

Hans Brügelmann (2004n)

Die Rechtschreibleistung in Kurztexten von SchülerInnen der achten Klasse in der internationalen IEA-Lesestudie (1991)

Die Bezugsstudie der Sekundärauswertung

Die vergleichende Lesestudie der International Association for the Evaluation of Educational Achievement (IEA) wurde am Ende des Schuljahres 1990/91 mit 9-Jährigen (in Deutschland dritte Klassen) und 14-Jährigen (Deutschland achte Klassen), also jeweils etwas früher als die aktuellen Studien IGLU (4. Klasse) und PISA (15-Jährige) durchgeführt¹. Für Deutschland war neben dem internationalen Vergleich vor allem – kurz nach der Wende 1989 – der Vergleich Ost-West interessant. Die wichtigsten Befunde²:

- Auf beiden Schulstufen lagen die Leseleistungen der deutschen SchülerInnen (insgesamt) im internationalen Mittelfeld³.
- Damit schnitten die DrittklässlerInnen 1991 also etwas schwächer ab als die ViertklässlerInnen 2001⁴, dagegen die 8.-KlässlerInnen deutlich besser als die 15-Jährigen im Jahre 2000⁵.
- Sowohl in den Mittelwerten als auch in den Streuungen unterschieden sich die beiden deutschen Stichproben (Ost vs. West) nicht bedeutsam⁶.

Das Textkorpus

Im Grundsatz waren die Leseaufgaben als Ankreuztest von Auswahlantworten zu Fragen angelegt, die sich auf vorgegebene Texte bezogen. Für die AchtklässlerInnen gab es jedoch zwei Aufgaben (die Texte „Okashi“ und „Sardine“), in denen auch freie Antworten formuliert werden konnten. Diese Möglichkeit hat die Mehrheit der SchülerInnen (mindestens einmal)

¹ Vgl. zum internationalen Vergleich: Elley, W.B. (1992): How in the world do students read? International Association for the Evaluation of Educational Achievement: s'Gravenhage; dt. Zusammenfassung der wichtigsten Befunde in: Balhorn/ Brügelmann (1993, 86-89).

² Lehmann, R.H., u. a. (1995): Leseverständnis und Lesegewohnheiten deutscher Schülerinnen und Schüler. Beltz: Weinheim/ Basel; Lehmann, R. H. (1993): Zum Leseverständnis von SchülerInnen und Schülern im vereinigten Deutschland. In: Balhorn, H./ Brügelmann, H. (Hrsg.) (1993): Bedeutungen erfinden -- im Kopf, mit Schrift und miteinander. Lesen und Schreiben als individuelle und soziale Konstruktion von Wirklichkeiten. DGLS-Jahrbuch Bd. 5. Faude: Konstanz, 160-167.

³ Vgl. Elley (1992, 14, 24, 10, 109).

⁴ Vgl. Bos, W., u. a. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Waxmann: Münster u. a., 122 ff.

⁵ Baumert, J., u. a. (Hrsg.) (2001): PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Leske + Budrich: Opladen.

⁶ Lehmann, R.H., u. a. (1995): Leseverständnis und Lesegewohnheiten deutscher Schülerinnen und Schüler. Beltz: Weinheim/ Basel, 144-148, 215-220. Vgl. zu analogen Befunde für die Rechtschreibung—zumindest in freien Texten – Brügelmann, H., Lange, I., u. a. (1994g): "Schreibvergleich BRDDR" 1990/91. In: Brügelmann/ Richter Brügelmann, H./ Richter, S. (Hrsg.) (1994): Wie wir recht schreiben lernen. Zehn Jahre Kinder auf dem Weg zur Schrift. Libelle Verlag: CH-Lengwil (2. Aufl. 1996), 129-134.

genutzt, so dass kurze Texte vorliegen. Diese Schreibproben haben wir genutzt, um die Rechtschreibung der SchülerInnen zu überprüfen. Dabei interessierte uns vor allem die Teilstichprobe NRW, da wir aus dem Jahre 2001 Texte einer ebenfalls repräsentativen Teilstichprobe der Studie NRW-KIDS⁷ zum Vergleich der Rechtschreibleistung⁸ heranziehen konnten. Insgesamt lagen uns 3.692 Testbögen vor⁹. Sie verteilen sich wie folgt (s. Tab. 1):

Tab 1: Stichproben-Größe¹⁰	ALLE	nur NRW¹¹	nicht NRW¹²
Zahl der Testbögen insgesamt ¹³	3.692	1.033	2.659
Antworttexte zur Aufgabe „Okashi“	2.263	614	1.649
Antworttexte zur Aufgabe „Sardine“	3.202	927	2.275
Antworttexte zu beiden Aufgaben zusammen	2.168	599	1.569

Danach lagen in den einzelnen Fällen folgende Daten vor¹⁴:

383 gar keinen Text verfasst, davon:

- 219 keine der beiden Aufgaben bearbeitet und deshalb gar kein Text,
- 90 einer der beiden Aufgaben bearbeitet, aber keinen Text,
- 74 beide Aufgaben bearbeitet, aber keinen Text verfasst;

1.141 nur einen Text verfasst, davon:

- 32 ein Text zu einer Aufgabe, zweite Aufgabe fehlt
- 748 eine Aufgabe bearbeitet und dazu einen Text verfasst,
- 361 beide Aufgaben bearbeitet, aber nur einen Text verfasst;

2.168 beide Aufgaben bearbeitet und einen Text verfasst

Repräsentativität der Stichprobe

Die Gesamtstichprobe des IEA-Lesevergleichs war als Zufallsstichprobe für die BRD repräsentativ. Zur Einschätzung einer möglichen Verzerrung der Teilstichprobe, von der Texte vorliegen, haben wir überprüft, ob sich die Leistungen in diesen Teilgruppen unterscheiden¹⁵: SchülerInnen, die nur einen Text verfassen, machen in diesem mehr Fehler (5-8%) als SchülerInnen, die zu beiden Aufgaben Texte verfassen (3-4%). Für die Stichprobe NRW liegen zusätzlich Werte aus dem Leseverständnistest vor¹⁶: Danach unterscheiden sich die Gruppen,

⁷ Zinnecker, J., u. a. (2002): null zoff & voll busy. Die erste Jugendgeneration des neuen Jahrhunderts. Leske + Budrich: Opladen.

⁸ Vgl. Brügelmann, H. (2003l): Der Karawanen-Effekt beim Rechtschreiblernen: Schwache SchülerInnen lernen genauso schnell, aber von einem anderen Startpunkt aus. Bericht aus der Reanalyse der freien Texte von 4.- bis 10-Klässlern in der NRW-Kids Studie 2001. Vervielf. Ms. FB2 der Universität: Siegen; Brügelmann, H. (2004m): Textmenge und Rechtschreibfehler in freien Texten (Klasse 4 bis 12). Eine Sekundärauswertung der Studie NRW-KIDS. FB 2 der Universität: Siegen.

⁹ Teilgenommen haben 4.348 SchülerInnen (vgl. Lehmann u. a. 1995, 23), die Differenz lässt sich nach der Wanderung der Test-Pakete über verschiedene Zwischenstationen nicht mehr aufklären.

¹⁰ IEA-TEXT-NRW.desc.04-02-07

¹¹ if (~ sysmis(sstratum))

¹² if (sysmis(sstratum))

¹³ s. Anm. 9.

¹⁴ Vgl. Grundtabellen 1a und 1b im statistischen Anhang

¹⁵ Vgl. Grundtabelle 2a

¹⁶ Vgl. Grundtabelle 2b

die keinen oder nur einen Text schreiben (-0.3 bis 0.6 Punkte¹⁷), deutlich von denen, die beide Texte schreiben (1.6 Punkte). Alle Indikatoren sprechen also dafür, dass Letztere eine positive Auslese darstellen.

Wir haben zusätzlich an der Teilstichprobe NRW untersucht, ob es eine Rolle spielt, warum jemand keinen Text zu der Aufgabe verfasst hat. Unterschieden werden diejenigen, die die Leseaufgabe gar nicht bearbeitet haben (= 1), von denjenigen, die nur die Auswahlfragen zum Text beantwortet und keine Erläuterung geschrieben haben, und denjenigen die zusätzlich einen erläuternden Text geschrieben haben¹⁸:

Tab. 2: Kennwerte der unterschiedlichen Antworttypen¹⁹	Fehleranteil	Leseleistung (nur NRW)
Sardine „0“/ Okashi „0“ (N = 62 / 299 / 54 / 106)	5.6 % / 7.2%	0.5 / -0.1
Sardine „1“/ Okashi „1“ (N = 16 / 732 / 29 / 286)	5.1% / 3.2%	0.5 / 0.8
Sardine „9“/ Okashi „9“ (N = 2168 / 3202 / 608 / 917) (N = 2263 / 2168 / 608 / 917)	3.6% / 3.9 % 3.7% / 3.2 %	1.5 / 1.2

Es bestätigt sich, dass SchülerInnen, die zu beiden Aufgaben geschrieben haben, die besseren RechtschreiberInnen sind. Zwischen den anderen beiden Teilgruppen zeigten sich in der einen Aufgabe Unterschiede, in der anderen aber nicht, so dass wir keinen systematischen Grund für das unterschiedliche Antwortverhalten vermuten können.

In einem weiteren Schritt haben wir die Teilstichprobe Nordrhein-Westfalen mit der Reststichprobe BRD (West) verglichen. Tab. 3a gibt einen Überblick über die Zahl der Wörter pro Text:

Tab. 3a: Text-Umfang²⁰	ALLE	nur NRW²¹	nicht NRW²²
Antworttexte zur Aufgabe "Okashi"	25.8 (.3)	27.1 (.6)	25.3 (.3)
Antworttexte zur Aufgabe "Sardine"	23.4 (.5)	23.0 (.4)	23.5 (.6)
gemittelt über beide Texte	24.6	25.1	24.4
Summe, falls beiden Aufgaben zusammen	49.3 (.7)	50.4 (.9)	48.9 (.9)

¹⁷ Hierbei handelt es sich um eine standardisierte Skala, was den negativen Mittelwert erklärt.

¹⁸ Vgl. Grundtabellen 2c und 2d

¹⁹ Vgl. Grundtabellen 2a und 2b

²⁰ Vgl. Grundtabellen 3a-c

²¹ Teildatei: if (~ sysmis(ssstratum))

²² if (sysmis(ssstratum))

In Tab 3b. sind die Fehlerquoten der Rechtschreibung zusammen gestellt:, wobei zunächst nur die Buchstabenfehler ausgezählt wurden:

Tab. 3b: Buchstabenfehler-Quote²³	ALLE	nur NRW	nicht NRW
Antworttexte zur Aufgabe "Okashi"	3.7 (.2)	3.3 (.3)	3.9 (.3)
Antworttexte zur Aufgabe "Sardine"	3.9 (.1)	4.3 (.3)	3.7 (.2)
gemittelt über beide Texte	3.8	3.8	3.8
Summe, falls beide Aufgaben zusammen	3.3 (.1)	3.5 (.2)	3.3 (.1)

Härter ist das Kriterium orthografischer Richtigkeit, wenn auch Fehler in der Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung berücksichtigt werden (Tab. 3c):

Tab. 3c: Gesamt-Wortfehler-Quote²⁴	ALLE	nur NRW	nicht NRW
Antworttexte zur Aufgabe "Okashi"	7.3 (.4)	6.4 (.7)	7.6 (.5)
Antworttexte zur Aufgabe "Sardine"	5.9 (.2)	6.4 (.3)	5.7 (.2)
gemittelt über beide Texte	6.6	6.4	6.6
Summe, falls beide Aufgaben zusammen	5.7 (.2)	5.9 (.4)	5.7 (.2)

Der Vergleich der verschiedenen Indizes zeigt, dass die Teilstichprobe NRW geringfügig (um 3%) längere Texte schreibt als die Teilstichprobe „Übrige Bundesländer (West, ohne NRW)“, dass ihre Rechtschreibleistung je nach Text einmal leicht besser, das andere Mal leicht schlechter ausfällt (plus/ minus 3-6%).

Insgesamt erweist sich unsere NRW-Stichprobe also als repräsentativ für das Bundesgebiet West zum damaligen Zeitpunkt. Diese Stichprobe selbst dürfte für die Rechtschreibkompetenz allerdings eher eine positive Auslese darstellen, da schwächere RechtschreiberInnen häufiger keinen Text geschrieben haben.

Vergleich der Leistungen in verschiedenen Schulformen

Für Nordrhein-Westfalen liegen als Bezugswerte die Ergebnisse des Leseverständnistests vor, deren arithmetische Mittel für die verschiedenen Schulformen den Erwartungen entsprechen (vgl. Tab. 4a):

²³ Reine Buchstabenfehler (ohne Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung); vgl. Grundtabellen 3a-c

²⁴ Alle Wortfehler (inkl. Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung); vgl. Grundtabellen 3d-f

Tab. 4a: Leseverstehen arithmetisches Mittel (N)	nur NRW
Hauptschule	.5 (355)
Gesamtschule	.9 (94)
Realschule	1.3 (294)
Gymnasium	2.2 (240)

An einer Teilstichprobe von N = 593 (NRW) können wir die Lese- und Rechtschreibleistungen korrelieren²⁵. Dabei zeigt sich eine mäßiger Zusammenhang mit dem Umfang der Texte (r = .25** bis .32**) und mit den Fehlerquoten (r = .19** bis .24**).

Die Summenwerte für die Rechtschreibung (vgl. Tab. 4b) liegen nur in Haupt- und Gesamtschule auf dem erwarteten Niveau. In NRW überrascht das schlechte Ergebnis der Realschulen, in den übrigen Bundesländer (zusammen genommen) das schlechte Ergebnis der Gymnasien.

Tab. 4b: FEHLSUMP²⁶ arithmetisches Mittel (N)	nur NRW	nicht NRW
Hauptschule	8.4 % (142)	10.1 % (338)
Gesamtschule	5.1 % (45)	5.0 % (92)
Realschule	8.2 % (185)	3.8 % (444)
Gymnasium	2.5 % (223)	4.9 % (661)

Da die Abweichungen mit den zum Teil doch recht kleinen Stichproben zusammenhängen können²⁷, haben wir in einer weiteren Auswertung die Buchstabenfehler in den Texten zur Aufgabe „Sardine“ einbezogen²⁸:

Ta. 4c: SFEHLALLP²⁹ arithmetisches Mittel (N)	nur NRW	nicht NRW
Hauptschule	6.2 (305)	6.9 (668)
Gesamtschule	3.6 (87)	3.3 (140)
Realschule	4.9 (284)	2.2 (661)
Gymnasium	1.7 (235)	2.4 (770)

²⁵ Vgl. Grundtabelle 6

²⁶ Alle Wortfehler (inkl. Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung) für beide Texte zusammen; vgl. Grundtabellen 4a und 4b

²⁷ Immerhin liegen die Standardfehler in der Stichprobe NRW zwischen .22 und 1.20, in der Stichprobe restliche BRD zwischen .29 und .61 (vgl. Grundtabellen 4a und 4b)

²⁸ Die Fehlerquote SFEHLP korreliert zwar nur zu .48** mit der Fehlerquote OFEHLALP im zweiten Text, ist aber ein guter Indikator für die Fehlerquote beider Texte zusammen (FEHALP), also mit der größeren Textmenge: r = .81** (vgl. Grundtabelle 6).

²⁹ Alle Buchstabenfehler im Wort in der Aufgabe „Sardine“ (ohne Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung); vgl. Grundtabellen 4a und 4b

Diese Auswertung ist zwar schon weniger störanfällig, aber wenn man bedenkt, dass nicht einzelne SchülerInnen, sondern ganze Klassen die Einheiten der Stichprobenziehung waren, sind es in den Realschulen NRW nur 13 Klassen gewesen, in den Gesamtschulen sogar nur vier, so dass spezifische Unterrichtsbedingungen zu einer Verzerrung der Stichproben geführt haben können. Da sich eingangs gezeigt hat, dass die Gesamtstichprobe NRW der Reststichprobe BRD sehr ähnlich ist, werden wir für die folgenden Analysen auf die robusteren Werte der Gesamtstichprobe „BRD-West“ zurückgreifen, deren zentrale Kennwerte in Abb. 5 zusammengefasst sind:

Tab 5: Zentrale Kennwerte³⁰ nach Schulformen (ALLE)				
	SWORTALL	SFEHLALP	FEHLSUMP	WORTALLE
Hauptschule	19.3 (973)	6.7 %	9.6 %	36.8 (480)
Gesamtschule	20.9 (227)	3.4 %	5.0 %	43.2 (137)
Realschule	22.8 (945)	3.0 %	5.1 %	49.3 (629)
Gymnasium	28.7 (1005)	2.2 %	4.3 %	56.8 (884)

Die beiden unterschiedlichen Zugänge führen zu denselben Ergebnissen:

- Mit ansteigendem Niveau der Schulform nimmt der Textumfang zu; die Texte der GymnasiastInnen sind um rund 50% länger als die der HauptschülerInnen, die Texte aus der Gesamt- und aus der Realschule liegen zwischen diesen Extremwerten.
- Mit ansteigendem Niveau der Schulform sinkt die Quote der Rechtschreibfehler; HauptschülerInnen machen etwa zwei- bis dreimal so viele Rechtschreibfehler wie GymnasiastInnen, Gesamt- und RealschülerInnen liegen wiederum in der Mitte.

Vergleich der Leistungen von 1991 mit 2001

Vergleiche von Schulleistungen über die Zeit hinweg sind schwierig:

- Die Stichproben sind nicht immer vergleichbar, insbesondere wenn es um einen Vergleich von Schulformen geht, deren Rekrutierung sich durch die wachsende Bildungsbeteiligung verändert.
- Die Instrumente bzw. Untersuchungsbedingungen sind nicht gleichartig.
- Im Bereich der Rechtschreibung haben sich Inhalte (Grundwortschatz!) und Materialien verändert, so dass man sogar nicht davon ausgehen kann, dass dieselbe Aufgabe (z. B. ein bestimmtes Diktat) wirklich „dieselbe“ Aufgabe darstellt.

Mit diesen Vorbehalten werden im Folgenden die Ergebnisse der Schreibproben unserer beiden eingangs vorgestellten Sekundärauswertungen aus der IEA-Studie von 1991 und aus der NRW-KIDS-Studie NRW 2001 einander gegenüber gestellt. Zu beachten ist, dass die Erhebungen

³⁰ Vgl. Grundtabellen 5a und 5b; die Kennwerte FSUMP und WORTALLE wurden wegen des größeren Wortumfangs der kombinierten Texte einbezogen, die Kennwerte SFEHLALP und SWORTALL wegen der größeren Zahl der Personen, die den Text „Sardine“ kommentiert haben.

1991 im Frühjahr, also gegen Ende des Schuljahres durchgeführt wurden, während die Erhebung 2001 im Herbst, also am Anfang des Schuljahres stattfand. Wir werden deshalb für 2001 sowohl die Werte der 8. als auch der 9. Klasse heranziehen, um zu einem fairen Vergleich zu kommen.

Das Kriterium WORTALLE in Tab. 6a zeigt zunächst, dass die Texte 1991 kürzer waren, was aus der anderen Aufgabenstellung (Kommentar zu einer von mehreren Aufgaben vs. freier Text zu einem umfassenderen Thema) plausibel ist.

Tab. 6a : Quote der Rechtschreibfehler: 1991 vs 2001				
arithmetisches Mittel NRW (BRD ALLE)				
	N =	WORTALLE	FEHLALLP	FEHLSUMP
„Okashi“ ³¹	614 (2.263)	27.1 (25.8)	3.3% (3.7%)	6.4% (7.3%)
8. IEA ³² 1991 „Sardine“ ³³	927 (3.202)	23.0 (23.4)	4.3% (3.9%)	6.4% (5.9%)
gemittelt		25.1 (24.6)	3.8% (3.8%)	6.4% (6.6%)
Gesamt³⁴	599 (2168)	50.4 (49.3)	3.5% (3.3%)	5.9% (5.7%)
8. NRW 2001	142	127.4	2.7 %	6.1%
9. NRW 2001	141	154.2	1.6 %	4.5 %
NRW 2001 8./ 9. gemittelt		140.8	2.2 %	5.3 %

Die Quote der Buchstabenfehler im Wort (FEHLALLP) war nach jedem Indikator 1991 höher als 2001. Das gilt sowohl, wenn man die beiden Stichproben NRW vergleicht, als auch wenn für 1991 auf die größere Stichprobe BRD insgesamt zurückgegriffen wird. Zudem gilt es sowohl für die 9.- als auch sogar für die 8.-Klässler 2001. Bezieht man aber auch Fehler in der Groß-/ Kleinschreibung und in der Getrennt-/ Zusammenschreibung mit ein (FEHLSUMP), schrumpft der Vorsprung der beiden Gruppen 2001 etwas.

Dieser Befund verdient eine differenziertere Betrachtung, da wir für die 4. Klassen 2001 eine etwa gleiche Leistung wie im Schreibvergleich BRDDR von 1991 festgestellt haben³⁵.

Beispielsweise ist aufgrund einer veränderten Bildungsbeteiligung mit einer anderen Verteilung der SchülerInnen auf die verschiedenen Schulformen zu rechnen. Unter diesem Aspekt sind die Daten deshalb in Tab. 6b weiter aufgeschlüsselt worden, wobei in diesem Vergleich das härtere Kriterium (*alle* Wortfehler = FEHLSUMP) angelegt wurde.

³¹ Alle Buchstabenfehler im Wort in der Aufgabe „Sardine“ (ohne Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung);

³² Aus den bereits erwähnten Gründen werden für 1991 die Werte „NRW“ und „BRD ALLE“ herangezogen, um die wegen ihrer geringen Größe störanfällige NRW-Stichprobe zu ergänzen

³³ S. Anm. 31

³⁴ alle Wortfehler (inkl. Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung) für beide Texte zusammen

³⁵ Vgl. Brügelmann (2004m). Bei einer früheren Auswertung hatten wir die unterschiedlichen Termine im Schuljahr nicht bedacht und deshalb fälschlicherweise eine Verschlechterung der Leistungen von 1991 bis 2001 unterstellt, vgl. noch: Brügelmann, H. (2003b): Rechtschreibung in freien Texten am Ende der Grundschule. Teilauswertung der NRW-KIDS-Studie 2001. In: Panagiotopoulou, A./ Brügelmann H. (Hrsg.) (2003): Grundschulpädagogik *meets* Kindheitsforschung: Zum Wechselverhältnis von schulischem Lernen und außerschulischem Erfahrungen im Grundschulalter. Leske+Budrich: Opladen, 173-178, hier: 176-177.

Tab. 6b : Quote der Rechtschreibfehler (FEHLSUMP) nach Schularten : 1991 vs 2001 arithmetisches Mittel NRW (BRD ALLE)								
	Hauptschule		Gesamtschule		Realschule		Gymnasium	
	„Okashi“ ³⁶	4.9 (6.3)	2.3 (3.3)	4.8 (2.9)	1.2 (2.9)			
8. IEA ³⁷ 1991	„Sardine“ ³⁸	6.2 (6.7)	3.6 (3.4)	4.9 (3.0)	1.7 (2.2)			
	gemittelt	5.6 (6.5)	3.0 (3.4)	4.9 (3.0)	1.5 (2.6)			
	Gesamt ³⁹	8.4 (9.6)	5.1 (5.1)	8.2 (5.1)	2.5 (4.3)			
8. NRW 2001	36/30/36/40	9.6	8.9	4.1	2.5			
9. NRW 2001	39/18/48/36	8.6	5.4	2.8	2.0			
	NRW 2001 8./9. gemittelt	9.1	7.2	3.5	2.3			

Bezogen auf die einzelnen Schulformen deutet sich eine stabile Situation an – außer für die Gesamtschule, wo wir eine Verschlechterung beobachten. Beziehen wir uns in 1991 auf die größeren Stichproben (Text „Sardine“), so sind die Leistungen 2001 in Haupt- und Gesamtschule schlechter. Beziehen wir uns auf die größere Textmenge 1991 (beide Texte zusammen = „Gesamt“), so sind die Ergebnisse 2001 im Vergleich mit der NRW-Stichprobe 1991 in allen Schulformen außer in der Gesamtschule gleich oder besser.

Für den Grundschulbereich zeigen die wenigen Befunde zur Rechtschreibung in *Aufsätzen* für frühere Perioden einen eher negativen Trend⁴⁰. Nach Menzel⁴¹ lag die Quote der Rechtschreibfehler in freien Texten in einer niedersächsischen Stichprobe 3./4. Klasse Anfang der 80er Jahre bei 7-8% – schon damals etwas höher als in einer noch früheren Stichprobe Anfang der 70er Jahre mit 6-7%. Im Schreibvergleich BRDDR 1990/91 erhielten wir für Wortfehler Ende vierter Klasse in der Stichprobe-West Quoten um 12%. In einer aktuellen bayerischen Stichprobe von Richter⁴² betrug die analoge Fehlerquote 14-15% (in 3. Klassen 18.2% und in 4. Klassen 11.7%) – etwa so hoch wie in unserer Sekundärauswertung der NRW-KIDS-Studie von 2001 mit 12-16%⁴³. Damit scheint die deutliche Fehlerzunahme von 1970 bis 1990 in den 90er Jahren etwas abgeflacht zu sein.

Daten aus standardisierten *Diktaten* zeigen allerdings kaum Veränderungen der Fehlerquote: Von 1982 bis 1995 kann das an der Eichung des DRT 3 belegt werden⁴⁴. Dasselbe gilt für den

³⁶ Alle Buchstabenfehler im Wort in der Aufgabe „Sardine“ (ohne Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung);

³⁷ Aus den bereits erwähnten Gründen werden für 1991 die Werte „NRW“ und „BRD ALLE“ herangezogen, um die wegen ihrer geringen Größe störanfällige NRW-Stichprobe zu ergänzen

³⁸ A. Anm. 31

³⁹ Alle Wortfehler (inkl. Groß-/ Kleinschreibung und Getrennt-/ Zusammenschreibung) für beide Texte zusammen

⁴⁰ Allerdings stellt Vogel (1988, s. Anm. 46) für 1970/74 vs. 1982/84 keinen Unterschied mehr fest, wenn er die gewachsene Textmenge und Wortschwierigkeit berücksichtigt.

⁴¹ Menzel, W. (1985): Rechtschreibunterricht -- Praxis und Theorie. Aus Fehlern lernen. Beiheft zu Praxis Deutsch Nr. 69. Friedrich Verlag: Seelze, 5.

⁴² Richter, S. (2002): Schreibwortschatz von Grundschulkindern. Bayerisches Korpus. S. Roderer Verlag: Regensburg.

⁴³ Vgl. zusammenfassend Brügelmann (2004m) gegen (2003b, 175).

⁴⁴ Vgl. Müller (1977), zit. nach dem Ms. von Roßbach, H.-G./ Wellenreuther, M. (2002): Empirische Forschungen zur Wirksamkeit von Methoden der Leistungsdifferenzierung in der Grundschule, für: Heinzl, F./ Prengel, A.

AST-3 von 1969/70 bis 1985/86⁴⁵ und sogar von 1959 bis 1984 in einer rheinland-pfälzischen Erhebung⁴⁶. Daten zur Neunormierung der HSP von Peter May⁴⁷ bestätigen diese Stabilität von Anfang bis Ende der 90er Jahre⁴⁸.

Der interessante Widerspruch: Wir beobachten also von Anfang der 70er Jahre bis 2001 eine Zunahme der Rechtschreibfehler in der Aufgabe „freier Text“ (von 6-7% auf 14-15%), parallel dazu aber keine Zunahme in den Diktaten. Die Rechtschreibleistung scheint unter unterschiedlichen Bedingungen sehr unterschiedlich auszufallen, und die Differenzen sind in verschiedenen Gruppen nicht unbedingt gleichartig⁴⁹.

Insofern wird man mit der verbreiteten These vom „Verfall“ der Rechtschreibleistungen sehr vorsichtig umgehen müssen. Quantitative und qualitative Reanalysen von Daten aus weiteren Quellen (z. B. aus der Neu-Normierung von verschiedenen Tests und aus breiten Erhebungen z. B. von Rekruten oder von Lehrlingen bei Eingangsprüfungen der Kammern⁵⁰) sind unbedingt erforderlich.

Vor allem die kontinuierlichen Zuwächse der Rechtschreibleistungen von Klasse 5 bis 12 in der Sekundarstufe⁵¹ stimmen optimistisch und machen deutlich, dass eine zu frühe Forderung hoher Rechtschreibsicherheit in der Grundschule nicht notwendig ist, um eine funktionale Rechtschreibfähigkeit bis zum Abschluss der Schulzeit zu gewährleisten.

(Hrsg.) (2002): Heterogenität, Integration und Differenzierung in der Primarstufe. Jahrbuch Grundschulforschung 6. Leske + Budrich: Opladen, 44-59 (im Ms. S. 3, im Haupttext entfallen)

⁴⁵ Vgl. Roßbach/ Tietze (1996), zit. ebda.

⁴⁶ Vgl. Vogel, R. (1988): Rechtschreibanforderungen und -Leistungen früher und heute. Eine vergleichende Untersuchung von Schularaufsätzen und -Diktaten. Franzbecker: Bad Salzdetfurth. Vogel berichtet auch die Ergebnisse einer Studie von Ingenkamp, der 1962 die Schulleistungen von 6.-Klässlern mit denen von 1949 verglichen hat. Entgegen den Urteilen der beteiligten LehrerInnen hat die Leistungsstreuung von 1949 bis 1962 abgenommen, sind die Leistungen angestiegen; vgl. Ingenkamp, K. (1967): Schulleistungen -- damals und heute. Beltz: Weinheim, S. 86 ff., 148, 152

⁴⁷ Pers. Mitteilung e-mail v. 7.10.02

⁴⁸ Vgl. zur Widersprüchlichkeit der Daten bzw. zu möglichen Verzerrungen und zu den Schwierigkeiten ihrer angemessenen Interpretation: Brügelmann, H. (Hrsg.) (1999): Was leisten unsere Schulen? Erhard-Friedrich-Verlag: Seelze, 10-20, insbesondere: 17-19

⁴⁹ Vgl. zum „Laufstalleffekt“ des DDR-Rechtschreibunterrichts im Bereich der Diktate: Brügelmann, H., u. a. (1994g): „Schreibvergleich BRDDR“ 1990/91. In: Brügelmann, H./ Richter, S. (Hrsg.) (1994): Wie wir recht schreiben lernen. Zehn Jahre Kinder auf dem Weg zur Schrift. Libelle Verlag: CH-Lengwil (2. Aufl. 1996), 129-134; und: Brügelmann, H., u. a. (1994c): Richtig schreiben durch freies Schreiben? Rechtschreibentwicklung in Schweizer Klassen, die nach "Lesen durch Schreiben" unterrichtet wurden. In: Brügelmann/ Richter (1994, 135-148).

⁵⁰ Allerdings müssen solche Daten sehr sorgfältig interpretiert werden; vgl. zu unzulässigen Vereinfachungen: Brügelmann (1999, 11-14).

⁵¹ Vgl. die Daten aus der NRW-KIDS-Studie 2001 in Brügelmann (2004m).