

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

*Carl Ludwig Naumann/Karl-Ludwig
Herné*

**WARUM IST DIE TOMATE
LEICHTER ALS DAS
FAHRRADSCHLOSS?
Überlegungen zu Interkomplexität
und Kompetenzmodellierung in
der Rechtschreibung**

In: Didaktik Deutsch. Jg. 15. H. 28. S. 5-21.

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.

Vorbemerkung

Vor kurzem sind Ergebnisse von Testungen erschienen, bei denen es um die Normierung der Bildungsstandards im Fach Deutsch für die Grundschule ging. Carl Ludwig Naumann und Karl-Ludwig Herné setzen sich mit zentralen Aspekten der Testung im Bereich Rechtschreibung auseinander. Katrin Böhme und Albert Bremerich-Vos replizieren und Naumann und Herné wiederum kommentieren diese Antwort in knapper Form. Deshalb wurde die Replik von Katrin Böhme und Albert Bremerich-Vos um römische Ziffern ergänzt. Auf die damit bezeichneten Stellen bezieht sich Naumanns und Hernés Kurzkommentar.

Carl Ludwig Naumann/Karl-Ludwig Herné

WARUM IST DIE TOMATE LEICHTER ALS DAS FAHRRADSCHLOSS?

Überlegungen zu Itemkomplexität und Kompetenzmodellierung in der Rechtschreibung

1 Ausgangslage und Zielsetzung

„Lässt man die einzelnen Bereiche des Deutschunterrichts Revue passieren, dann zeigt sich, dass die Arbeit an Kompetenzmodellen unterschiedlich weit gediehen ist. [...] Die Domäne »Rechtschreiben« dürfte am intensivsten analysiert sein.“ (Bremerich-Vos 2009, 31) Böhme & Bremerich-Vos (2009) legen jetzt eine umfangreiche Untersuchung vor, die u.W. als erste offenlegt, jedenfalls partiell, welche diagnostischen Kategorien mit Hilfe welcher Items operationalisiert sind. Diese Offenlegung ist eigentlich für die wissenschaftliche Diskussion eine Selbstverständlichkeit – weil man sonst Ergebnisse von Untersuchungen nur glauben, aber nicht prüfen kann. Sie wird jedoch in ähnlichen Untersuchungen oft der Absicht geopfert, die entwickelten Verfahren einsetzen zu können, ohne dass potenzielle Probanden sie unterlaufen, indem sie die Items trainieren.¹ (Im Jargon würden die Items damit „verbrannt“.) Böhme & Bremerich-Vos suchen mit der partiellen Veröffentlichung (20 von 80 Testwörtern eines Lückentests im Rahmen der Evaluation der Bildungsstandards Deutsch, der mit 3480 Schülerinnen und Schülern durchgeführt wurde) einen Kompromiss zwischen praktischen und wissenschaftlichen Erfordernissen. (Vgl. auch Böhme & Bremerich-Vos, Fußnote 10.) Als Gesamtergebnis halten sie fest:

¹ Offiziell nur der Fachöffentlichkeit zugänglich waren und sind die Daten zu einem Teil der Rechtschreibtests seit den 1960er Jahren, die als Vorarbeiten zu den derzeitigen Kompetenzmessungen betrachtet werden können.

Die Befundlage lässt erkennen, dass es gelungen ist, Testinstrumente zu konstruieren, die unter den Bedingungen des Large-scale Assessments eine reliable Messung der globalen Rechtschreibfähigkeit von Schülerinnen und Schülern der dritten und vierten Jahrgangsstufe gestatten.“ (350)

Böhme & Bremerich-Vos haben außerdem untersucht, „inwieweit sich für ein orthographietheoretisch fundiertes neundimensionales Modell der Rechtschreibkompetenz empirische Belege finden lassen.“ (Böhme & Bremerich-Vos, 351) Das neundimensionale Modell enthält folgende Kategorien, in Anlehnung an Herné/Naumann 2002:

SG = spezielle Grapheme und Graphemverbindungen, wie <x, qu, st>.

VL+ = Vokallänge in der Mehrheit der Fälle: Einfacher Vokalbuchstabe bzw. <ie>...

VL- = Vokallänge in der Minderheit der Fälle: ... sonst das stumme <h>.

VK = Vokalkürze: Folgt im Stamm auf einen kurzen betonten Vokal ein einzelner Konsonant, wird der Buchstabe für diesen Konsonanten (fast immer) doppelt hingeschrieben.

HM = häufige Morpheme.

MG = Morphemgrenze: „Treffen gleiche oder hinreichend ähnliche Laute an einer Morphemgrenze zusammen, verschmelzen sie im Mündlichen zu einem Laut. Im Geschriebenen werden solche Reduktionen meistens vermieden. (Bei Fahrrad z. B. hört man nur ein /r/, man schreibt aber – nach Zerlegung in zwei Morpheme – zwei <r>.) Zur Kategorie MG werden hier auch Fälle gezählt, bei denen Fugenelemente wie <s> zu schreiben sind.“

VA = vokalische Ableitung: *Bäcker* z. B. wegen *backen*.

KA = konsonantische Ableitung: „Am Ende eines Morphems oder in bestimmten Konsonantenverbindungen kann die Schreibung einiger Konsonanten nicht immer aus der Lautung erschlossen werden. In sehr vielen Fällen hilft hier das Verlängern (um eine Silbe). Schreibt man *lieblich* mit *ch* oder mit *g*? Die Verlängerung *lieblicher* sorgt für Klarheit. (*-lich* könnte auch als häufiges Morphem – HM – verbucht werden. Führt die Verlängerungsprobe zum Ziel, wird im Folgenden immer „KA“ und nicht „HM“ gewählt.)“

G/K = Groß- und Kleinschreibung: „Hier geht es vor allem um „Konkreta“; Abstrakta und Nominalisierungen sind selten.“...

AF = anderer Fehler: Restklasse.

(Böhme & Bremerich-Vos 2009, 338f.; vollständig zitiert sind hier nur die Erläuterungen zu MG und KA, weil auf sie im Einzelnen zurückzukommen ist.)

Hinsichtlich dieser Kategorien zeigte sich jedoch:

Eine Prüfung der Mehrdimensionalität auf Fehlerebene gemäß den neun orthografie-theoretisch definierten Fehlerkategorien mittels [mehrerer statistischer Verfahren] so-

wie mithilfe kognitiver Diagnosemodelle ergibt keine Evidenz für neun Dimensionen. (330)

Allerdings merken Böhme & Bremerich-Vos am Schluss an:

Der Befund hoher positiver Zusammenhänge zwischen den neun postulierten Dimensionen impliziert somit, dass nicht von einem differenziellen Erwerb bestimmter Aspekte ausgegangen werden kann, sondern dass der Erwerb vielmehr kumulativ oder parallel breit über alle Teilbereiche hinweg erfolgt. Daraus zu schließen, dass eine thematische Differenzierung im Rechtschreibunterricht überflüssig ist, wäre allerdings abwegig.

Es sollte auch bedacht werden, dass alle hier vorgestellten Ergebnisse vor dem Hintergrund des interindividuellen Paradigmas interpretiert werden müssen und keine Aussagen über intraindividuelle Leistungsentwicklungen gestatten. (353)

Für die Lernenden und die Lehrkräfte muss es aber gerade um „intraindividuelle Leistungsentwicklungen“ gehen. Daher wäre es misslich, wenn Lernstandsbestimmung und Lernwege nicht aufeinander bezogen werden könnten, wie es das Zitat nahelegt.² Es erscheint uns nötig, die von Bremerich-Vos und Böhme 2009 angesetzten Kategorien zu erweitern, sodass eine Eigenschaft der Testwörter angemessener berücksichtigt wird: die Itemkomplexität, vorläufig beschrieben mit der Frage: Wie viel Nachdenken erfordert ein Wort?

Denn aufgrund unserer diagnostischen und förderpraktischen Erfahrung halten wir die zitierte Modellierung insoweit für unzureichend, als gerade dem Aspekt der Itemkomplexität nicht angemessen Rechnung getragen wird. Wir vermuten, dass die in den empirischen Daten vorgefundenen hohen Zusammenhänge zwischen den einzelnen rechtschreiblichen Phänomenbereichen z. T. auf einer anderen Operationalisierung der Fehler-Kategorien beruhen, z. T. auch den Besonderheiten des gewählten Wortmaterials geschuldet sind. Wir wollen zeigen, dass komplexe Testwörter, offenbar aus Ökonomie- und/oder Normierungsgründen, gewissermaßen selbstverständlich geworden sind, was sich hier als begrenzt sinnvoll zeigt. – Im 2-Wege-Modell, Abb. 1, ist also der Bezug zwischen den beiden Gedächtnis-Aspekten „Wort- bzw. morphem-spezifisches Wissen“ und „Orthographische Regeln und Strukturen“ näher zu bestimmen.

2 Wir möchten uns nicht damit zufrieden geben, dass hier eine Verschiedenheit der Einsatzbereiche von Überprüfungen sichtbar würde, etwa des Systemmonitorings oder der Schulevaluation, die anderes verlange als die „Individualdiagnostik und Förderung einzelner Schülerinnen und Schüler: Mit dem Einsatz von Tests verbindet sich im pädagogischen Alltag die Erwartung, Aussagen über spezifische Stärken und Schwächen und damit den Förderbedarf einzelner Schülerinnen und Schüler zu machen.“ (Vgl. Klieme 2003, 82f.)

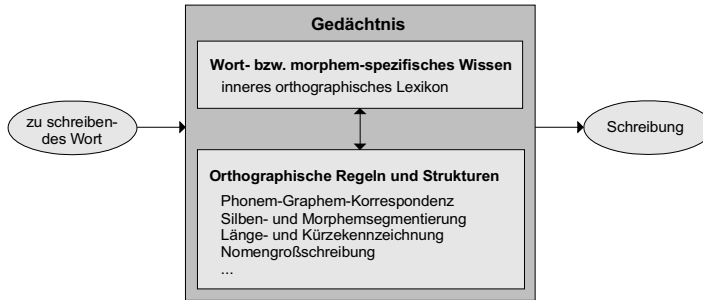


Abbildung 1: Zwei Gedächtnis-Aspekte im 2-Wege-Modell der Rechtschreibung (in Anlehnung an Scheerer-Neumann 2004)

2 Itemkomplexität

Bei der Erfassung der Rechtschreibkompetenz ist es zunächst unerlässlich, nicht nur auf ganze Wörter zu schauen, sondern auch auf die einzelnen Fehler, von denen es in einem Wort mehrere geben kann. Das spiegelt die vorherrschende und von Böhme & Bremerich-Vos geteilte Ansicht, dass nicht allein Wörter als ganze gelernt werden, sondern auch Regeln. (Vgl. auch Abb.1) Die gefundenen Fehler werden ins Verhältnis gesetzt zu den Regeln, gegen die sie verstoßen: Schreibt ein Schüler z. B. <Briftreger>, so hat er erstens nicht beachtet, dass /i:/ in der Regel als <ie> zu schreiben ist und zweitens, dass die morphologische Verwandtschaft mit <tragen> die Schreibung <-träger> erklärt.

Man bestimmt also „Lupenstellen“, d. h. Stellen im Wortmaterial, an denen Fehler einer Kategorie vorkommen könnten, weil hier eine bestimmte Regel anzuwenden wäre. Itemkomplexität ist demnach in einer ersten Näherung zu bestimmen als die Anzahl der Lupenstellen.

Zweitens geht es um die Bestimmung der Regeln, gegen die an jeder Lupenstelle verstoßen wird. Auf die graphematisch-theoretischen Erörterungen dazu gehen wir hier nicht ein, sondern betonen den prinzipiellen Konsens mit Böhme & Bremerich-Vos.³

Drittens kommen – wenn auch nicht häufig – alternative Deutungen vor:

Schreibt ein Schüler zum Beispiel *<sule> statt <Schule>, wird die fehlerhafte Verschriftung des Wortanfangs in dreifacher Hinsicht [...] signiert:

³ Zum Streit über das Gewicht von silbenstrukturellen Aspekten bzw. Regelformulierungen vgl. Hinney u. a. 2008, 118f., bes. Tab 1, und jetzt Weinhold 2009, bes. Tab. 5; Naumann 2008. Die verschiedenen Auffassungen über die Fundierung der satzinternen Großschreibung sind hier nicht relevant.

- GA: Die Kennzeichnung als GA-Fehler trägt dem Umstand Rechnung, dass ein Graphem ausgewählt wurde, das in dieser Umgebung keine lauttreue Verschriftung des Anfangsphonems darstellt.
- SG+: Da <sch> als häufiges spezielles Graphem behandelt wird, ist der registrierte Fehler auch bei SG+ zu kennzeichnen.
- GK–: Die fehlerhafte Kleinschreibung des Wortes ist außerdem bei GK– zu signieren. (Herné & Naumann 2002, 19)

Der Fehler <s> statt <sch> kann z. B. darauf beruhen, dass der Schüler /s/ und /ʃ/ nicht unterschieden hat, was in AFRA mit GA signiert wird. Wobei offen bleibt, ob z. B. Aussprache- oder Hör(analyse)probleme zugrunde liegen; oder darauf, dass er das spezielle Graphem nicht (sicher genug) zu verwenden weiß, was mit SG+ signiert wird, wobei wieder die Erklärung offen bleiben kann, ob z. B. ein anderes Graphemsystem, etwa das türkische, oder eine Übergeneralisierung von <st> oder <sp> her anzunehmen ist. Außerdem ist der erste Buchstabe fälschlich klein geschrieben.⁴ – Böhme & Bremerich-Vos verfahren hier ‚dezisionistisch‘; vgl. oben ihre Erläuterung zu KA. Für die Förderung muss allerdings an jeder einzelnen Stelle offen bleiben, welche der Zuordnungen plausibler ist. Entscheidungen und Ursachenannahmen ergeben sich u.U. bei einer Häufung gleich kategorisierter Fehler.

Viertens sind Interaktionen zwischen Lupenstellen zu beachten. Schreibt ein Schüler z. B. <Wantafel>, so ist anzunehmen, dass eine morphologische Segmentierung in [vant] und [ta:fl] bzw. [ta:fəl] ihn leichter zur Ableitung des <d> im ersten Teil des Wortes geführt hätte. Ähnlich <Manschaft>: Die morphologische Segmentierung hätte leichter zur Geminatbildung <nn> geführt. Wir sprechen von Interaktion der Lupenstellen. Besonders häufig treten, wie in unseren Beispielen, Interaktionen zwischen der Morphemgrenze und dem vorausgehenden Stammmorphem auf, also der Konsonantengeminatbildung und der ‚Auslautverhärtung‘. (Einzuschließen sind hier <ck> und <tz> sowie die g-Spirantisierung.) In der Förderung muss ggf. zunächst z. B. die Konsonantengeminatbildung in morphologisch unkritischen Wörtern gesichert werden, bevor zu komplexeren übergegangen wird (z. B. zuerst <fallen>, <Locke> usw., später <Falltür, fällst>, <Lockvogel> usw.). Bei dieser Interaktion der Lupenstellen liegt unser hauptsächlicher Dissens mit Böhme & Bremerich-Vos; vgl. unten Abschnitt 5.

Wir nehmen an, dass die Wahrscheinlichkeit der Richtigschreibung eines Items (d. h. die Itemschwierigkeit) maßgeblich beeinflusst wird von der Anzahl der kritischen Lupenstellen, die das betreffende Item enthält. Diese betrachten wir als Operationalisierung der Itemkomplexität.

4 Fehlerkategorisierungen wie die AFRA trennen also zwischen der Beschreibung, d. h. einer Interpretation des Fehlers als Regelverstoß, und der Erklärung des Verstoßes, wie sie hier angedeutet wurden.

3 Zur Forschungslage

3.1 Haase (1979) stellt fest:

Fehleranalysen haben ergeben, dass ein Teil der Fehler daraus resultiert, dass der Schüler nicht genügend in die Wortstruktur eindringt, das heißt die Morpheme in Konstruktionen nicht erkennt. (44)

Sie führt dazu eine Untersuchung an einer 5. und einer 6. Leipziger Klasse an, in der morphologisches Wissen und orthographische Entscheidungen verlangt wurden. Besonders schlecht wurden Stämme in Zusammensetzungen gefunden, nämlich mit 60 % Fehlern, gefolgt von Suffixen. Am besten wurden Stämme in Ableitungen gefunden (33 % Fehler) und Präfixe (39 %).

Bei Ableitungen wurden vor allem die Konstruktionen, die sowohl Präfixe als auch Suffixe enthalten, weniger durchschaut (z. B. Entscheidung, gelehrig), besonders dann, wenn der Stamm als Morphemvariante auftritt (z. B. bei Errungenschaft durch die Ableitung vom Partizip II). (45)

Am Beispiel <verräterisch> demonstriert Haase, dass die einzelnen Probleme für den Schüler nicht neu sind, sondern dass „ein Eindringen in die Struktur“ stattfinden muss, eine Vernetzung von Überlegungen.

3.2 May (1993): Aus Voruntersuchungen zur HSP führt May folgende Ergebnisse von Hamburger Viertklässlern an, und zwar unter der Überschrift „III. Integration komplexer Teilprozesse“.

Wort(teil)schreibungen in unterschiedlich komplexen Aufgabenstellungen

	Wort(teil) richtig (ohne Großschreibung)
„das <u>Verkehr</u> sschild“	33,3 %
„Polizisten regeln den <u>Verkehr</u> “	48,1 %
„Das Polizist knackt das <u>Fahrrad</u> -schloss“	52,6 %
„Wimpel flattern am <u>Fahrrad</u> .“	70,1 %

(Auszug aus Tab. 3, 282)

[...] Eingebunden in die Komposita, die einen wesentlich höheren Analyseaufwand erfordern und einen Teil der kognitiven Aktivität vom betreffenden Teilwort abziehen, werden die Teilwörter »Fahrrad« und »Verkehr« wesentlich weniger häufig richtig geschrieben als bei der Einzelwortschreibung,⁵ wo die Kinder ihre Aufmerksamkeit auf eine geringere Zahl von schwierigen Stellen konzentrieren können.

⁵ Linguistisch-terminologisch muss man <Fahrrad> ebenfalls als Kompositum bezeichnen, das beeinträchtigt aber nicht Mays Argument.

[...] Die Geordnetheit der operativen Abläufe wird durch die Häufung der notwendigen Steuer- und Kontrollprozesse beeinträchtigt. (282f.)

[Ein weiteres] Beispiel belegt, daß Schreibern, wenn sie um orthographische Strukturen wissen, nicht isolierte Regeln auf die einzelnen Wortstellen anwenden, sondern komplexere sprachliche Einheiten aufeinander zu beziehen versuchen und dabei die Einzelentscheidungen voneinander abhängig machen. (285)

May resümiert,

dass formal gleich scheinende orthographische Regelungen in verschiedenen Wörtern und Aufgabenstellungen [...] unterschiedlich schwer zu rekonstruieren sind und dass die Komplexität der gesamten Aufgabe den kognitiven Aufwand auch für die Entscheidung bei einer einzelnen Phonembezeichnung bestimmt. (287)

3.3 Scheele (2006) kommt zu ähnlichen Ergebnissen wie May. Sie untersucht die morphologischen und mit ihnen zusammenhängende Fehlerkategorien, die sie in einem Pseudolängsschnitt mit 225 Schülerinnen und Schüler in 3. bis 6. Klassen eines Schulzentrums gewonnen hat. Auch sie kann nachweisen, dass das Erfordernis von Morphemzerlegung einen statistisch signifikanten Einfluss auf die Fehlerquote hat. „Lupenstellen, die vor der Ableitung die Zerlegung eines Items erfordern, werden zu durchschnittlich 15 % falsch geschrieben; Lupenstellen, die keine Zerlegung erfordern, zu durchschnittlich 5,3 %.“ (177) Bezogen auf die untersuchten Jahrgänge nehmen die Mittelwertsdifferenzen mit steigender Klassenstufe ab. Ähnlich stellen sich die verschiedenen Leistungsgruppen dar: „In der stärksten Leistungsgruppe sind die Mittelwertsdifferenzen [...] so gering, dass kein Einfluss der Notwendigkeit von Morphemsegmentierung nachgewiesen werden kann.“ (177f.) – Zu den Beispielen *Fernsehprogramm*, *sieht*, *seht* führt Scheele aus:

Auch im Vergleich dieser drei Items zeigt sich der Einfluss der Komplexität der Wortform auf die betroffene Lupenstelle. Mit 16,9 % Fehlschreibungen bei $n = 178$ ist *Fernsehprogramm* am fehlerträchtigsten. 10,7 % von $n = 225$ schreiben die Lupenstelle in *sieht* falsch, 3,6 % in *seht*. Einzige Fehlschreibung ist jeweils fehlendes <h>. (Scheele 2006, 187)

Und zur Differenzierung nach Leistungsgruppen zeigt sie: Während die rechsreibstärksten 25 % der Schüler ab Klasse 4 <Fahrradschloß> korrekt zerlegen, zeigen 42,9 % der schwächsten Schwierigkeiten mit der korrekten Zerlegung. – Im Fazit zur Frage nach einer Erwerbsfolge der Kategorien sagt sie:

Die Einzelwortanalysen zu allen Fehlerkategorien haben aber ergeben, dass innerhalb jeder Kategorie weiter nach der Häufigkeit der Items bzw. Morpheme, den in einem Item zusätzlich vorhandenen Lupenstellen, den an den Lupenstellen betroffenen Graphemen, den Umgebungsbedingungen der Lupenstellen usw. unterschieden werden muss. (Scheele 2006, 375)

3.4 Fay (2009) misst die Rechtschreibkompetenz in Aufsätzen und Tests (HSP) von 60 Zweit- und Viertklässlern mit dem Ziel, die ökologische Validität von Tests zu überprüfen. Sie berichtet beim Vergleich der Schreibungen in Aufsatz und Test u. a. über die <ie>-Schreibungen eines Probanden: Alle neun Auftretensfälle (5 types) im betrachteten Aufsatz sind insoweit normgemäß (Originalschreibungen): <fliegende,

Korier, Korier, wieder, rief, rief, rief, vierten, rief>. Demgegenüber werden 2 der 3 Items aus der HSP 4/5 nicht mit <ie> geschrieben: <Briefträger, Giskanne, Schizrichter>.⁶

4 Test-Praxis in der HSP

In der aktuellen Fassung des Manuals zur HSP erwähnt May zwar seine oben angeführten Ergebnisse von 1993 und zitiert u. a. seine oben wiedergegebene Tabelle (vgl. May 2002, 39–41). Jedoch gibt es in der bis Mitte Klasse 5 anzusetzenden HSP 4/5 nicht einmal halb so viele einfache wie komplexe Formen. Im Vergleich mit HSP 3 :

	morphologisch einfache	morphologisch komplexe Formen ⁷
HSP 3	19	19
HSP 4/5	11	29

May begründet den Verzicht auf die Abfrage einfacher neben komplexen Formen so: „Die Wortauswahl in der HSP erfolgte nach Schwierigkeit und Trennschärfe. Unter den ‚normalen‘ deutschstämmigen Wörtern gab es in unserem Erprobungspool (fast) nur Komposita, die in der Sekundarstufe noch hinreichend schwierig waren.“⁸ Die zur Kompetenzmodellierung und Förderdiagnostik interessante Unterscheidung von einfachen vs. komplexen Items fällt also offenbar der Normorientierung und der Einfachheit der Auswertung zum Opfer. (Zur Lupenstellenauswahl vgl. ferner May 2002, 42f; zur Kritik vgl. Herné 2003, 891.)

6 Vgl. Kapitel 4.4: Einzelfallanalyse. Eine informelle Sichtung weiterer Beispiele, die Frau Fay freundlicherweise vorgenommen hat, führt zu einem ähnlichen Beispiel. Ein anderer Viertklässler schrieb im Aufsatz <Befiel> statt <befiehlt>; aber <verschieben> (2-mal); in der HSP < Briefträger, Gieskanne>, aber <Schizrichter> und auch <Ferkerschielt> statt <Verkehrsschild>. Sicher kann auch <befiehlt> mit dem seltenen <ieh> als komplexes Item betrachtet werden. – Alle Formwörter, das sind hier <die> und <sie>, wurden in Aufsatz und Test richtig geschrieben und werden nicht angeführt. Ob man angesichts des erst kursorischen Charakters dieser Sichtung in den Beispielen mehr als eine Illustration sehen kann, bleibt offen.

7 Als morphologisch komplexe Formen fassen wir hier zusammen: affigierte Formen (auch mehrfach) wie <knackt, bekommt>; Komposita (auch mit mehr als zwei Bestandteilen) wie <Gießkanne, Fahrradschloss>; Komposita mit Affigierungen wie <Briefträger>. Zur Vereinfachung ist hier das Fugen-s als Affix mitgezählt.

8 Mail von Peter May vom 27. April 2009; Zitiererlaubnis vom 25. Juni 2009.

5 Itemkomplexität in den Daten von Bremerich-Vos und Böhme

In Tabelle 1. haben wir nach den Beschreibungen der Kategorien bei Böhme & Bremerich-Vos (s.o.) die Lupenstellen aufgelistet und ihre Summen pro Item mit der Itemschwierigkeit korreliert. (Siehe dort Tab. 3: Prozentuale Lösungshäufigkeiten [...], 344.) Tab. 1, die u.E. die Itemkomplexität nicht ausreichend berücksichtigt, zeigt keine Auswirkung auf die Itemschwierigkeit. Es ergeben sich Korrelationen von -0,06 für die dritten und -0,05 für die vierten Klassen, die als nicht signifikant zu betrachten sind.

Kategorie	SG	VL +/-	VK	MG	HM	VA	KA	G/K	Anz. krit. Lupenstel- len	Klasse 3	Klasse 4
gesperrt	sp		err						2	10,5	35,5
empfehlen		eh							1	20,7	43,3
Schlittschuh- läufer		uh	itt		er	äu		S	5	22,2	40,4
verabschiedete		ie	ab		ver				3	23,7	50,8
unübertrefflich		üb	un, eff				ch		4	28,4	48,9
Verkehr	V	eh						V	3	32,9	64
verstaubten	st				ver		b		3	34,6	51,7
Kartoffelsuppe			off, upp					K	3	44,5	69,4
vorspielen	sp	ie			vor				3	55,2	73,7
Geburtstag		ag		t's			g	G	4	55,6	69,4
Schlüsselloch			üss	l'l				S	3	57,6	70,4
versteckt	st		eck		ver				3	60,8	79,7
eigentlich							ch		1	63,3	74,7
Staubsauger	St				er		b	S	4	63,8	80,5
Wohnzimmer		oh	imm					W	3	63,8	86,4
Lehrerin		eh			in			L	3	67	85,5
Vogelfutter	V	og	utt					V	4	69,2	86,7
Geschenke	nk							G	2	71,9	83,2

Korrelation	-0,06	-0,05
Signifikanz	0,59	0,57

Tabelle 1: Itemkomplexität nach Böhme & Bremerich-Vos

Anders Tab. 2. Für Tab. 2. haben wir uns in den meisten Kategorien nur z. T. an die zur Verfügung gestellte Codieranleitung gehalten, sondern entsprechend Herné & Naumann 2002 andere – überwiegend zusätzliche – Zuordnungen vorgenommen, im Interesse einer angemessenen Berücksichtigung der Itemkomplexität. Dabei wurden Mehrfachklassifikationen vermieden, um nicht eine ungewollte implizite Gewichtung der Lupenstellen vorzunehmen. Zwar sind 18 Items eine eher kleine Basis, aber mit einer Korrelation von -0,44 für die dritten Klassen bzw. -0,52 für die vierten, signifikant auf 5 %-Niveau, kann man die so bestimmte Itemkomplexität als recht gute Vorhersage für die Itemschwierigkeit ansehen.⁹

AFRA-Kategorie	SG+/-, SV+/-	LV-	KV _D +, KV-	MS	UM	VA	KA+ /-	GK -	Anz. krit. LupSt.	Klasse 3	Klasse 4
gesperrt	sp, r		err	e'sp, r't	ge	e	t		8	10,5	35,5
empfehlen	f	eh			emp		p		4	20,7	43,3
Schlittschuhläufer	sch, sch, äu, f	uh	itt	t's, h'l		äu	t	S	11	22,2	40,4
verabschiedete	v, sch	ie	ab				b		5	23,7	50,8
unübertrefflich	r, f		un, eff	f'l	lich	e			7	28,4	48,9
Verkehr	V, k, r	eh						V	5	32,9	64
verstaubten	v, st, au			r'st, b't			b		6	34,6	51,7
Kartoffelsuppe	r, s		off, upp					K	5	44,5	69,4
vorspielen	v, r, sp	ie		r'sp					5	55,2	73,7
Geburtstag	r			t's	ge, s		t, g	G	7	55,6	69,4
Schlüsselloch	sch, ch		üss	l'l				S	5	57,6	70,4
versteckt	v, st		eck	r'st, ck't		e	k, t		8	60,8	79,7
eigentlich	ei				lich		t		3	63,3	74,7
Staubsauger	St, au, au, r			b's			b	S	7	63,8	80,5
Wohnzimmer	z	oh	imm					W	4	63,8	86,4
Lehrerin	r	eh			in			L	4	67	85,5
Vogelfutter	V		utt					V	3	69,2	86,7
Geschenke	sch, nk				ge	e		G	5	71,9	83,2

Korrelation	-0.44	-0.52
Signifikanz	0.03	0.01

Tabelle 2: Itemkomplexität nach Herné & Naumann 2002

⁹ Eine Rangkorrelation nach Spearman wurde gewählt, weil nicht davon auszugehen ist, dass unsere Schätzung der Itemkomplexität anhand der Anzahl der kritischen Lupenstellen intervallskaliert ist. Beide Korrelationen sind signifikant auf dem 5 %-Niveau.

Zwei Ausreißer haben wir vorab gestrichen: <flitzten>, das laut Fußnote zu Tab. 3 bei Böhme & Bremerich-Vos teilweise fehlerhaft diktiert worden war, und <kaputt>, dessen nicht nur unter den vorliegenden Testwörtern ungewöhnliche Struktur als besondere Fehlerfälle gelten darf.¹⁰ Um derartige Eigenheiten bei der Zusammenhangsanalyse zu berücksichtigen, müssten die verschiedenartigen Lupenstellen gewichtet werden. An der Gewichtungsfrage zeigt sich ein Unterschied zwischen den Zielsetzungen einer Großuntersuchung und der Förderdiagnostik, vgl. auch Fn. 3 und in Abschnitt 2 die „Alternativen Deutungen“: Was eine sachliche Notwendigkeit für die Diagnostik ist, würde in der Großuntersuchung zu einer Verzerrung führen.

SG+/-&SV+/-:¹¹ U.E. stellen *alle* von uns angeführten Kategorien auch für Viertklässler potenzielle Probleme dar.

VL/LV-: Nur die *graphisch markierte* Länge des Vokals haben wir als Lernproblem aufgeführt.

MG/MS: In Herné und Naumann 2002 heißt es zur Morphemgrenzenbehandlung: „Die Zerlegung zusammengesetzter, flektierter oder abgeleiteter Wortformen in ihre morphologischen Bestandteile kann die Voraussetzung sein für die Anwendung verschiedener rechtsschreiblicher Regeln.“ (13) Bei der Musteranalyse zu ‚Jana‘ sind u. a. deren Schreibungen <grüst, nent, zangt, ferstekt, besprüt ...> aus dem DRT 3 als Beispiele angeführt. (52) Die Bewertung der Items für Tab. 2. folgt unserer weitergehenden Auffassung.. Böhme & Bremerich-Vos 2009 operationalisieren hingegen ihre Kategorie MG eingeschränkter, s.o. ihre Erläuterung. Dies könnte dem Wunsch entsprechen, Lupenstellen nur ‚lokal‘ zu definieren (vgl. Böhme & Bremerich-Vos 362f.), sei es, um die Codierung zu vereinfachen, sei es, um Lupenstellen als unabhängig auffassen zu können. Damit wird aber die Kompetenz insoweit unzureichend erfasst.¹²

HM/UM: Die Codieranleitung führt <ver> in <Verkehr> nicht als HM an, wohl aber in <versteckt, verabschiedete, verstaubten>. Wir haben einheitlich <v> nur in der Kategorie SG erfasst. Entsprechend sind wir mit <-er> und <r> verfahren. Für eine (nicht gewichtende) Zählung der Lupenstellen kann die Frage nach der Angemessenheit der jeweiligen Zuordnung offenbleiben. Das Präfix <ge-> haben wir aufgeführt, jedoch – wie Böhme & Bremerich-Vos – kein <-t>, weil wir beide Vorkommensfälle unter KA anführen.

10 Prototypische indigene Wörter enthalten, falls sie mehr als eine Silbe haben, mindestens ein <e>, meist für [ə] bzw. als Teil der Reduktionssilbe. In der Regel ist von zwei Silben die erste betont. Dass drittens kurze/ungespannte Vokale nur unter Betonung mit Konsonantengemination oder <h> etc. markiert werden können, ist (nicht nur) Schülern weitgehend unbekannt.

11 Soweit erforderlich, sind den Kategorienamen von Böhme & Bremerich-Vos 2002 diejenigen aus Herné und Naumann nachgestellt.

12 Einen psychometrisch gangbaren Weg zur Berücksichtigung der Komplexität sähen wir analog zu May 1993: Testung komplexer und einfacher Items an den gleichen Probanden. Ob das den Rahmenbedingungen eines Tests in einer Großuntersuchung entspricht, ist uns unklar.

VA: Für diese Kategorie haben wir im Gegensatz zu Böhme & Bremerich-Vos nicht allein diejenigen Fälle angesetzt, in denen das <ä> aus einer Form mit <a> abzuleiten ist. Aus der Lernerperspektive gehört hierhin *jedes* /ɛ/ bzw. /oi/, weil die Frage nach einer Ableitbarkeit immer gestellt werden muss, also nicht nur für <Schlittschuhläufer>, sondern auch für <gesperrt, unübertrefflich, versteckt, Geschenke>.¹³

KA: Analog zu VA haben wir auch z. B. /p/ und /t/ in <empfehlen> und <Schlittschuhläufer> als KA kategorisiert, weil der Lerner sich überlegen muss, ob er den gehörten Laut ‚direkt‘ verschriften kann oder ihn durch den Versuch der ‚Verlängerung‘ überprüfen muss - auch wenn eine Verlängerung nicht zu einer Veränderung des kritischen Lautes führt (<gesperrt>, <Geburtstag>, <Schlittschuhläufer>). (Vgl. Herné & Naumann 2002, 15)

6 Zur Itemkomplexität in zwei anderen Großuntersuchungen

Ohne die beiden aktuellen Großuntersuchungen zur Rechtschreibkompetenz von 4.-Klässlern derjenigen von Böhme & Bremerich-Vos 2009 in aller Breite gegenüberzustellen, soll auf ihre z. T. andersartigen Ergebnisse zur Itemkomplexität eingegangen werden. Erwähnt werden muss die missliche Tatsache, dass zu diesen Untersuchungen (IGLU-Auswertungen von Löffler & Meyer-Schepers zusammen mit unterschiedlichen Autoren) und Voss/Blatt/Kowalski (2007) nur wenige Testitems und deren Klassifikation vorgelegt wurden.

6.1 Zu IGLU haben Löffler und Meyer-Schepers die Rechtschreibleistungen von 3391 4.-Klässlern feinkategorial ausgewertet. Wir finden zwar keine Mehrfachklassifikationen von Fehlern.¹⁴ Aber bei der Auswertung ihrer untersten Leistungsgruppe (das sind die schlechtesten 25 %, also 849 Schüler) stellen sie fest: „In der Gruppe des unteren Viertels sind Falschschreibungen verbreitet, die sich im oberen Viertel nicht finden. Es handelt sich um *Kombinationen von Einzelfehlern*.“ [Hervorhebungen im Original].¹⁵ „Eine vergleichbare Mischung von Einzelfehlern findet sich im unteren Viertel mehr oder weniger bei allen DoSE-Wörtern.“ (Löffler & Meyer-Schepers 2005, 98f.)

Sie fügen dem die Interpretation an: Die vielen Schreibvarianten „offenbaren, dass diese Schreiberler [...] diverse orthographische Vorschriften kennen gelernt haben, ohne sie jedoch sinnvoll anwenden zu können. Es sieht so aus, dass [...] die Desinte-

¹³ Nicht berücksichtigt wurde die norddeutsch-umgangssprachliche Nichtunterscheidung zwischen /e:/ und /ɛ:/, sodass <empfehlen> z. B. nicht als VA kodiert ist.

¹⁴ Mangels ausführlicher Beschreibungen in Löffler/Meyer-Schepers 2005 stützen wir uns auch auf Löffler/Meyer-Schepers 1992. Vgl. auch Bremerich-Vos (2004, 98), der z. B. kritisiert, dass das <k> in <sink> neben der elementaren Lautanalyse grammatische Überlegungen verlangt, nämlich eine morphologische Zerlegung.

¹⁵ Beispiel: <verköhlt> wird von 77 % dieser Schüler ohne <h> geschrieben. 40 % fügen aber weitere Einzelfehler hinzu, sodass Formen resultieren wie <verköllt, Ferkölt, ferkölt, vergölt, ferkölt, vergüllt, ferküllt>.

gration der verschiedenen Schreibregeln [...] ein wesentlicher Grund hierfür ist.“ (102)

Die auszugsweise angeführten Testwörter scheinen nicht ganz so komplex zu sein wie die von Böhme & Bremerich-Vos veröffentlichten; eine Berechnung wie in Tab. 1/2 ist aber natürlich nicht möglich.

6.2 Voss/Blatt/Kowalski (2007) haben die eben erwähnten IGLU-Daten sekundär ausgewertet, mit einem anderen Fehleranalyseverfahren, das leider nicht an Beispielen vorgeführt wird. Sie setzen fünf Anforderungsbereiche an, siehe Abb. 2:

AB	1	2	3	4	5
Orientierung an Prinzipien	Phonographisches und silbisches Prinzip im Kernbereich	Morphologisches Prinzip im Kernbereich	Peripheriebereich	Prinzipien der Wortbildung	Wortübergreifendes Prinzip
Teilfähigkeiten	Bezug herstellen zwischen Schrift- und Lautstruktur unter Berücksichtigung der silbenstrukturellen Informationen (Silbenanfangs u. -endrand u. Silbenschnitt)	Vererbte silbenschriftliche Informationen in flektierten + abgeleiteten Formen herleiten; Flexionsmorpheme kennen u. anwenden	Markierungen in offenen Silben setzen und vererbte Schreibweisen herleiten; Transfer bei Sonderfällen und Lernwörtern; Fremdwortschreibung	Wortarten u. Wortbildungsmorpheme kennen und in Ableitungen und Komposita produktiv anwenden	Wortarten u. Syntaxstrukturen kennen u. für Groß-/Klein-, dass-Schreibung u. Kommasetzung anwenden

Abb. 2: Anforderungsbereiche bei Voss/Blatt/Kowalski 2007, 17

Ob bei der Zuordnung von Schreibungen die Mehrfachklassifikation eine Rolle spielt, wird uns nicht erkennbar. Bemerkenswert ist allerdings, dass AB 2 die Fähigkeit enthält, „vererbte silbenschriftliche Informationen in flektierten + abgeleiteten Formen her[z]u[le]iten“, und AB 3 „Markierungen in offenen Silben setzen und vererbte Schreibungen herleiten“. Wegen der sehr lückenhaften Darstellung ist nicht zu sehen, wie bei denkbaren Überschneidungen operationalisiert wurde. Ähnliches ist zu fragen, soweit es um die Überschneidungen mit AB 3 und 4 geht.

Die immerhin in Auszügen mitgeteilten Items (23 von 112 erprobten und 103 verwendeten, vgl. 22 und v.a. Tab. 4, 26) lassen aber erkennen, dass neben komplexen auch einfache Items verwendet wurden und dass dies zur Itemschwierigkeit beiträgt:

Eine vergleichende Analyse ergibt Hinweise auf mögliche Ursachen für die Unterschiede im Grad der Aufgabenschwierigkeit. Zum einen ist die Aufgabenschwierigkeit ein und desselben Phänomens höher, wenn dieses in einer flektierten oder abgeleiteten Wortform vorkommt (kommen – Schnurr-, Reh – fröhlich, süße – schließlich, nach – hochleben). (26)

Ein Fazit der AutorInnen (25) zur Frage der Differenzierung nach Teilkompetenzen lautet:

Aus analytischer Sicht kann das komplexere fünfdimensionale orthographische Kompetenzmodell die in den erfassten Schülerdaten enthaltenen Informationen besser darstellen als ein Generalfaktormodell, in dem nur nach Richtig- und Falschschreibung unterschieden wird. Dieser Unterschied ist signifikant.

Den dargestellten Analysen zufolge ergibt sich also eine empirische Evidenz für das differenzielle Rechtschreibkompetenzmodell, wobei noch nicht alle theoretisch möglichen Modellvarianten berechnet wurden.

Der letzte Satz ist dunkel, denn was ‚Modellvarianten‘ sein könnten, wird nicht gesagt – möglicherweise ein Modell mit nicht gerade 5 Anforderungsbereichen? Deswegen Begründung (16–18) ist eher lapidar. Und die angeblich erhärtete Fünfdimensionalität wird durch folgenden Satz aus dem Fazit wieder fraglich: „Die Wortschreibungen im Kernbereich [oben Spalten 1 und 2] stellen ein und dieselbe Kompetenzleistung dar,“ (29).

6.3. Böhme & Bremerich-Vos versuchen abschließend (im Abschnitt 3.5.9, Diskussion, bes. 352), ihre Ergebnisse zur Eindimensionalität des „orthografischen“ Schreibens in die Diskussion über Rechtschreibkompetenz-Modellierung einzuordnen. Dabei ziehen sie den letzten zitierten Satz aus Voss/Blatt/Kowalski heran, gehen aber nicht auf den Widerspruch zum Zitat aus S. 25 ein.

Schlagend scheint uns auch nicht ihr Bezug auf Klicpera u. a. (2003) und auf Notenboom/Reitsma (2003), weil diese von Modellen ausgehen, die am Englischen gewonnen wurden.¹⁶ – Interessant ist der Bezug auf Scheele, deren Fazit Böhme & Bremerich-Vos ebenfalls als Beleg für Eindimensionalität lesen. Scheele betont allerdings als Ergebnis ihrer Untersuchung, dass es nach dem Übergang aus der lautorientierten (,alphabetischen‘) in die lautergänzende (,orthographische‘) Stufe keine weitere ‚Stufung‘ i.e.S. gebe.¹⁷ Vielmehr „entwickeln sich die Zugriffsweisen im

16 Notenboom & Reitsma 2003 stellen zwar die prinzipielle Frage nach der Übertragbarkeit von Rechtschreibkompetenzmodellen aus dem Englischen auf andere Sprachen. Aber die spezifische Frage, was die systematischen Eigenschaften der deutschen Orthographie sind, stellen sie gerade nicht. – Dass Klicpera et al. (2003) in gleicher Weise den am Englischen entwickelten Vorstellungen verpflichtet sind, ist anzunehmen (obwohl sie die größere Transparenz der deutschen Orthographie erwähnen): Sie fordern als Grundlagen einer Legasthenie-Subtypisierung auch für das Deutsche eine Differenzierung gemäß ‚phonologischer Bewusstheit‘ und ‚orthographischem Wissen‘ (die sie in der deutschen Literatur nicht finden, vgl. jedoch explizit Mann (1998 und öfter), implizit die Auffassungen von LRS als Schrifterwerbsverzögerung, z. B. Scheerer-Neumann 1989; May 2002, 47–54).

17 Mit Ausnahme der Großschreibung (Scheele 2006, 391), die aber Böhme & Bremerich-Vos – für die Grundschule nachvollziehbar – nicht in der gleichen Breite wie Scheele behandeln.

fortgeschrittenen Erwerb parallel zueinander“ (Scheele 2006, 392; mit ausdrücklichem Bezug auf Eichler 1986)

Dass sie mit der zeitlichen Staffelung aber nicht zugleich die Verschiedenheit der ‚parallelen Zugriffsweisen‘ [Plural!] in Frage gestellt sieht, erhellt sich wenig später: „Der Rechtschreiberwerb auf der ‚orthographischen Stufe‘ ist somit kein ungeordnetes Konglomerat von Regelerwerbsprozessen“ (ebd.).

7 Fazit

Frühere Untersuchungen haben gezeigt, dass die Integration von einzelnen Komponenten des Wissens über Rechtschreibung erforderlich ist und ggf. gesondert gelernt werden muss. Diagnostische, unterrichtliche und lerntherapeutische Erfahrung kennen diese Anforderung als typische potenzielle Hürde am Ende der Grundschule; die Beobachtungen, besonders zu leistungsschwächeren Schülern, von Scheele und Löffler & Meyer-Schepers weisen in die gleiche Richtung. Die entsprechende Formulierung als Itemkomplexität in Kompetenzuntersuchungen und die Entwicklung eines Maßes dafür ist erst begonnen.

Literatur

- Bremerich-Vos, Albert (2009) Kap. 3 – Die Bildungsstandards Deutsch. In: Albert Bremerich-Vos/ Dietlinde Granzer/Ulrike Behrens/Olaf Köller (Hg.) Bildungsstandards für die Grundschule: Deutsch konkret. Cornelsen 2009, S. 14–42.
- Böhme, Katrin & Bremerich-Vos, Albert (2009) Diagnostik der Rechtschreibkompetenz in der Grundschule – Konstruktionprüfung mittels Fehler- und Dimensionsanalysen. In: Granzer, Dietlinde; Köller, Olaf; Bremerich-Vos, Albert u. a. (Hg.) Bildungsstandards Deutsch und Mathematik. Weinheim/Basel: Beltz 2009, S. 330–356.
- Bremerich-Vos, Albert (2004) Bildungsstandards, Rechtschreibkompetenz und IGLU – einige Anmerkungen. In: Ders.; Löffler, Cordula; Herné, Karl-Ludwig (Hg.): Neue Beiträge zur Rechtschreibtheorie und -didaktik. Freiburg: Fillibach 2004, 85-104.
- Eichler, Wolfgang (1986): Zu Uta Frith’ Dreiphasenmodell des Lesen (und Schreiben) Lernens. Oder: Lassen sich verschiedene Modelle des Schriftspracherwerbs aufeinander beziehen und weiterentwickeln? In: Augst, Gerhard (Hg.): New Trends in Graphemics and Orthography. Berlin/New York: Walter de Gruyter, 234-247.
- Fay, Johanna (2009) Die Entwicklung der Rechtschreibkompetenz beim Textschreiben. Eine empirische Untersuchung in Klasse 1 bis 4. Diss. Lüneburg.
- Haase, Christel (1979): Zur Bedeutung wesentlicher Tätigkeiten im Bereich der Wort- und Formenbildung für die Herausbildung orthographischen Könnens. In: Deutschunterricht Jg. 32/H. 1, 44-48.
- Herné, K.-L. (2003) Rechtschreibtests. In: Bredel, Ursula/Günther, Hartmut u. a. (Hg.): Didaktik der deutschen Sprache. Ein Handbuch. 2. Teilband. Paderborn: Schöningh. 883–897.

- Herné, K.-L./ Naumann, C.L. (2002) AFRA–Aachener Förderdiagnostische Rechtschreibfehler-Analyse. Systematische Einführung in die Praxis der Fehleranalyse. Handbuch mit Auswertungshilfen zu insgesamt 33 standardisierten Testverfahren als Kopiervorlagen. Aachen: Alfa Zentaurus
- Hinney, G.; Huneke H.-W., Müller, A., Weinhold, S. (2008) Definition und Messung von Rechtschreibkompetenz. In: Didaktik Deutsch Sonderheft. Beiträge zum 16. Symposium Deutschdidaktik. 2008, 107–126.
- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Tenorth, H.-E./Vollmer, H.J. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Bonn: Bundesministerium für Bildung und Forschung
- Klicpera, C. u. a.(2003) Legasthenie. Modelle, Diagnose, Therapie und Förderung. München: Reinhardt.
- Löffler, Ilona & Ursula Meyer-Schepers (1992). DoRA Dortmunder Rechtschreibfehler-Analyse zur Ermittlung des Schriftsprachstatus rechtschreibschwacher Schüler. Ein Arbeitsbuch für die Hand des Lehrers. Dortmund.
- Löffler, Ilona & Ursula Meyer-Schepers (2005) Orthographische Kompetenzen: Ergebnisse qualitativer Fehleranalysen, insbesondere bei schwachen Rechtschreibern. In: Bos, W. u. a. (Hg.): IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien. Münster: Waxmann 2005, 81–108.
- Mann, Christine (1998) Legasthenie. In: Günter Lange, Karl Neumann, Werner Ziesenis (Hg.): Taschenbuch des Deutschunterrichts, Bd. 1. Baltmannsweiler: Schneider Hohengehren, 6., vollst. überarb. Aufl. 1998, 305–318.
- May, Peter (1993) Vom Umgang mit Komplexität beim Schreiben. In: Balhorn, Heiko/Brügelmann, Hans: Bedeutungen erfinden – im Kopf, mit Schrift und miteinander. Konstanz: Faude, 277-289. (= libelle: wissenschaft lesen und schreiben, 5).
- May, Peter (2002): Hamburger Schreib-Probe HSP 1-9. Diagnose orthographischer Kompetenz zur Erfassung der grundlegenden Rechtschreibstrategien. Handbuch. Unter Mitarbeit von U. Vieluf und V. Malitzky. 6., aktualisierte und erweiterte Auflage. Hamburg: Verlag für pädagogische Medien.
- Notenboom, A./Reitsma, P. (2003): Investigating the dimensions of spelling ability. In: Educational and Psychological Measurement 63, S. 1039–1059.
- Scheele, Veronika (2006) Entwicklung fortgeschrittener Rechtschreibfertigkeiten. Ein Beitrag zum Erwerb der ‚orthographischen‘ Strategien. Frankfurt: Lang 2005.
- Scheerer-Neumann, Gerheid (1989/2003): Rechtschreibschwäche im Kontext der Entwicklung. In: Naegele, Ingrid M./Valtin, Renate (Hg.): LRS – Legasthenie in den Klassen 1-10. Handbuch der Lese-Rechtschreibschwierigkeiten. Band 1: Grundlagen und Grundsätze der Lese-Rechtschreib-Förderung. 6. Aufl. Weinheim/Basel: Beltz Verlag, 45–65 (Erste Auflage 1989).
- Scheerer-Neumann, Gerheid (2004): „Ich rede so im Kopf, wie man es schreibt.“ Mitteilungen von Kindern zum wortspezifischen Orthographieerwerb. In: Albert Bremerich-Vos, Cordula Löffler, Karl-Ludwig Herné (Hg.): Neue Beiträge zur Rechtschreibtheorie und -didaktik. Freiburg im Breisgau: Fillibach, 105–123.

- Weinhold, Swantje (2009) Effekte fachdidaktischer Ansätze auf den Schriftspracherwerb in der Grundschule. Ergebnisse zu Lese- und Rechtschreibleistungen in den Jahrgangsstufen 1-4. In: Didaktik Deutsch. Heft 27, 52–75.
- Voss, Andreas/Inge Blatt /Kerstin Kowalski (2007) Zur Erfassung orthographischer Kompetenz in IGLU 2006: Dargestellt an einem sprachsystematischen Test auf Grundlage von Daten aus der IGLU-Voruntersuchung. In: Didaktik Deutsch. Heft 23, 15–33.

Anschriften der Verfasser:

Carl Ludwig Naumann, Leibniz Universität Hannover, Deutsches Seminar, Königsworther Platz 1, 30167 Hannover

Karl-Ludwig Herné, Beratungsstelle für Lese-Rechtschreib-Schwäche e.V., Franzstraße 32, 52064 Aachen