

Bibliographischer Hinweis sowie Verlagsrechte bei den online-Versionen der DD-Beiträge:



**Halbjahresschrift für die Didaktik
der deutschen Sprache und
Literatur**

<http://www.didaktik-deutsch.de>
23. Jahrgang 2018 – ISSN 1431-4355
Schneider Verlag Hohengehren
GmbH

Melanie Bangel / Astrid Müller

**STRUKTURORIENTIERTES
RECHTSCHREIBLERNEN.
ERGEBNISSE EINER
INTERVENTIONSSTUDIE ZUR
WORTSCHREIBUNG IN KLASSE 5
MIT BLICK AUF SCHWACHE
LERNER/-INNEN**

In: Didaktik Deutsch. Jg. 23. H. 45. S. 29-49.

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder andere Verfahren – reproduziert oder in eine von Maschinen, insbesondere von Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. – Fotokopien für den persönlichen und sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopien hergestellt werden.

Melanie Bangel & Astrid Müller

STRUKTURORIENTIERTES RECHTSCHREIBLERNEN. ERGEBNISSE EINER INTERVENTIONSSTUDIE ZUR WORTSCHREIBUNG IN KLASSE 5 MIT BLICK AUF SCHWACHE LERNER/-INNEN

Zusammenfassung

Mit der vorliegenden Interventionsstudie sollte überprüft werden, ob und wie sich die Rechtschreibleistung von Fünftklässler/-innen durch einen Unterricht entwickelt, der sich an der Struktur des Schriftsystems orientiert, wie sie durch graphematische Forschungen fundiert wurde.

Im Schuljahr 2016/17 wurden 15 Klassen nach einem Konzept unterrichtet, das die phonographisch-silbischen und morphologischen Regularitäten der Wortschreibung berücksichtigte. Als Vergleichsklassen dienten zehn Klassen, die schulbuchorientiert an der Wortschreibung arbeiteten. Im Mittelpunkt des Beitrags steht die Leistungsentwicklung der 25% schwächsten Rechtschreiber/-innen der Stichprobe. Die Ergebnisse zeigen, dass die schwachen Rechtschreiber/-innen der Interventionsgruppe ($n = 50$) sowohl direkt im Anschluss an die Intervention als auch vier Monate später signifikant größere Lernfortschritte aufweisen als die vergleichbare Gruppe der Kontrollklassen ($n = 40$). Die Effektstärken liegen im mittleren bis großen Bereich (Nachtest: $\eta^2 = .08-.12$, verzögerter Nachtest: $\eta^2 = .07-.13$). Das untere Leistungsquartil der Interventionsgruppe profitiert damit stärker als alle anderen Gruppen von dem Lernangebot.

Abstract

The purpose of the present intervention study was to examine whether and how the spelling performance of fifth graders develops through teaching based on the structure of the graphic system, as established by graphemic research.

In the school year 2016/17, 15 classes were taught according to a concept that took into account the phonographic-syllabic and morphological regularities of word spelling. Ten classes that followed a schoolbook-oriented approach served as a control group.

The focus of the article is on the performance of the weakest 25% of students in the sample. The results show that the weak writers in the treatment group ($n = 50$) had a significantly greater learning progress both directly after the intervention and four months later than comparable students in the control group ($n = 40$). The effect sizes are in the medium to large range (post-test: $\eta^2 = .08-.12$, delayed post-test: $\eta^2 = .07-.13$). Consequently, the lower quartile of the treatment group showed a greater benefit than any other group in the sample.

1. Einleitung

Um die orthographischen Leistungen von Schülerinnen und Schülern an deutschen Schulen steht es nach wie vor nicht gut. So liegt laut IQB-Länderbericht 2016 der Anteil der Schüler/-innen in Jahrgang 4, die den Regelstandard im Kompetenzbereich Orthographie erreichen oder übertreffen, bei nur 54%. Damit fällt ihr Anteil um zehn Prozentpunkte schlechter und signifikant geringer als im ersten Länderbericht 2011 aus. Zudem hat sich der Anteil derjenigen Viertklässler/-innen, die den Mindeststandard nicht erreichen, zwischen dem ersten und dem zweiten IQB-Länderbericht mit 8% ebenfalls signifikant vergrößert und liegt deutschlandweit bei fast 21% (vgl. Weirich et al. 2017).

Differenzierte und evidenzbasierte Aussagen dazu, wie die Entwicklung orthographischer (Teil-)kompetenzen wirkungsvoll unterstützt werden kann, liegen aktuell nur vereinzelt vor (vgl. z.B. Wahl et al. 2017). Ergebnisse aus der graphematischen Forschung und darauf basierende Modellierungen zur Schriftaneignung legen nahe, dass eine schriftstrukturbasierte Ausrichtung des Unterrichts gerade denjenigen Lerner/-innen zugutekommt, die aufgrund eingeschränkter sprachlicher und/oder kognitiver Voraussetzungen nicht zur eigenständigen Entdeckung der Schriftstruktur als Grundlage für erfolgreiches Rechtschreiblernen in der Lage sind (vgl. u. a. Bredel 2016, Hinney 2014).

Diese Lücke in der empirischen orthographiedidaktischen Forschung will die Interventionsstudie,¹ aus der wir hier einzelne Ergebnisse berichten, für den Bereich der Wortschreibung schließen helfen. Der Fokus des Beitrags liegt auf der Leistungsentwicklung der schwachen Rechtschreiber/-innen, denn wir gehen davon aus, dass gerade diese Schülergruppe besonders von sachangemessenen Vorstrukturierungen des Lerngegenstands *Schriftsprache* profitiert. In einer Vorgängerstudie konnten wir bereits für die Fähigkeit zum Dekodieren komplexer Wörter zeigen, dass schriftstrukturorientierte Zugänge zu einem besonders großen Lernzuwachs in der Gruppe der schwächeren Lerner/-innen (hier besonders des 2. Quartils) führten (vgl. Bangel/Müller 2015).

2. Wortstruktur als Lerngegenstand

Die graphematische Forschung hat entscheidend zu einem veränderten Blick auf das Schriftsystem des Deutschen beigetragen (vgl. u. a. Eisenberg seit 1983, Maas seit 1992). Als wesentliche Erkenntnis für orthographiedidaktische Überlegungen lässt sich daraus mit Eisenberg (2016) zusammenfassen: „Der Schreiber kann die Orthografie seiner Sprache nicht nur beherrschen, er kann sie auch verstehen“

¹ Vgl. DFG-Projekt „Zur Effektivität strukturorientierter Zugänge zum Schriftsystem als Grundlage für orthographisches Lernen. Interventionsstudie im Rechtschreibunterricht in fünften Klassen“, Geschäftszeichen: MU 2832/4-1, Laufzeit: 10/2015 bis 09/2019 (<http://gepris.dfg.de/gepris/projekt/279112504>).

(ebd.: 65). Unter der Perspektive einsichtsvollen Lernens sind für die Wortschreibung u. a. folgende Ergebnisse der sprachwissenschaftlichen Forschung relevant:

Der Kernbereich der Wortschreibung ist durch das Zusammenwirken von phonographisch-silbischen und morphologischen Regularitäten bestimmt, wobei die musterbildende Basisform die zweisilbige Wortform ist, die aus Haupt- und Reduktionssilbe besteht (vgl. Eisenberg 2013, Bredel 2010, Hinney 2017):

Hauptsilbe			Reduktionssilbe		
AR	K	ER	AR	K	ER
l	a		d	e	n
l	a	n	d	e	n

AR = Anfangsrand
K = Kern
ER = Endrand

↓
Silbengrenze

Abb. 1: Aufbau der Schreibsilbe im Kernbereich

Die Hauptsilbe umfasst i. d. R. den größten Teil des Wortstamms, die Reduktionssilbe enthält im Anfangsrand den „Stammrest“ und ist im Kern auf das Graphem <e> beschränkt. Im Gesprochenen weisen diese Wörter ein trochäisches Betonungsmuster auf. Diese Basisform lässt sich in vier bzw. fünf grundlegende Baumuster differenzieren:

- Zu Baumuster I gehören Wörter, in denen der Silbenendrand der Hauptsilbe nicht besetzt ist (*malen, schreiben, reden*). Die Leseinstruktion für diese Wörter mit offener Hauptsilbe lautet: Das Vokalphonem der betonten Silbe wird gespannt und lang gesprochen.
- Zu Baumuster II gehören Wörter, deren Silbenendrand durch ein Konsonantgraphem besetzt ist (*bremsen, rechnen, finden*). Die Hauptsilbe ist geschlossen. Das Vokalphonem der betonten Silbe wird folglich ungespannt und kurz gelesen.
- Zu Baumuster III gehören Wörter mit Silbengelenkschreibung (*rennen, fallen, hoffen, backen, ritzen*). Für das Lesen erhalten wir aus der geschriebenen Form einen Hinweis auf die Ungespanntheit und Kürze des Vokalphonems der betonten Silbe, wobei der doppelte Konsonantbuchstabe nicht doppelt artikuliert wird, denn im Gesprochenen liegt ein ambisilbischer Konsonant vor (vgl. Eisenberg 2013).
- Baumuster IV bilden Wörter mit einem <h> im Anfangsrand der geschriebenen Reduktionssilbe (*sehen, gehen, fliehen*). Das <h> dient der besseren visuellen Strukturierung des Wortes. Es wird in der Standardaussprache nicht artikuliert.
- Zum erweiterten Kernbereich (Baumuster V) gehören Wörter mit Dehnungs-h (*fahren, dehnen, zählen*), das die Funktion hat, die Hauptsilbe optisch zu verlängern. Es tritt ausschließlich (aber nur in ca. 50% der möglichen Fälle) dann auf, wenn die Reduktionssilbe mit <l, m, n, r> beginnt (vgl. Hinney 2017).

Die morphologischen Regularitäten des deutschen Schriftsystems zeigen sich u. a. darin, dass die phonographisch-silbischen Informationen in flektierten, auch einsilbigen, Wortformen (*du rennst, er steht*) sowie in Komposita (*Schwimmflügel, Gehweg*) und Derivaten (*lieblich, sonnig*) erkennbar bleiben. Die zweisilbige Basisform ist somit u. a. dafür verantwortlich, dass die Auslautverhärtung nicht verschriftet wird (*Bad, Weg*). Morphologisch erklärbar sind außerdem die *ä/äu*-Schreibung (*Bäder, Träume* wegen *baden, Traum*), die Beibehaltung der Silbengelenkschreibung, der *ß*-Schreibung und des silbeninitialen *h* in einsilbigen und komplexen Wortformen (*rennt, Rennrad, grüßt, Grußwort, geht, Gehweg*) sowie die Konstant-schreibung von Morphemen an der Morphemgrenze (*annehmen, auffassen*).

Die beschriebenen Regularitäten bilden gleichzeitig den Hintergrund für didaktisch-methodische Entscheidungen in der zu berichtenden Interventionsstudie: Da der Aufbau der Schreibsilbe in der zweisilbigen Basisform in erster Linie für das Lesen funktional ist, „müssen die Regularitäten aus Sicht des Lesers verstanden werden“ (Hinney 2017: 263). Die unterrichtliche Basis bilden folglich Wörter aus dem (erweiterten) Kernbereich des deutschen Schriftsystems, die im Hinblick auf ihre lesesteuernde Funktion untersucht werden. Die Baumuster geben die Lernprogression für die Erschließung der phonographisch-silbischen Regularitäten vor, der Einsatz unterschiedlicher Analysestrategien für das gelenkte Entdecken dieser Regularitäten erfolgt ausgehend vom zweisilbigen Baumuster zu einsilbigen bzw. wortbildungsmorphologisch komplexen Wortformen:

- (1) zweisilbige Basisform untersuchen (malen, finden, fallen),
 - (2) zweisilbige Basisform bilden (z. B.: Weg ↪ Wege, Fuß ↪ Füße, kommt ↪ kommen),
 - (3) einsilbige Form ableiten (Bäume ↪ Baum),
 - (4) Wortbausteine segmentieren und Stämme und Affixe identifizieren
(z. B. dünn|häut|ig)
- (vgl. Abschnitt 4).

3. Zielsetzung und Hypothesen²

Der Rechtschreibunterricht in den Interventionsklassen des Projekts orientierte sich an den beschriebenen Regularitäten der deutschen Wortschreibung. Das Projekt sollte die Frage beantworten helfen, ob ein an der Struktur der Schrift orientierter Zugang zur Wortschreibung die Leistungsentwicklung in der Wortschreibung besser unterstützt als ein phänomen- und regelorientierter Unterricht. Ein zweiter Teil der Intervention sollte der Entwicklung der satzinternen Großschreibung durch

² Die Beschreibung des Forschungsdesigns (Punkt 3 und 4) findet sich ausführlich in einem bereits veröffentlichten Beitrag in der Zeitschrift „Unterrichtswissenschaft“ (vgl. Bangel/Müller 2018a).

einen syntaxbezogenen Zugang dienen. Da die satzinterne Großschreibung aus zeitlichen Gründen nicht in allen Klassen entsprechend der Projektplanung behandelt werden konnte, wird darauf in diesem Beitrag nicht eingegangen.

Für den Unterricht zur Wortschreibung wurde eine entsprechende Konzeption entwickelt, die die Basis für Lehrerfortbildungen sowie für den Umgang mit dem unterrichtsbegleitenden Schülerarbeitsheft bildete. Die Konzeption ist so angelegt, dass sie im regulären Deutschunterricht umgesetzt werden kann. Forschungsleitend war die folgende Hypothese:

Fünftklässlerinnen und Fünftklässler, die über einen Zeitraum von ca. 16 Wochen nach einem an den Regularitäten des deutschen Schriftsystems orientierten Konzept unterrichtet werden, erzielen in einem Test zur Erfassung der Leistungen in der *Wortschreibung* signifikant höhere Leistungszuwächse als Schülerinnen und Schüler, die einen „herkömmlichen“ Rechtschreibunterricht erhalten haben.

Darüber hinaus wurden qualitative Daten mit einer Teilstichprobe erhoben, mit denen überprüft werden sollte, ob das erworbene Wissen auf Anwendungssituationen transferierbar ist und ob sich die Intervention auf das metakognitive Wissen der Lernenden auswirkt. Besonders relevant und interessant sind, wie eingangs begründet, die Ergebnisse der schwachen Rechtschreiber/-innen. Sie stehen im Mittelpunkt dieses Beitrags.

4. Methode

4.1 Stichprobenbeschreibung: Gesamtstichprobe

Die Intervention wurde im Schuljahr 2016/2017 in 15 fünften Klassen aus drei Hamburger Stadtteilschulen durchgeführt, die vergleichbare sozioökonomische Voraussetzungen aufweisen. Als Kontrollgruppe dienten zehn Klassen aus drei Hamburger Stadtteilschulen mit vergleichbaren sozioökonomischen Voraussetzungen. Die Schulen wurden im Frühjahr 2016 entweder als Interventions- oder Kontrollschule angefragt. Bei der Zuordnung der Schulen wurde beachtet, dass die Gruppen hinsichtlich der Rahmenbedingungen vergleichbar sind (Erfahrung der Lehrkräfte, Belastung durch andere Projekte). Um einschätzen zu können, ob die motivationalen Voraussetzungen bzgl. des Orthographieunterrichts bei den Lehrkräften der Interventions- und Kontrollschulen vergleichbar sind, fand vor Beginn der Intervention an jeder Schule ein Treffen zwischen den Deutschlehrer/-innen der zukünftigen fünften Klassen und der Projektleiterin statt. Es zeigte sich, dass alle am Projekt beteiligten Lehrkräfte dem Orthographieunterricht einen hohen Stellenwert einräumen und neuen Ansätzen offen gegenüberstehen.

Die Klassengröße variierte zwischen 19 und 26 Schüler/-innen ($M = 23.00$, $SD = 2.00$). Für die statistische Auswertung wurden nur Lernende einbezogen, von denen Testergebnisse zu allen drei Testzeitpunkten vorliegen. Da vier der Interventionsklassen aufgrund unzureichender Treatmentintegrität aus der Analyse her-

ausgenommen wurden (vgl. 4.5), umfasst der Datensatz Testwerte von 364 Schüler/-innen aus insgesamt 21 Klassen (elf Interventions- und zehn Kontrollklassen): 186 Jungen und 178 Mädchen mit einem Durchschnittsalter von 10,28 Jahren. 58% der Schüler/-innen der Interventionsgruppe (IG) gab an, dass in ihren Familien ausschließlich eine andere Sprache oder neben Deutsch noch eine andere Sprache gesprochen wird, in der Kontrollgruppe (KG) traf dies auf 47% zu. Der Unterschied ist nicht signifikant, er wurde jedoch als Hintergrundvariable *Sprachstatus* im Rahmen einer Kovarianzanalyse bei der Wirksamkeitsprüfung berücksichtigt. Für die Einschätzung des *sozioökonomischen Status* wurde die Angabe zur Anzahl der Bücher im Haushalt hinzugezogen. Für diese Variable liegen signifikante Gruppenunterschiede mit einer leicht höheren Ausprägung aufseiten der IG vor ($\chi^2(4, N = 364) = 15.74, p = .003$). Diese Unterschiede wurden statistisch kontrolliert, indem die Variable *häuslicher Bücherbesitz* als Kovariate bei der Wirksamkeitsprüfung modelliert wurde.

Um zu prüfen, ob fehlende Werte im Nachtest und im verzögerten Nachtest als zufällig fehlend betrachtet werden können, wurde der Stichprobenausfall in den vier Leistungsquartilen zwischen der Interventions- und der Kontrollgruppe verglichen. Dabei zeigten sich für keine der abhängigen Variablen systematische Unterschiede in der Anzahl der fehlenden Werte.

4.2 Stichprobenbeschreibung: unteres Leistungsquartil

Die 25% schwächsten Rechtschreiber/-innen der Gesamtstichprobe (Prozentrangband: 0.3 bis 24.7, $M = 12.9$, $SD = 7.32$) gehören mit Blick auf die Ergebnisse in der Hamburger Schreibprobe (HSP 4/5) zu den unteren 19% der Gesamtpopulation (Prozentrangband: 1.4 bis 19, $M = 8.99$, $SD = 6.06$). Die Gruppe setzt sich so zusammen: Aus der IG gehören 50 Schüler/-innen (21 Mädchen und 29 Jungen aus allen elf Klassen) zu dieser Gruppe, in der KG sind es 40 (27 Mädchen und 13 Jungen aus allen zehn Klassen). Davon gaben 60% an, zu Hause nur eine andere Sprache oder neben Deutsch noch eine andere Sprache zu sprechen. Hinsichtlich der Variablen *Geschlecht* und *Sprachstatus* unterscheiden sich die Gruppen nicht signifikant. Auch der *häusliche Bücherbesitz* ist in den Gruppen vergleichbar. Dennoch wurden diese Hintergrundvariablen bei der Wirksamkeitsprüfung statistisch kontrolliert.

4.3 Inhalt und Durchführung der Intervention

Die Intervention wurde von den Deutschlehrkräften durchgeführt. Sie gestalteten den Unterricht auf der Grundlage des Konzepts in Abhängigkeit von den konkreten Lernbedürfnissen der Klasse. In den Fortbildungen erhielten die Lehrkräfte Anregungen zur Arbeit mit dem Schülerarbeitsheft.³ Fortbildungsinhalte (vier Termine à 2,5 Stunden) waren deshalb: phonographisch-silbische und morphologische Regularitäten der Wortschreibung, Wortbildungstypen und morphologischer Auf-

³ Das Schülerarbeitsheft liegt in leicht veränderter Form vor: vgl. Bangel/Müller (2018b).

bau komplexer Wörter sowie Regularitäten der satzinternen Großschreibung (vgl. Tab. 1). Darüber hinaus wurden Anregungen zum entdeckenden und einsichtsvollen Lernen gegeben. Dazu gehörten z. B. Aufgaben zum Ordnen und Vergleichen von Wörtern nach bestimmten Kriterien (Anzahl der Silben, offene und geschlossene Silben usw.) und Aufgaben zum Erklären von Wortschreibungen (z. B. Erklären von Silbengelenkschreibungen).

Es sollte ca. 45 Minuten pro Woche nach dem Konzept gearbeitet werden. Der Stand der Arbeit sowie eventuell aufgetretene Schwierigkeiten wurden durch studentische Hilfskräfte, die den Unterricht begleiteten, protokolliert. Die Lehrkräfte der KG wurden parallel über Thema und zeitliche Struktur der Intervention informiert. Zudem wurde erfragt, womit im Rechtschreibunterricht in Klasse 5 gearbeitet wird und warum das Material ausgewählt wurde. Die Unterrichtsschwerpunkte in der KG wurden durch den Einsatz von Lehrerfragebögen kontrolliert (vgl. 4.5) und sind in Tabelle 1 zusammengefasst. In allen Kontrollklassen wurde entweder mit dem Schulbuch „Doppel-Klick“ und/oder dem Arbeitsheft „Abschreiben erwünscht“ gearbeitet. Da sich die Auswahl der Lerninhalte sowohl in diesen Materialien als auch in dem in der Intervention eingesetzten Arbeitsheft am Hamburger Bildungsplan orientiert, konnten nach vergleichender Sichtung keine Unterschiede in den zu behandelnden Themen gefunden werden. Der wesentliche Unterschied besteht in der Modellierung des Lerngegenstands. Während in der KG an einzelnen Rechtschreibphänomenen gearbeitet wurde (Wörter mit *ck*, Wörter mit *mm*, *nn* und *ss*), wurde der Lerngegenstand Schrift für die Schüler/-innen der IG so aufbereitet, dass Einsicht in die grundlegenden Regularitäten der Wortschreibung möglich war. Bei der Erarbeitung besonderer Markierungen (z. B. Silbengelenkschreibung oder silbeninitials h) wurde jeweils die Basisform herangezogen und anschließend durch Kontrolllesen überprüft, ob die Schreibung eine Markierung verlangt (vgl. Abschnitt 2). In den Materialien der KG findet hingegen keine systematische Orientierung an den Regularitäten der deutschen Wortschreibung statt, es werden stattdessen einzelne Phänomene isoliert und mit Blick auf den Kern- und Peripheriebereich durchmischt behandelt (z. B. Wörter mit *ll*, *nn*, *ss*: *denn*, *dass*, *sollten*). Rechtschreibstrategien wie das Ableiten zur Herleitung der *ä/äu*-Schreibung (z. B. *Baum* – *Bäume*) oder das Verlängern zur Herleitung der Nichtverschriftlichung der Auslautverhärtung (z. B. *Berg* – *Berge*) werden zwar auch in den Materialien der KG eingeführt, es finden sich aber keine Aufgaben zur Übertragung dieser Strategien auf die Schreibung von wortbildungsmorphologisch komplexen Wörtern. Es dominieren Aufgaben zum Abschreiben und Merken, zur Sprechsilbengliederung und zu Wortfamilien. Zudem wird die Rückführung des Einsilbers auf die zweisilbige Basisform nicht für die Überprüfung der Silbengelenkschreibung, der *ß*-Schreibung oder des silbeninitials h genutzt.

Die Befragung der Lehrkräfte der KG nach der Intervention ergab, dass die Klassen in ungefähr gleichem Umfang an der Wortschreibung arbeiteten wie die Klassen der IG. Der gravierendste Unterschied zwischen beiden Gruppen besteht im Hinblick

auf die dominierenden Aufgabenformate darin, dass in der KG vorrangig an Aufgaben zum Nachvollziehen von Regeln gearbeitet wurde („Nach einem kurzen Vokal (Selbstlaut) folgen meist zwei Konsonanten (Mitlaute): alle, denn, das Schloss“; Doppel-Klick 5, vgl. Feustel et al. 2009: 226). Daran schließen sich Aufgaben zum Schreiben von Wörtern, die zur Regel passen, an. In der IG hingegen sollten Regularitäten selbstständig entdeckt werden („Was ist eine Silbengelenkschreibung?“; vgl. Bangel/Müller 2018b: 15).

Tab. 1: Überblick über das Forschungsdesign

	IG (15 Klassen)	KG (10 Klassen)
April 2016	Auswahl der Schulen und Klassen	
Mai–Juni 2016	Lehrerfortbildungen	Information über Inhalt der Intervention
Sept. 2016	Vortest: Rechtschreibtest (HRT 5), Wortschatztest des CFT 20-R, freie Schreibproben (Geschichten), Interviews zu Schreibungen mit einer Teilstichprobe der IG und KG, Schülerfragebogen, Lehrerfragebogen (IG)	
Sept.–Okt. 2016	<p>Intervention Teil A: morphologisch einfache Wörter (Simplizia) untersuchen</p> <p>(zweisilbige Schlüsselwörter untersuchen und bilden)</p> <p>(ca. 8-10 Unterrichtsstunden)</p> <p>Ermitteln der Silbengrenze und der Silbenanzahl in vorgegebenen Wörtern Ermitteln der Silbenstruktur (vokalischer Silbenkern, Silbenanfangs- und -endründer) in Basiswörtern Unterscheiden von offenen und geschlossenen Silben in Basisformen Entdecken von „vererbten“ silbischen Informationen durch Überführen von einsilbigen und komplexen Wortformen in die prototypische zweisilbige Form (<i>fällt</i> → <i>fallen</i>)</p>	<p>schulbuchorientierte Arbeit an Rechtschreibphänomenen in morphologisch einfachen Wörtern und Vermitteln von Strategien und Arbeitstechniken mit Büchern und Arbeitsheften wie „Doppel-Klick 5“, „Abschreiben erwünscht“ und anderen Materialien zu diesen Schwerpunkten:</p> <p>„Sprechen – Hören – Schreiben“ (Doppel-Klick 5, S. 218f.) Sprechen, schreiben, merken: „Wörter mit ie“ (ebd., S. 224) Schreiben, hören, Basisform suchen: „Wörter mit ll, nn, ss“ (ebd., S. 226f.) Sprechen, schreiben: „Wörter mit ß“ (ebd., S. 232) Wörter ableiten und verlängern (ebd., S. 246f) Wörter nachschlagen, Training mit Wörterlisten, Faltdiktat, Laufdiktat, Rechtschreibkartei (ebd., S. 238f.)</p>
Nov.–Feb. 2017	<p>Intervention Teil B: zusammengesetzte und abgeleitete Wörter untersuchen</p> <p>(ca. 8 Unterrichtsstunden)</p> <p>Bilden von Wortfamilien und Untersuchen des Wortstammes (<i>Traum, verträumt, Traumschiff</i>) Ermitteln der Wortbausteine in Komposita sowie in prä- und suffigierten Wörtern (<i>großherzig</i>) Entdecken von „vererbten“ silbischen Informationen durch Bilden der Basisform (<i>großherzig: groß</i> → <i>größer, Herz</i> → <i>Herzen</i>)</p>	<p>Arbeit an Rechtschreibphänomenen in morphologisch komplexen Wörtern mit Materialien wie „Doppel-Klick 5“, „Abschreiben erwünscht“:</p> <p>„Wortfamilien mit ie“ (Doppel-Klick 5, S. 225) „die Wortfamilie fallen“ (ebd., S. 227) „Adjektive mit -ig und -lich“ (ebd., S. 228) „Wörter mit ver-, vor- und ent-“ (ebd., S. 236)</p>
Nov.–Feb. 2017	<p>Intervention Teil C: satzinterne Großschreibung</p> <p>(ca. 8 Unterrichtsstunden in einzelnen Klassen)</p> <p>Ermitteln der erweiterbaren Kerne von Nominalgruppe durch die Erweiterungsprobe</p>	<p>Arbeit an der satzinternen Großschreibung mit Materialien wie „Doppel-Klick 5“, „Abschreiben erwünscht“</p>
Feb. 2017	Nachtest: Rechtschreibtest (HRT 5)	
Feb.–Juni 2017	regulärer Deutschunterricht	
Juni 2017	verzögerter Nachtest: Rechtschreibtest (HRT 5), freie Schreibproben (Geschichten), Interviews über einzelne Schreibungen mit einer Teilstichprobe der IG und KG, Lehrerfragebogen (IG und KG)	

4.4 Erhebungsinstrumente

Da es zurzeit noch keinen frei zugänglichen standardisierten Rechtschreibtest gibt, der die unter 2 beschriebene Modellierung der Wortschreibung hinreichend abbildet, wurde ein Rechtschreibtest („Hamburger Rechtschreibtest: HRT 5“) entwickelt. Im Vergleich zu anderen Tests wie z.B. der „Hamburger Schreibprobe“ (vgl. May 2010) enthält der Test ausschließlich native Wörter aus dem Kernbereich der Wortschreibung sowie Wörter mit dem sogenannten Dehnungs-h, die hier zum erweiterten Kernbereich gezählt werden. Der Basistest, der zu allen drei Testzeitpunkten eingesetzt wurde, besteht aus zwei Teilen: Teil A: Wortschreibung, Teil B: satzinterne Großschreibung. Die Leistung im Bereich Wortschreibung wurde zum ersten Testzeitpunkt anhand eines bildunterstützten Wordiktats mit 26 wortbildungsmorphologisch komplexen Wörtern (z. B. *Bergziege*) erfasst (A1), da sich diese im Gegensatz zu Simplizia als Items mit einem idealen mittleren Schwierigkeitsniveau erwiesen haben. Zur genaueren Differenzierung im unteren Leistungsbereich wurde zum zweiten und dritten Testzeitpunkt ein Testteil mit einem Lückentext, in den Simplizia (z. B. *treffen*, *gräbt*) eingesetzt werden mussten, ergänzt (A2). Das Auswertungsraster für die Wortschreibung orientiert sich an dem Vorgehen der HSP und ermöglicht eine Ergebnispräsentation auf drei Ebenen:

1. Gesamtanzahl der richtigen Grapheme bzw. Graphemkombinationen,
2. Anzahl der richtigen Grapheme bzw. Graphemkombinationen, getrennt nach den Baumustern der Wortschreibung (vgl. Abschnitt 2),
3. Anzahl der korrekten Verschriftung von im Vorfeld als besonders fehleranfällig identifizierten Ankerstellen.

Diese Auswertungsdimensionen sind in Tabelle 2 für den Subtest *A1: komplexe Wörter* konkretisiert, wobei hier aus Platzgründen nur eine exemplarische Darstellung der insgesamt 24 Ankerstellen (Types), die in den Testwörtern jeweils ein- bis viermal vorkommen (Gesamtzahl der Token = 69), erfolgen kann. Eine detailliertere Auflistung findet sich in Bangel/Müller (2018a). Die Auswertung des Subtests *A2: einfache Wörter* ist vergleichbar aufgebaut, wobei dieser Testteil 13 Ankerstellen (Types) enthält, die jeweils zwei- bis viermal vorkommen (Gesamtzahl der Token = 31). Wie in Punkt 2 ausgeführt, sind viele Informationen in der Basisform kodiert und werden an einsilbige Formen vererbt. Da damit unterschiedliche kognitive Anforderungen verbunden sind, wurde diese Unterscheidung in das Kategoriensystem aufgenommen, wobei jeweils die beiden unmittelbaren Konstituenten des komplexen Wortes als Orientierungseinheiten dienen (z. B. *Staffel/stab*: *Staffel* (Basisform) + *Stab* (vererbte Form); *zer/reißen*: *zer* (Präfix) + *reißen* (Basisform)).

Tab. 2: Rahmenmodell der Testkonzeption *Wortschreibung* (Subtest A 1: komplexe Wörter)

		Auswertungsdimensionen		
		richtige Grapheme/Graphemkombinationen	Ankerstellen	erwartbare Fehler
phonographisch-silbische und stammbezogene morphologische Schreibungen	Basisform	Baumuster 1 z e r r e i ß e n	zerreißen	s oder ss statt ß
		Baumuster 2 H o l z b a l k e n	Holzbalcken	ck statt k
		Baumuster 3 S t a f f e l s t a b	Staffelstab	f statt ff
		Baumuster 4 a n n ä h e n	annähen	kein h
	vererbte Form	Baumuster 1 s c h m i e r i g	schmierig	i statt ie
		Baumuster 2 H o l z b a l k e n	Holzbalcken	tz statt z
		Baumuster 3 G e h s t o c k	Gehstock	k statt ck
		Baumuster 4 G e h s t o c k	Gehstock	kein h
	baumusterübergreifende morphol. Schreibungen		Schrittzähler	e statt ä
	Affixschreib.	Präfix V o r p l a t z	Vorplatz	for- statt vor-
		Suffix s c h m i e r i g	schmierig	-ich statt -ig
	Peripherie	W a c k e l z ä h n	Wackelzahn	kein h
Summenvariablen	richtige Grapheme: komplexe Wörter (Max = 228)	Ankerstellen: komplexe Wörter (Max = 69)		

Anmerkung: ■ = eine Auswertungseinheit in der entsprechenden Auswertungsdimension

Um zu überprüfen, ob Fehler in den Ankerstellen auf fehlerhafte Einsicht in die Regularitäten der Wortschreibung zurückzuführen sind (z. B. **Verstek*) oder diesen Fehlern diffuse Schreibungen und/oder lexikalische „Fehlgriffe“ zugrunde liegen (z. B. **Verschegt*), wurden die Fehler in dieser Dimension auf der Grundlage der Pilotierung als *erwartbare* bzw. *nicht erwartbare Fehler* kategorisiert.

Zur Einschätzung der Reliabilität des Tests wurde dieser in vier 5. Klassen ($N = 86$) im Schuljahr 2015/16 pilotiert. Auf der Grundlage dieser Ergebnisse wurden drei der insgesamt 26 Items des Subtests *A1: komplexe Wörter*, die von mehr als 80% der Schüler/-innen richtig gelöst wurden, durch neue Items ersetzt. Deshalb beziehen sich die folgenden Angaben zu den Testgütekriterien des Subtests *A1: komplexe Wörter* auf die Ergebnisse des Vortests der Projektstichprobe ($N = 364$). Für die Reliabilitätsprüfung des Subtests *A2: einfache Wörter* wurden die Ergebnisse des Nachtests herangezogen.

Als Subskalen für die Wortschreibung wurden sowohl die Skalen *richtige Grapheme* als auch *Ankerstellen* herangezogen, sodass die Reliabilitätsprüfung für fünf Subskalen erfolgte (*richtige Grapheme: komplexe Wörter*, *richtige Grapheme: einfache Wörter*, *Ankerstellen: komplexe Wörter*, *Ankerstellen: einfache Wörter*, *satzinterne Großschreibung*). Die Subskalen weisen eine interne Konsistenz von $\alpha = .81 - .89$ auf, wobei die Trennschärfen der Items zwischen $.13$ und $.70$ liegen. Die Itemschwierigkeiten wurden ebenfalls für jede Subskala getrennt berechnet. Sie liegen zwischen $p_i = .46 - .80$. Für die Subskala *richtige Grapheme: komplexe Wörter* wurde zusätzlich anhand einer Teilstichprobe von 94 Schüler/-innen, von denen die Ergebnisse aus der in den Klassen etwa zeitgleich eingesetzten HSP 4/5 vorlagen, die Paralleltest-Reliabilität berechnet. Diese liegt mit $r = .85$ im guten Bereich.⁴ Neben dem Rechtschreibtest wurde zur Einschätzung der Allgemeinbildung und der verbalen Verarbeitungskapazität der Wortschatztest des CFT 20-R zum ersten Messzeitpunkt in allen Klassen eingesetzt (vgl. Weiß 2007).

4.5 Überprüfung des orthographietheoretischen und -didaktischen Wissens der Lehrkräfte und Kontrolle der Treatmentintegrität

Um zu erfassen, ob die Intervention in allen Klassen konzeptkonform durchgeführt wurde, wurde ein zweischrittiges Kontrollverfahren eingesetzt, das das orthographietheoretische und -didaktische Professionswissen (und seine Entwicklung) der Lehrkräfte der IG und die klassenspezifische Umsetzung der Intervention (Treatmentintegrität) fokussiert. An dieser Stelle sollen nur die Ergebnisse dieses Kontrollverfahrens skizziert werden. Eine genauere Beschreibung zum methodischen Vorgehen findet sich in Bangel/Müller (2018a).

Das Fachwissen wurde in der ersten Fortbildung und nach Abschluss der Intervention in Form eines Fragebogens erhoben. Die Auswertung mithilfe eines Kodierleitfadens ergab, dass lediglich eine Lehrkraft nach Abschluss der Intervention nicht über das Wissen verfügte, das sie zu einer konzeptkonformen Umsetzung der Intervention befähigt hätte. Die Klasse wurde deshalb aus der Wirksamkeitsprüfung der Intervention ausgeschlossen.

Für die Überprüfung der Treatmentintegrität wurde ein Beobachtungsbogen für die studentischen Hilfskräfte erarbeitet, die den Unterricht in den Klassen begleiteten

⁴ Zur Korrelation der Subskalen vgl. Bangel/Müller (2018a).

(i. d. R. einmal wöchentlich) und die sich bei den Lehrkräften und Schüler/-innen über nichtbeobachtete Stunden informierten. Die Ausprägung der Kriterien (tatsächlicher zeitlicher Umfang der Intervention, konkrete methodische Umsetzung, Einstellung der Lehrkraft zum Ansatz) wurde auf einer Ratingskala (trifft zu – trifft überwiegend zu – trifft in Ansätzen zu – trifft nicht zu) bewertet und verbale Einschätzungen zum Unterricht gegeben. Unzureichende Erfüllung der Treatmentintegrität traf auf vier der 15 Interventionsklassen aus unterschiedlichen Gründen zu: ausschließliches Abarbeiten des Hefts ohne Austausch, inhaltlicher Ausstieg aus dem Konzept, ablehnende Haltung gegenüber dem Ansatz (vgl. auch Bangel/Müller 2018a). Der mögliche Einfluss der Treatmentintegrität auf die Testergebnisse in der IG wurde mit einer zweifaktoriellen Varianzanalyse mit Messwiederholung überprüft. Für die fünf Hauptvariablen ließen sich zwar keine signifikanten Interaktionseffekte zwischen dem Faktor Testzeitpunkt und dem Faktor Treatmentintegrität nachweisen, dennoch wurden für die weiteren Analysen die vier Klassen ausgeschlossen, die die Kriterien der Treatmentintegrität nicht ausreichend erfüllten, um gewährleisten zu können, dass die gefundenen Effekte tatsächlich auf das Treatment zurückgeführt werden können.

4.6 Inhalte des Deutschunterrichts in den Kontrollklassen

In der KG wurde mit einem Lehrerfragebogen zum zweiten und dritten Testzeitpunkt erhoben, in welchem Umfang und mit welchen Schwerpunkten zur Rechtschreibung gearbeitet wurde (vgl. Tab. 1). Außerdem informierten sich die studentischen Hilfskräfte, die die Testungen in der KG durchführten, zum Nachtest und zum verzögerten Nachtest bei den Lehrkräften nach dem konkreten Inhalt, dem zeitlichen Umfang und besonderen Bedingungen des Orthographieunterrichts. Ähnlich wie in der IG variierte die durchschnittliche wöchentliche Arbeitszeit zu rechtsschriftlichen Themen zwischen 30 und 60 Minuten. Während des Interventionszeitraums wurden v. a. Aufgaben aus dem Schulbuch „Doppel-Klick“ und dem Arbeitsheft „Abschreiben erwünscht“ genutzt.

4.7 Versuchsablauf

Zu Schuljahresbeginn (September 2016) wurde der Rechtschreibtest in allen Klassen durch geschulte Hilfskräfte durchgeführt. Zum ersten Testzeitpunkt wurde außerdem der Wortschatztest des CFT-20 eingesetzt. Die Testbedingungen in den einzelnen Klassen wurden protokolliert, um ggf. Abweichungen und Störfaktoren identifizieren zu können.

Nach der Intervention wurde der Rechtschreibtest als Nachtest (Februar 2017) und zum Schuljahresende (Juni 2017) als verzögerter Nachtest eingesetzt. Zu diesem Zeitpunkt wurden außerdem mittels eines Lehrerfragebogens die Inhalte des Deutschunterrichts im gesamten Schuljahr erhoben.

4.8 Statistische Analysen

Um die Wirksamkeit der Intervention auf die zentralen abhängigen Variablen (= Summenvariablen) *richtige Grapheme: komplexe Wörter*, *Ankerstellen: komplexe Wörter*, *richtige Grapheme: einfache Wörter* und *Ankerstellen: einfache Wörter* zum zweiten und dritten Testzeitpunkt zu prüfen, wurde eine multivariate Kovarianzanalyse (MANCOVA) gerechnet.⁵ Zur Vermeidung einer Alphafehler-Kumulierung bei der differenzierten Betrachtung der einzelnen abhängigen Variablen wurde eine Bonferroni-Korrektur vorgenommen. Das Treatment sowie die Leistungsquartile wurden als between-subject-Faktor modelliert. Die Vortestwerte auf den Kriteriumsvariablen *richtige Grapheme: komplexe Wörter* und *satzinterne Großschreibung* wurden als Kovariaten berücksichtigt, wobei sich mit Blick auf alle Vortestvariablen lediglich für die Variable *satzinterne Großschreibung* signifikante Unterschiede zwischen der IG und der KG (mit einer signifikanten Überlegenheit der IG) nachweisen lassen ($t(362) = 2.5, p = .013, d = .26$). Darüber hinaus wurden der Sprachstatus, der häusliche Bücherbesitz und die verbale Verarbeitungskapazität (Ergebnisse des CFT 20-R) als Kovariaten statistisch kontrolliert. Für eine differenzierte Wirksamkeitsprüfung der Intervention, bezogen auf unterschiedliche Leistungsgruppen, wurde das oben beschriebene Modell für alle vier Leistungsquartile noch einmal separat gerechnet, wobei hier lediglich das Treatment als between-subject-Faktor modelliert wurde. Zur Identifizierung der Hauptfehlerquellen wurden die Lösungshäufigkeiten (LH) für jede Ankerstelle der komplexen Wörter herangezogen. Um zu differenzierten Aussagen bezüglich der Wirkung des Treatments auf einzelne Fehlerschwerpunkte zu gelangen, wurden in einem nächsten Schritt in Anlehnung an das oben beschriebene Modell mehrere Kovarianzanalysen mit den als Fehlerschwerpunkte identifizierten Ankerstellen (vgl. Abb. 2) als abhängige Variablen gerechnet.⁶

5. Ergebnisse

In Tabelle 3 sind die Mittelwerte und Standardabweichungen (Vortest (t1) → Nachtest (t2), Nachtest (t2) → verzögerter Nachtest (t3), Vortest (t1) → verzögerter Nachtest (t3)) der IG und KG zu den drei Testzeitpunkten zu den zentralen Kriteriumsvariablen (= Summenvariablen) *richtige Grapheme: komplexe Wörter*, *Ankerstellen: komplexe Wörter*, *richtige Grapheme: einfache Wörter*, *Ankerstellen: ein-*

⁵ Bei der Interpretation der Ergebnisse muss berücksichtigt werden, dass die Variablen *richtige Grapheme* und *Ankerstellen* nicht unabhängig voneinander sind: Jeder Fehler in den Ankerstellen ist gleichzeitig ein Fehler im Bereich *richtige Grapheme*. Der erreichte Wert in der Summenvariable *richtige Grapheme* dient der Einschätzung der allgemeinen Rechtschreibleistung, wohingegen mit der Summenvariable *Ankerstellen* spezifische Fehlerschwerpunkte fokussiert werden.

⁶ Zur Begründung, warum auf eine mehrebenenanalytische Auswertungsstrategie verzichtet wurde, vgl. Bangel/Müller (2018a).

fache Wörter für die Gesamtgruppe und getrennt nach Leistungsquartilen dargestellt.

Tab. 3: Mittelwerte und Standardabweichungen der Vor-, Nach- und verzögerten Nachtestmessung

		Gesamtgruppe		Quartile							
		Interventionsgruppe				Kontrollgruppe					
		IG	KG	1.	2.	3.	4.	1.	2.	3	4
		<i>n</i> = 207	<i>n</i> = 157	<i>n</i> = 50	<i>n</i> = 50	<i>n</i> = 55	<i>n</i> = 52	<i>n</i> = 40	<i>n</i> = 39	<i>n</i> = 32	<i>n</i> = 46
		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
richtige Grapheme: kompl. Wörter (Max = 228)	t1	205.52 (13.44)	204.76 (13.81)	186.32 (8.32)	202.24 (2.71)	211.98 (2.21)	220.29 (3.38)	185.53 (8.56)	202.18 (2.38)	210.75 (2.21)	219.52 (2.55)
	t2	214.14 (9.81)	211.01 (11.10)	201.88 (9.43)	213.16 (5.55)	217.82 (4.05)	223.17 (3.20)	197.28 (9.10)	208.59 (6.13)	216.56 (4.12)	221.15 (3.43)
	t3	215.75 (8.41)	213.76 (9.34)	205.98 (7.66)	213.26 (5.23)	219.38 (3.50)	223.69 (3.46)	202.55 (7.96)	212.15 (5.46)	217.72 (4.37)	222.11 (3.52)
Ankerstellen: kompl. Wörter (Max = 69)	t1	51.64 (9.75)	50.62 (10.31)	38.12 (5.64)	49.26 (3.81)	56.01 (2.66)	62.30 (3.15)	37.40 (6.47)	47.35 (2.98)	55.12 (4.39)	61.73 (2.43)
	t2	58.06 (7.04)	55.52 (8.67)	49.84 (6.00)	56.70 (4.76)	60.29 (3.74)	64.92 (2.76)	45.12 (7.67)	53.51 (4.52)	59.65 (3.61)	63.36 (2.94)
	t3	59.24 (6.43)	57.04 (7.78)	52.00 (5.29)	57.08 (4.36)	61.69 (2.96)	65.69 (2.76)	47.97 (7.23)	55.89 (4.95)	60.00 (3.59)	63.84 (2.80)
richtige Grapheme: einf. Wörter (Max = 126)	t2	119.56 (5.50)	117.43 (6.94)	112.64 (5.18)	118.98 (3.53)	122.15 (2.27)	124.06 (1.91)	109.40 (6.26)	115.87 (4.70)	120.78 (3.10)	123.41 (1.85)
	t3	119.89 (6.11)	119.54 (5.75)	113.04 (7.35)	119.34 (3.42)	122.56 (3.07)	124.19 (2.17)	112.70 (5.28)	118.95 (3.89)	121.94 (2.86)	124.33 (1.53)
Ankerstellen: einf. Wörter (Max = 31)	t2	26.65 (3.79)	25.20 (5.27)	22.06 (3.72)	26.08 (2.52)	28.47 (1.66)	29.67 (1.30)	19.47 (6.03)	24.28 (3.06)	27.56 (2.19)	29.30 (1.50)
	t3	27.09 (4.00)	26.20 (4.73)	22.70 (4.79)	26.72 (2.28)	28.70 (2.28)	29.94 (1.34)	20.75 (5.26)	25.89 (2.90)	28.21 (1.84)	29.80 (1.20)

5.1 Gesamtgruppe

Bei der Überprüfung des Gesamtmodells zeigen sich signifikante Effekte für die Kovariaten *verbale Verarbeitungskapazität* (Pillai-Spur = .04, $F(8, 344) = 2.09$, $p = .03$, $\eta^2 = .04$) und *richtige Grapheme: komplexe Wörter (t1)* (Pillai-Spur = .27, $F(8, 344) = 16.32$, $p < .001$, $\eta^2 = .27$), für die Zwischensubjektfaktoren *Treatment* (Pillai-Spur = .15, $F(8, 344) = 7.86$, $p < .001$, $\eta^2 = .15$) und *Leistungsquartile* (Pillai-Spur = .11, $F(8, 344) = 1.75$, $p = .01$, $\eta^2 = .04$) sowie für die Interaktion *Treatment*Leistungsquartile* (Pillai-Spur = .10, $F(8, 344) = 1.61$, $p = .03$, $\eta^2 = .04$).

Zum zweiten Testzeitpunkt zeigen sich signifikante Effekte für alle vier abhängigen Variablen im Bereich *Wortschreibung*. Die Effektstärken liegen alle im mittleren Bereich: *richtige Grapheme: komplexe Wörter* ($F(1, 351) = 17.55$, $p < .001$, $\eta^2 = .05$), *Ankerstellen: komplexe Wörter* ($F(1, 351) = 20.11$, $p < .001$, $\eta^2 = .05$), *richtige Grapheme: einfache Wörter* ($F(1, 351) = 27.36$, $p < .001$, $\eta^2 = .07$) und *Ankerstellen: einfache Wörter* ($F(1, 351) = 19.98$, $p < .001$, $\eta^2 = .05$). Zum dritten Testzeitpunkt lassen sich noch signifikante Effekte mit einer Effektstärke im kleinen bis mittleren Bereich für drei der vier zentralen Kriteriumsvariablen nachweisen: *richtige*

Grapheme: komplexe Wörter ($F(1, 351) = 7.48, p = .007, \eta^2 = .02$), *Ankerstellen: komplexe Wörter* ($F(1, 351) = 17.73, p < .001, \eta^2 = .05$), *Ankerstellen: einfache Wörter* ($F(1, 351) = 7.71, p = .006, \eta^2 = .02$).⁷

5.2 Unteres Leistungsquartil

Für die 25% schwächsten Rechtschreiber/-innen der Stichprobe lassen sich ebenfalls signifikante Effekte für alle vier abhängigen Variablen im Bereich Wortschreibung nachweisen. Die Interventionseffekte liegen im mittleren Bereich: *richtige Grapheme: komplexe Wörter* ($F(1, 83) = 7.12, p = .01, \eta^2 = .08$), *Ankerstellen: komplexe Wörter* ($F(1, 83) = 11.12, p = .001, \eta^2 = .12$), *richtige Grapheme: einfache Wörter* ($F(1, 83) = 8.98, p = .01, \eta^2 = .08$) und *Ankerstellen: einfache Wörter* ($F(1, 83) = 8.07, p = .01, \eta^2 = .09$). Zum dritten Testzeitpunkt lassen sich noch signifikante Effekte mit einer Effektstärke im mittleren bis großen Bereich für drei der vier zentralen Kriteriumsvariablen nachweisen: *richtige Grapheme: komplexe Wörter* ($F(1, 83) = 7.38, p = .01, \eta^2 = .08$), *Ankerstellen: komplexe Wörter* ($F(1, 83) = 12.16, p = .001, \eta^2 = .13$), *Ankerstellen: einfache Wörter* ($F(1, 83) = 5.82, p = .02, \eta^2 = .07$).

Mit Blick auf die Lösungshäufigkeiten (LH) der einzelnen Ankerstellen der komplexen Wörter lassen sich vor der Intervention folgende Hauptfehlerquellen für das erste Leistungsquartil identifizieren: β in der Basisform (IG: LH = 10.7%, KG: LH = 7.7%), Silbengelenkschreibung in der vererbten Form (IG: LH = 18.9%, KG: LH = 23.1%), \ddot{a}/\ddot{au} (IG: LH = 25.0%, KG: LH = 25.6%), Suffixschreibung *-ig* (IG: LH = 34.4%, KG: LH = 34.6%), Doppelkonsonantschreibung an der Morphemgrenze (IG: LH = 48.8%, KG: LH = 55.1%) (vgl. Abb. 2).

In vier der fünf Fehlerschwerpunkte haben sich die Schüler/-innen der IG im Posttest signifikant stärker verbessert als die Schüler/-innen der KG. Die Effektstärken liegen im mittleren Bereich (β in der Basisform: $F(1, 83) = 4.99, p = .02, \eta^2 = .06$, Silbengelenkschreibung in der vererbten Form: $F(1, 83) = 6.98, p = .01, \eta^2 = .08$), \ddot{a}/\ddot{au} : $F(1, 83) = 8.19, p = .01, \eta^2 = .09$, Doppelkonsonantschreibung an der Morphemgrenze: $F(1, 83) = 6.86, p = .01, \eta^2 = .08$). Für die Suffixschreibung *-ig* zeigt sich im Posttest eine signifikante Überlegenheit der KG gegenüber der IG ($F(1, 83) = 4.28, p = .04, \eta^2 = .05$) (vgl. Abb. 2).

⁷ Zum Varianzanteil der abhängigen Variablen, der durch die fünf weiteren unabhängigen Variablen erklärt wird, vgl. Bangel/Müller (2018a).

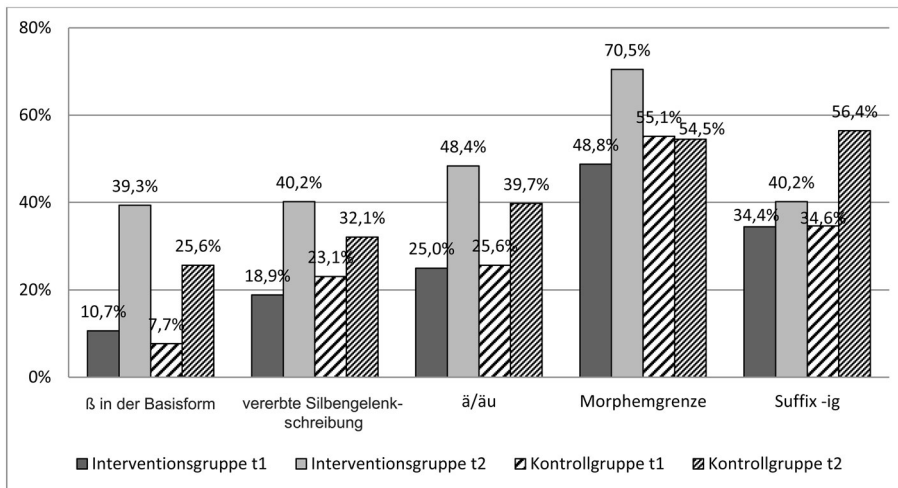


Abb. 2: Lösungshäufigkeiten: Fehlerschwerpunkte

Für vier der fünf Fehlerschwerpunkte finden sich zum zweiten Testzeitpunkt in über 85% der jeweiligen Fehlschreibungen die als erwartbar klassifizierten Fehler. Beim Suffix *-ig* ist der Anteil an erwartbaren Fehlern mit 75% etwas geringer.

Auch für die wortbildungsmorphologisch einfachen Wörter ergeben sich die größten Interventionseffekte in der *ä/äu*-Schreibung ($F(1, 83) = 9.43, p < .001, \eta^2 = .10$), der vererbten Silbengelenkschreibung ($F(1, 83) = 9.34, p < .001, \eta^2 = .10$) und der β -Schreibung in der Basisform ($F(1, 83) = 5.84, p = .02, \eta^2 = .06$).

5.3 Zweites bis viertes Leistungsquartil

Für das zweite Leistungsquartil (Prozentrang 26 bis 50) zeigen sich zum zweiten Testzeitpunkt ebenfalls signifikante Interventionseffekte für alle vier abhängigen Summenvariablen im Bereich Wortschreibung. Die Effektstärken liegen im großen Bereich: *richtige Grapheme: komplexe Wörter* ($F(1, 82) = 11.57, p = .001, \eta^2 = .12$), *Ankerstellen: komplexe Wörter* ($F(1, 82) = 10.40, p = .002, \eta^2 = .11$), *richtige Grapheme: einfache Wörter* ($F(1, 82) = 16.04, p < .001, \eta^2 = .16$) und *Ankerstellen: einfache Wörter* ($F(1, 82) = 11.16, p = .001, \eta^2 = .12$). Zum dritten Testzeitpunkt ergeben sich für dieses Leistungsquartil keine signifikanten Effekte mehr für die Summenvariablen. Für das dritte Leistungsquartil (Prozentrang 51 bis 75) finden sich weder zum zweiten noch zum dritten Testzeitpunkt signifikante Unterschiede für die abhängigen Summenvariablen zwischen der IG und der KG. Für die stärksten Rechtschreiber/-innen zeigt sich zum zweiten Testzeitpunkt ein signifikanter Interventionseffekt für die Summenvariable *richtige Grapheme: komplexe Wörter* ($F(1, 91) = 5.85, p = .01, \eta^2 = .06$) und zum dritten Testzeitpunkt ein signifikanter Über-

legenheitseffekt der IG für die Variable *Ankerstellen: komplexe Wörter* ($F(1, 91) = 6.03, p = .01, \eta^2 = .06$).

6. Diskussion

Die Schüler/-innen der IG produzieren direkt nach der Intervention weniger Fehler als die der KG, im Posttest lassen sich für die IG Interventionseffekte mittlerer Größe ($\eta^2 = .05 - .07$) nachweisen. Zum dritten Testzeitpunkt holt die KG auf. Die IG ist weiterhin signifikant überlegen, insbesondere in der Auswertungskategorie *Ankerstellen: komplexe Wörter*. Für diese Kategorie zeigt sich noch im verzögerten Nachtest ein Interventionseffekt mittlerer Größe ($\eta^2 = .05$). Bei den Ankerstellen der wortbildungsmorphologisch einfachen Wörter gibt es nur noch eine leichte Überlegenheit der IG.

Es zeigt sich, dass sich die Interventionseffekte in den unterschiedlichen Leistungsquartilen signifikant unterscheiden (Pillai-Spur = .10, $F(8, 344) = 1.61, p = .03, \eta^2 = .04$). Die nach Leistungsgruppen differenzierte Auswertung belegt, dass die schwächsten Rechtschreiber/-innen der IG am meisten von der Intervention profitiert haben. Die Größe des Überlegenheitseffekts für die Schreibung der komplexen Wörter bleibt auch drei Monate nach Abschluss der Intervention stabil und liegt im mittleren bis großen Bereich (*richtige Grapheme*: $\eta^2 = .07$, *Ankerstellen*: $\eta^2 = .13$). Bei differenzierter Betrachtung lassen sich mit Ausnahme der Suffixschreibung *-ig* für alle Fehlerschwerpunkte (β -Schreibung in der Basisform, \ddot{a}/\ddot{au} -Schreibung, vererbte Silbengelenkschreibung, Doppelkonsonantschreibung an der Morphemgrenze) signifikante Interventionseffekte von mittlerer Größe nachweisen. Auch für die Schreibung der wortbildungsmorphologisch einfachen Wörter zeigen sich in der β -Schreibung, der \ddot{a}/\ddot{au} -Schreibung und der *vererbten Silbengelenkschreibung* mittlere bis große Interventionseffekte.

Da es sich vor allem bei den morphologisch komplexen Testwörtern um niedrigfrequente Wörter handelt, die mit hoher Wahrscheinlichkeit nicht zum Sichtwortschatz der Schüler/-innen gehören, lassen die Interventionseffekte in diesem Bereich darauf schließen, dass die Intervention dazu beigetragen hat, dass die Schüler/-innen der IG (zumindest implizit) sicherer auf die grundlegenden Baumuster des Kernbereichs der deutschen Wortschreibung zugreifen können. Das vergleichsweise schlechtere Abschneiden der Schüler/-innen der Kontrollgruppe im Posttest sowie im verzögerten Nachtest ist insofern aufschlussreich für rechtschreibdidaktische Entscheidungen, als diese Klassen ebenfalls intensiv an der Schreibung von Basis- und vererbten Wortformen des Kernbereichs gearbeitet haben (vgl. Tab. 1). Ein wesentlicher Unterschied zwischen den Herangehensweisen beider Gruppen liegt darin, dass es im Unterricht der Interventionsklassen vordergründig um Struktureinsicht ging, indem die Erschließung der grundlegenden Baumuster nativer Wörter (vgl. Abschnitt 2) durch Angebote zum entdeckenden und einsichtsvollen Lernen ermöglicht wurde. Die Schüler/-innen sollten die Struktureinsichten

auf die Schreibung vergleichbarer Wörter übertragen und sie sich so selbstständig erschließen. Dies kann zur kognitiven Entlastung beitragen, da sie sich nicht mehr Einzelwörter und Einzelphänomene merken müssen, sondern vier (bzw. fünf) Baumuster als Basis für grundlegende Schreibentscheidungen ausreichen. In den Kontrollklassen wurden (i. d. R. unterrichtsleitend) Materialien verwendet, in denen die orthographischen Auffälligkeiten einzelner Wörter (*Wörter mit ie, Wörter mit ll, mm, ss, Wörter mit ß*, vgl. Tab. 1) anhand von Wörtern aus dem Kern- und Peripheriebereich thematisiert wurden, ohne dass für das deutsche Schriftsystem prototypische Strukturen erschlossen und durch Analogiebildung auf vergleichbare Wortschreibungen übertragen werden sollten. Jedes Wort bildet i. d. R. eine eigene Lerneinheit, der Eindruck von Zufälligkeit wird durch die häufig wenig systematische Gruppe an einzuprägenden Wörtern verstärkt. Die Materialien ermöglichen zwar gleichfalls den Erwerb und die Anwendung von Strategien, da diese aber zum einen – wahrscheinlich aus Gründen der vermeintlichen Reduktion der kognitiven Anforderungen – auf bereits in der Grundschule praktizierte Strategien wie genaues Hören, Mitsprechen und Merken (vgl. Tab. 1) ausgerichtet sind, wiederholen sie Lernangebote, die gerade schwächeren Rechtschreiber/-innen keine wesentliche Unterstützung bieten. Das Hörbarmachen von Silbengelenkschreibungen oder des silbeninitialen h z. B. verlangt von den Schülerinnen und Schülern eine schriftinduzierte Überlautung, über die sie nur verfügen können, wenn sie bereits auf die korrekten Wortschreibungen zugreifen können. Diese schriftgeprägte Lautung hat häufig mit der Standardaussprache der Wörter wenig zu tun.⁸ Zum anderen erfolgt das Einführen von Rechtschreibstrategien in den Materialien der Kontrollklassen kaum, um sie als Problemlöseangebot zu nutzen und sie auf vergleichbare Wörter zu übertragen.

In den Interventionsklassen wurden die sprachlich-kognitiven Ausgangsbedingungen der schwächeren Rechtschreiber/-innen insofern berücksichtigt, als die Erarbeitung von Strukturen und Strategien an einer beschränkten Anzahl an Wörtern aus dem Kernbereich der Baumuster 1 bis 4 (bzw. 5) erfolgte. Dadurch, dass die Schüler/-innen an einer überschaubaren Gruppe von Wörtern arbeiteten, indem sie sie im Hinblick auf ihre Struktur verglichen, sie ordneten, die Schreibungen erklärten, konnten sie die so gewonnenen Einsichten auf weitere, ähnlich strukturierte Wörter

⁸ Auf eine kritische Auseinandersetzung mit den Evaluationsergebnissen aus Trainingsprogrammen, die das Mitsprechen fokussieren (vgl. Neubauer/Kirchner 2013), kann im Rahmen dieses Beitrags nicht eingegangen werden. Es sei an dieser Stelle lediglich darauf verwiesen, dass es sich bei diesen Wirksamkeitsstudien ausschließlich um die Evaluation additiver Förderprogramme handelt. Dies schränkt die Vergleichbarkeit der Ergebnisse erheblich ein. Darüber hinaus betonen Neubauer/Kirchner auf der Basis ihrer Evaluationsergebnisse zum Konzept von Reuter-Liehr, dass dies vor allem den Kindern zugutekommt, die noch erhebliche Schwierigkeiten in der phonologischen Wortanalyse sowie in der Beherrschung grundlegender Phonem-Graphem-Korrespondenzen aufweisen (vgl. Neubauer/Kirchner 2013). Bei den Fünftklässlerinnen und Fünftklässlern der hier berichteten Interventionsstudie konnten i. d. R. keine gravierenden Schwierigkeiten in diesen Bereichen mehr festgestellt werden.

übertragen. Im Mittelpunkt standen Strategien wie die Ermittlung der Silbenstruktur im prototypischen Zweisilber, die Herleitung der zweisilbigen Basisform durch Verlängern einsilbiger Wortformen, wobei diese häufig zunächst als Bestandteile eines wortbildungsmorphologisch komplexen Wortes ermittelt werden mussten.

Die Studie zeigt, dass vor allem die sehr schwachen Rechtschreiber/-innen von einer sachangemessenen Vorstrukturierung des Lerngegenstands profitieren konnten, wie der Vergleich der Ergebnisse mit denen der übrigen Leistungsquartile belegt. Für die Schüler/-innen des zweiten Quartils lassen sich zwar die größten Interventioneffekte nachweisen, zum dritten Testzeitpunkt holen die Lerner/-innen der KG allerdings so weit auf, dass sich keine signifikanten Unterschiede mehr nachweisen lassen. Diese Schüler/-innen scheinen demzufolge mit einer auf Einsicht in das System ausgerichteten Konzeption schneller zu lernen, aber letztendlich die Regularitäten der Wortschreibung auch selbstständig entdecken zu können. Bei den stärkeren Rechtschreiberinnen und Rechtschreibern (drittes und viertes Quartil) zeigt sich dies noch deutlicher. Als zentrales Ergebnis bleibt festzuhalten, dass die gewählte Vermittlungspraxis im Rechtschreibunterricht vor allem für die schwachen Rechtschreiber/-innen äußerst relevant ist. Die Interventioneffekte lassen den Schluss zu, dass diejenigen Lerner/-innen der IG, die mit hoher Wahrscheinlichkeit zu der Gruppe gehören, die den Mindeststandard im Bereich *Orthographie* in Klasse 4 nicht erreicht haben (vgl. Abschnitt 1), durch ein schriftstrukturorientiertes Vorgehen wirkungsvoll unterstützt werden können. Ihre gleichaltrigen rechtschreibstärkeren Mitschüler/-innen scheinen weitaus besser und ohne entsprechende Lernangebote in der Lage zu sein, sich eigenaktiv die richtige Schreibung von Wörtern herzuleiten und in unterschiedlichen Schreibsituationen anwenden zu können. Die moderate Entwicklung dieser Lerner/-innen belegt dennoch, dass diese vom gewählten Ansatz ebenfalls profitieren können bzw. zumindest nicht in ihrem Lernfortschritt gebremst werden.

Die dargestellten Ergebnisse geben selbstverständlich in erster Linie Auskunft über das prozedurale Wissen (Können) der getesteten Fünftklässler/-innen in der Domäne *Wortschreibung*. Wie sich die Intervention konkret auf das Lernen der Schüler/-innen ausgewirkt hat, kann mit den vorliegenden Testdaten nicht beantwortet werden. Darüber sollen die in Vorbereitung befindlichen qualitativen Auswertungen der im Projekt erhobenen Interviewdaten Aufschluss geben. In den Interviews sollten die Schüler/-innen fehlerhafte Schreibungen korrigieren und ihre Korrekturentscheidungen begründen. Von den qualitativen Daten sind demnach Antworten darauf zu erhoffen, welche Strategien die Schüler/-innen zur Kontrolle und Begründung von Schreibungen anwenden, um Rückschlüsse darauf ziehen zu können, wie die Lernangebote die Hypothesen der Schüler/-innen über die Schreibung von Wörtern und ihr Verhältnis zur Lautung beeinflussen.

Erste Analysen zusätzlich erhobener freier Schreibproben deuten darauf hin, dass der Transfer der dargestellten Effekte auf das selbstständige Textschreiben im Rahmen dieser Intervention nicht ermöglicht werden konnte. Dazu muss (neben

einer intensiven Erarbeitung von Strukturen) sicher eine reflektierte Schreibpraxis, die unterschiedliche Anwendungssituationen einschließt, treten. Dazu würden z. B. Lernangebote zur Arbeit an tatsächlich verwendeten Wörtern in eigenen Texten gehören.

Diese Einschränkungen machen grundlegende Forschungsdesiderate sichtbar: Wir wissen z. B. wenig darüber, welche Faktoren eines schriftstrukturorientierten Rechtschreibunterrichts wie auf die Lernprozesse der Schülerinnen und Schüler einwirken, wie die Automatisierung von Schreibungen erfolgen kann und wie gleichzeitig erworbene Strategien als Kontrollmöglichkeiten im Schreibprozess genutzt werden können. Darüber hinaus benötigen wir unterrichtsbezogene Studien für den schriftsprachlichen Anfangsunterricht, in dem Lesen- und Schreibenlernen stärker als bislang üblich aufeinander bezogen werden. Zum einen ist der Lerngegenstand *Wortschreibung* von Anfang an zentral im schriftsprachlichen Lernprozess, zum anderen wäre ein Blick auf Ergebnisse eines Lernprozesses, der von Anfang an die Schriftstruktur im hier dargelegten Sinn berücksichtigt, von erheblichem orthographiedidaktischem Gewinn. Welche Rolle die Instruktionen der Lehrkräfte und ihre Modellierungen des Lerngegenstands im Lernprozess für den Lernerfolg spielen und was für sie und ihre Schüler/-innen dabei handlungsleitend ist, wartet ebenfalls auf empirische Erhebung und Überprüfung.

Literatur

- Bangel, Melanie/Müller, Astrid (2015): Einsichten in morphologische Strukturen als Grundlage für die Entwicklung der Dekodierfähigkeit. In: Rautenberg, Iris/Reißig, Tilo (Hg.): Lesen und Lesedidaktik aus linguistischer Perspektive. Frankfurt a. M. u. a.: Peter Lang Verlag. S. 17–54.
- Bangel, Melanie/Müller, Astrid (2018a): Orthographisches Lernen durch Zugang zur Schriftstruktur. Ergebnisse einer Interventionsstudie in Klasse 5. In: Unterrichtswissenschaft 2018. <https://doi.org/10.1007/s42010-018-0015-y>.
- Bangel, Melanie/Müller, Astrid (2018b): Wörtern und Sätzen auf der Spur: Mein Rechtschreibheft. Seelze: Klett und Kallmeyer.
- Bredel, Ursula (2010): Der Schrift vertrauen. Wie Wörter und ihre Strukturen entdeckt werden können. In: Praxis Deutsch, H. 221. S. 14–21.
- Bredel, Ursula (2016): Schriftspracherwerb. In: Domahs, Ulrike/Primus, Beatrice: Handbuch Laut, Gebärde, Buchstabe. Berlin u. a.: de Gruyter. S. 436–454.
- Eisenberg, Peter (1983): Orthografie und Schriftsystem. In: Günther, Klaus B./Günther, Hartmut (Hg.): Schrift, Schreiben, Schriftlichkeit. Arbeiten zur Struktur, Funktion und Entwicklung schriftlicher Sprache. Tübingen: Niemeyer. S. 41–68.
- Eisenberg, Peter (2013): Grundriss der deutschen Grammatik. 4. Aufl. Stuttgart: Metzler.
- Eisenberg, Peter (2016): Phonem und Graphem. In: Duden. Grammatik der deutschen Gegenwartssprache. 9. Aufl. Mannheim. S. 19–94.
- Feustel, Filiz/Theuer, Stephan/Hergesell, Dirk/Placzek, Katrin/Krull, Renate (2009): Doppelklick. Das Sprach- und Lesebuch 5. 2. Aufl. Berlin: Cornelsen.

- Hinney, Gabriele (1997): Neubestimmung von Lerninhalten für den Rechtschreibunterricht. Ein fachdidaktischer Beitrag zur Schriftaneignung als Problemlöseprozeß. Frankfurt a. M.: Lang.
- Hinney, Gabriele (2014): Lesen- und Schreibenlernen mit der Silbe – ein sprachdidaktischer Fortschritt? In: Wrobel, Dieter/Müller, Astrid (Hg.): Bildungsmedien für den Deutschunterricht. Vielfalt – Entwicklungen – Herausforderungen. Bad Heilbrunn: Klinkhardt. S. 143–169.
- Hinney, Gabriele (2017): Wortschreibung. In: Baurmann, Jürgen/Kammler, Clemens/Müller, Astrid (Hg.): Handbuch Deutschunterricht. Seelze: Klett & Kallmeyer. S. 263–267.
- Jacobs, August-Bernhard (2005): Abschreiben erwünscht: 5./6. Schuljahr – Texte zum Abschreiben, Üben, Diktieren: Trainingsheft mit Lösungen. Nachdruck. Berlin: Cornelsen.
- Maas, Utz (1992): Grundzüge der deutschen Orthographie. Tübingen: de Gruyter.
- May, Peter (2010): HSP 1–9. Diagnose orthografischer Kompetenz. Zur Erfassung der grundlegenden Rechtschreibstrategien mit der Hamburger Schreibprobe. Hamburg: vpm.
- Neubauer, Skadi/Kirchner, Sabine (2013): Rechtschreibförderung von Grundschulern unter Berücksichtigung individueller Lernvoraussetzungen. Zwei Evaluationsstudien im Vergleich. In: Didaktik Deutsch. Jg. 18, Heft 35. S. 44–61.
- Wahl, Stefan/Rautenberg, Iris/Helms, Stefanie (2017): Evaluation einer syntaxbasierten Didaktik zur satzinternen Großschreibung. In: Didaktik Deutsch. H. 42. S. 32–52.
- Weirich, Sebastian/Wittig, Julia/Stanat, Petra (2017): Kompetenzstufenbesetzung im Fach Deutsch. In: Stanat, Petra/Schipolowski, Stefan/Rjosk, Camilla/Weirich, Sebastian/Haag, Nicole (Hg.): IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich. Münster: Waxmann. S. 129–139.
- Weiß, Rudolf H. (2007): Wortschatztest und Zahlenfolgentest – Revision. WS/ZF-R; Ergänzungstests zum CFT 20-R; Manual. Göttingen: Hogrefe.

Anschrift der Verfasserinnen:

Melanie Bangel und Astrid Müller, Universität Hamburg, Fakultät für Erziehungswissenschaft, Fachbereich 4: Didaktik der deutschen Sprache und Literatur, Von-Melle-Park 8, D-20146 Hamburg.

*Melanie.Bangel@uni-hamburg.de,
Astrid.Mueller@uni-hamburg.de*