachfach verbinden lassen, welches die natur- und geisteswissenschaftlichen Fächer der Sekundarstufe I vorbereitet.

4 Beschreibung des dramagrammatischen Modells mit Beispielablauf

Evens Modell der Drama-Grammatik (2003), das mit expliziter Grammatikvermittlung und Sprachreflexion das dreistufige Modell von Schewe (1993) um zwei Komponenten erweitert, gab den Anstoß für ein dramagrammatisches Fünfphasenmodell für Grundschulkinder mit DaZ (Bryant, 2012). Im Gegensatz zu erwachsenen Fremdsprachenlernenden, auf die Evens Modell zugeschnitten war – und bei denen eine gewisse Freiwilligkeit zur Beschäftigung mit Sprache angenommen werden kann (Bryant, 2012, S. 39) –, musste das Phasenmodell für Grundschulkinder auch Schüler:innen motivieren, die teils mit auffälligen Verhalten, teils mit Abneigung gegenüber Sprach- und Schriftarbeit in die Theater-AG kamen (Bryant, 2012). Eine attraktive Gestaltung des Angebots erwies sich als maßgeblich, um die Aufmerksamkeit der

[2] Beschrieben sind die Merkmale ebenfalls in Vogel (in Vorb.).

Kinder nicht umgehend zu verlieren (Bryant, 2012). Das Tübinger Modell ähnelt dem von Even (2003), enthält jedoch eine durchgängigere Fokussierung auf eine Zielstruktur (Bryant, 2012, S. 40):

1. Aufwärmphase mit strukturiertem Input und Wiederholungsübungen, z.B. durch Kreis-, Rotations- oder Improvisationsspiele	*2. Motivation der Zielstruktur*	*3. Strukturvermittlung*	*4. Strukturanwendung*	*5. Sprachreflexion und Präsentation*
← gleiche Förderstruktur für circa vier bis neun Fördereinheiten →				
← gleiches Sachfachthema für circa fünf bis zehn Fördereinheiten →				

Tabelle 1: Übersicht über die fünf Phasen des dramagrammatischen Modells für Kinder mit DaZ in den Theater-AGs

Ausgewählt wurden verschiedene Techniken aus dem breiten Repertoire der Theater- und Dramamethoden je nach ihrer Nutzbarkeit zur Vermittlung der grammatischen Struktur. Das übergeordnete Sachfachthema rahmte gleichzeitig Theater- und Spracharbeit. Besonderer Wert wurde außerdem auf einen auf das Sprachniveau abgestimmten Einbezug von Operatoren gelegt. Diese handlungsinitiierenden Verben sind zum einen nicht leicht zu verstehen, zum anderen sind Kindern mit sprachlichem Förderbedarf oft das Vorgehen und die sprachlichen Muster, nach denen diese Operatoren verlangen, nicht geläufig (z.B. *benenne, beschreibe, vergleiche*; Bryant & Rummel, 2015, S. 19). Eine Übersicht über die Förderthemen bezogen auf die jeweiligen Sprachniveaus findet sich in Abbildung 2:

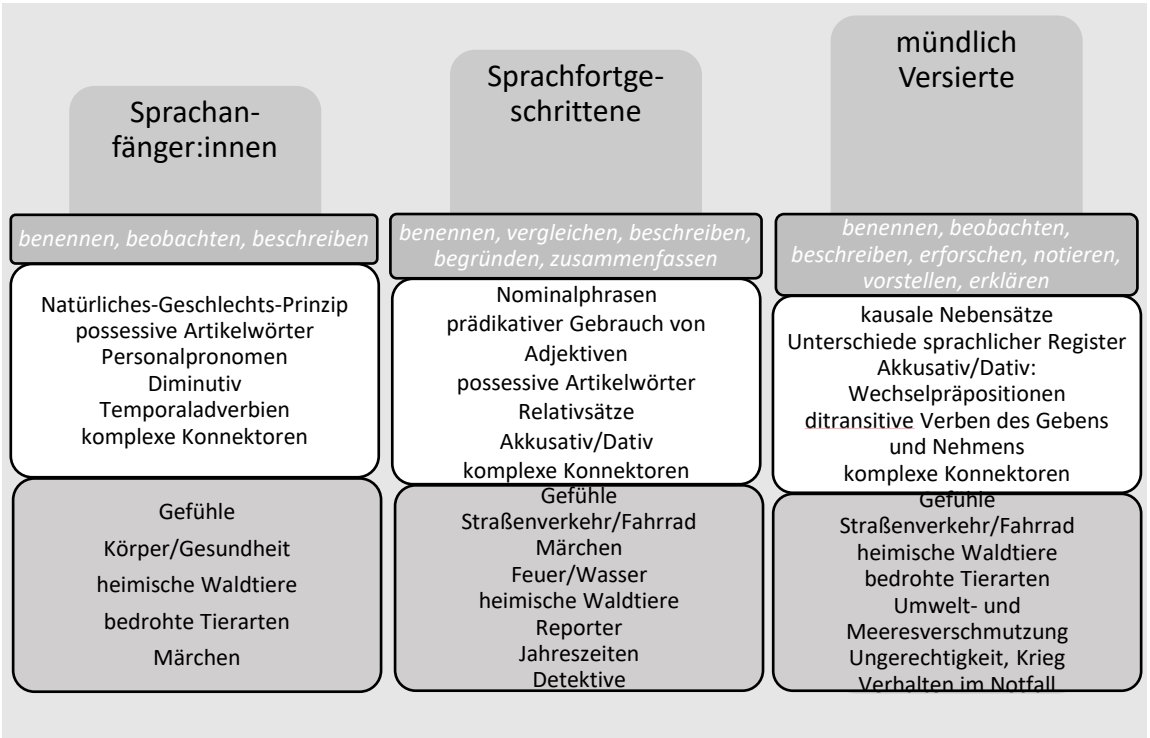


Abbildung 1: Übersicht über die Sprach- und Sachfachförderthemen sowie Operatoren unterteilt nach Sprachniveau

Im Folgenden soll beispielhaft eine Sprachfördereinheit aus dem sprachfortgeschrittenen Niveau der Theater-AGs einen Einblick in den konkreten Ablauf des dramagrammatischen Fünfphasenmodells geben. Dabei handelte es sich um die siebte Sitzung zum Thema Märchen[3] mit dem grammatischen Schwerpunkt Akkusativdeklination. Der Fokus lag auf Akkusativobjekten mit unbestimmtem Artikel im Vergleich zu Nominalphrasen im Nominativ. Innerhalb der Übungen kam es jedoch zu einer gewissen Varianz, sodass auch stellenweise Akkusativobjekte mit bestimmten und possessiven Artikelwörtern auftauchten. Sowohl Akkusativ- als auch Dativdeklination sind für Kinder eines fortgeschrittenen Sprachniveaus vor allem deshalb noch ein Förderthema, weil die Formen bei Maskulina im Akkusativ (*einen*/*den*) den Dativformen (*einem*/*dem*) sehr ähnlich sind und sich beim Hören nicht gut unterscheiden lassen. Jede der vorangegangenen Sitzungen behandelte ein anderes Märchen. In der siebten Sitzung kommen erstmals unterschiedliche Märchenfiguren in einem „Märchenfundbüro“ zusammen.

4.1 Aufwärmphase

Zu Beginn der Aufwärmphase finden sich alle im Kreis ein. In der Mitte liegt zu jeder Seite ein Raptext aus. Der Text wird zweimal zusammen rhythmisch gesprochen und dann in zwei Untergruppen unterteilt (Holl, 2011):

> A: Ich habe einen Helm (Taler, Korb/eine Krone, Kugel, Katze/ein Feuer, Käppchen).
>
> B: Ein/e/n was?
>
> A: Ein/e/n _____ .
>
> B: Ah, du hast ein/e/n _____ .
>
> A (emotional gesprochen, nicht gerappt): Ja, und der/die/das ist jetzt weg.

Im Anschluss bewegen sich die Kinder zu leiser, stimmungsvoller, instrumentaler Musik durch den Raum. Dort sind Bilder von Märchenfiguren als Ausdruck (Bild mit Beschriftung) an verschiedenen Stellen angebracht. Nun beginnt die Förderkraft, kleine Situationen zu schildern, welche die Kinder als Märchenfiguren erleben. Die Kinder spielen die Angaben der Förderkraft nach. Dazu wird die Musik noch etwas leiser gedreht:

> Du bist die Prinzessin. Du trägst ein Kleid. Ah, toll! Es steht dir gut. Wenn du gehst, fühlst du dich sehr schön. Du hast eine Krone. – Sehr schön! Du hast einen Gürtel mit Glitzersteinen – Wie edel! Oh, aber… oh nein! Du hast keine

[3] Die Sachfachthemen wurden vor der Förderung selbst von den Sachfachlehrkräften der Kinder vorgeschlagen.

> goldene Kugel. Du suchst deine Kugel. Eine Prinzessin braucht eine goldene Kugel.
>
> Du bist Rumpelstilzchen. Da ist ein Hut. Den setzt du auf. Du wolltest schon immer einen Hut. – Sehr schick! Im Königsschloss hat niemand so einen Hut. Ach, nein, ein Sturm! Der Sturm nimmt dir deinen Hut weg.
>
> Du bist der Wolf. Du bist im Haus der Großmutter. Alles ist jetzt deins: Das Bett und das weiße Nachthemd. Zieh es mal an! Jetzt hast du ein Nachthemd. Das sieht aber lustig aus. Du trägst jetzt ein Nachthemd und dazu ein Käppchen auf dem Kopf. Da kommt der Jäger. Schnell weg! Er kriegt dich nicht. Aber du verlierst dein Käppchen.

Zum Schluss sollen sich alle Kinder dem Bild derjenigen Märchenfigur zuordnen, die sie gern im Folgenden verkörpern möchten.

4.2 Motivation der Zielstruktur

Die Förderkraft berät sich mit den Schüler:innen. Sie hebt hervor, dass nun jeder Märchenfigur etwas fehlt und überlegt mit den Kindern, ob jemand eine Idee hat, wo man nach Verlorenem fragen könnte. Nach kurzem Abwarten erklärt sie das Konzept des Fundbüros. Dort gibt es Formulare, die ausgefüllt, und Mitarbeiter:innen, die richtig angesprochen werden müssen. Besonders als Märchenfigur gelte es, das Vertrauen des Mitarbeitenden zu gewinnen, da wegen einerseits unaufrichtiger Märchen- und Geschichtenkolleg:innen wie Pinocchio und andererseits des großen zauberhaften Wertes Märchengegenstände schwer zurückzubekommen seien. Der richtige Auftritt sei daher maßgeblich.

4.3 Strukturvermittlung

Die Förderkraft betont weiter, dass - damit das Zurückerhalten wirklich gelingt - nun zum einen ein passendes Formular ausgefüllt werden muss, und zum anderen eine Vorbereitung für das Gespräch mit dem Fundbüropersonal stattfinden muss. Mithilfe eines vorbereiteten Plakats sollten die Gegenstände, die fehlen, sowohl erinnert werden (und sollen in Lücken auf dem Plakat eingetragen werden), als auch vorbereitet werden, welche davon man vermisst.

Das ist der ein _____ Hut / Helm / etc.	Ich habe/vermisse den/einen _____ .
Das ist die/eine _____ Kugel / Krone / etc.	Ich habe/vermisse die/eine _____ .
Das ist das/ein Käppchen / Nachthemd / etc.	Ich habe/vermisse das/ein _____ .

Abbildung 2: Beispielplakat für die Strukturvermittlungsphase

Die Schüler:innen werden befragt, ob sie Unterschiede zwischen der linken Seite und der rechten Seite des Plakates bemerkt haben, und wenn ja, welche diese sind. Gemeinsam wird erörtert, woher die Veränderungen beim maskulinen Artikel kommen könnten, und ob es über *haben/vermissen* hinaus Verben gibt, bei denen den Kindern aufgefallen ist, dass sich der maskuline Artikel von *der/ein* zu *den/einen* verändert.

4.4 Strukturanwendung

Durch das Plakat, das gut sichtbar für alle im Raum aufgehängt wird, sind die Kinder nun in der Lage, eine kleine Szene im Fundbüro zu entwickeln. Dazu finden sie sich in Kleingruppen als die anfangs gewählten Märchenfiguren zusammen. Der Rap der ersten Phase soll nun kreativ bearbeitet werden. Jedes Kind erhält dazu einen dem Plakat ähnlichen Lückentext, worauf der zur Figur passende Gegenstand noch einmal notiert werden muss. Auf der zweiten Seite befindet sich ein kleiner Ablaufplan, auf dem der eigene Sprechtext und Regiestichpunkte notiert werden sollen. Es können beispielsweise die Verben *haben* oder *vermissen* gewählt werden. Der Rap darf allein, zu zweit oder alles zusammen gesprochen werden, es kann geklatscht und gestampft oder gesungen werden. Außerdem gilt es, die Rolle des/der Fundbüromitarbeitenden zu vergeben. Für die Rolle wird eine passende, amüsante Reaktion auf die vor ihm stehenden Märchenfigurengruppe überlegt. Im Anschluss übt die Gruppe ihre Szene einige Male.

4.5 Präsentation und Sprachreflexion

Jede Gruppe führt ihre Rapversion und die Reaktion des Fundbüropersonals auf, was von den Schüler:innen in Publikumsrolle mit tosendem Applaus honoriert wird. Diese erhalten vorab den Auftrag, sich zu notieren, wie sie den Auftritt fanden und den Schauspielenden dafür zu danken (*Danke für den schönen/lustigen Auftritt!).* Gemeinsam wird noch einmal überlegt, warum auch bei diesem Feedbacksatz *der* zu *den* geworden ist.

5 Fragestellungen

Das dramagrammatische Fünfphasenmodell für Grundschulkinder war bereits vor 11 Jahren Gegenstand eines Artikels dieses Journals (Bryant, 2012). In diesem Beitrag ging Bryant sowohl auf die Konzeption und den Tagesablauf im Theatercamp-Alltag ein als auch auf die Ergebnisse der Evaluation des ersten Camps. Das Anschlussprojekt der Theater-AGs (siehe Bryant & Rummel, 2015) sollte die dramagrammatische Methode auf ihre sprachförderliche Wirksamkeit für in sprachhomogenen Gruppen geförderte Grundschulkinder über den Zeitraum von zwei Jahren überprüfen. Die mit dieser Intervention verknüpfte empirische Untersuchung, die 2023 als Dissertationsschrift eingereicht wurde, evaluierte die Eignung der

auf das jeweilige Sprachniveau bezogenen, dramapädagogischen Grammatikvermittlung. Mit Blick auf die vom „Expertenrat für Herkunft und Bildung“ (Expertenrat H&B, 2011) bereits 2011 geforderten und jüngst durch die IGLU-2021-Studie (McElvany et al., 2023) wiederholte, nachdrückliche Empfehlung, die nachmittäglichen Zeitkontingente im Ganztagsschulbetrieb zur disparitätsausgleichenden, additiven Förderung zu nutzen, entsprach man mit Einrichten der AGs. Denn ohne solche zusätzlichen Angebote sei eine ausreichende (Sprach-)Förderung für förderbedürftige Kinder kaum zu erreichen (Baumert, 2011, S. 18). In der Dissertationsschrift ging es außerdem darum, die grammatischen Kompetenzzuwächse der drei Interventionsgruppen untereinander, aber auch mit der jeweils sprachlich vergleichbaren Kontrollgruppe ohne dramagrammatische AG am Nachmittag (sondern Koch-AG, Ballspiel-AG, Draußen-AG etc.) deskriptiv und inferenzstatistisch zu untersuchen. Außerdem sollte die Strahlkraft der Methode betrachtet werden: In einer Theater-AG liest man naheliegenderweise auch Texte und im Falle der Generierung eines eigenen Stückes produziert man zudem Schriftstücke. Obwohl die Fertigkeiten *Lesen* und *Schreiben* nicht im Hauptfokus der Förderung standen, galt es herauszufinden, ob es in diesen Fertigkeiten größere Veränderungen als in den Kontrollgruppen zu verzeichnen gab, die sich dementsprechend nicht auf den Schulunterricht, sondern die Intervention zurückführen ließen. Wie auch die Entwicklungen in der Grammatik wurden die Unterschiede in den Diagnostikleistungen im *Lesen* und *Schreiben* vor, während und nach der Intervention ebenfalls auf statistische Signifikanz und Effektstärken untersucht. Von Interesse war zudem die Förderdauer: Konnte eine Zeitspanne als besonders effektive Förderzeit ausgemacht werden, das heißt, konnten deutliche Veränderungen innerhalb zweier Schuljahre, innerhalb eines Schuljahres oder auch in kürzerer Zeit gefunden werden? Außerdem ging es darum, herauszufinden, ob sich eine Grundschulklassenstufe für einen Eintritt in die Theater-AG in besonderem Maße eignete. Diese Fragestellungen standen unter der Prämisse, dass sich die personenbezogenen Merkmale der Interventions- und Kontrollgruppen hinsichtlich der Sprachbiografie und dem Familienhintergrund innerhalb der einzelnen Sprachniveaus vergleichen ließen, weshalb vorab eine zweifaktorielle, multivariate Varianzanalyse genutzt wurde (*mANOVA*, Grum & Zydatiß, 2016, S. 321). Hierbei sollte ausgeschlossen werden, dass Gruppenunterschiede bereits vor der Intervention bestanden, und sich damit die gefundenen sprachlichen Veränderungen nicht auf die Intervention zurückführen ließen.

6 Methode

Die dramagrammatische Fördermethode unterliegt, wie jede Förderung, die mit performativen Mitteln arbeitet, der Schwierigkeit des Ansetzens empirischer Untersuchungsmöglichkeiten und damit evidenzbasierter Aussagen über ein wirksames

Fremd-/Zweitsprachenlernen (Küppers & Walter, 2012, S. 1). Dabei handelt es sich um ein „allgemeines Methodenproblem der Dramapädagogikforschung“ (Wirag, 2021, S. 8). Es kann bislang nicht gemessen werden, in welchem Ausmaß die performative Methode beispielweise zu sprachlichen Veränderungen führt, sondern nur die Veränderung auf dem Gebiet der sprachlichen Fertigkeiten selbst sind messbar. Für diese Problematik ist bis dato kein angemessener Lösungsansatz gefunden worden (Wirag, 2021, S. 8). Zwar kann ein anspruchsvolles Design, das sich Mixed-Methods Ansätzen bedient und zudem so weit wie möglich die Maximen der Objektivität, Validität, Reliabilität und Kausalität befolgt, zu aufschlussreichen Ergebnissen kommen (Wirag, 2021, S. 8; Bortz & Döring 2016, S. 215; Klempin & Wirag 2020, S. 108; Lutzker 2016, S. 67). Allerdings mindert dies nicht die Herausforderung, der sich Forscher:innen im Bereich performativer Methoden wie auch im Bereich wirksamer Sprachförderung gegenüber sehen, nämlich dass Ursache-Wirkungs-Aussagen nur unter bestimmten Voraussetzungen zu treffen sind (Klempin & Wirag, 2020, S. 96, Bortz & Döring, 2016, S. 215). Das Design einer Interventionsstudie, die sowohl Ansprüchen der empirischen Bildungswissenschaft als auch dem Ziel, Kinder individuell fördern zu können, entsprechen will, muss daher einen beinahe unmöglichen Spagat bewerkstelligen. Zu welchen Ergebnissen die Abwägungen zum Design geführt haben, wird im nächsten Abschnitt in Kürze dargestellt.[4]

7 Untersuchungsdesign

Für das Theater-AG-Untersuchungsdesign kamen aufgrund der Vereinbarungen mit dem Financier und Kooperationspartner, der Stadt Tübingen, nur der Einsatz einer maßgeschneiderten Intervention (*tailored/ tailor-made intervention*) infrage. So war zum Beispiel eine randomisierte Gruppenzuordnung zu den Faktoren *Gruppe (Intervention/ Kontrolle)* sowie *Sprachniveau* nicht möglich – denn Kinder aus dem Theatercamp sollten mit der gleichen Methode in AGs weiter sprachhomogen gefördert werden. Durch die Vorgaben des Geldgebers sollten zudem Plätze, die durch Abschiebungen entstanden waren, wieder besetzt werden. Insgesamt sah sich die Intervention entfernt von idealen und in diesem Sinne laborähnlichen Bedingungen. Diesen Limitationen (siehe Abschnitt 9) nahm sich die das Projekt evaluierende Dissertationsschrift in den Auswertungen sehr genau an. Nach der Feststellung der Vergleichbarkeit der Gesamtheit der Interventions- und Kontrollkinder wurde jedes der beiden Interventionsschuljahre 2015-16 und 2016-2017 einzeln auf sprachliche Entwicklungen von Interventions- sowie Kontrollkindern betrachtet, um dem Umstand leicht veränderter Gruppenzusammensetzungen Rechnung zu tragen. Die Entwicklungen konnten mittels standardisierter Tests für drei Sprachbereiche (Grammatik, Textproduktion,

[4] Dieser Abschnitt (6) findet sich ebenfalls in der Dissertationsschrift (Vogel, in Vorb.).

Leseverständnis) zu drei Zeitpunkten (Anfang: T1, Mitte: T2, Ende: T3) erfasst werden. Außerdem wurde der Spracherwerb von elf Teilnehmenden (je zwei/drei Interventions- und einem Kontrollkind pro Sprachniveau) über zwei Jahre hinweg detailliert in allen drei oben genannten Sprachbereichen betrachtet. Hinzu kamen die Fragebogendaten von Kindern und Lehrkräften zu familiärem Hintergrund sowie Einstellung, Motivation, Anstrengungsbereitschaft und Selbstkonzept im Sachfach und Deutsch.

7.1 Gruppenbildung

Die drei zu erstellenden Interventionsgruppen mussten qua Design unterschiedliche Sprachstände aufweisen. Pro Niveau (Sprachanfänger:innen, Sprachfortgeschrittene, mündlich Versierte) gab es maximal 12 Förderplätze. Um einen individuellen Blick auf jedes Kind gewährleisten zu können, durften die Gruppen nicht größer sein. Die Kontrollgruppenkinder wurden von den Lehrkräften aufgrund der subjektiven Einschätzung der Vergleichbarkeit (Sprachstand, personenbezogene Merkmale) vorgeschlagen. Sie nahmen ebenfalls zu den drei Diagnostikzeitpunkten (Anfang: T1, Mitte: T2, Ende: T3) an den drei Diagnostikteilen (Grammatik, Textproduktion, Lesefertigkeit) teil und füllten auch die oben genannten Fragebögen aus. Das Zustandekommen der Gruppen kann in Abbildung 3 nochmals nachvollzogen werden.

7.2 Instrumente und Erhebung

Die Wirksamkeit der dramagrammatischen Sprachförderung wurde anhand standardisierter Sprachstandsdiagnostiken in den Bereichen Textproduktion (*Der Sturz ins Tulpenbeet*), Leseverständnis (*ELFE* 1-6) und Grammatik (*TROG-D*) evaluiert. Die ersten beiden Testungen sind Gruppendiagnostiken, der *TROG-D* muss als Einzeltestung durchgeführt werden. Die Sprachstandsdiagnostik teilte sich in drei Testzeiträume auf: 1) Nach der Kennenlernphase im Winter 2015/16 (T1), 2) im Sommer 2016 (T2) und 3) im Sommer 2017 (T3) (Abb. 4). In jedem Testzeitraum wurden alle drei Messinstrumente in Gruppen- und Einzelsitzungen in jeder der drei Interventions- und jeder der drei Kontrollgruppen durchgeführt, woraus sich 490 Diagnostikbögen ergaben.

Hinzu kamen Schulkind- und Lehrkraftfragebögen zum persönlichen Hintergrund und der Sprachbiografie der Kinder, den kognitiven Fähigkeiten sowie Einstellung, Selbstkonzept und Mitarbeit in den Fächern Deutsch und Sachunterricht. Wie in Abbildung 1 zu sehen war, unterschieden sich logischerweise die Sprachförderthemen nach Sprachniveau bei einer entwicklungsproximal geplanten Förderung. In jedem Sprachniveau wurden Interventions- und Kontrollkinder zusätzlich zu den Gesamtergebnissen der Diagnostiken verglichen, indem

man aus denjenigen Unterabschnitten der Diagnostiken solche gesondert betrachtete, die im jeweiligen Sprachniveau bearbeitet wurden.

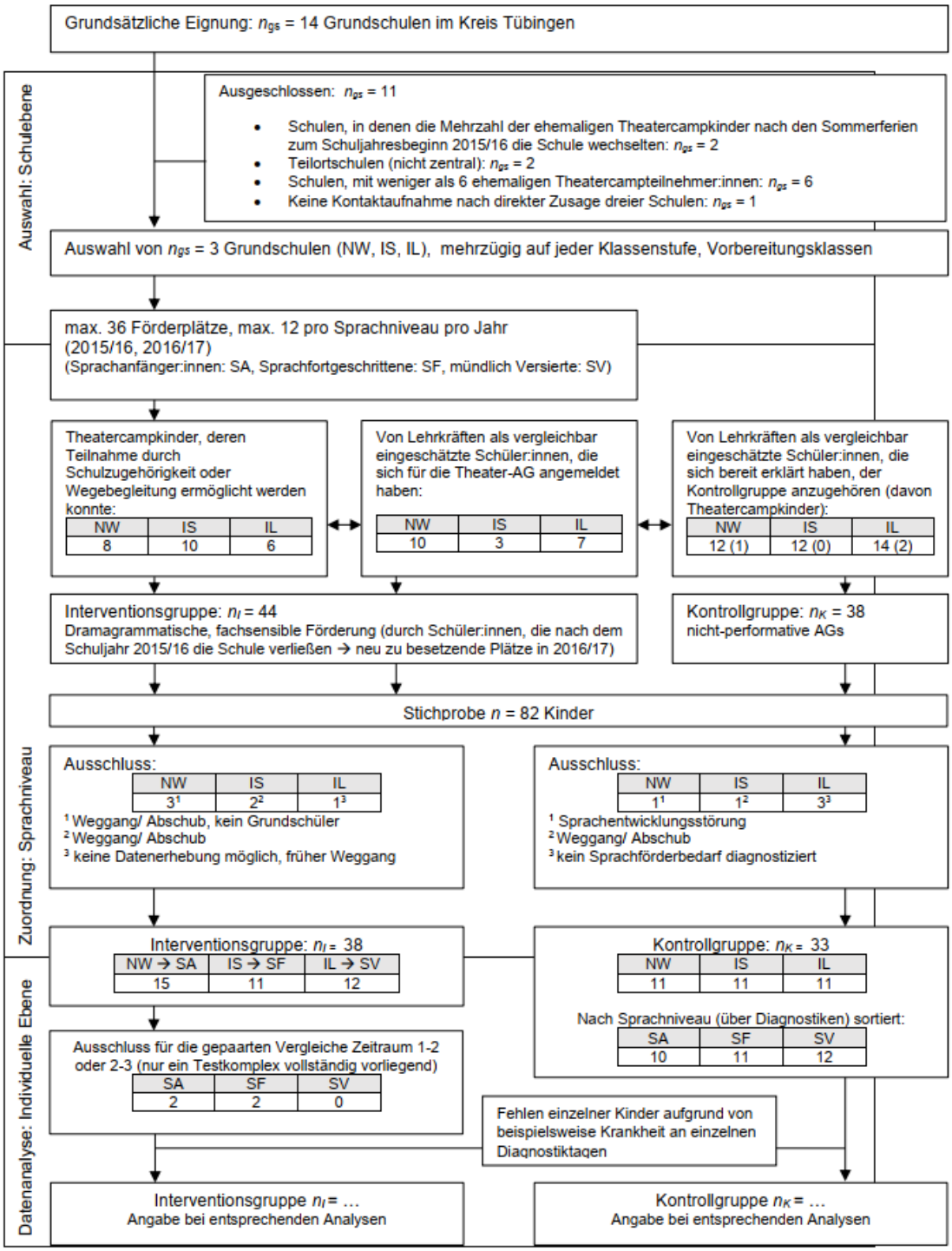


Abbildung 3: Zustandekommen und Überblick über die Stichprobe der Interventionsstudie aus Vogel (in Vorb.), adaptiert nach CONSORT Flow Diagram (Schulz et al., 2011, S. e20, Siegmund, 2022, S. 129)

Ein einziges grammatisches Förderthema fand sich aber in allen drei Gruppen mit einer identisch geplanten Förderreihe: der zweigliedrige, negativ-additive Konnektor *weder … noch*. Hierzu gab es eine zusätzliche Prä-Post-Diagnostik mit rezeptivem und produktivem Teil.

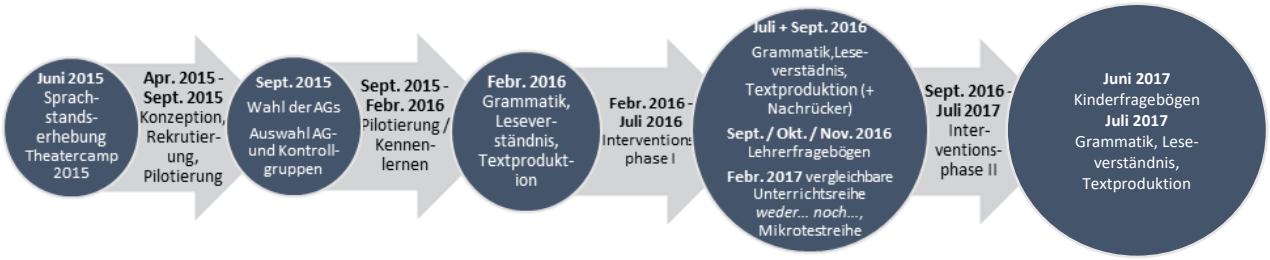


Abbildung 4: Zeitliche Einordnung der Evaluationsteile,
hellgrau: Interventionsphase, dunkelgrau: Erhebungszeitpunkte und Art der einzelnen Testungen und Fragebögen

Pro Förderjahr erhielt jede der drei Theater-AG-Gruppen 31 bis 33 Fördereinheiten à 90 Minuten. Dies bedeutete für mich (S.V.) 288 Zeitstunden Förderung neben Vor- und Nachbereitungen, Diagnostik- sowie Gestaltungszeit, die für die insgesamt neun Theateraufführungen (drei pro Gruppe) anberaumt werden musste. Dieses Pensum für mich als einzelne Test- und Studienleitung muss als Grund genannt werden, warum nicht mehrere Gruppen pro Sprachniveau eingerichtet werden konnten, die eine größere Reichweite an Daten hätten abbilden können.

7.3 Vergleichbarkeit der Kinder

Die Kontrollkinder zeigten weder Theateraufführungen, noch produzierten sie wöchentliche Texte innerhalb der Studie. Sie nahmen zwar ebenfalls nachmittags im Ganztagsschulbetrieb an AGs wie Koch-, Mädchen-, Draußen-, Tanz- oder Ballsport-AG teil und hatten dort wie die Interventionskinder deutschsprachigen Kontakt, jedoch enthielten diese AGs kein dramagrammatisches Förderkonzept. Erfolge der Interventionskinder in den Sprachfördergebieten sollten durch den Ausschluss performativ-sprachförderlicher AGs der Kontrollkinder auf die Intervention zurückzuführen sein. Anhand einer zweifaktoriellen, multivariaten Varianzanalyse wurde im Voraus ausgeschlossen, dass eventuelle Unterschiede schon vor der Intervention bestanden. Sechzehn abhängige Variablen mit den Faktoren *Gruppe (Intervention/ Kontrolle)* und *Niveau (Sprachanfänger:innen, Sprachfortgeschrittene, mündlich Versierte)* wurden in die Analyse einbezogen (*Geschlecht, Kontaktdauer mit dem Deutschen, Interventionseintrittsalter, Anfangsklasse, Teilaspekte der Leseverständnisdiagnostik: Wort-, Satz- und Textverständnis, Teilaspekte der Textproduktion: Anzahl der Wörter, Aussagen mit Vollverb, mittlere Satzlänge, Textbewältigungspunktzahl, Anzahl der Satzverbindungen, Anzahl bildungssprachlicher Elemente, Teilaspekte der Grammatikdiagnostik: Personalpronomen, Akkusativ-Dativ).* Die Analyseergebnisse der

mANOVA zeigen keinen signifikanten Haupteffekt auf die Varianz der sechzehn Variablen ($F(32, 39) = 1.16$, $p = 0.29$).[5]

8 Ergebnisse aus dem Anfänger:innenniveau

Aus den 150 Seiten deskriptiver und inferenzstatistischer Auswertung der Fragebögen und Sprachstandsdiagnostiken der Dissertationsschrift (Vogel, in Vorb.) sollen an dieser Stelle lediglich die Ergebnisse des Anfänger:innenniveaus in allen drei getesteten Sprachbereichen skizziert werden. Das betrifft Hypothese 2 und 4 der sieben Hypothesen der Dissertationsschrift, und zwar:

- H2: Das Interventionsprojekt unterstützt sprachliche Veränderungen in den von ihm geförderten Gebieten. Das heißt, die Interventionskinder unterscheiden sich im Nachgang der Förderung bezüglich ihrer sprachlichen Fertigkeiten von den Kontrollkindern des gleichen Niveaus.
- H4: Die Interventionskinder unterscheiden sich hinsichtlich der in der Intervention nicht explizit geförderten a) Schreib- und b) Lesekompetenz von den Kontrollkindern innerhalb ihres Sprachniveaus nach der dramagrammatischen Förderung nicht.[6]

8.1 Vergleich der Ergebnisse aus der Grammatiktestung TROG-D

Die hier betrachtete Hypothese 2 (H2) der Dissertationsschrift besagt, dass das Interventionsprojekt sprachliche Veränderungen in den geförderten Gebieten unterstützt, was für die Interventionskinder des Sprachanfänger:innenniveaus bedeutet, dass sie sich nach der Förderung von den Kontrollkindern im Gebrauch von Personalpronomen im Nominativ, Adjektiven und in der Verwendung von Komplementsätzen mit *dass* im ersten Jahr unterscheiden (T1→T2). Im zweiten Jahr sollten die sprachlichen Veränderungen der Interventionskinder im Gegensatz zu den Kontrollkindern stärker ausfallen im Bereich der Personalpronomen im Nominativ, der Temporaladverbien, des komplexen Konnektors *weder…noch,* der Syntax bei Topikalisierungen sowie in der Verwendung von Komplementsätzen mit *dass* (T2→T3).

[5] Eine detaillierte Übersicht über die abhängigen Variablen findet sich in der Dissertationsschrift (Vogel, in Vorb.).
[6] Die Hypothesen und Fragestellungen sind adaptiert nach dem entsprechenden Kapitel der Dissertationsschrift (Vogel, in Vorb.).

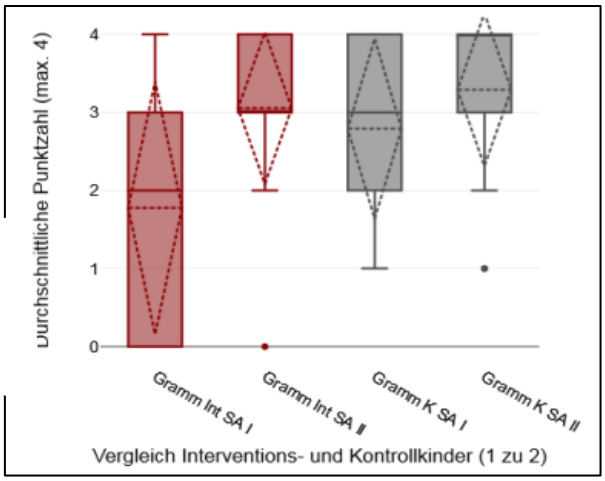


Abbildung 3: Vergleich der Interventions- und Kontrollkinder (rot und grau) der Sprachanfänger:innen (TROG-D) zwischen T1 und T2

	Interventionsgruppe		**Kontrollgruppe**	
	T1	**T2**	**T1**	**T2**
Anzahl gültiger Werte	18		24	
M *SD*	1.78 1.67	3.06 1.00	2.79 1.18	3.29 1.00
Δ(M) *t* *p* *Cohen's d*	1.28 -4.42 <0.001*** -1.30		0.50 -2.63 0.015* -0.64	

Tabelle 2: Arithmetisches Mittel (M), Standardabweichung (SD) und Mittelwertdifferenz (Δ(M)) der Sprachanfänger:innen im Vergleich zwischen T1 und T2 in den durch den TROG-D abbildbaren grammatischen Fördergebieten sowie Ergebnisse der t-Tests für abhängige Stichproben neben Effektstärke d

Anhand der Abbildung 5 wird sichtbar, dass die Interventionskinder in den Testblöcken des *TROG-D*, die die Förderthemen des ersten Jahres abbilden können (T1→T2), anfangs deutlich niedrigere Punktwerte als die Kontrollkinder erreichen (T1: Gramm Int SA I, Gramm K SA I). Sie gleichen sich aber in der zweiten Testung dem zunächst sehr viel höheren Niveau der Kontrollkinder an (T2: Gramm Int SA II, Gramm K SA II). Obwohl sich die Spannweite der *TROG-D*-Punktzahl in der Interventionsgruppe in der ersten Testung deutlich größer als in der Kontrollgruppe zeigt, kann das untere Leistungsspektrum hier enorm aufholen (T1→T2), sodass zur zweiten Testung der erreichte Mittelwert der Interventionskinder nur noch wenig unter dem der Kontrollkinder liegt (Tab. 2). Zudem erweist sich die Mittelwertdifferenz der Interventionsgruppe mehr als doppelt so groß wie die der Kontrollgruppe zwischen der ersten und zweiten Erhebung (Tab. 2: T1→T2; $\boldsymbol{\Delta}(M_{I,K})$ = 1.28, 0.50). Damit jene auf statistische Signifikanz überprüft werden konnten, wurde jeweils ein *t*-Test für abhängige Stichproben durchgeführt (Tab. 2), der für die Interventionsgruppe statistisch signifikant ausfällt ($t_I(17)$ = -4.42, p_I = <0.001***, 95%-CI [-1.89, -0,67]). Als sehr groß kann ebenfalls die Effektstärke (d_I = -1.30) eingeschätzt werden. Jenes trifft sogar zu mit den auf die Forschung im Zweit- und Fremdsprachenbereich von Plonsky & Oswald (2014) vorgeschlagenen, höheren Effektstärkebereichen im Vergleich zu den eher niedrigeren in der Psychologie oder Bildungswissenschaft üblichen: „*d* values in the neighborhood of .40 should be considered

small, .70 medium, and 1.00 large“ (Plonsky & Oswald 2014, S. 12). Auch die Kontrollgruppe verzeichnet eine Verbesserung in den in der Interventionsgruppe grammatisch geförderten Gebieten im Bereich statistischer Signifikanz. Jedoch kann sie mit d_K = -0.64 nach Plonsky & Oswald (2014) einen noch kleinen, nahezu mittelstarken Effekt erzielen, keinen sehr großen wie die Interventionskinder.

Im zweiten Testzeitraum (T2→T3) erreichten Interventions- und Kontrollkinder zunächst nahezu gleiche Mittelwerte (T2, $M_{I,K}$: 1.51, 1.52, Tab. 3). Die Mittelwertdifferenz der Interventionsgruppe nach der Förderung zeigt sich erneut doppelt so groß wie die der Kontrollgruppe (T3; $\Delta(M_{I,K})$: 1.22, 0.60). An Abbildung 6 lässt sich weiterhin ablesen, dass sich die mittleren 50%-Werte aus der Interventionsgruppe (quadratische rote Box ohne t-Whisker nach oben oder unten) enorm nach oben verschieben (T2→T3, Gramm Int SA II→ Gramm Int SA III). Im Gegensatz dazu bleiben in der Kontrollgruppe die niedrigeren Werte ohne Veränderung (quadratische graue Box bleibt auf der X-Achse).

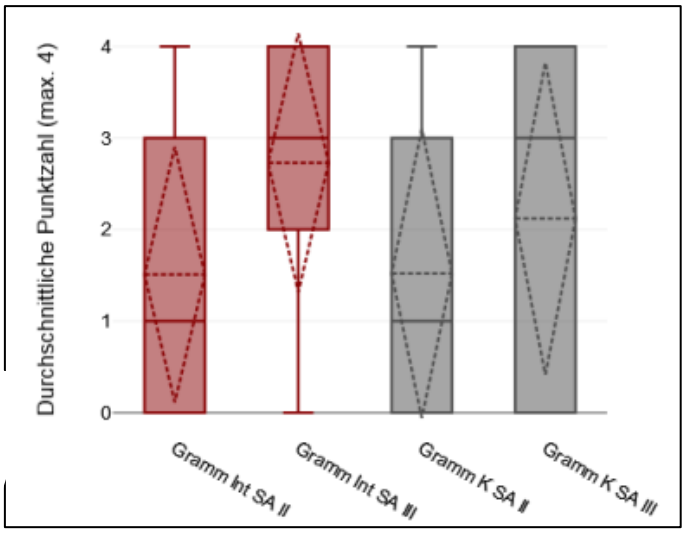


Abbildung 4: Vergleich der Interventions- und Kontrollkinder (rot und grau) der Sprachanfänger:innen (TROG-D) zwischen T2 und T3

	Interventionsgruppe		Kontrollgruppe	
	T2	**T3**	**T2**	**T3**
Anzahl gültiger Werte	59		25	
M *SD*	1.51 1.41	2.73 1.42	1.52 1.61	2.12 1.74
Δ (M) *t* *p* *Cohen's d*	1.22 -7.06 <0.001*** -0.98		0.60 -2.52 0.019** -0.71	

Tabelle 3: Arithmetisches Mittel (M), Standardabweichung (SD) und Mittelwertdifferenz (Δ(M)) der Sprachanfänger:innen im Vergleich zwischen T2 und T3 in den durch den TROG-D abbildbaren grammatischen Fördergebieten sowie Ergebnisse der t-Tests für abhängige Stichproben neben Effektstärke d

Die Unterschiede zwischen T2 und T3 sind für beide Gruppen statistisch signifikant (*t*-Test für abhängige Stichproben: $t_I(58)$ = -7.06, p_I = <0.001***, 95%-CI [-1.57, -0.87]; $t_K(24)$ = -2.52, p_K =

0.019*, 95%-CI [-1.09, -0.11]). Es liegt für die Theater-AG-Gruppe ein fast großer Effekt vor, für die Kontrolle ein gerade mittelstarker ($d_{I,\,K}$ = -0.98, -0.71).

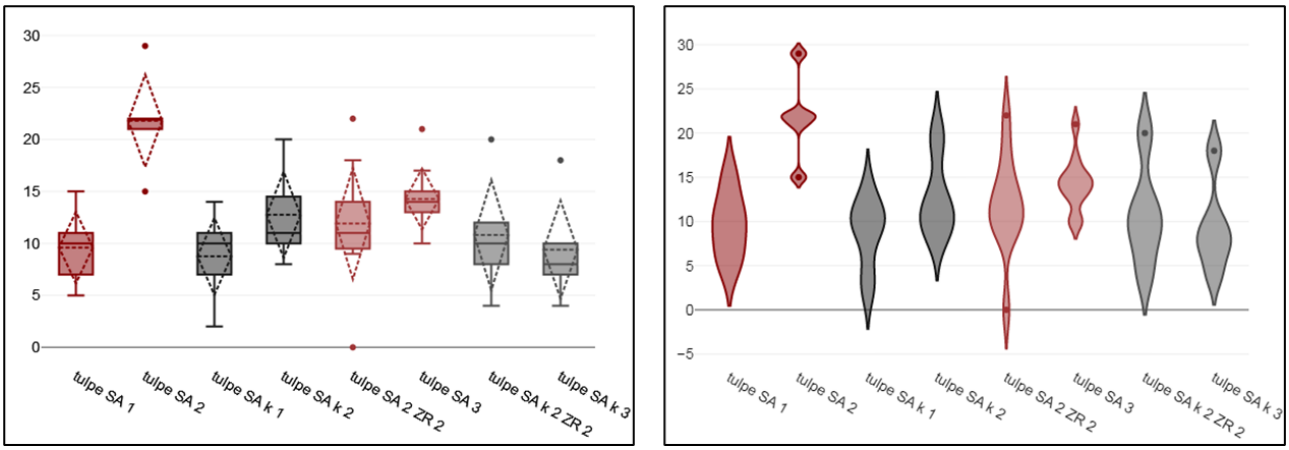
Abschließend kann zu den Grammatikergebnissen der Sprachanfänger:innen zusammengefasst werden, dass sich die Förderung mit dramagrammatischen Fünfphasenmodell als profitable Methode eignet. Kontrollkinder, die im Ganztagsschulbetrieb gleichermaßen wie die Interventionskinder anwesend waren, aber an anderen AGs ohne sprachförderlichen Inhalt teilnahmen, konnten weniger deutliche Erfolge erzielen. Hiermit wird der in der Zweitsprachenforschung hinlänglich bekannte Konsens unterstrichen, dass für Kinder dieses Niveaus regelhaft ein motivierendes Programm mit explizit sprachförderlichen Inhalten angeboten werden muss (Expertenrat H&B, 2011; McElvany et al., 2023), damit sie ihren Schulalltag so schnell es geht für sie zufriedenstellend bestreiten können. Die Hypothese H2 kann infolge der inferenzstatistischen Untersuchungen für beide Förderjahre **angenommen** werden.

8.2 Vergleich der Ergebnisse der Textproduktion

Die Hypothese H4a) geht davon aus, dass es im Bereich der nicht explizit im Förderfokus stehenden Textproduktion keine unterschiedliche Entwicklung von Interventions- und Kontrollkindern zu verzeichnen gibt. Anders als die Hypothese es annimmt, ließen sich jedoch bei den Interventionskindern des Sprachanfänger:innenniveaus erstaunliche Fortschritte feststellen.[7] Im ersten Zeitraum (T1) starteten beide Gruppen des Niveaus ungefähr gleich (Abb. 7a): T1: tulpe SA 1, tulpe SA k 1), wobei die unteren 25% der Kontrollkinder (im Boxplot: graue t-förmige Whisker nach unten bei tulpe SA k 1) leicht niedriger angesiedelt waren als die der Interventionskinder. Beide Gruppen entwickelten sich in ihrer Textproduktion zwischen T1 und T2 statistisch signifikant (*t*-Test für abhängige Stichproben: $t_I(4) = -3.38$, p_I = 0.028*, 95%-CI [-22.22, -2.18]; $t_K(7) = -3.16$, p_K = 0.016*, 95%-CI [-7.00, -1.00]).

Die Interventionsgruppe erreichte eine überaus große Effektstärke (d_I = 2.76), die Kontrollgruppe einen fast großen Effekt (d_K = 0.95) (nach Plonsky & Oswald, 2014). Die Quartilswerte (Abb. 7a – die jeweils untere sowie obere Linie der Boxen von tulpe SA 1 → 2 / tulpe SA k 1 → 2) sowie Mittelwertunterschiede von Interventions- und Kontrollgruppe (**Δ(**$M_{I,K}$): 12.2, 4.0, Tab. 4) veranschaulichen die besonders starke Entwicklung in der Interventionsgruppe. Besser noch als im Boxplot lässt sich diese Entwicklung im Violinenplot (Abb. 7b) nachvollziehen. Zu erkennen ist dort, dass sich die Interventionskinder in beiden Förderzeiträumen auf einem hohen Niveau einander angleichen, was für die Kontrollkinder nur in geringerem Maße gilt.

[7] Nicht alphabetisierte Kinder konnten in der Textproduktionsdiagnostik ihre Geschichte diktieren.



Abbildungen 7a) und b): Erreichte Punktwerte im standardisierten Textproduktionstest „Der Sturz ins Tulpenbeet" der Sprachanfänger:innen dargestellt als Boxplot (links) und Violinenplot (rechts).
T1 →T2: Interventions- (dunkelrot, tulpe SA 1 ⍰ tulpe SA 2) und Kontrollkinder (dunkelgrau, tulpe SA k 1 ⍰ tulpe SA k 2)
T2 →T3: Interventions- (hellrot, tulpe SA 2 ZR 2 ⍰ tulpe SA 3) und Kontrollkinder (hellgrau, tulpe SA k 2 ZR 2 ⍰ tulpe SA k3)

	Intervention T1→T2		**Kontrolle T1→T2**		**Intervention T2→T3**		**Kontrolle T2→T3**	
	T1	**T2**	**T1**	**T2**	**T2**	**T3**	**T2**	**T3**
n	6		8		11		5	
M	9.60	21.80	8.75	12.75	11.91	14.27	10.80	9.40
Md	10	22	10	11	11	14	10	8
SD	3.85	4.97	3.96	4.43	5.61	3.07	5.93	5.27
Min.	5	15	2	8	0	10	4	4
Max.	15	29	14	20	22	21	20	18
Quartil 1	7	21	7	10	9.5	13	8	7
Quartil 3	11	22	11	14.5	14	15	12	10
Δ(M)	12.2		4.0		2.36		-1.4	
t	-3.38		-3.16		-1.46		3.50	
p	0.028		0.016		0.175		0.025	
Cohen's d	2.76		0.95		0.52		-0.25	
CLES[8]	0.98		0.75		0.64		0.57	
d_{ppc2}	1.96				0.62			

Tabelle 4: Stichprobe (n), arithmetisches Mittel (M), Median (Md), Standardabweichung (SD), Minimum und Maximum, Quartil 1 und 3 sowie Mittelwertdifferenz (Δ(M)), Ergebnisse des t-Tests, Cohen's d, CLES, d_{ppc2} der Gesamtergebnisse der Textproduktion „Tulpenbeet" zwischen den Interventions- und Kontrollkindern des Sprachanfängerniveaus zwischen T1 →T2 sowie T2 →T3

Im zweiten Zeitraum (T2→T3) starteten die Interventionskinder mit minimal höheren Werten als die Kontrollkinder, nähern sich aber deutlicher einander an und verbessern sich dabei. Dies trifft auf die Kontrollkinder jedoch nicht zu. Trotz der Verbesserung auf Seiten der Interventionskinder ist der Unterschied statistisch nicht signifikant ($t_I(10) = -1.46$, $p_I = 0.175$, 95%-CI [-5.97, 1.24]), jedoch gibt es nach Plonsky & Oswald (2014) einen kleinen Effekt von $d_I = 0.52$ zu verzeichnen. In der Kontrollgruppe findet sich mit $d_K = -0.25$ kein nennenswerter Effekt, wobei hier der Vorher-Nachher-Vergleich statistisch signifikant ist – jedoch in Richtung

[8] Als non-parametrische Effektstärke wird die *Common Language Effect Size* (CLES; McGraw/ Wong 1992) ausgegeben, um die Wahrscheinlichkeit anzugeben, dass ein zufällig aus der einen Stichprobe gezogener Fall einen höheren Wert hat als ein zufällig gezogener Fall aus der anderen Stichprobe. Im Psychometrica-Rechner wird die Wahrscheinlichkeit aus Sicht der Gruppe mit dem höheren Mittelwert berichtet (https://www.psychometrica.de/effektstaerke.html).

eines schlechteren Ergebnisses ($t_K(4) = 3.5$, $p_K = 0.025$*, 95%-CI [0.29, 2.51]). Wie im Boxplot (Abb. 7a: tulpe SA k 2 ZR 2 → tulpe SA k3) zu sehen ist, verschlechtert sich demnach die Kontrollgruppe in ihrer Textproduktion. Dies wirft die Frage auf, ob die Kontrollkinder in den vorherigen Testungen möglicherweise deshalb höhere Werte erzielen konnten, weil die Testsituation für sie spannend und neu war (ein sogenannter Novitätseffekt, Bortz & Döring, 2016, S. 101) und nun in der dritten Testung mit demselben Material (*Der Sturz ins Tulpenbeet*) sich wenig Anreiz für weitere Bemühungen bot. Gleichzeitig ist es möglich, dass sich die Interventionskinder weiter verbesserten,[9] weil sie vertrauter mit der Versuchsleiterin waren, der gleichzeitigen AG-Leitung, und ihren von ihnen vermuteten Ansprüchen gerecht werden wollten (Rosenthal-/Versuchsleiter(erwartungs)effekt, Bortz & Döring, 2016, S. 927).

Die Hypothese H4a), die besagt, dass Interventionskinder sich in ihrer Entwicklung im Schreiben nicht maßgeblich von den Kontrollkindern unterscheiden, wird abgelehnt, da sich die Sprachanfänger:innen in der AG im ersten Jahr sehr viel deutlicher als die Kontrollkinder entwickeln (T1→T2). Im zweiten Jahr entwickeln sie sich zwar nicht statistisch signifikant (T2→T3), verschlechtern sich aber im Gegensatz zur Kontrollgruppe auch nicht statistisch signifikant und unterscheiden sich demnach ebenfalls.

Die Prüfung auf statistische Signifikanz erfolgte mit dem Wissen, dass kleine Stichproben eher zu größeren *p*-Werten führen (Bortz & Döring, 2016, S. 1027). Es ist also durchaus denkbar, dass sich in einer größeren Stichprobe viel eher die hier deskriptiv sichtbaren, für die Interventionsgruppe anhand der Grafen sehr viel positiveren Unterschiede auch statistisch abbilden ließen.

8.3 Vergleich der Ergebnisse der Lesefertigkeit

Hypothese H4b) nimmt an, dass sich die Interventionskinder in der Fertigkeit *Lesen* nicht von den Kontrollkindern in ihrer Entwicklung abheben, weil diese mit Anwendung der drama*grammatischen* Methode ebenfalls nicht im Mittelpunkt der Förderung stand.
Ein echter Vergleich der Lesefähigkeiten im ersten Zeitraum (T1→T2) war für die Interventionsgruppe nicht möglich, weil aufgrund fehlender primärer und mitunter sekundärer Literalisierung in vielen Fällen in Absprache mit den Lehrkräften keine Lesediagnostik vorgenommen wurde. In Abbildung 8 ist die erste Box in helllila deshalb andersfarbig markiert. Denn es konnten in diesem Niveau nur zwei erste Lesediagnostiken (T1) bei den Interventionskindern vorgenommen werden. Die anderen neun Kinder wurden

[9] Auch wenn diese Verbesserung der Interventionsgruppe im zweiten Jahr statistisch nicht signifikant ist, lässt sie sich an den Mittelwertdifferenzen zwischen T2 und T3 ($\Delta(M_{T2\rightarrow T3})$ und dem Violinenplot (Abb. 7b) ablesen.

hier mit dem hypothetischen Wert null versehen, obwohl sie eventuell im Wortverständnis einige wenige Punkte hätten erreichen können.

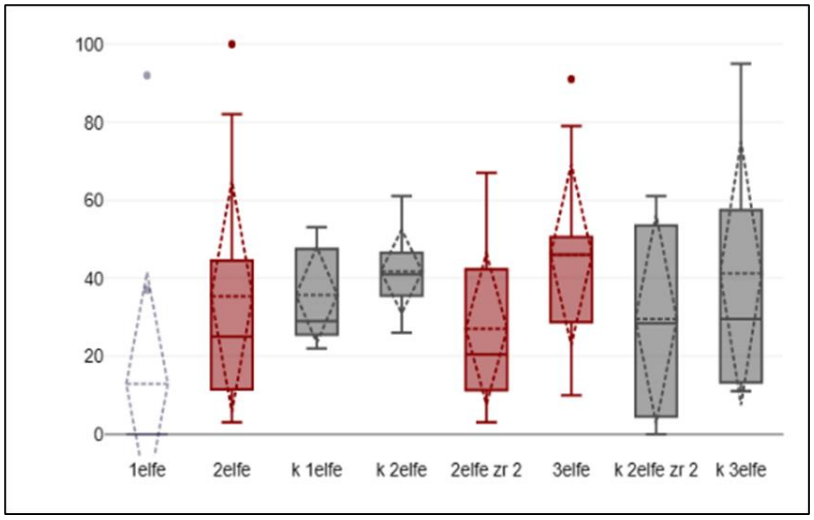


Abbildung 8: Erreichte Punktwerte (X/120) im Leseverständnistest ELFE 1-6 der Sprachanfänger:innen: T1→T2: Interventions- und Kontrollkinder (1elfe, 2 elfe und k1elfe, k 2elfe) sowie T2→T3: Interventions- und Kontrollkinder (2elfe zr 2, 3 elfe und k2 elfe zr 2, k3 elfe)

Echte paarweise Vergleiche waren somit nur mit Kindern möglich, von denen jeweils am Anfang und Ende eines Förderzeitraumes für den entsprechenden Testdaten vorlagen.

Für die standardisierte Lesefertigkeitsdiagnostik *ELFE* 1-6 existieren zwei unterschiedliche Testbögen, deren zweiter Teil im letzten, dritten Diagnostikzeitraum genutzt wurde. Der Boxplot (Abb. 8) zeigt, dass sich die Interventionskinder des Sprachanfängerniveaus zwischen der zweiten und dritten Testung deutlich verbesserten (T2→T3), obwohl hier noch einmal sechs Anfänger:innen in die Gruppe aufgenommen wurden. Bei den Kontrollkindern des gleichen Sprachniveaus kamen im gleichen Zeitraum nur zwei neue Kinder hinzu, sodass die sichtbar schlechteren Anfangswerte (die unteren, grauen 25 % von k2elfe zr2 reichen bis zur x-Achse) im Gegensatz zum ersten Testzeitraum Fragen aufwerfen. Möglicherweise ist diese Veränderung durch die nun sehr kleine Anzahl von vier Kontrollkindern zu erklären. Zudem führten alle Kinder im letzten Testzeitraum den zweiten, ihnen bisher unbekannten *ELFE*-Testbogen durch. Allerdings kann dies schwerlich als Erklärung für die eher geringe Verbesserung der Kontrollkinder dienen (T3). Denn einerseits waren die Testbögen der *ELFE* 1-6 standardisiert und somit auf ihre Vergleichbarkeit hin geprüft (Lenhard & Schneider, 2006), andererseits konnten die Interventionskinder mit Mittelwertdifferenzen von immerhin 19.00 Punkten eine enorme Entwicklung vorweisen (T2→T3, Tab. 5).

Die angenommene Entwicklung der Interventionskinder scheint im ersten Jahr (T1→T2) sehr stark zu sein (**Δ***(M$_I$):* 23.63). Hätte man die nicht alphabetisierten Kinder tatsächlich vor einen Testbogen gesetzt, den sie vermutlich nicht hätten lösen können, hätte dies wahrscheinlich zu

einem signifikanten Unterschied belegt durch den Friedman-Test bei nichtparametrischen Daten führen können.[10]

	Intervention T1→T2		Kontrolle T1→T2		Intervention T2→T3		Kontrolle T2→T3	
	T1	T2	T1	T2	T2	T3	T2	T3
n	11		7		10		4	
M	11.73	35.36	35.71	41.71	27	46	29.5	41.25
Md	0	25	29	41	20.5	46	28.5	29.5
SD	28.84	31.14	13.29	11.51	20.4	24.44	30.97	38.99
Min.	0	3	22	26	3	10	0	11
Max.	92	100	53	61	67	91	61	95
Quartil 1	0	11.5	25.5	35.5	11.25	28.75	4.5	13.25
Quartil 3	0	44.5	47.5	46.5	42.25	50.5	53.5	57.5
Δ(M)	23.63		6.00		19.00		11.75	
t	-4.80		-1.22		-5.06		-1.42	
p	<0.001***		0.27		<0.001***		0.25	
Cohen's d	-2.68		-0.44		-3.02		-1.70	

Tabelle 5: Arithmetisches Mittel (M), Median (Md), Standardabweichung (SD), Minimum, Maximum, Quartil 1 und 3 sowie Mittelwertdifferenz (Δ(M)) der erreichten Punktwerte (X/120) im ELFE 1-6, Ergebnisse des t-Tests, Cohen's d zwischen den Interventions- und Kontrollkindern des Sprachanfänger:innenniveaus zwischen T1 → T2 sowie T2 → T3

Hinsichtlich ethischer Fragen wurde dies jedoch nicht in Betracht gezogen. Solch ein möglich signifikanter Unterschied findet sich in der Kontrollgruppe mit einer Differenz von **Δ***(M$_K$):* 6.00 nicht. Wählt man hingegen einen parametrischen Test und führt dann pro Gruppe einen *t*-Test für abhängige Stichproben durch (T1→T2), ergibt sich ein statistisch signifikanter Unterschied ($t_I(10) = -4.8$, $p_I = 0.001$**, 95%-CI [-34.62, -12.65]). [1] Die Kontrollgruppe erreicht mit $d_K = -0.44$ im ersten Jahr einen kleinen Effekt. Laut *t*-Test für abhängige Stichproben ist der Unterschied zwischen T1 und T2 hier statistisch nicht signifikant ($t_K(6) = -1.22$, $p_K = 0.268$, 95%-CI [-18.04, 6.04]).

Zwischen T2 und T3 wird in der Interventionsgruppe, nun mit vollständig möglicher Datenerhebung, ein sehr großer Effekt ermittelt ($d_I = -3.02$) bei mittels *t*-Test für abhängige Stichproben berechnetem statistisch signifikanten Unterschied ($t_I(9) = -5.06$, $p_I = 0.001$**, 95%-CI [-27.49, -10.51]). Das Betrachten der Mittel- und Quartilswerte sowie der Mediane macht aber auch ohne *t*-Tests aussagekräftige Veränderungen deutlich (Tab. 5). Die schwächsten 25 % der Interventionsgruppenergebnisse liegen zur letzten Testung auf dem Niveau der unteren 50 % der vorherigen (Abb. 8: Intervention, rot, T2→T3). Das sichtbare Angleichen der schwächeren Leistungen an die unteren 50 % des mittleren Leistungsspektrums in der Lesefertigkeit könnte auf die zusätzliche Beschäftigung mit dem Lesen in der Theater-AG zurückzuführen sein, die so vermutlich in wenig anderen der angebotenen AGs für die Mitschüler:innen in den Kontrollgruppen auftrat. Allerdings erreichte ebenfalls die Kontrollgruppe einen großen Effekt von $d_K = -1.70$, wobei die

[10] Ein nichtparametrischer Test wurde wegen der Stichprobengröße von unter zwanzig Teilnehmer:innen gewählt, um eine robustere statistische Aussagekraft zu erhalten.

untersuchten Mittelwertunterschiede statistisch nicht signifikant sind ($t_K(3) = -1.42$, $p_K = 0.251$, 95%-CI [-38.13, 14.63]).

Für die Interventionskinder ergibt sich ein angenehmer und in seiner Deutlichkeit im Sprachanfänger:innenniveau überraschender Nebeneffekt, der nicht als Hauptförderschwerpunkt fokussiert worden war. Die Hypothese H4b), die besagt, dass sich Interventionskinder bezüglich der Fertigkeit *Lesen* nicht maßgeblich von den Kontrollkindern unterscheiden, kann für die Sprachanfänger:innen im ersten Zeitraum nur unter Vorbehalt der Testunmöglichkeit bei Analphabeten **abgelehnt** werden (T1→T2). Für den zweiten Zeitraum lässt sie sich klar **ablehnen**, denn Interventionskinder unterscheiden sich hier statistisch signifikant mit deutlich besseren Werten von den Kontrollkindern (T2→T3).

9 Limitationen

Einige Punkte sollen vor dem Fazit als Einschränkungen benannt werden.[11]

1. Wie im Studiendesign vorgesehen, wurden Theatercampkinder nach dem Camp 2015 weiter gefördert. Das Nachhaltigkeitsprogramm fand innerhalb des AG-Nachmittages dreier Grundschulen statt. Aus diesen Vorgaben heraus ergab sich ein quasi-experimentelles Studiendesign ohne Möglichkeit zur Randomisierung (Bortz & Döring, 2016, S. 215). Insgesamt waren Interventions- und Kontrollkinder hinsichtlich ihrer sprachlichen und persönlichen Merkmale vergleichbar (siehe 7.3). Ein Vorteil des Designs zeigt sich zum einen im natürlichen Setting des regulären Nachmittags(förder)programms (Bortz & Döring, 2016, S. 106), zum anderen in der freiwilligen Teilnahme aller.
2. Hinsichtlich des Vorhandenseins von Treatment- und Kontrollgruppen, der Auswahl der breit gefächerten Testinstrumente und der außergewöhnlich langen Förderdauer von zwei Jahren besteht eine überaus solide Datengrundlage. Nichtsdestotrotz müsste sie größer sein, um bildungswissenschaftlichen Standards zu genügen.
3. Obwohl alle Paper-Pencil-Tests (*Grammatik, Schreiben, Lesen*) und die Erläuterungen standardisiert erfolgten, besteht die Möglichkeit eines Rosenthal-/ Versuchsleiter(erwartungs)effekts (Bortz & Döring, 2016, S. 927), denn die Versuchsleiterin war den Interventionskindern durch die AG-Stunden wesentlich vertrauter als den Kontrollkindern. Dadurch könnten die Bemühungen der Interventionskinder größer gewesen sein, um den der Versuchsleiterin zugeschriebenen Erwartungen zu entsprechen.

[11] Limitationen der Interventionsstudie werden in dieser Weise ebenfalls in der Dissertationsschrift benannt (Vogel, in Vorb.).

4. Mitunter zeigten sich Kontrollkinder in den ersten Testungen erfolgreicher als in den folgenden. Hier sollte ein Novitätseffekt in Betracht gezogen werden (Bortz & Döring, 2016, S. 101), der die Teilnehmenden aufgrund der Neuheit einer Erfahrung dazu anregen kann, sich mehr Mühe zu geben oder länger bei der Sache zu bleiben.
5. Rückschlüsse können durch das Vorhandensein messbarer Effekte auf eine erfolgreiche Art dramagrammatischer Sprachförderung gezogen werden, erlauben aber keine generelle Aussage über die dramagrammatische Methode. Die Ergebnisse bieten demnach Aufschluss über die sprachlichen Veränderungen in einem dramagrammatischen Angebot im Gegensatz zu anderen Angeboten, die Kontrollkinder im Ganztagsschulbetrieb am Nachmittag wahrnehmen ‚müssen', aber deren Inhalt freiwillig gewählt wird. Diese konnten zwar performativer, durften aber nicht sprachförderlicher Natur sein.

10 Fazit

Mit dem Studiendesign wurde versucht, die wissenschaftlichen Gütekriterien bestmöglich einzuhalten, wozu der Einsatz standardisierter Messverfahren beitrug. Gleichwohl ist es in Fragen der Objektivität herausfordernd, wenn ein Interventionsprojekt Förderkraft und Forscherin in einer Person vereint. Dem Ausgang einer sprachförderlichen Intervention ist sie nicht neutral gegenüber eingestellt, sondern die Interventionsgruppen sollen sich in ihrem Sinne sprachlich entwickeln. Jedoch kann durch diese einzelne Person eine vergleichbare Qualität und Quantität des sprachförderlichen Inhalts gesichert werden. Standardisierte Sprachstandsmessverfahren können dabei als valide betrachtet werden, sodass davon ausgegangen wird, dass sie Sprachzuwächse sicher erfassen können. Zudem ist in diesen Diagnostiken von einer hohen Reliabilität durch die Messgenauigkeit der quantitativen Daten auszugehen. Weiterhin ist durch das Formen von Interventions- und Kontrollgruppen sowie die dreifachen Messwiederholungen von einer akzeptablen Kausalität auszugehen, obwohl keine Möglichkeit zur Randomisierung der Kinder bestand. Hier half es, dass sich durch die unterschiedliche Schul- und Klassenherkunft der Teilnehmer:innen verzerrende Effekte ausschließen ließen, die sich beispielsweise bei einer Kohorte aus einer Regelklasse im selben Fachunterricht zeigen können. Die Ergebnisse lassen sich aus diesem Grund für die Fremd- und Zweitspracherwerbsforschung nutzen, denn sie speisen sich aus einer annehmbar großen Stichprobe mit sowohl quantitativen als auch qualitativen Daten unterschiedlicher Diagnostiken, die über zwei Jahre hinweg mit drei Messwiederholungen generiert wurden. Diese Ergebnisse nicht nur des hier skizzierten Anfänger:innenniveaus, sondern der gesamten Dissertationsschrift tragen dazu bei, die Forschungslücke in Bezug auf Langzeitwirkungen additiver dramagrammatischer Förderung zu schließen. Es konnte in diesem Artikel gezeigt

werden, dass sich Sprachanfänger:innen in der Theater-AG deutlicher als die Kontrollkinder im Bereich der Grammatik entwickeln (kleinere p-Werte, größere Effekte) und darüber hinaus auch im Schreiben und Lesen. Die besonders großen Entwicklungen die Lesefertigkeit betreffend, sowohl bei den Anfänger:innen als auch in den übrigen Niveaus, war nicht erwartet worden (Vogel, in Vorb.). Sie sprechen insbesondere seit den aktuellen Befunden zur Lesekompetenz von Grundschüler:innen in der Bundesrepublik (Ludewig et al., 2022, IGLU 2021, McElvany et al., 2023).[12] für den Einsatz einer zusätzlichen dramapädagogischen Förderung – ganz allgemein, aber im Besonderen für Kinder mit sprachlichem Förderbedarf in sprachhomogenen Gruppen. Mit additiven Ansätzen wie dem oben beschriebenen Interventions-Theater-AG-Projekt verbindet sich die Hoffnung, kompensierend auf ungleiche Bildungschancen zu wirken sowie einen daraus resultierenden Matthäus-Effekt zu vermeiden, sodass letztlich sprachlich schwächeren Schulkindern entgegengekommen wird.[13] Dabei ist zu bedenken, dass kurzfristige Effekte von Fördermaßnahmen zwar mitunter in ebenso kurzer Dauer zu erzielen sind, es aber im Gesamten mit dramapädagogischen Ansätzen

> [s]o schnell […] nicht [geht], im Gegenteil: Kulturbildende Prozesse sind langwierig und am Anfang uneffizient, sie erfordern von den Beteiligten Geduld, Sorgfalt und Beharrungsvermögen und verlangen nach einem mit Brutwärme vergleichbaren spezifischen Mikroklima. (Huber, 2003, S. 73)

Bibliografie

Apeltauer, E. (2013). Neue Perspektiven sprachlicher Frühförderung. *Flensburger Papiere zur Mehrsprachigkeit und Kulturenvielfalt im* Unterricht (59/60).

Arndt, P. A. & Sambanis, M. (2017). *Didaktik und Neurowissenschaften – Dialog zwischen Wissenschaft und Praxis.* Narr.

Baumert, J., Baden-Württemberg/Ministerium für Kultus, Jugend und Sport (2011). *Expertenrat „Herkunft und Bildungserfolg". Empfehlungen für Bildungspolitische Weichenstellungen in der Perspektive auf das Jahr 2020* (BW2020). MfKJS Stuttgart.

Becker, T. (2011). *Schriftspracherwerb in der Zweitsprache. Eine qualitative Längsschnittstudie.* Schneider.

Berendes, K., Dragon, N., Weinert, S., Heppt, B.& Stanat, P. (2013). Hürde Bildungssprache? Eine Annäherung an das Konzept „Bildungssprache" unter Einbezug aktueller empirischer

[12] Ein Viertel der Viertklässler:innen erreicht die Mindeststandards im Lesen nicht (McElvany et al., 2023, S. 17).
[13] Dass Kinder mit höherem sozioökonomischen Status oder einsprachig aufwachsende Schüler:innen sowohl mit besseren Voraussetzungen zur Schule kommen als auch gleichzeitig mehr dazulernen als Kinder mit niedrigerem sozioökonomischen Status bzw. mehrsprachig aufwachsende Kinder, wird unter anderem unter dem Matthäus-Effekt verstanden und von Heppt & Stanat (2020), Volodina et al. (2020) sowie Volodina et al. (2021) belegt.

Forschungsergebnisse. In A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik: interdisziplinäre Perspektiven* (S. 17-41). Waxmann.

Bernstein, N. (2021). Bericht: Zeit für die Künste, eine führende Rolle zu übernehmen: Zur Podiumsdiskussion „Performatives Lehren und Lernen“ im Rahmen der „Sommerlichen FaDaF-Literaturtage“. *Scenario, 15*(2), 115-124. https://doi.org/10.33178/scenario.15.2.9

Boorsma, G. & Moraitis, A. (2007). *„Zeitsprung". Ein Sprachförder- und Theaterprojekt für Essener Kinder mit Migrationshintergrund. Eine Kooperation mit der Stadt Essen, der Zeche Carl e.V., der Universität Duisburg-Essen und dem Theaterpädagogischem Zentrum Ruhr e.V.*

Bora, S. F. (2017). *Beyond the mere word. Exploring the language of Drama through text- and performance-based approaches for developing L2 oral skills*. Dissertation University of Essex.

Bortz, J. & Döring, N. (unter Mitarbeit von S. Pöschl) (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Springer. https://doi.org/10.1007/978-3-642-41089-5

Bryant, D. & Zepter, A. L. (2022). *Performative Zugänge zu Deutsch als Zweisprache (DaZ). Ein Lehr- und Praxisbuch.* Narr Francke Attempto. https://doi.org/10.24053/9783823395133

Bryant, D. & Rummel, S. C. (2015). Nachhaltige dramapädagogische Sprachförderung für Grundschulkinder mit DaZ. *Scenario, 9*(2), 7-32. https://doi.org/10.33178/scenario.9.2.2

Bryant, D. (2012). DaZ und Theater. Der dramapädagogische Ansatz zur Förderung der Bildungssprache. *Scenario, 6*(1), 27-55. https://doi.org/10.33178/scenario.6.1.3

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N.H. Hornberger (Hrsg.), *Encyclopedia of language and education*, *Volume 2: Literacy* (2. Aufl., S. 71-83). Springer.

Dragović, G. (2019a*). Fremdsprachenunterricht (ent-)dramatisieren. Eine empirische Untersuchung zur Effizienz des dramapädagogischen Ansatzes im schulischen DaF-Unterricht mit speziellem Fokus auf Grammatik*. Dissertation Universität Freiburg i. d. Schweiz.

Dragović, G. (2019b). Zwischen Anspruch und Wirklichkeit: Dramapädagogik im schulischen Fremdsprachenunterricht. In S. Giebert & E. Göksel (Hrsg.), *Dramapädagogik-Tage 2018/Drama in Education Days 2018 – Conference Proceedings of the 4th Annual Conference on Performative Language Teaching and Learning* (S. 33-42). https://dramapaedagogik.de/wp-content/uploads/proceedings2018/final.pdf

Delius, K. (2020). *Förderung der Sprachkompetenz durch Synthese von generischem Lernen und Dramapädagogik. Eine Design-Based Research-Studie im Englischunterricht.* J.B. Metzler. https://doi.org/10.1007/978-3-662-62738-9

Domkowsky, R. & Walter, M. (2012). Was kann Theater? Ergebnisse empirischer Wirkungsforschung. *Scenario, 6*(1), 104-127. https://doi.org/10.33178/scenario.6.1.7

Even, S. (2003). *Drama Grammatik. Dramapädagogische Ansätze für den Grammatikunterricht Deutsch als Fremdsprache.* Iudicium.

Fox, A. V. (2013). *TROG-D. Test zur Überprüfung des Grammatikverständnisses.* Schulz-Kirchner.

Grum, U. & Zydatiß, W. (2016). Statistische Verfahren – Einleitung. In D. Caspari, F. Klippel, M. Legutke & K. Schramm (Hrsg.), *Forschungsmethoden in der Fremdsprachendidaktik. Ein Handbuch* (S. 319-324). Narr Francke Attempto.

Haberzettl, S. (2007). Progression im ungesteuerten und im gesteuerten Erwerb. In B. Ahrenholz (Hrsg.)*, Kinder mit Migrationshintergrund. Spracherwerb und Fördermöglichkeiten* (2. Aufl., S 203-220). Fillibach.

Hensel, A. (2020*). Fremdsprachenunterricht als Ereignis. Zur Fundierung einer performativ-ästhetischen Praxis* (Edition SCENARIO Bd. IV). Schibri.

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,* 903-928. https://doi.org/10.1007/s11218-022-09704-8

Heppt, B. & Stanat, P. (2020). Development of academic language comprehension across primary school: Findings for German monolinguals, bilinguals, and second language learners. *Virtual paper presentation at the Spring Workshop of the College for Interdisciplinary Educational Research* (CIDER).

Holl, E. (2011): *SPRACH-FLUSS – Theaterübungen für Sprachunterricht und interkulturelles Lernen.* Hueber.

Huber, R. (2003). *Im Haus der Sprache wohnen. Wahrnehmung und Theater im Fremdsprachenunterricht.* Max Niemeyer. https://doi.org/10.1515/9783110918380

Janzen Ulbricht, N. (2022). Ein Theaterexperiment oder ein Experiment im Theater? In S. Giebert & E. Göksel (Hrsg.), *Dramapädagogik-Tage 2020. Conference proceedings of the 6th annual conference on performative language teaching and learning* (S. 118-133). https://doi.org/10.25656/01:25321

Jeuk, S. (2018). Schriftspracherwerb und Alphabetisierung in der Zweitsprache im Grundschulalter. In Grießhaber, W., Schmölzer-Eibinger, S., Roll, H. und Schramm, K. (Hrsg.), *Schreiben in der Zweitsprache Deutsch*, (S. 49-62). De Gruyter/ Mouton. https://doi.org/10.1515/9783110354577

Kinze, J. (2012). Das HamburgerTheaterSprachCamp – Methoden und Ergebnisse der Evaluation. *Scenario, 6*(1), 85-102. https://doi.org/10.33178/scenario.6.1.6

Klempin, C. & Wirag, A. (2020). Wie wirkt Dramapädagogik im Fremdsprachenunterricht? Ein Leitfaden zum Aufbau, zur Durchführung und Auswertung von Experimentalstudien. *Scenario, 14*(2), 96-115. https://doi.org/10.33178/scenario.14.2.6

Klempin, C. (2019): *Reflexionskompetenz von Englischlehramtsstudierenden im Lehr-Lern-Labor-Seminar. Eine Interventionsstudie zur Förderung und Messung*. J.B. Metzler. https://doi.org/10.1007/978-3-476-05120-2

Kniffka, G. & Roelcke, T. (2015). *Fachsprachenvermittlung im Unterricht.* UTB.

Kniffka, G. (2013). Deutsch für Kinder. Früher Fremd- und Zweitsprachenerwerb Deutsch. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*, *48*, 3-12. https://doi.org/10.37307/j.2194-1823.2013.48.03

Kniffka, G. & Siebert-Ott, G. (2012). *Deutsch als Zweitsprache. Lehren und Lernen* (3. aktual. Aufl.). UTB.

Küppers, A. & Walter, M. (2012). Theatermethoden auf dem Prüfstand der Forschung. *Scenario – Zeitschrift für Drama- und Theaterpädagogik in der Fremd- und Zweitsprachenvermittlung 2012/1, 1-9.* https://doi.org/10.33178/scenario.6.1.1

Küppers, A. (2011). The DICE Consortium (2010), The DICE has been cast. *Scenario*, *5*(1), 107-112. https://doi.org/10.33178/scenario.5.1.12

Lenhard, W. & Lenhard, A. (2016). *Berechnung von Effektstärken*. Abgerufen unter: https://www.psychometrica.de/effektstaerke.html. Psychometrica.

Lenhard, W. & Schneider, W. (2006*). ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler.* Hogrefe.

Ludewig, U., Schlitter, T., Lorenz, R., Kleinkorres, R., Schaufelberger, R., Frey, A. & McElvany, N. (2022). *Die COVID-19 Pandemie und Lesekompetenz von Viertklässler*innen. Ergebnisse der IFS-Schulpanelstudie 2016-2021.*

Lutzker, P. (2016). Die Wiederentdeckung der Erfahrung beim Lernen und Lehren von Fremdsprachen. In S. Even & M. Schewe (Hrsg.), *Performatives Lehren – Lernen – Forschen* (S. 67-86). Schibri.

McElvany, N., Lorenz, R., Frey, A., Goldhammer, F., Schilcher, A. & Stubbe, T. C. (Hrsg.) (2023). *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre*. Waxmann. https://doi.org/10.31244/9783830997009

McGraw, K. O. & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin, 111*(2), 361-365. https://doi.org/10.1037/0033-2909.111.2.361

Morris, S.B. (2008). Estimating Effect Sizes From Pretest-Posttest-Control Group Designs. *Organizational Research Methods,11*(2), 364-386. https://doi.org/10.1177/1094428106291059

Oelschläger, B. (2008). Theaterpädagogik und Sprachförderung - im Nachklang zweier Tagungen. Theaterpädagogische Methoden und Spracherwerb - Fachtagung an der Bundesakademie für kulturelle Bildung Wolfenbüttel, 14.-15.5.2007; Spielen Sprechen Lernen – Theaterpädagogik und Sprachförderung, Theaterpädagogisches Fachforum Sichten IX, 16.-17.11.2007, Freizeit- und Erholungszentrum/FEZ-Berlin. *Scenario, 2*(1)*,* 112-117. https://doi.org/10.33178/scenario.2.1.7

Piazzoli, E. (2011): Process drama: the use of affective space to reduce language anxiety in the additional language learning classroom. *RiDE: The Journal of Applied Theatre and Performance 16(4)*, 557-573. https://doi.org/10.1080/13569783.2011.617104

Plonsky, L. & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning 10/2014* Language Learning Research Club, University of Michigan, 1-35. https://doi.org/10.1111/lang.12079

Reich, H. H., Roth, H.-J.& Gantefort, C. (2008). *Das Tulpenbeet. Ein profilanalytisches Instrument am Übergang vom Primar- in den Sekundarbereich.* Universität Hamburg.

Rösch, H. (2007). Das Jacobs-Sommercamp – neue Ansätze zur Förderung von Deutsch als Zweitsprache. B. Ahrenholz (Hrsg.), *Kinder mit Migrationshintergrund. Spracherwerb und Fördermöglichkeiten* (S. 287-302). Fillibach.

Schappert, P. (2020). *Mit Dramagrammatik zu grammatikalischer Kompetenz. Eine empirische Untersuchung im Rahmen von Alphabetisierungskursen für Erwachsene*. Tectum. https://doi.org/10.5771/9783828874886

Schewe, M. (1993). *Fremdsprache inszenieren.* Zur Fundierung einer dramapädagogischen Lehr- und Lernpraxis. Carl von Ossietzky Universität Oldenburg.

Schmölzer-Eibinger, S. (2018). Literalität und Schreiben in der Zweitsprache. In W. Grießhaber, S. Schmölzer-Eibinger, H. Roll & K. Schramm (Hrsg.), *Schreiben in der Zweitsprache Deutsch* (S. 3-16). https://doi.org/10.1515/9783110354577

Schulz, K. F., Altman, D. G. & Moher, D. (2011). CONSORT 2010: Aktualisierte Leitlinie für Berichte randomisierter Studien im Parallelgruppen-Design. *DMW - Deutsche Medizinische Wochenschrift 136/8*, e20-e23. https://doi.org/10.1055/s-0031-1272980

Siegmund, B. (2022*). Sprachbildung im naturwissenschaftlichen Sachunterricht. Eine Interventionsstudie zur Wirksamkeit fachintegrierter Sprachbildung nach dem Scaffolding-Ansatz und mit Focus-on-Form-Strategien*. UB Verlag. Reihe: Sprachlich-literarisches Lernen und Deutschdidaktik. https://doi.org/10.46586/SLLD.253

Sirisrimangkorn, L. & Suwanthep, J. (2013). The effects of integrated drama-based role play and student teams achievement division (STAD) on students' speaking skills and affective involvement. *Scenario, 7(2),* 64-78. https://doi.org/10.33178/scenario.7.2.5

Stanat, P., Baumert, J. & Müller, A. (2005). Förderung von deutschen Sprachkompetenzen bei Kindern aus zugewanderten und sozial benachteiligten Familien. Evaluationskonzeption für das Jacobs-Sommercamp Projekt. *Zeitschrift für Pädagogik, 51*(6), 856-875.

Vogel, S. C. (in Vorb.). Einsatz und Evaluation dramapädagogischer, fachsensibler Sprachförderung in schulischer Theater-AG für Grundschulkinder mit DaZ. Eine longitudinale Interventionsstudie.

Volodina, A., Heppt, B. & Weinert, S. (2021). Relations between the comprehension of connectives and school performance in primary school. *Learning and Instruction, 74*(2021), 101-430. https://doi.org/10.1016/j.learninstruc.2020.101430

Volodina, A. & Weinert, S. (2020). Comprehension of connectives: Development across primary school age and influencing factors. *Frontiers in Psychology, 11:118,* 1-15. https://doi.org/10.3389/fpsyg.2020.00814

Volodina, A., Weinert, S. & Mursin, K. (2020). Development of academic vocabulary across primary school age: Differential growth and influential factors for German monolinguals and language minority learners. *Developmental Psychology, 56*, 922-936. https://doi.org/10.1037/dev0000910

Walter, M. (2013). In diesem Ansatz steckt noch viel mehr drin! *Scenario, 7(2),* 112-117. https://doi.org/10.33178/scenario.7.2.10

Weyer, D. (2019). *Investigating trust in a multilingual theatre project: Potentials for a Humanising Pedagogy.* Dissertation Université du Luxembourg. http://hdl.handle.net/10993/38971

Wirag, A. & Surkamp, C. (2022). Boon or burden? Drama pedagogy elements and their relation to foreign-language anxiety in EFL drama clubs. In S. Giebert & E. Göksel (Hrsg.), *Dramapädagogik-Tage 2020. Conference proceedings of the 6th annual conference on performative language teaching and learning* (S. 69-85)

Wirag, A. (2021). Hensel, Alexandra. 2020.Fremdsprachenunterricht als Ereignis. Zur Fundierung einer performativ-ästhetischen Praxis. *Zeitschrift für Angewandte Linguistik*, *2021*(75), 454-462. https://doi.org/10.1515/zfal-2021-2075