

**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>  
15. Jahrgang 2010 – ISSN 1431-4355  
Schneider Verlag Hohengehren  
GmbH

*Johanna Fay*

**KOMPETENZFACETTEN IN  
DER  
RECHTSCHREIBDIAGNOSTIK  
Rechtschreibleistung im Test und  
im freien Test**

In: Didaktik Deutsch. Jg. 15. H. 29. S. 15-36.

---

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.

Johanna Fay

## KOMPETENZFACETTEN IN DER RECHTSCHREIBDIAGNOSTIK

### Rechtschreibleistung im Test und im freien Text

#### 0 Ausgangspunkt und Fragestellung

Ausgangspunkt der Studie „Die Entwicklung der Rechtschreibkompetenz beim Textschreiben“ (Fay 2010) ist eine alltägliche Beobachtung von Lehrerinnen und Lehrern im Schreibunterricht: Was ein Kind im Diktat oder Lückentext schon richtig schreiben kann, gelingt im frei verfassten Text plötzlich nicht mehr. Die Konzentration auf inhaltliche und sprachliche Aspekte des Textschreibens geht – so die Vermutung – offenbar zu Lasten der Aufmerksamkeit für die Orthographie. Damit stellen die Lehrkräfte implizit fest, dass der Handlungszusammenhang, in dem richtig geschrieben werden soll, spezifische Anforderungen stellt und so einen Einfluss auf die Rechtschreibleistung der Kinder hat. Zentrales Anliegen ist daher die Abhängigkeit der Rechtschreibkompetenz vom Handlungszusammenhang empirisch zu untersuchen. Im Mittelpunkt stehen die Fragen:

1. Wie entwickelt sich die Rechtschreibleistung beim Textschreiben in den Klassen 1 bis 4?
2. Wie unterscheidet sich dies von der Rechtschreibleistung im Schreiben diktierter Wörter und Sätze eines Rechtschreibtests?

In den folgenden Ausführungen geht es um die zweite Frage. Damit wird ein Ausschnitt der Studie betrachtet: Gezeigt wird der Vergleich der Rechtschreibleistungen im selbst verfassten Text und in Einzelwort- und Satzschreibungen eines Rechtschreibtests von 30 Viert-Klässlern in zehn ausgesuchten silbischen Fehlerkategorien (s. Abschnitt 3). Ein solcher Vergleich ist von besonderem Interesse. Einerseits spielt das Abschneiden in rechtschreibdiagnostischen Tests eine wesentliche Rolle bei der Entscheidung über Maßnahmen zur Rechtschreibförderung, bei der Notendifindung und in Empfehlungen zur weiteren Schullaufbahn. Andererseits wird in der rechtschreibdidaktischen Forschung eine textsortenspezifische Abhängigkeit der Rechtschreibleistung angenommen, die allerdings noch nicht annähernd geklärt ist (vgl. Risel 2008, S. 51). Inwiefern die Leistungen in Rechtschreibtests also auf andere Schreibformen übertragbar sind, ist bislang eine offene Forschungsfrage. Ursachen für die Inkonsistenz von Rechtschreibleistungen in verschiedenen Schreibformen werden auf unterschiedlichen Ebenen gesucht: Für ein besseres Abschneiden im eigenen Text können z. B. Fehlervermeidungsstrategien durch eine bewusste Auswahl einfacher oder bekannter Wörter verantwortlich gemacht werden. Ferner können automatisierte Handlungsabläufe im Rechtschreiben dann gestört werden, wenn sie nicht mehr ‚nebenbei‘ passieren, sondern bewusst gesteuert werden. Dies kann – neben psychischen Belastungen wie Angst und Stress – eine Erklärung für

schwache Leistungen in der Situation eines Rechtschreibtests sein, aber durchaus auch in anderen Schreibzusammenhängen auftreten. Diese Einflussfaktoren wurden in der Untersuchung nicht systematisch verfolgt (vgl. aber Fußnote 13).

In der vorliegenden Studie sind die *Handlungsanforderungen* des Schreibens Ausgangspunkt für die Untersuchung des Verhältnisses von Rechtschreibleistung und Schreibzusammenhang. Das Theorem Otto Ludwigs (1995) bildet für diese Betrachtung die Grundlage und zeigt, dass sich die Handlungsanforderungen in verschiedenen „Schreibdimensionen“ (Ludwig 1995, S. 275f.) wesentlich unterscheiden.

Deshalb wird im Folgenden zunächst das Konstrukt ‚Rechtschreibkompetenz‘ und seine Modellierung bezüglich des Handlungszusammenhangs differenziert (1). Daran anschließend geht es um die Fragen, welche Verfahren der Rechtschreibdiagnostik es gibt und welche sich im Besonderen für eine quantitativ und qualitativ differenzierte vergleichende Messung der Rechtschreibleistung im Text und im Test eignen (2). In einer empirischen Untersuchung in Klasse 4 werden die Leistungen im Text und im Test kontrastiv analysiert (3 und 4) und abschließend Konsequenzen für Didaktik und Diagnostik gezogen (5).

## 1 Facetten der Rechtschreibkompetenz

Rechtschreibkompetenz ist im Vergleich zu anderen Domänen der Deutschdidaktik, wohl aufgrund des hohen normativen Charakters der Orthographie, zunächst einfach zu erklären: Sie bedeutet, keine oder wenig Rechtschreibfehler zu machen. Uneinigkeit besteht im wissenschaftlichen Diskurs hingegen über die Auswahl und Relevanz der Lerninhalte, die Modellierung von Entwicklungs- und Niveaustufen und über die Differenzierung in Teilkompetenzen (vgl. Böhnisch 2008). Im Einzelnen:

Die *Lerninhalte* sind geprägt von den zugrunde liegenden schrifttheoretischen Annahmen, wie z. B dem Verhältnis von gesprochener und geschriebener Sprache, den zentralen Einheiten der Schrift oder der Begründung einzelner orthographischer Phänomene (vgl. u. a. die Kontroverse von Eisenberg 2004 und Maas 1992). Für das in der vorliegenden Studie betrachtete Verhältnis von Rechtschreibkompetenz und dem Handlungszusammenhang spielen diese Aspekte eine untergeordnete Rolle, da nicht interessiert, auf welcher theoretischen Basis Schreibstrukturen hergeleitet werden, sondern wie dies in unterschiedlichen Kontexten gelingt.

Die *Entwicklung* von Rechtschreibkönnen zu untersuchen, hat sich in der Schreibforschung der letzten 30 Jahre fest etabliert (vgl. u. a. Frith 1986, Günther 1986, Scheerer-Neumann 2003, Brügelmann/Brinkmann 1994 und Valtin 1988). Neu ist seit der Erarbeitung nationaler Bildungsstandards 2004, diese Forschungsarbeiten zu ‚Kompetenzmodellen‘ auszuweiten. Dazu gehört u. a. das Verständnis des Untersuchungsgegenstandes, des Rechtschreibens, dahingehend zu modifizieren, dass er dem Kompetenzbegriff von Weinert (2001) gerecht wird. Inwiefern es sich dabei tatsächlich inhaltlich um ein Novum handelt, oder ob die Neuerung vornehmlich terminologisch anzusiedeln ist, darf vorsichtig hinterfragt werden. Weiterhin wird

die Entwicklung von Rechtschreibkompetenz auf *Niveaustufen* umgelegt. Ergebnis dieser Bemühungen sind z. B die Arbeiten von Löffler/Meyer-Schepers u. a. (2003) und Voss/Blatt/Kowalski (2007). Löffler/Meyer-Schepers u. a. ordnen auf Basis der IGLU-Daten orthographische Phänomene einer lautanalytischen oder grammatischen Kompetenz und einer darin verorteten elementaren oder erweiterten Stufe zu. Voss/Blatt/Kowalski ordnen auf Basis schriftsystematischer Überlegungen orthographische Phänomene einem Kernbereich und einem Peripheriebereich entsprechend ihrer Regelhaftigkeit zu. Neben einigen zu diskutierenden inhaltlichen Zuordnungen (vgl. u. a. die Kritik von Bremerich-Vos 2004) sind die bestehenden Niveaustufenmodelle auch um den hier betrachteten Gedanken zu ergänzen: Der Schwierigkeitsgrad in der Rechtschreibung bezieht sich nicht nur auf einzelne mehr oder weniger komplexe Schriftstrukturen, sondern eben auch auf die Handlungsanforderungen des jeweiligen Schreibzusammenhanges. Worin sich die Schwierigkeit des Rechtschreibens im eigenen Text und im Rechtschreibtest im Einzelnen manifestiert und wie sich dies in der Rechtschreibleistung niederschlägt, ist Thema dieses Beitrages.

Die *Differenzierung in Teilkompetenzen*, bei Klieme u. a. (2003) „Kompetenzfacetten“ genannt, wird in der Rechtschreibforschung auf ganz unterschiedlichen Ebenen vorgenommen. Hinney unterscheidet 1997 in Anlehnung an Mandl u. a. aus lernpsychologischer Perspektive vier Wissensarten und bezieht diese auf den Rechtschreiberwerb. Auf derselben Grundlage ist auch das mittlerweile etablierte Strukturmodell von Ossner (2006) aufgebaut. May (2007) modelliert Kompetenzfacetten hingegen nach strategischen Zugriffsweisen beim Schreiben, die sich vier Ebenen des Schriftsystems, also Bereichen der Orthographie, zuordnen lassen.

In der vorliegenden Untersuchung spielt die Segmentierung des Konstrukts ‚Rechtschreibkompetenz‘ in Facetten eine entscheidende Rolle. Dabei bezieht sich das Unterscheidungskriterium allerdings auf den Handlungszusammenhang des Schreibens. Dieser geht bei Ludwig (1995) in zwei Dimensionen auf, die mit unterschiedlichen Anforderungen an den Schreiber verknüpft sind. Das Theorem Ludwigs lässt sich auf die Facetten der Rechtschreibkompetenz übertragen (vgl. Abbildung 1).

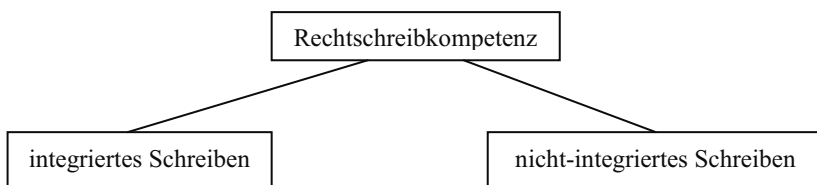


Abbildung 1: Kompetenzfacetten in Anlehnung an Ludwig (1995).

„Integriertes Schreiben“ ist die Schreibdimension, in der das Schreiben in die Textproduktion eingebettet ist. Dem gegenüber steht das „nicht-integrierte Schreiben“. Es liegt vor, wenn beispielsweise ein Lückentext ausgefüllt oder nach Diktat geschrieben wird, wenn also der Prozess des Schreibens von dem der Textproduktion abgetrennt ist. Dies sind, laut Ludwig, „zwei grundverschiedene Schreibpraxen“, die

„Lösungen unterschiedlicher Probleme darstellen“ (Ludwig 1995, S. 275). Somit sind auch für den Rechtschreiber die Handlungsanforderungen im integrierten Schreiben wesentlich höher als die im nicht-integrierten, da er sich – neben dem Beachten der orthographischen Norm – in erster Linie auf Teilprozesse der Textproduktion, wie das Finden von Ideen, das Ordnen der Gedanken und das Formulieren von Sätzen konzentrieren muss und diese „Kettenglieder des Schreibens“, so nennt es Weigl aus psycholinguistischer Sicht, „miteinander rivalisieren“ (Weigl 1975, S. 167). Für die Kompetenzdefinition bedeutet diese „Grundverschiedenheit“, dass eine Differenzierung nach der Schreibdimension sinnvoll ist. Eine Entsprechung bietet das Kompetenzverständnis von Groeben. Er folgt dem interaktionstheoretischen Ansatz, indem er konstatiert, dass Kompetenz eine Wechselwirkung aus dispositionellen Fähigkeiten einerseits und der konkreten Situation andererseits sei (vgl. Groeben 2006, S. 13f.). Rechtschreibkompetenz ist also im Kontext der konkreten Situation, nämlich dem Handlungszusammenhang des Schreibens, zu verstehen.

Es sind aber nicht nur die unterschiedlichen Handlungsanforderungen, die eine solche Differenzierung der Rechtschreibkompetenz erforderlich machen. Ein weiterer wichtiger Faktor ist die funktionale Anwendung der Rechtschreibung: Im Gegensatz zu vielen Formen des nicht-integrierten Schreibens hat die Orthographie im selbst verfassten Text die tatsächliche Funktion einer schriftsprachlichen Norm. Sie ist ein Service für den Leser der Texte und hilft, dem Geschriebenen schnell und sicher Sinn zu entnehmen. Rechtschreibkompetenz im integrierten Schreiben hat also im Hinblick auf den praktischen Gebrauch eine höhere Relevanz als ihr Äquivalent im nicht-integrierten Schreiben. Diese an der realen Kommunikation orientierte Gewichtung der Kompetenzfacetten ist im wissenschaftlichen Diskurs nicht neu. Besonders in den 1990er Jahren wurde im Zuge der verstärkten Kritik an Diktaten, also einer Form des nicht-integrierten Schreibens, auf den hohen Stellenwert des Rechtschreiblernens im selbstständigen Verfassen von Texten aufmerksam gemacht (vgl. u. a. Bartnitzky/Brügelmann/Erichson 1998).

Inwiefern diese theoretische Unterscheidung von Kompetenzfacetten und ihre unterschiedliche Relevanz für den alltäglichen Schreibgebrauch in der Diagnostik berücksichtigt werden, ist im Folgenden die leitende Frage.

## 2 Verfahren der Rechtschreibdiagnostik

Versteht man Kompetenz als theoretisches Konstrukt einer Disposition, die sich in messbaren Leistungen manifestiert, so ist Rechtschreibdiagnostik der Versuch, Leistungen mittels differenzierter Aufgaben so präzise zu messen, dass konkrete Rückschlüsse auf die Rechtschreibkompetenz möglich sind. Diagnostische Intention ist entweder die Selektion in Leistungsniveaus („Selektionsdiagnostik“) oder das Gewinnen von differenzierten Informationen über bereichsspezifische Stärken und Schwächen zur anschließenden Kompetenzförderung („Förderdiagnostik“) (vgl. Herné 2003, S. 884). Bewährt hat sich in beiden Fällen die quantitative und qualitative Analyse von Falschschreibungen. Dass es dabei keinen eindeutig linearen

Zusammenhang von der gemessenen Leistung zu der dahinter stehenden Kompetenz gibt, liegt in der Natur der Sache. So erklärt Ossner: „Es gibt keine Methode, um das grundsätzlich Prekäre theoretischer Begriffe zu minimieren. [...] Man [kann] aus der Fehlerzahl nicht mit Notwendigkeit auf das Vorliegen einer bestimmten Rechtschreibkompetenz schließen“ (Ossner, 2006, S. 7). Dieser Einwand ist grundsätzlich und muss bedacht werden, wenn es im Folgenden um die Empirie von Rechtschreibkompetenz geht.

## 2.1 Rechtschreibtests

Zu den gängigen, also in Schulpraxis und Forschung viel verwendeten, Verfahren der Rechtschreibdiagnostik gehören an erster Stelle standardisierte Rechtschreibtests wie z. B die Hamburger Schreibprobe, HSP (May u. a. 2007), der Diagnostische Rechtschreibtest, DRT (Müller 2004), der Weingartener Grundwortschatz Rechtschreibtest, WRT (Birkel 2007) und der Deutsche Rechtschreibtest, DERET (Stock/Schneider 2008).<sup>1</sup> Des Weiteren werden in der didaktischen Rechtschreibforschung Schreibaufgaben konzipiert und angewandt, die zunächst nicht für die schulische Diagnosepraxis gedacht sind, so z. B in IGLU-E (vgl. Valtin u. a. 2003). All diese Diagnoseverfahren haben gemeinsam, dass sie nicht-integrierte Schreibungen in Form von diktierten Einzelwörtern und -sätzen zur Basis haben.<sup>2</sup> Testökonomisch hat dieses Vorgehen einige bedeutende Vorteile:

- Das Wortmaterial kann gezielt nach orthographischen Gesichtspunkten ausgewählt werden, um rechtschriftliche Fähigkeiten punktuell überprüfen zu können.
- Die Leistungen der Probanden können direkt und schnell miteinander verglichen werden, weil alle dieselben Items zu bearbeiten haben.
- Die Leistungen des Einzelnen können i. d. R. auf Grundlage einer repräsentativen Stichprobe beurteilt und eingeordnet werden.
- Der zeitliche Aufwand in Durchführung und Auswertung ist begrenzt.

Problematisch ist, dass zugunsten der Testökonomie nicht *die* Rechtschreibkompetenz als gesamtes Konstrukt, sondern lediglich *eine* Kompetenzfacette erhoben wird. Dabei handelt es sich um jene Facette, die bezogen auf die reale Kommunikation eine weit untergeordnete Rolle spielt. Die wirklich wichtige Facette der Rechtschreibkompetenz im integrierten Schreiben wird auf diese Weise nicht erreicht. Besonders unter förderdiagnostischen Gesichtspunkten ist diese einseitige Testung

---

1 Zu dieser Auswahl an paper-pencil-Tests kommt eine ständig wachsende Zahl an Online-Diagnoseverfahren, wie z. B die Dortmunder Schriftkompetenzermittlung (DoSE) unter [www.gutschrift-institut.de](http://www.gutschrift-institut.de) und Löffler/Meyer-Schepers 2008; die Münsteraner Rechtschreibanalyse, MRA unter [www.lernserver.de](http://www.lernserver.de) und Schönweiss 2004; fördern@cornelsen unter [www.foerdern.cornelsen.de](http://www.foerdern.cornelsen.de) und lerndesign unter [www.lerndesign.com](http://www.lerndesign.com).

2 Mitgedacht werden darf an dieser Stelle auch die Rechtschreibmessung mittels klassischem Diktat. Diese Form der Diagnostik ist vor allem in den 1990er Jahren in der Fachdidaktik ausführlich diskutiert worden (vgl. auch jüngste Forschungsergebnisse von Birkel 2009).

heikel: gefördert werden soll ja schließlich eine Kompetenz, die alltagstauglich ist; getestet wird aber etwas anderes. Ob die dadurch entstandene eingeschränkte Validität der Rechtschreibtests den Testautoren und den Testbenutzern bewusst ist, bleibt meistens undeutlich; schließlich wird in den Testmanualen i. d. R. angegeben *die* Rechtschreibfähigkeit bzw. -leistung zu erheben (vgl. z. B. die Manuale des DRT 2004, S. 7 und des WRT 1994, S. 7). Dass es sich dabei lediglich um einen Ausschnitt handelt, erkennt May, indem er feststellt: „Gerade in [...] Alltagsschreibungen zeigt sich jedoch die Fähigkeit der Kinder, die erworbene orthographische Kompetenz auch in selbst konstruierten Sätzen und Texten, deren Wortwahl in erster Linie ihrem Kommunikationsbedürfnis entspricht, zu realisieren“ (May 1998, S. 280).

## 2.2 Testunabhängige Verfahren

Neben den populären Rechtschreibtests gibt es eine Reihe testunabhängiger Verfahren, die ihren Schwerpunkt auf die differenzierte qualitative Analyse integrierter Schreibungen legen. An erster Stelle sind hierbei Fehlerschlüssel wie z. B. die Aachener förderdiagnostische Rechtschreibfehleranalyse, AFRA (Naumann/Herné 2005), die Dortmunder Rechtschreibfehler-Analyse, DoRA (Löffler/ Meyer-Schepers 1992) und die Oldenburger Fehleranalyse, OLFA (Thomé/ Thomé 2004) zu nennen, die auf jegliche Schreibprodukte angewendet werden können.<sup>3</sup> Zwar sind diese Verfahren im Gegensatz zu standardisierten Rechtschreibtests nicht empirisch normiert, dafür können sie die Kompetenzfacette im integrierten Schreiben messen und damit die „Ernstsituation“ des Schreibens (Brügelmann; u. a. 1992, S. 7) erfassen. Dies hat außerdem den Effekt, dass diese Form der Diagnostik Wortmaterial zur Grundlage hat, das dem individuellen Sprachstand der Kinder gerecht wird. Geschrieben wird, was dem Ausdruckswunsch entspricht, was zum Schreiben motiviert. Ein solches Schreibprodukt orthographisch zu analysieren bedeutet also auch der pädagogischen Forderung nach ‚individuellem Lernen und Lehren‘ gerecht zu werden. Mindestens genauso wichtig ist jedoch die Tatsache, dass mit Hilfe von Fehlerschlüsseln qualitativ besonders differenziert gemessen werden kann. In der Förderdiagnostik interessiert schließlich nicht nur, in welchem Ausmaß Fehler auftreten (quantitative Analyse), sondern vor allen Dingen, welcher Art die Recht-

---

3 Zu ergänzen sind Verfahren, die gebunden an bestimmte Materialien zum Schriftspracherwerb sind, wie z. B. die Lernstandsdiagnostik des „Individuellen Rechtschreibtrainings“ von Leßmann (2007) für Klasse 1 und 2 und die Beobachtungsbögen des Arbeitshefts „Ich und meine Geschichten“ des Konzepts „ABC Lernlandschaft“ von Brinkmann (2008). Sie analysieren ebenfalls freie Schreibprodukte nach vorgegebenen qualitativen Schriftmerkmalen. Für den frühen Beginn des Schriftspracherwerbs haben Dehn/ Hütts-Graff (2006) die Schulanfangsbeobachtung mit dem „Leeren Blatt“ entwickelt. Hierbei werden die Schreibversuche allerdings ohne vorgegebenen Kriterienkatalog analysiert. May (1998) überträgt die Kategorien der HSP auf die Analyse freier Texte, indem er erst die dazugehörigen Lupenstellen in dem geschriebenen Wortmaterial sucht und diese hinsichtlich der Richtigschreibung prüft.

schreibfehler sind (qualitative Analyse). Besonders die qualitative Veränderung von Fehlern zeigt, „dass die Kinder Fortschritte machen, neue Hypothesen entwickeln und beim Schreiben erproben“ (Tophinke 2005, S. 118). Diese Sorte Lernstandserhebung gelingt mithilfe von Fehlerkategorisierungen, die orthographietheoretisch fundiert sind und Ursachenvermutungen wie beispielsweise die viel diskutierten Kategorien „Wahrnehmungs- und Flüchtigkeitsfehler“ so weit wie möglich vermeiden (vgl. u. a. DRT für 3. Klassen, 2004; Kritik z. B. bei Löffler/ Meyer-Schepers 1992, S. 14).

In der Anwendung von Fehlerschlüsseln auf freie Schreibprodukte sind zwei Dinge zu berücksichtigen:

### *1. Die Möglichkeit der Mehrfachsignierung*

Da es sich bei der Fehleranalyse um eine Zuordnung mit interpretativem Charakter handelt, ist zu bedenken, dass ein Rechtschreibfehler in all jenen Kategorien signiert werden muss, in die er eingeordnet werden kann. So wird verhindert, dass eine Eindeutigkeit suggeriert wird, die es in einer Fehleranalyse nicht geben kann (vgl. Herné/Naumann 2005, S. 19).

### *2. Die Analyse der theoretischen Fehlerverlockung des Wortmaterials*

Die alleinige Anzahl der Rechtschreibfehler, seien sie auch noch so differenziert kategorisiert, sagt wenig über die Rechtschreibleistung aus, wenn das Wortausgangsmaterial bei der Analyse nicht berücksichtigt wird. So ist beispielsweise die Aussagekraft der viel zitierten Fehlerstatistiken von Menzel (1985) und Riehme/Heidrich (1981) kritisch zu betrachten: Auf Basis von mehreren tausend Rechtschreibfehlern wurde hierbei u. a. festgestellt, dass die Kategorie Groß- und Kleinschreibung im Ranking weit vorne liegt (23,3% in der Liste von Menzel und 20,6% bei Riehme/Heidrich). Unbeachtet bleibt dabei die Tatsache, dass es sich dabei um einen Fehlertyp handelt, der in jedem (!) Wort auftreten kann; dessen ‚Auftrittschance‘ also sehr viel höher ist als beispielsweise von silbischen oder morphematischen Kategorien. Dies bemerkt Eichler bereits 1983, indem er die Fehlerstatistik von Riehme/Heidrich einer Statistik über „Fehlermöglichkeiten“ gegenüberstellt und feststellt: „Länge und Kürze [bezogen auf die Kennzeichnung der Vokalquantität; JF] bewirken relativ mehr aktuelle Fehler als theoretisch vorherzusagen sind, die Groß- und Kleinschreibung weniger“ (Eichler 1983, S. 636). Rückschlüsse auf den Schwierigkeitsgrad von orthographischen Phänomenen und darauf basierende Einordnung von Rechtschreibleistungen in Niveausstufen sind mit derlei Verfahren somit höchst problematisch (vgl. dazu auch die Kritik von Risel 2008, S. 50). Die Verfasser der AFRA greifen diese Schwierigkeit auf und weisen darauf hin, dass es – insbesondere bei der vergleichenden Messung von Rechtschreibleistung in frei verfassten Schreibprodukten – unerlässlich ist, die Rechtschreibfehler im Verhältnis zu den theoretischen Fehlermöglichkeiten jeder Kategorie zu analysieren (vgl. Herné/Naumann 2005, S. 17). Diese Erkenntnis ist leitend bei der Entwicklung des Analyseverfahrens der vorliegenden Studie (s. 3.2).



Ein dritter – hier untergeordneter – Aspekt, der bei der Anwendung von Fehler-schlüsseln zu berücksichtigen ist, ist die orthographietheoretische Basis, auf die die Konstruktion der Fehlerkategorien aufbaut. Sie kann bestimmte schriftsystematische Ebenen in der Analyse hervorheben und die Fehlerzuordnung beeinflussen (vgl. bspw. die Zuordnung von Fehlern in der Dehnungs- und Schärfungsmarkierung in der DoRA – phonologisch-silbische Ebene – und der AFRA – phonologisch-morphematische Ebene). Diesbezüglich wurde in der vorliegenden Untersuchung ein Fehlerschlüssel entwickelt, in dem der Fokus auf einer breiten silbischen Ebene der Orthographiesystematik liegt, in der auch solche Kategorien auftreten, die in gängigen Rastern nicht zu finden sind (S. 3.2).

### 3 Methodik der Studie

Die zugrunde liegende *Hypothese* entspricht dem Ausgangspunkt der Untersuchung und dem Theorem der Handlungsanforderung in den beiden Schreiddimensionen: „Die Rechtschreibleistung ist im integrierten Schreiben wegen der höheren Handlungsanforderungen schwächer als im nicht-integrierten Schreiben.“

#### 3.1 Untersuchungsdesign

##### *Sequenzdesign*

Die Untersuchung ist als Sequenzdesign konzipiert und in Tabelle 1 dargestellt.

Klasse	1. Erhebung	2. Erhebung
A	1. Klasse →	2. Klasse
B	1. Klasse →	2. Klasse
C	1. Klasse →	2. Klasse
D	3. Klasse →	4. Klasse
E	3. Klasse →	4. Klasse
F	3. Klasse →	4. Klasse

*Tabelle 1: Darstellung des Sequenzdesigns.*

In den Schuljahren 2003/2004 und 2004/2005 wurden in sechs verschiedenen Klassen aus Niedersachsen und Nordrhein-Westfalen jeweils in den letzten drei Wochen des Schuljahres Datenerhebungen durchgeführt. Zum ersten Erhebungszeitpunkt waren die Kinder in der 1. und in der 3. Klasse; ein Jahr später wurden dieselben Kinder wieder untersucht, dann waren sie in der 2. und in der 4. Klasse. Für die Analyse der Entwicklung der Rechtschreibleistung von Klasse 1 bis 4 wurden die Daten beider Erhebungszeitpunkte berücksichtigt. Die hier betrachtete vergleichende Analyse basiert auf den Daten aus der 2. Erhebung mit einem besonderen Schwerpunkt auf Klasse 4.

### *Stichprobenziehung*

Die Gesamtstichprobe umfasst 170 Kinder, gearbeitet wurde mit einer Teilstichprobe von 60 Kindern ( $N = 60$ , davon 30 w und 30 m). Das entspricht 30 Kindern pro Jahrgang und 10 Kindern aus jeder Klasse; die hier betrachtete Teiluntersuchung basiert also auf *30 Kindern*. Bei der Teilstichprobe handelt es sich um eine nicht-probabilistische Stichprobenziehung. Da in erster Linie durchschnittliche Leistungen einer Jahrgangsstufe betrachtet werden, um Aussagen über eben diese Entwicklungsstufe zu machen, ist es wichtig, dass die Leistungen auf einer Stichprobe basieren, die von spezifischen Merkmalen unabhängig ist. Konstant gehalten wurden daher:

- das Geschlecht
- das unterrichtliche Konzept zum Schriftspracherwerb (gleich viele Probanden wurden jeweils mit den Konzepten „Fibel“, „Lesen durch Schreiben“ und „Silbenanalytische Methode“ unterrichtet)<sup>4</sup>
- die Ausgangslage in der Rechtschreibleistung.

Da in der Analyse die Mittelwerte der Rechtschreibleistungen in den Dimensionen ‚Test‘ und ‚Text‘ in jeder Jahrgangsstufe verglichen werden, ist insbesondere eine gleiche oder ähnliche Ausgangslage in der Rechtschreibleistung der Kinder vonnöten. Diese wurde mit Hilfe eines *matched samples*, also einer gestuften Parallelisierung der Probandengruppen, erreicht. Aus den Klassen wurden jeweils 10 Kinder so ausgewählt, dass insgesamt pro Jahrgang gleich viele unterdurchschnittliche, durchschnittliche und überdurchschnittliche Rechtschreiber vertreten waren. Grundlage für diese Bestimmung sind die Prozenträge der Kategorie ‚Graphemtreffer‘ in der regulären Auswertung der HSP-Testungen.<sup>5</sup>

Die gezogenen Probanden repräsentieren also nicht die durchschnittliche Leistung ihrer Klasse, sondern bringen – im Gegenteil – in ihrer Ausgangslage eine klassenspezifische ausgeglichene Leistung mit.

Der sozioökonomische Hintergrund der Probanden wurde zum Teil über eine behördliche Einstufung des Bildungsgrades der Elternschaft der teilnehmenden Schulen und zum Teil auf Basis von Einschätzungen der Klassen- und Schulleitung erhoben. In dem vorliegenden Teil der Studie wird dieser Faktor aber nicht weiter berücksichtigt (vgl. dazu Fay 2010).

### *Datenbasis*

Die Datenbasis der Studie umfasst 120 Schülertexte; der vorliegenden Fragestellung liegen *30 Texte* zugrunde. Bei der Erhebung der Textdaten las die Untersuchungslei-

---

4 Näheres zu den Konzepten und ihrer Wirkung auf den Schriftspracherwerb vgl. Weinhold 2006.

5 Die Einteilung der Prozentranggruppen 0-40, 41-60 und 61-100 in unter-, über- und durchschnittliche Leistung entspricht nicht einer diagnostischen Interpretation der Prozenträge, sondern lediglich einer einfachen mathematischen.

terin den Kindern in einem Stuhlkreis das Bilderbuch „Eine Schlacht? Die könnt ihr haben!“ von Andreas Röckener vor. Anschließend wurde den Kindern eine Auswahl von Schreibblättern mit Motiven aus der Geschichte bzw. leere Schreibblätter angeboten. Die Schreibaufgabe lautete: „Schreib deine eigene Geschichte dazu“. Auf diese Weise konnten freie Texte entstehen, da die Kinder in ihrem Schreibprozess nicht durch Vorgaben wie z. B. Bildfolgen, Schlüsselwörter oder Themenumrisse eingeschränkt worden sind.

Zur Analyse der nicht-integrierten Schreibungen liegen 60 Hamburger Schreibproben (HSP 2 und HSP 4/5) vor; für die vorliegende Fragestellung werden 30 HSP 4/5 verwendet. Damit wird eine besondere Form des nicht-integrierten Schreibens gewählt, die beispielhaft für gängige Rechtschreibdiagnoseverfahren steht und sich sowohl in der didaktischen Forschung wie auch im schulischen Bereich etabliert hat (vgl. u. a. Weinhold 2009). Im Vergleich der Rechtschreibleistungen in Text und Test spielt die reguläre Auswertung der HSP jedoch keine weitere Rolle, da beide Schreibdimensionen mit demselben – eigens entwickelten – Verfahren analysiert werden; die HSP liefert hierbei also nur das Wortmaterial der nicht-integrierten Schreibungen.

### 3.2 Analyseinstrument

Für die qualitative Analyse der Schreibungen wurde ein eigenes Rechtschreibanalyseraster entwickelt. Das Raster umfasst 36 Fehlerkategorien, die vier schriftsystematischen Ebenen zugeordnet werden. Sie sind teilweise aus bestehenden Fehlerschlüsseln wie z. B. der AFRA und der DoRA übernommen, insbesondere in den Bereichen silbischer und syntaktischer Phänomene sind sie aber ergänzt bzw. verfeinert. Aus Gründen der Datenreduktion wurden in der Studie nicht alle Kategorien verwendet, sondern es wurde ein Ausschnitt gewählt. Dabei handelt es sich um 10 silbische Kategorien (vgl. Tabelle 2).<sup>6</sup>

Diese Kategorien sind deshalb besonders interessant, weil die silbische Struktur in ihrer Gesamtheit in der Fehleranalyse bislang wenig erforscht ist. Einige klassische Fehlerkategorien sind auch in gängigen Rastern wiederzufinden, wurden aber noch nicht im Vergleich der Schreibdimensionen untersucht (z. B. <ie>, LV, KV). Die *Hyperkorrekturen durch überflüssige orthographische Elemente* werden nach May u. a. (2007) nach legaler und illegaler Position in Bezug auf die Betonung und die Position zum Vokal unterschieden (vgl. May u. a. 2007, S. 55). Sie ermöglichen wichtige Einblicke in den Lernprozess der Schreiber, da sie zeigen, welche schriftstrukturellen Markierungen schon bekannt sind, auch wenn sie noch nicht richtig eingesetzt werden können. Die Reduktionssilbe wird in drei Kategorien zur <e>-*Verschriftung des Reims* und in einer Kategorie zum *Silbeninitialen-<h>* untersucht.

---

6 Orthographietheoretisch sind die silbischen Kategorien an der Systematik von Maas (1992) orientiert (vgl. auch das Raster von Thelen 2006). Ergänzt sind die Kategorien der Hyperkorrektur in Anlehnung an May u. a. (2007).

Diese treten in dieser Form in keinem gängigen Kategoriensystem auf,<sup>7</sup> obwohl es spezifisches silbenbezogenes Wissen braucht, um sie richtig zu verschriften, da sie nicht über die Umgangssprache herleitbar ist.

Kürzel	Kategorie	Beispiel
<ie>	<ie>-Schreibung (/i:/ in betonter Silbe, Regelfall)	*spilen
<i>	<i>-Schreibung (/i:/ in betonter Silbe, Ausnahme)	*Tieger
LV	Prominente Silbe, Kennzeichnung des langen Vokals (Dehnungs-h, Doppelvokal)	*faren
KV	Prominente Silbe, Kennzeichnung des kurzen Vokals (Geminate bzw. <ck, tz>)	*komen
ÜOE-L	Überflüssiges orthographisches Element in legaler Position	*Köhnig
ÜOE-IL	Überflüssiges orthographisches Element in illegaler Position	*Bläter
RS-<e>	Reim der Reduktionssilbe mit <e>	*Heft (für <Hefte>)
RS-<er/ern/ert>	Reim der Reduktionssilbe mit <er/ern/ert>	*Kinda
RS-<en/em/el/eln>	Reim der Reduktionssilbe mit <en/em/el/eln>	*schlafn, *Pinsl
RS-<h>	Silbeninitiales-h	Sie *gen (für <gehen>)

Tabelle 2: Ausschnitt aus dem Rechtschreibkategorienraster von Fay 2010.

Die Ergebnisse der Untersuchung beziehen sich daher explizit auf diese orthographe-systematische Ebene. Ob sich dieselben Ergebnisse auch in den übrigen Kategorien zeigen, muss erst noch überprüft werden.

### 3.3 Analyseverfahren

In der Anwendung dieses Kategorienrasters auf die beiden Schreibdimensionen müssen in quantitativer Hinsicht zweierlei Dinge berücksichtigt werden: Zum einen unterscheiden sich die Schülertexte untereinander und die Texte und die Tests teilweise sehr in Umfang und in der Komplexität des geschriebenen Wortmaterials, was den direkten Vergleich der Leistungen erschwert. Zum anderen soll dem Umstand Rechnung getragen werden, dass eine differenzierte Analyse der Rechtschreibleistung nur dann sinnvoll ist, wenn die theoretischen Fehlermöglichkeiten der Schreibprodukte einbezogen werden. Deshalb werden die tatsächlich gemachten Fehler im Verhältnis zu den möglichen Fehlern für jede Kategorie prozentual bestimmt. Dieses Verfahren gliedert sich in drei Schritte:

1. Zunächst werden in dem Schreibprodukt die theoretischen Fehlermöglichkeiten je Kategorie gezählt. Damit wird die so genannt „Basisrate“ (BR) ermittelt (vgl. Naumann/Herné 2005, S. 18). Z. B: Wie oft müsste im vorliegenden Schreibprodukt ein <ie> geschrieben werden?

<sup>7</sup> Ausnahme bildet das Kategorienraster von Thelen, das explizit silbenbasiert ist. Vgl. u. a. Thelen 2006.

2. Im zweiten Schritt werden die tatsächlich gemachten Fehler, die „Fehlerrohwer-te“ (FR), analysiert. Z. B: Wie oft wurde dabei das erforderliche <ie> nicht (oder anders, z. B als einfaches <i>) verschrifet?
3. Diese werden im dritten Schritt im Verhältnis zu ihren Basisraten ausgedrückt. So entsteht ein „Fehlerquotient“ (FQ) für jede Kategorie und jedes Schreibprodukt. Z. B: Von sechs zu schreibenden <ie> wurden fünf nicht korrekt verschrifet → Fehlerquotient 83%.

Diese Form der Analysearbeit ist umfangreich: Bei 10031 zu analysierenden Wörtern und 10 Fehlerkategorien mussten insgesamt 300.930 einzelne Analyseschritte ‚zu Fuß‘ gegangen werden.

Bei einigen Kategorien ist die Bestimmung der Basisraten nicht ganz einfach. Im Gegensatz zu Kategorien wie beispielsweise der <ie>-Schreibung, bei der lediglich die <ie>-Grapheme des korrekten Wortmaterials gezählt werden müssen, muss z. B in den Kategorien *Hyperkorrektur durch überflüssige orthographische Elemente, legal und illegal* eine komplexere Voranalyse durchgeführt werden. Basisraten sind hier die Anzahl der Stellen im Wort, in denen derlei Fehler auftreten können. In der legalen Position beispielsweise nach dem Vokal der prominenten und untergeordnet prominenten Silbe. Dies führt dazu, dass die Basisraten in diesen Kategorien sehr viel höher liegen als die Basisraten der Kategorien, in denen das orthographische Phänomen an sich zu zählen ist. Um die Darstellung und Vergleichbarkeit zwischen den Kategorien zu erleichtern, ist es daher sinnvoll, die Kategorien nach der Höhe ihrer durchschnittlichen Basisraten und damit einhergehend auch nach der Bestimmbarkeit des orthographischen Phänomens zu gruppieren. Eine Einteilung in die Gruppen  $BR \geq 15$  und  $BR > 15$  hat sich nach ersten Voruntersuchungen als günstig erwiesen.

Dennoch hat die Höhe der Basisrate auch innerhalb der so gebildeten Gruppen – das bringt das Ausdrücken in Prozent mit sich – einen hohen Einfluss auf das Ergebnis. Der Fehlerquotient allein sagt wenig über die Leistung aus, wenn der Wert der Basisrate nicht berücksichtigt wird. Taucht ein Phänomen beispielsweise nur einmal in einem Text auf und wird falsch verschrifet, so liegt der Fehlerquotient bei 100. Im Vergleich zu häufig auftretenden Phänomenen würde dieses Ergebnis falsch interpretiert, nämlich überbewertet werden, wenn die Höhe der Basisrate nicht mitbedacht wird. Deshalb wurde für die Studie eine Darstellung der Rechtschreibleistung entwickelt, in der der Fehlerquotient im Verhältnis zu seiner Basisrate veranschaulicht wird. Für die Beantwortung der ersten Frage, die nicht Gegenstand dieses Beitrages ist, wird in dem Diagramm zugleich die Leistungsentwicklung von Klasse 1 bis 4 sichtbar gemacht (vgl. Abbildung).<sup>8</sup>

---

8 Da hier der Vergleich der Schreibdimensionen ‚Text vs. Test‘ in der 4. Klasse im Vordergrund steht, spielt die Diagrammdarstellung im Weiteren keine Rolle. Die Analyse von Entwicklungsverläufen der Rechtschreibleistung auf Basis freier Texte, in denen diese Darstellung eine wichtige Grundlage bildet, ist bei Fay (2010) nachzulesen.

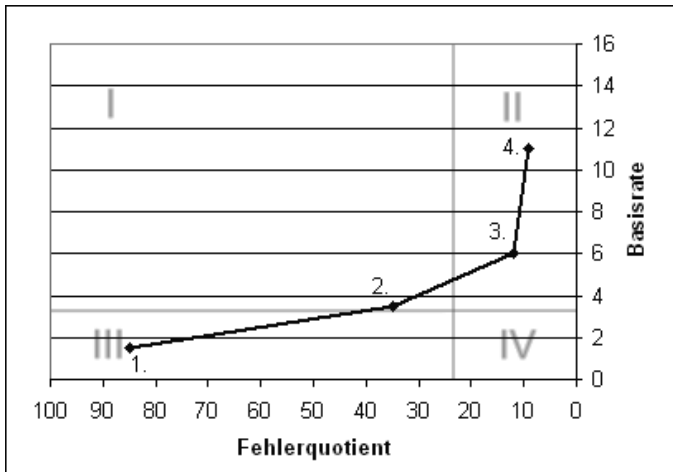


Abbildung 2: Darstellung der Rechtschreibleistung in Fehlerquotient und Basisrate.

Der Graph in dem hier dargestellten Beispiel (es handelt sich um die Kategorie „KV; Kennzeichnung des kurzen Vokals“) zeigt eine typische Leistungsentwicklung von Klasse 1 bis 4 im integrierten Schreiben. Dabei sind zwei Entwicklungen zu beachten: Das in der Fehlerkategorie betrachtete Rechtschreibphänomen wird im Laufe der vier Jahre häufiger verwendet; die Basisrate steigt. Gleichzeitig ist zu sehen: Während zu Beginn des Schriftspracherwerbs noch besonders viele Fehler gemacht werden, der Graph also gemäß hohem Fehlerquotienten weit links auf der x-Achse beginnt, wandert er mit zunehmender Rechtschreibkompetenz weiter nach rechts; der Fehlerquotient sinkt. Eine zusätzliche Hilfe bei der Beschreibung und Deutung der Ergebnisse ist die Einteilung des Diagrammes in Felder – gekennzeichnet durch die grauen senkrechten bzw. waagerechten Linien im Diagramm. Diese Einteilung wird auf Basis der Mittelwerte aller Fehlerquotienten und aller Basisraten vorgenommen. Die Leistung in Abbildung 2 wäre z. B im 3. und 4. Schuljahr vergleichsweise positiv einzuschätzen, denn in dem Feld II werden Phänomene häufig verwendet und dabei auch häufig richtig geschrieben.

## 4 Ergebnisse

Im Folgenden werden die Ergebnisse des Vergleichs der beiden Schreibdimensionen aus dem 4. Jahrgang vorgestellt. Sie beziehen sich ausschließlich auf die zehn ausgewählten oben vorgestellten silbischen Kategorien.

### 4.1 Rechtschreibleistung im Vergleich

Um zu überprüfen, ob die unterschiedlichen Anforderungen in den beiden Schreibdimensionen gemäß der Hypothese zu schwächeren Ergebnissen in den frei

verfassten Texten führen, werden die Gesamtfehlerquotienten (GesamtFQ) pro Schreibdimension in jeder Kategorie verglichen.<sup>9</sup> Sie sind im Folgenden tabellarisch einander gegenüber gestellt (vgl. Tabelle 3).

	GesamtFQ Text		GesamtFQ Test	Chi2-Wert	p-Wert	df	Kontingenz koeffizient- Wert	N	Signifikant?
<ie>	14,99	<	25,00	43,983	0,021	27	0,776	29	*
<i>	0,00	<	6,06	/	/	/	/	3	
LV	24,09	<	26,49	19,513	0,724	24	0,670	24	
KV	16,59	>	16,55	1,731	0,128	153	0,923	30	
ÜOE-L	0,89	<	1,54	51,773	0,766	60	0,796	30	
ÜOE-IL	0,09	<	0,34	1,014	1,000	15	0,181	30	
RS-<e>	1,42	>	1,11	0,667	0,955	4	0,147	30	
RS-<er/ern/ ert>	0,81	<	5,22	29,000	0,000	4	0,701	29	*
RS-<en/em/ el/eln>	1,61	<	4,88	5,000	0,416	5	0,378	30	
RS-<h>	0,00	<	20,00	/	/	/	/	15	

Tabelle 3: Vergleich der Gesamtfehlerquotienten Text vs. Test; Jahrgang 4.

Da bei diesem Vergleich jedoch die Schwierigkeit besteht, dass hierbei nicht-metrische Häufigkeiten mit unterschiedlichen Basisraten miteinander verglichen werden, wird diese zunächst einfache quantitative Betrachtung statistisch überprüft. Der Chi-Quadrat-Test, der im Gegensatz zum T-Test nicht der ungeprüften Annahme folgt, dass das Maß ‚Fehler‘ ein metrisches sei, gibt Auskunft über die Signifikanz des Größenverhältnisses. Zusätzlich ist der mit dem jeweiligen Chi-Quadrat-Test assoziierte Kontingenzkoeffizient angegeben. Die analysierten Fälle N variieren je nach Anzahl der Kinder, die in ihrem Schreibprodukt das betrachtete Phänomen verwendet bzw. nicht verwendet haben, deren Basisrate also = 0 ist, so dass kein Fehlerquotient berechnet wird.

Die Gegenüberstellung zeigt: Das Verhältnis der Gesamtfehlerquotienten entspricht nur in zwei von zehn Fehlerkategorien den Erwartungen (in der Tabelle grau unterlegt); die Leistung ist hier im Text durchschnittlich schwächer als im Test.<sup>10</sup> In allen übrigen Kategorien ist das Leistungsverhältnis umgekehrt; in den Kategorien <ie> und RS-<er/ern/ert> ist es mit  $p = .021$  und  $p = .000$  sogar hoch signifikant. Das bedeutet: *Die Hypothese zum Einfluss der Handlungsanforderungen kann für die hier betrachteten Viert-Klässler nicht bestätigt werden.* Trotz der höheren Anforderungen im Textschreiben sind die Leistungen in den silbischen Kategorien im Rechtschreibtest schwächer.

9 Der „Gesamtfehlerquotient“ setzt sich aus der Summe aller Basisraten und der Summe aller Fehlerrohwerte der jeweiligen Kategorie zusammen und bezieht sich auf die Leistung des gesamten Jahrgangs (N = 30). Dieser Wert ist deshalb ein geeignetes Maß zur Feststellung der durchschnittlichen Jahrgangsleistung, da das – mitunter näherliegende – Mitteln der Fehlerquotienten durch die fehlende Intervallskalierung nicht möglich ist.

10 Allerdings sind die Unterschiede in den Kategorien KV und ÜOE-IL so minimal, dass die Leistungen in Text und Test als annähernd gleich beschrieben werden kann.

Ein kurzer Exkurs in die Ergebnisse aus Klasse 2 erhärtet diese Feststellungen, denn in der Entwicklung der Leistungen von Jahrgang 2 zu Jahrgang 4 zeigt sich der Unterschied zwischen den Schreibdimensionen besonders prägnant:

In allen drei Kategorien zur Schwa-Verschription in der Reduktionssilbe ist im Rechtschreibtest eine Verschlechterung der Leistung festzustellen; der Fehlerquotient steigt (vgl. Abbildung 3). Im frei verfassten Text hingegen zählt die Reduktionssilbe zu den Kategorien, die bereits ab dem Ende der 1. Klasse vergleichsweise wenig Probleme bereiten (vgl. Fay 2010 und Röber 2006). Gleiches ist bei der Entwicklung in der Kategorie *Hyperkorrektur durch überflüssige orthographische Elemente, legal* festzustellen. Während sich die Leistung im Text vom 2. zum 4. Jahrgang erwartungsgemäß verbessert, verschlechtert sie sich im Rechtschreibtest. Die Kinder setzen hier also vermehrt überflüssige orthographische Elemente (vgl. Abbildung 3: die durchgezogenen, ansteigenden Graphen sind die Leistungen im Test, die gestrichelten, abfallenden Graphen die im Text. Die Wertetabelle zeigt die Gesamtfehlerquotienten der einzelnen Kategorien in Klasse 2 und 4).

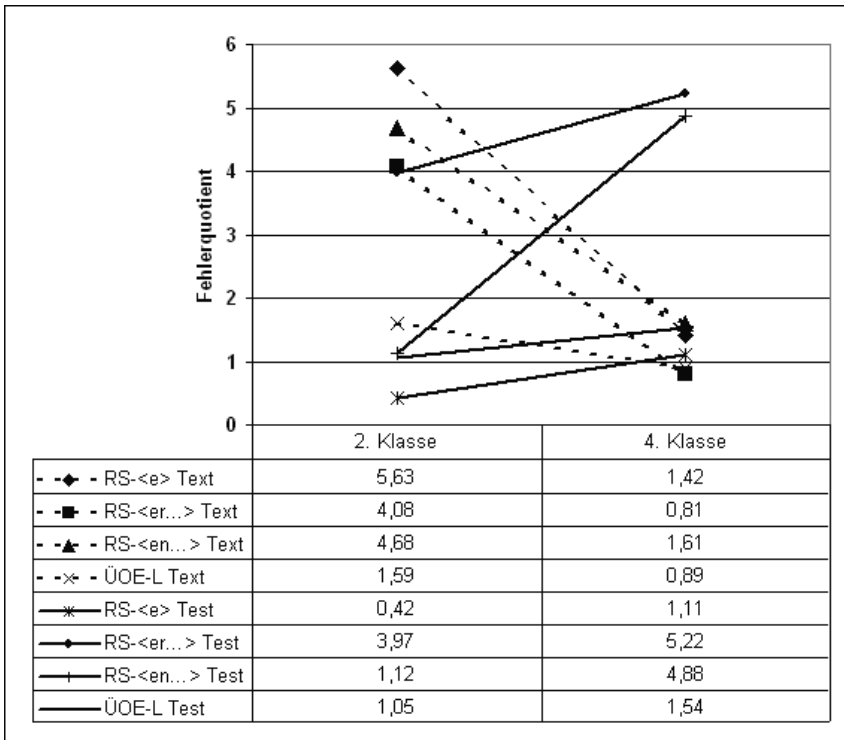


Abbildung 3: Entwicklung der Rechtschreibleistung von Klasse 2 zu 4; Text vs. Test.



May (1993) stellt ebenfalls fest, dass sich die Anzahl der Hyperkorrekturen in der HSP in diesem Zeitraum erhöht. Er bringt dieses Ergebnis allerdings nicht mit einer Leistungsentwicklung im integrierten Schreiben in Verbindung.

Diese Leistungsunterschiede in Text und Test zeigen, dass die Schreiblerner offenbar nicht nur insgesamt am Ende der Grundschulzeit schwächere Leistungen im Rechtschreibtest erbringen, sondern auch in der Entwicklung ihres orthographischen Strukturwissens im Rechtschreibtest verunsichert sind.

Wie ist dieser Leistungsunterschied zwischen den beiden Schreibdimensionen erklärbar? Der Schlüssel liegt im ersten der drei Analyseschritte (s. 3.3), in der Analyse der Basisraten.

## 4.2 Basisraten im Vergleich

Die Höhe der Basisraten spiegelt gewissermaßen die orthographischen Anforderungen eines Schreibproduktes wider. Sie gibt die Anzahl möglicher Falschreibungen innerhalb einer Analysekategorie an und es gilt: Je höher die Basisrate desto mehr ‚Fehlerfällen‘ liegen vor.

I. d. R. korrelieren die Basisraten mit der Anzahl der geschriebenen Wörter in den frei verfassten Texten erwartungsgemäß ( $r = 0,9$ ). Es gilt: Je länger ein Text ist, desto höher sind die Basisraten. Wird nun die durchschnittliche Anzahl der geschriebenen Wörter in Text und Test miteinander verglichen, ist festzustellen, dass die Kinder in ihren eigenen Texten wesentlich mehr schreiben, nämlich mehr als doppelt so viel wie im Rechtschreibtest (durchschnittlich 110,8 Wörtern und Satzzeichen im Text und 47 in der HSP 4). Es liegt daher zunächst die Vermutung nahe, dass die Basisraten in den meisten Fehlerkategorien im Text durchschnittlich höher sind als im Test. Der Vergleich zeigt jedoch, dass dem nicht so ist (vgl. Tabelle 4).

	Mittelwert BR Text	Standard- abweichung Text	Verhältnis BR Text vs. BR Test	Mittelwert BR Test	Standard- abweichung Test
<ie>	4,00	2,32	>	2,93	0,25
<i>	0,20	0,66	<	1,00	0,00
LV	2,77	2,75	<	3,90	0,31
KV	7,83	4,79	<	14,90	0,31
ÜOE-L	74,87	27,07	>	62,83	0,79
ÜOE-IL	110,13	41,32	>	88,90	0,76
RS-<e>	11,73	6,51	>	9,03	0,18
RS-<er/ern/ert>	4,10	3,42	<	8,93	0,37
RS-<en/em/el/eln>	14,50	10,13	>	4,10	0,31
RS-<h>	0,83	0,99	<	3,00	0,00

Tabelle 4: Mittelwerte und Standardabweichung der Basisraten im Text und im Test; Jahrgang 4.

Die Tabelle zeigt die Mittelwerte der Basisraten für jede Kategorie im Vergleich.<sup>11</sup> Grau unterlegt sind jene Größenverhältnisse, die der vorformulierten Vermutung entsprechen; die Basisraten sind im Text durchschnittlich höher als im Test. In den Kategorien <i>, LV, KV, RS-<er...> und RS-<h> ist dieses Verhältnis jedoch umgekehrt. Was bedeutet das?

Im Rechtschreibtest liegt eine *Verdichtung orthographischer Phänomene* vor; auf wenige Wörter verteilt treten gebündelt ‚Fehlerfallen‘ auf. Testkonzeptionell ist das gewollt, schließlich „müssen [die Wörter] solche Phänomene in genügender Anzahl enthalten, die die wichtigsten alphabetischen, orthographischen und morphematischen Zugriffe prüfen [...]“ (May u. a. 2007, S. 15). Dies wird erreicht, indem besonders komplexes Wortmaterial benutzt wird, meistens in Form von Komposita (in der HSP 4 werden – je nach Zählung – über 15 Komposita verlangt). Auf diese Weise können innerhalb eines Items gleich mehrere Rechtschreibstrukturen überprüft werden. Demzufolge sind die orthographischen Anforderungen im Rechtschreibtest weitaus höher als im frei verfassten Text.<sup>12</sup>

Dem Einflussfaktor ‚Handlungsanforderungen‘ ist also ein konträr wirkender Faktor ‚orthographische Anforderungen‘ entgegengesetzt. Die Rechtschreibleistung ist beim nicht-integrierten Schreiben eines Rechtschreibtests in den untersuchten Kategorien schlechter als beim integrierten Schreiben eines frei verfassten Textes, da die ‚Dichte der Fehlerfallen‘ im Rechtschreibtest wesentlich höher ist. Dies führt offenbar nicht nur in der Summe sondern auch in der Relation zu möglichen Richtigschreibungen (Fehlerquotient) bei denselben Schreibern zu eklatanten Leistungsunterschieden.<sup>13</sup>

---

11 Die zusätzlich angegebene Standardabweichung zeigt das grundsätzlich Prekäre, wenn heterogene Schreibprodukte wie hier die 30 Texte der untersuchten Viert-Klässler zu einem Wert zusammengefasst werden, um einen Durchschnitt darzustellen: Die jeweiligen Basisraten in den Texten streuen teilweise hoch, so dass Rückschlüsse auf die gesamte Kohorte nur behutsam gezogen werden dürfen.

12 May weist selber die Auswirkungen der hohen Fehlerfallendichte des Wortmaterials der HSP nach, indem er zeigt, dass dieselben Lupenstellen oder Morpheme als Ganzes im Kontext eines Kompositums öfter und länger falsch geschrieben werden als in Simplicia. Vgl. May 1993. Naumann/Herné greifen jüngst das Thema „Fehlerfallendichte“ auf und untersuchen ihrerseits den Einfluss von Itemkomplexität auf die Rechtschreibleistung (vgl. Naumann/Herné 2010).

13 Die Vermutung, die Rechtschreibleistung im eigenen Text sei deshalb besser, weil die Schreiber das Wortmaterial bewusst nach orthographischen Kriterien aussuchen können, wurde zwar nicht systematisch überprüft, ist aber nach gründlicher Sichtung der Texte höchst unwahrscheinlich: Die Probanden haben in ihren Texten fast ausschließlich die als Schreiben Anlass vorgelesene Geschichte nacherzählt. Das verwendete Wortmaterial folgt daher der Semantik dieser Geschichte, eingeschlossen orthographisch komplexe Wortformen wie z. B. <ein fliegender Kurier>, <Äpfel zerstampfen> und <Schneeballschlacht>. Psychische Einflussfaktoren wie Angst und Stress in der Prüfungssituation des Rechtschreibtests sowie eine Verunsicherung durch das Bewusstmachen von mitunter automatisierten Rechtschreibprozessen (im Hinblick auf die Erhebung von Verbaldaten

## 5 Schluss

Ausgangspunkt der Studie war „die allgemeine Erfahrung, dass – bei vergleichbar schwierigen Wörtern – beim Schreiben von Texten meist erheblich mehr Rechtschreibfehler vorkommen als beim Schreiben diktierter Wörter und Sätze“ (May u. a. 2007, S. 77). Durch die Integration der Orthographie in konzeptionelle Textproduktionsprozesse steigen die Handlungsanforderungen und damit – so die Vermutung – auch die Rechtschreibfehler. Nach empirischer Untersuchung der Leistungen von 30 4.-Klässlern, kann diese Tendenz für den Vergleich freier Schreibungen und diktierter Wörter und Sätze aus der ‚Hamburger Schreibprobe‘ im Bereich silbischer Schreibungen jedoch nicht bestätigt werden. Die ‚Dichte der Fehlerfallen‘ im Rechtschreibtest wiegt offenbar schwerer als die konzeptionellen Anforderungen des Textschreibens: Die Leistungen der Kinder ist in acht von zehn silbischen Kategorien schwächer als im Text.

Welche Konsequenzen sind aus diesem Ergebnis zu ziehen?

*Für die Kompetenzmodellierung* bedeutet es, dass das Theorem ‚Rechtschreibkompetenz‘ nach Schreibdimensionen differenziert werden muss. Die Kompetenzfacetten im integrierten Schreiben eines Textes und im nicht-integrierten Schreiben eines Tests sind hinsichtlich der Handlungsanforderungen und der orthographischen Anforderungen unterschiedlich anspruchsvoll; sie sind hinsichtlich der alltäglichen Schreibpraxis unterschiedlich relevant und sie führen – das hat diese Studie gezeigt – zu unterschiedlichen Rechtschreibleistungen.

*Für die Rechtschreibdiagnostik* offenbart sich ein Validitätsproblem: Was in gängigen Rechtschreibtests erfasst wird, ist tatsächlich eine Leistung, die auf die Kompetenzfacette im integrierten Schreiben schwerlich übertragbar ist. Bedenkt man, dass aus pädagogischer Sicht „doch das Ziel [bleibt], einen selbst verfassten Text orthographisch korrekt, d. h. für Leser angemessen, aufs Papier bringen zu können“ (Augst/Dehn 2002, S. 308), muss auch in der Rechtschreibdiagnostik der Blick hinsichtlich einer Differenzierung nach Kompetenzfacetten erweitert und ein Fokus auf die Analyse integrierter Schreibungen gelegt werden.

*Für die Lehrerbildung* bedeutet das, dass die Förderung von diagnostischer Kompetenz untrennbar verbunden ist mit der Förderung von Fachkompetenz. Nur durch ein breites Fachwissen in der Orthographiesystematik kann es Lehrkräften gelingen, mit geschultem Blick auf integrierte Schreibungen zu schauen und Fehler und Basisraten nach qualitativen Merkmalen einzuordnen. Beide Wissensbereiche – Fach- und diagnostische Kompetenz – schlagen sich direkt in den Rechtschreibleistungen der Schüler nieder: Es sind diejenigen Schüler vergleichsweise rechtschreibstark, deren Lehrer über hohes Sachwissen *und* hohe diagnostische Kompetenz verfügen (vgl. Roos/Schöler 2009, S. 189f.).

---

als „Beobachterparadox“ beschrieben; vgl. Labov 1972, S. 209) können in der vorliegenden Studie jedoch nicht vollends ausgeschlossen werden.

*Für die sprachdidaktische Forschung* bedeutet das, dass Verfahren, die quantitative und qualitative Methoden der Rechtschreibanalyse verknüpfen, weiterentwickelt werden müssen. Es gilt u. a. nach Basisraten zu gewichten und hohe Streuungen bei Gruppenvergleichen zu berücksichtigen. Ein Ausgangspunkt für derlei Modifikationen bietet das Verfahren der hier vorgestellten Studie.

## Literatur

- Bartnitzky, Horst/ Erichson, Christa/ Brügelmann, Hans: Fördert das Rechtschreiblernen – schafft die Klassendiktate ab. In: Grundschulverband aktuell, 1998, S. 3-5.
- Birkel, Peter: Rechtschreibleistung im Diktat – eine objektiv beurteilbare Leistung? In: Didaktik Deutsch, H. 27 (2009), S. 5-33.
- Birkel, Peter: Weingartener Grundwortschatz Rechtschreib-Test für erste und zweite Klassen. WRT 1+; Manual. Göttingen 2007.
- Blatt, Inge/ Voss, Andreas/ Kowalski, Kerstin: Zur Erfassung orthographischer Kompetenz in IGLU 2006. Dargestellt an einem sprachsystematischen Test auf Grundlage von Daten aus der IGLU-Voruntersuchung. In: Didaktik Deutsch, H. 23 (2007), S. 15-32.
- Böhnisch, Martin: Diskussionslinien innerhalb der Kompetenzdebatte. Ein Strukturierungsversuch. In: Didaktik Deutsch, Sonderheft 2008, S. 5-20.
- Bremerich-Vos, Albert: Rechtschreibstandards, Kompetenzstufen und IGLU – einige Anmerkungen. In: Ders. u. a. (Hrsg.): Neue Beiträge zur Rechtschreibtheorie und Didaktik. Freiburg 2004, S. 85-104.
- Brinkmann, Erika: ABC-Lernlandschaft. Seelze 2008.
- Brügelmann, Hans/ Brinkmann, Erika: Stufen des Schriftspracherwerbs und Ansätze zu seiner Förderung. In: Brügelmann, H./ Richter, S. (Hrsg.): Wie wir recht schreiben lernen. 10 Jahre Kinder auf dem Weg zur Schrift. Lengwil am Bodensee 1994, S 44-53.
- Brügelmann, Hans/ Spitta, Gudrun/ Lange, Inge: Zum ‚Laufstalleffekt‘ im Rechtschreibunterricht. Neue Ergebnisse aus dem ‚Schreibvergleich Ost-West‘ 1990 und 1991. In: Balhorn, H. (Hrsg.): Fibel ade? Lesen und Schreiben in der Grundschule. DGLS-Beiträge. Hamburg 1992, S. 3-13.
- Dehn, Mechthild/ Hüttis-Graff, Petra: Zeit für Schrift 2. Beobachtung und Diagnose. Berlin 2006.
- Eichler, Wolfgang: Kreative Schreibirrtümer. In: Diskussion Deutsch, H. 74 (1983), S. 629-640.
- Eisenberg, Peter: Das Wort. Grundriß der deutschen Grammatik. Stuttgart 2004.
- Fay, Johanna: Die Entwicklung der Rechtschreibkompetenz beim Textschreiben. Eine empirische Untersuchung in Klasse 1 bis 4. Frankfurt a. M. 2010.
- Frith, Uta: Beneath the Surface of Developmental Dyslexia. In: Patterson, K. u. a. (Hrsg.): Surface Dyslexia. Neuropsychological and Cognitive Studies of Phonological Reading. London 1985, S. 301-330.

- Groeben, Norbert: Zur konzeptionellen Struktur des Konstrukts „Lesekompetenz“. In: Ders./ Hurrelmann, B. (Hrsg.): Lesekompetenz. Bedingungen, Dimensionen, Funktionen. Weinheim 2006, S. 11-21.
- Günther, Klaus-B: Ein Stufenmodell der Entwicklung kindlicher Lese- und Schreibstrategien. In: Brügelmann, H. (Hrsg.): ABC und Schriftsprache. Rätsel für Kinder, Lehrer und Forscher. Lengwil 1986, S. 32-54.
- Herné, Karl-Ludwig: Rechtschreibtests. In: Bredel, U. u. a. (Hrsg.): Didaktik der deutschen Sprache. Paderborn 2003, S. 883-898.
- Herné, Karl-Ludwig/ Naumann, Carl Ludwig: Aachener Förderdiagnostische Rechtschreibfehler-Analyse. Systematische Einführung in die Praxis der Fehleranalyse mit Auswertungshilfen zu insgesamt 31 standardisierten Testverfahren als Kopiervorlagen. Aachen 2005.
- Hinney, Gabriele: Neubestimmung von Lerninhalten für den Rechtschreibunterricht. Ein fachdidaktischer Beitrag zur Schriftaneignung als Problemlöseprozess. Frankfurt a. M. 1997.
- Klieme, Eckhard u. a: Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Bonn 2007.
- Labov, William: Sociolinguistic Patterns. Pennsylvania 1972.
- Leßmann, Beate: Klassen 1 und 2. Individuelle Lernwege im Schreiben und Rechtschreiben. Teil 1. Heinsberg 2007.
- Löffler, Ilona/ Meyer-Schepers, Ursula: Dortmunder Rechtschreibfehler-Analyse zur Ermittlung des Schriftsprachstatus rechtschreibschwacher Schüler (DoRA). Dortmund 1992.
- Ludwig, Otto: Integriertes und nicht-integriertes Schreiben. Zu einer Theorie des Schreibens. Eine Skizze. In: Baurmann, J./ Weingarten, R. (Hrsg.): Schreiben. Prozesse, Prozeduren und Produkte. Opladen 1995, S. 273-288.
- Maas, Utz: Grundzüge der deutschen Orthographie. Tübingen 1992.
- May, Peter: Strategiebezogene Rechtschreibdiagnose – mit und ohne Test. Analyse von freien Schreibungen mit Hilfe der HSP-Kategorien. In: Balhorn, H./ Bartnitzky, H.; Büchner, I./ Speck-Hamdan, A. (Hrsg.): Schatzkiste Sprache 1. Von Wegen der Kinder in die Schrift. Hamburg 1998, S. 279-293.
- May, Peter: Vom Umgang mit Komplexität beim Schreiben. Herausbildung orthographischer Kompetenz als erweiterte Rekonstruktion sprachlicher Strukturen. In: Balhorn, H./ Brügelmann, H. (Hrsg.): Bedeutung erfinden – im Kopf, mit Schrift und miteinander. Lengwil 1993, S. 277-289.
- May, Peter/ Vieluf, Ulrich/ Malitzky, Volkmar: Diagnose orthographischer Kompetenz. Manual. Hamburger Schreib-Probe 1-9. Hamburg 2007.
- Menzel, Wolfgang: Rechtschreibfehler – Rechtschreibübungen. In: Praxis Deutsch H. 69 (1985), S. 9-11.
- Müller, Rudolf: Diagnostischer Rechtschreibtest für 3. Klassen (DRT 3). Manual. Göttingen 2004.

- Naumann, Carl Ludwig/ Herné, Karl-Ludwig: Warum ist die Tomate leichter als das Fahrradschloss. Überlegungen zur Itemkomplexität und Kompetenzmodellierung in der Rechtschreibung. In: Didaktik Deutsch, H. 28 (2010), S. 5-21.
- Ossner, Jakob: Kompetenzen und Kompetenzmodelle im Deutschunterricht. In: Didaktik Deutsch, H. 21 (2006), S. 5-19.
- Piirainen, Ilpo Tapani: Handbuch der deutschen Rechtschreibung. Grundlagen der Rechtschreibung und Methoden des Rechtschreibunterrichts. Bochum 1981.
- Risel, Heinz: Arbeitsbuch Rechtschreibdidaktik. Baltmannsweiler 2008.
- Röber, Christa: Zur Arbeit mit Häuschen. Freiburg o. J.
- Röber-Siekmeier, Christa: Schriffterwerb. In: Knapp, K. (Hrsg.): Angewandte Linguistik. Ein Lehrbuch. Tübingen 2004, S. 5-25.
- Röckener, Andreas/ Wiencirz, Gerlinde: Eine Schlacht? Die könnt ihr haben! München 1996.
- Roos, Jeanette; Schöler, Hermann (Hrsg.): Entwicklung des Schriftspracherwerbs in der Grundschule: Längsschnittanalyse zweier Kohorten über die Grundschulzeit. Wiesbaden 2009.
- Scheerer-Neumann, Gerheid: Stufenmodelle des Schriftspracherwerbs – Wo stehen wir heute. In: Balhorn, H. u. a. (Hrsg.): Schatzkiste Sprache 1. Von Wegen der Kinder in die Schrift. Hamburg 1998, S. 54-62.
- Schönweiss, Friedrich: Münsteraner Rechtschreibanalyse (MRA). Münster 2004.
- Stock, Claudia; Schneider, Wolfgang: Deutscher Rechtschreibtest für das erste und zweite [dritte und vierte] Schuljahr. Göttingen 2008.
- Thelen, Tobias: Praktische Möglichkeiten computergestützter Rechtschreibanalyse. In: Weinhold, S. (Hrsg.): Schriftspracherwerb empirisch. Baltmannsweiler 2006, S. 178-199.
- Thomé, Günther/ Thomé, Dorothea: OLFA. Oldenburger Fehleranalyse. Instrument und Handbuch zur Ermittlung der orthographischen Kompetenz ab Klasse 3. Oldenburg 2004.
- Topinke, Doris: Rechtschreiben. In: Lange, G./ Weinhold, S. (Hrsg.): Grundlagen der Deutschdidaktik. Sprachdidaktik, Mediendidaktik, Literaturdidaktik. Baltmannsweiler 2005, S. 101-127.
- Valtin, Renate: Vom Kritzelbrief zur verschrifteten Mitteilung. In: Dies./ Naegele, I. (Hrsg.): ‚Schreiben ist wichtig!‘. Frankfurt a.M. 1986, S. 54-66.
- Valtin, Renate/ Badel, Isolde/ Löffler, Ilona/ Meyer-Schepers, Ursula/ Voss, Andreas: Orthographische Kompetenzen von Schülerinnen und Schülern der vierten Klasse. In: Bos, W./ Lankes, E. (Hrsg.): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster 2003, S. 227-265.
- Weigl, Egon: Zur Schriftsprache und ihrem Erwerb – neuropsychologische und psycholinguistische Betrachtung. In: Eichler, W./ Hofer, A. (Hrsg.): Spracherwerb und linguistische Theorie. Texte zur Sprache des Kindes. München 1974, S. 94-173.

- Weinhold, Swantje: Entwicklungsverläufe im Lesen- und Schreibenlernen in Abhängigkeit verschiedener didaktischer Konzepte. Eine Longitudinalstudie in Klasse 1-4. In: Dies. (Hrsg.): Schriftspracherwerb empirisch. Konzepte, Diagnostik, Entwicklung. Baltmannsweiler 2006, S. 120-152.
- Weinhold, Swantje: Effekte fachdidaktischer Ansätze auf den Schriftspracherwerb in der Grundschule. In: Didaktik Deutsch, H. 27 (2009), S. 52-76.
- Winkler, Karin: Die Systematik einer silbenanalytischen Darstellung der Schrift im Anfangsunterricht. Ein Praxisbericht. In: Bredel, U. u. a. (Hrsg.): Schriftspracherwerb und Orthographie. Baltmannsweiler 2004, S. 22-30.
- [www.foerdern.cornelsen.de](http://www.foerdern.cornelsen.de) [Abrufdatum 05.03.2010]
- [www.gutschrift-institut.de](http://www.gutschrift-institut.de) [Abrufdatum 05.03.2010]
- [www.lerndesign.com](http://www.lerndesign.com) [Abrufdatum 05.03.2010]
- [www.lernserver.de](http://www.lernserver.de) [Abrufdatum 05.03.2010]

Anschrift der Verfasserin:

*Juniorprof. Dr. Johanna Fay, Pädagogische Hochschule Karlsruhe, Institut für deutsche Sprache und Literatur und ihre Didaktik, Bismarckstraße 10, 76133 Karlsruhe.*

*fay@ph-karlsruhe.de*