

Ein Sturz und seine Folgen. Zur Evaluation von Textkompetenz im narrativen Schreiben mit dem FÖRMIG-Instrument ‚Tulpenbeet‘

Einleitung

In diesem Beitrag werden vertiefende Analysen zu narrativen Schülertexten vorgestellt, die mit einer standardisierten Schreibaufgabe im Kontext der Evaluation des BLK-Programms im Abstand von einem Jahr erhoben wurden. Es hatte sich herausgestellt, dass mit der ursprünglichen Konzeption des Erhebungsinstruments nicht in allen der ausgewerteten Bereiche Lernfortschritte beobachtet werden konnten – z.T. waren durchaus auch signifikante Rückschritte festzustellen – bzw. konnten nicht alle Lernfortschritte der Schülerinnen und Schüler erfasst werden. In den folgenden Abschnitten werden mit einer Analyse der Erzählkompetenz einerseits textsortenspezifische Gestaltungsmittel näher betrachtet, zum anderen werden kurz Ergebnisse einer Analyse orthografischer und grammatischer Normabweichungen vorgestellt. Diese vertiefenden Beobachtungen beziehen sich auf eine Zufallsstichprobe von 84 Schülerinnen und Schülern aus einer nordrhein-westfälischen Region – insgesamt sind hier elf Schulen beteiligt – und sollen dazu dienen, die Auswertungskriterien der Schreibaufgabe gegebenenfalls zu modifizieren.

Im ersten Abschnitt wird zunächst die ursprüngliche Konzeption des Instruments vorgestellt. Der zweite Abschnitt befasst sich mit der Analyse der Erzählkompetenz als neuem Indikator sowie dem zugrunde liegenden Entwicklungsmodell. Im dritten Abschnitt wird zusätzlich untersucht, ob die Schülerinnen und Schüler ihre orthografischen und grammatischen Kompetenzen weiterentwickeln konnten.

1. Das ‚Tulpenbeet‘ als förderdiagnostisches Evaluationsinstrument

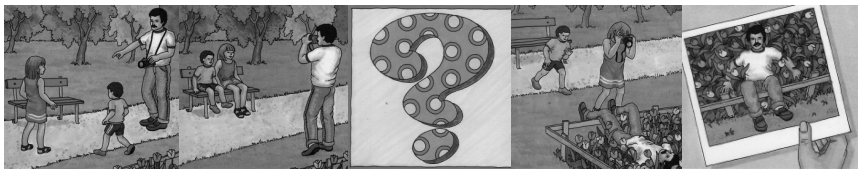
Im Rahmen des BLK-Programms FÖRMIG wurde die Schreibaufgabe ‚Der Sturz ins Tulpenbeet‘ unter einer doppelten Zielvorgabe entwickelt: Einerseits soll es ein förderdiagnostisches Instrument für die Hand der Lehrkräfte sein, um diese darin zu unterstützen, den Sprachstand vor

allem von Schülerinnen und Schülern mit Migrationshintergrund nach Teilkompetenzen differenziert sowohl in der Erst- als auch in der Zweitsprache zu bestimmen und auf dieser Grundlage sprachliche Förderung auf die individuellen Lernbedürfnisse der Schülerinnen und Schüler bezogen durchführen zu können. Andererseits dient das Instrument dazu, die im Programm erprobten Fördermaßnahmen im Rahmen der Gesamtevaluation auf ihre Wirksamkeit hin zu untersuchen. Zu diesen beiden Zwecken wurde die Schreibprobe in zwei Wellen erhoben, zwischen denen ein knappes Jahr der sprachlichen Förderung lag. Nachfolgend soll zunächst die erste Konzeption des Instruments kurz vorgestellt werden.

Die Konzeption des Instruments

Die Aufgabe der Schülerinnen und Schüler besteht darin, die auf einer Folge von fünf Bildern dargestellten Ereignisse (siehe Abb. 1) zu verschriftlichen, das heißt aus diesem Impuls einen erzählenden Text zu verfassen. Dem Bildimpuls liegt folgendes ‚Script‘ zugrunde: Bei einem Besuch im Park versucht ein Vater seine beiden Kinder zu fotografieren, dabei kommt es zu den in Abbildung 1 erkennbaren Komplikationen. Anstelle des dritten Bildes ist ein großes Fragezeichen abgebildet, so dass die Schüler gehalten sind, die ‚auslösende Komplikation‘ der Ereignisfolge eigenständig herauszuarbeiten, also sprachlich darzustellen, wie es zu dem Sturz ins Tulpenbeet gekommen ist.

Abbildung 1: Die Bildsequenz ‚Der Sturz ins Tulpenbeet‘

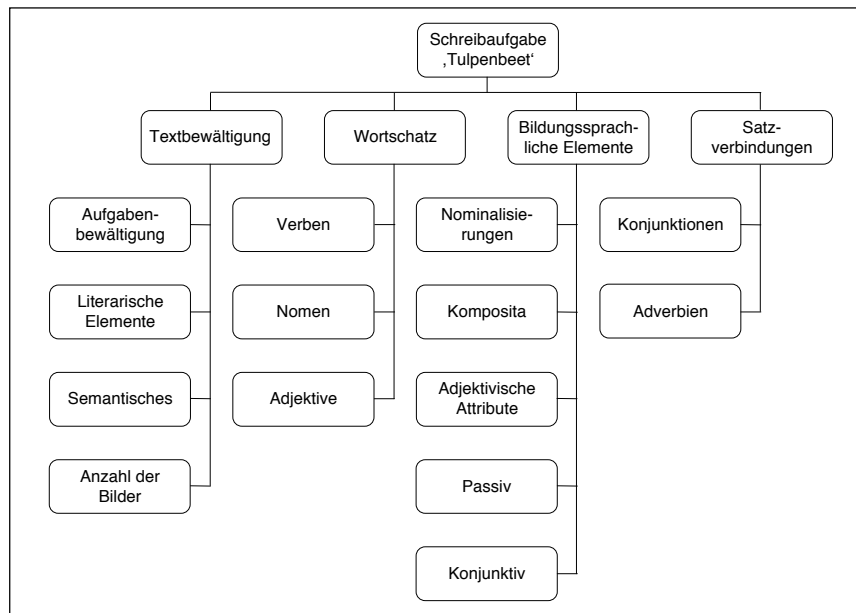


© SCHUBI. Bilderbox 1, Und dann ...? SCHUBI Lernmedien AG, Schaffhausen 1990.

Mit den von den Schülern verfassten Texten werden im Nachgang eine Reihe von Auswertungsprozeduren durchgeführt, welche im Folgenden dargestellt werden. Insgesamt richtet sich die Auswertung weniger auf grammatische Kompetenzen, wie es in Verfahren für den Übergang vom Elementarbereich in die Primarstufe häufiger der Fall ist. Im Vordergrund der Auswertung des ‚Tulpenbeets‘ stehen vielmehr textbezogene Kriterien, die zu Beginn und im weiteren Verlauf der Sekundarstufe I eine wichtige Rolle spielen. Zu diesem Zeitpunkt ihrer schulischen Laufbahn befinden sich die Lernenden nach abgeschlossener Alphabetisierung in einem Stadium fortgeschrittenen Schriftspracherwerbs, so dass es sinnvoll ist, die Textebene ins Zentrum sprachdiagnostischer Aktivität-

ten zu rücken. Sollten Lernende erst eine elementare Sprachkompetenz im Deutschen besitzen, so kann zusätzlich eine grammatische Profilanalyse unter Rückgriff auf die bekannten Erwerbssequenzen in der deutschen Sprache durchgeführt werden.¹ Abbildung 2 zeigt überblicksweise die sprachlichen Dimensionen, welche durch die Auswertung zugänglich werden.

Abbildung 2: Sprachliche Dimensionen des Instruments



Ein erster Block ist mit dem Ausdruck ‚*Textbewältigung*‘ überschrieben; damit soll auf die Bedeutung des handelnden Umgangs mit Texten hingewiesen werden. Kompetenzen – seien sie schulischen oder anderen Lernprozessen erwachsen – können immer nur indirekt erschlossen werden. Beobachten lässt sich der je aktuelle Umgang mit der Textförmigkeit im Schreiben selbst. Der Block umfasst folgende Indikatoren:

- Die *Aufgabenbewältigung* ist ein globales Maß für die schriftliche Lösung der gestellten Aufgabe, einen Erzähltext zu produzieren. Es wird erfasst, inwiefern es den Schülerinnen und Schülern gelingt, die in der Bildsequenz dargestellten Vorgänge, Aktanten, Gegenstände und Zusammenhänge in der jeweiligen Sprache zu verschriftlichen. Da die Auswertung des gesamten Blocks ‚*Textbewältigung*‘ nicht in

1 Eine Adaption liegt für die Herkunftssprachen Türkisch und Russisch vor.

Abhängigkeit von einzelsprachlichen Strukturen (Syntax, Lexik, Morphologie) erfolgt, ist es möglich, die im Deutschen und der Herkunftssprache erreichten Werte direkt miteinander zu vergleichen.

Die zweite Analysekategorie der *Textbewältigung* umfasst spezifische, von Erzähltexten zu erwartende *literarische Elemente*, dabei werden die Gestaltungsformen von Inhalt und Schluss in den Blick genommen. Solche narrativen Gestaltungselemente geben einen Einblick in die Textkompetenz der Schülerinnen und Schüler und schließen an in der Schule vermittelte Inhalte und Fähigkeiten an. Weiterhin wird die ausgelassene Szene (*‚Semantisches‘*) dahingehend untersucht, wie die Schülerinnen und Schüler mit der hier gegebenen Gestaltungsfreiheit umgehen.

- Zur *Textbewältigung* gehört es auch, ob die Texte alle Bilder umfassen; daher wird die *Anzahl der beschriebenen Bilder* gesondert erfasst. Angesichts der in geschriebenen Texten höheren Frequenz und Bedeutung von Inhaltswörtern (*‚lexical items‘*) kommt dem *‚Wortschatz‘* als zweitem Auswertungsblock eine besondere Bedeutung zu. Es ist zu erwarten, dass neben hoch frequenten Wörtern mit eher allgemeiner Bedeutung solche mit spezifischer, auf die Textsorte Erzählung bezogener Bedeutung Verwendung finden.

Aufgrund der Voruntersuchungen hat sich gezeigt, dass – im Gegensatz zur Einschätzung des Sprachstands in der gesprochenen Sprache – eine Erhebung der *Verben* allein nicht ausreicht. Gerade für bildungssprachliches Schreiben und die Einschätzung der Begriffsbildung ist der *nominale Wortschatz* von hoher Bedeutung. Einen Text von tieferer Differenziertheit zeichnet außerdem die Verwendung spezifizierender *Adjektive* aus.

Im Kontext der Schriftsprachlichkeit haben sich in vorangegangenen Untersuchungen *‚bildungssprachliche Elemente‘* (*Nominalisierungen, Komposita, adjektivische Attribute, Passiv und Konjunktiv*) als stabile Indikatoren eines entwickelten Sprachstands erwiesen (vgl. Roth, Neumann & Gogolin, 2007, S. 58 ff.). Diese werden im dritten Auswertungsblock des Instruments gesondert erfasst.

Weiterhin werden die *‚Satzverbindungen‘* verzeichnet, die einen Einblick in den Verknüpftheitsgrad eines Textes zulassen (Textkohäsion) und veranschaulichen, in welchen (semantischen) Relationen einzelne Teilsätze zueinander stehen. Solche Konnektoren spielen auch für die Erzählkompetenz eine herausragende Rolle, da vor allem mit adversativen Konnektoren Brüche in der Handlung verdeutlicht und somit einerseits die Erzählwürdigkeit der Ereignisfolge herausgestellt wird, zum anderen aber auch Übergänge zwischen den Makrokonstituenten der Textform markiert werden (vgl. Abschnitt 2).

Schließlich werden quantitative Maße wie die Anzahl der Wörter, der Sätze, die mittlere Satzlänge und die lexikalische Dichte erhoben.

In den Auswertungen und Analysen der deutschsprachigen Texte nach der zweiten Erhebungswelle stellte sich heraus, dass mit den aufgeführten, eher quantitativen Indikatoren eine Entwicklung der Textkompetenz der Schülerinnen und Schüler nicht hinreichend abgebildet werden kann. Die Gründe dafür liegen vor allem auf der Ebene der Durchführung: Anscheinend wurden die notwendigen Bedingungen der Erhebungsobjektivität nicht eingehalten. In der zweiten Erhebungswelle bestand für die Schüler – laut Aussagen der Lehrkräfte – mit einem knapperen Zeitrahmen bzw. geringerer motivationaler Unterstützung ein anderes Aufgabenumfeld (vgl. Hayes & Flower, 1980). Aus wissenschaftlicher Sicht könnte man nun berechtigt die weitere Analyse abbrechen und das vorgelegte Instrument auf einen kontrollierten Einsatz beschränken. Vielleicht liegt aber die größere Herausforderung darin, auch unter defizitären Bedingungen entstandene Daten auszuwerten und ein Instrument zu entwickeln, das zumindest in Teilen so robust ist, dass es motivationale und durchführungsbedingte Verzerrung verkraftet? Im Folgenden werden wir einen entsprechenden Ansatz vorstellen.

In diesem Zusammenhang muss außerdem auf ein grundlegendes Messproblem der (longitudinalen) Sprachstandsfeststellung hingewiesen werden, das sich aus einer möglichen Diskrepanz zwischen den tatsächlich produzierten Texten (Performanz) und der individuell bedingten höchstmöglichen Kompetenz ergibt: Nicht jede Performanz lässt auf die vorhandene Kompetenz zurückschließen – dieses ist nur im Kontext einer hinreichenden Menge an (in unserem Fall) schriftlichen Produktionen (und motivationaler Lage) möglich. Aufgrund der geschilderten Auswertungsproblematik erschien es uns sinnvoll, eine vertiefende Analyse der bereits im ursprünglichen Design des Instruments vorhandenen Kategorie der ‚literarischen Elemente‘ vorzunehmen, um einer Beobachtung des professionellen Auswerters nachzugehen, der sämtliche Texte der ersten und zweiten Welle bearbeitet hat: Der augenscheinliche Gesamteindruck der Texte der zweiten Welle wies – trotz gesunkener Textmenge – in eine Richtung, die eine Verbesserung der Textkompetenz greifbar werden zu lassen schien.

Um die Ebene der Textkompetenz stärker in den Blick zu bekommen, bot es sich an, auf das von Boueke und anderen (vgl. Boueke, Schüle, Büscher, Terhorst & Wolf, 1995) entwickelte Modell der Erzählfähigkeit zurückzugreifen. Einerseits berücksichtigen die Autoren explizit den Faktor Entwicklung, andererseits fußt das Modell auf einer empirischen Untersuchung, in der ebenfalls mit Bildsequenzen als Impulsen gearbeitet wurde. Als zentrale Kategorien des Modells, aus welchem die Indikatoren für die in Abschnitt 3 vorgestellte Stichprobe hergeleitet wurden,

sind neben der Segmentierung der Erzählung in die Ereignisstruktur widerspiegelnde Konstituenten vor allem ‚affektive Markierungen‘ zu nennen. Dies sind sprachliche oder inhaltliche Elemente, die vor allem dazu dienen, den Rezipienten einer Geschichte in die Handlung zu involvieren und somit einen (impliziten) Adressatenbezug herstellen.

Zurzeit wird von Christoph Gantefort die zweisprachige Entwicklung der Erzählkompetenz von Kindern eines von den Universitäten Köln und Hamburg wissenschaftlich begleiteten bilingualen Grundschulprojekts untersucht. Im Zentrum steht die Untersuchung der textuellen Fähigkeiten von Schülerinnen und Schülern zweier Sprachgruppen: die einen sind simultan bilingual mit den Sprachen Sorbisch und Deutsch aufgewachsen, die anderen sind sukzessiv über Zweitspracherwerb und mit Unterstützung von Bildungsinstitutionen (Kindergarten, Grundschule) an die Bilingualität herangeführt worden.

2. Analyse der Erzählkompetenz

In diesem ersten Auswertungsblock der vertiefenden Beobachtungen wird zunächst der theoretische Rahmen des Modells geklärt, bevor anschließend die Ergebnisse dargestellt und interpretiert werden.

2.1 Ein Modell zur Entwicklung der Erzählfähigkeit

Es werden zunächst kurz maßgebliche Einflüsse der bisherigen Erzählforschung auf die Arbeit von Boueke und anderen nachgezeichnet. Nach daran anschließenden Hinweisen zur Methodik werden die in unserer Untersuchung verwendeten Variablen aus den ‚Strukturtypen‘ zur Erzählfähigkeit hergeleitet.

Aus der an dieser Stelle nicht darzustellenden Vielfalt an Einflüssen aus der Erzählforschung sei lediglich auf das Werk von Labov und Waletzky (1973) verwiesen, welches sich sowohl auf das Modell von Boueke und anderen (1995) als auch auf viele andere Autoren entscheidend ausgewirkt hat. Die Autoren entwickelten eine ‚Normalform‘ von Erzählungen, die sich textstrukturell typischerweise aus den fünf Makrokonstituenten Exposition, Komplikation, Evaluation, Auflösung sowie der Coda zusammensetzt. In diesem linear-temporalen Höhepunktmodell dient die Exposition dazu, den Rezipienten mit der grundlegenden Situation der dargestellten Sequenz vertraut zu machen, aus der sich mit der Komplikation als ‚erregendem Moment‘ eine Problematik entwickelt, die die dargestellte Ereignisfolge ‚erzählwürdig‘ werden lässt. Es muss

also ein gewisser Bruch in der Handlung gegeben sein, um im Sinne von Labov und Waletzky von einer *Geschichte* reden zu können.

Die Autoren unterscheiden auf der elementaren Ebene zwischen freien und narrativen Teilsätzen. Während die narrativen Teilsätze dazu dienen, die Ereignisfolge darzustellen, sie demnach also nicht innerhalb der Erzählung vertauscht werden können, ohne die temporale Abfolge zu verändern, besitzen freie Teilsätze oft die Funktion, das Erlebte zu kommentieren und zu bewerten. An einem Punkt maximaler Komplikation finden sich gehäuft solche freien, evaluierenden Teilsätze und konstituieren damit die Evaluation. In einer späteren Veröffentlichung (vgl. Labov, 1980) revidiert Labov diese Annahme insofern, als diese evaluativen Elemente nicht nur zur Gestaltung des narrativen Höhepunkts, sondern auch in den anderen Teilen einer Erzählung vereinzelt auftauchen können.

In der Auflösung als viertem Teil der Normalform wird die Lösung der (problematischen) Situation dargestellt, mit welcher der oder die Aktanten der Geschichte konfrontiert waren.

Die Coda dient schließlich dazu, durch Äußerungen wie „*und so war das*“ den Abschluss der Diskurseinheit explizit zu markieren. Dadurch wird die temporäre Inanspruchnahme des Rederechts durch den ‚Erzähler‘ aufgelöst und der Fokus wieder auf die gegenwärtige Gesprächssituation gerichtet.

Boueke und andere beziehen dieses Strukturmodell auf die *Entwicklung* der Erzählfähigkeit von Kindern, indem sie anhand eines Korpus von transkribierten Erzählungen untersuchten, inwiefern die Kinder ihre Texte analog zu den genannten Makrokonstituenten gestalteten und evaluative Mittel verwendeten. Die Begriffe *Diskontinuität* und *Emotionalität* nehmen in diesem Kontext eine zentrale Stellung ein, sie beziehen sich auf die oben bereits angedeuteten Segmentierungen und affektiven Markierungen des Textes: Wie und wann gelingt es den Kindern, ihre Erzählungen zu strukturieren und dabei insbesondere die Komplikation als zentralen Bruch in der Handlung und als dasjenige Element, was eine Ereignisfolge erst erzählwürdig werden lässt, sprachlich zu realisieren? Inwiefern verwenden Kinder evaluative Elemente, mit denen die sprachlich dargestellte Ereignisfolge emotional qualifiziert und damit von reinem ‚Berichten‘ unterscheidbar wird?

Methodik

Um eine Antwort auf die zum Schluss des letzten Absatzes aufgeworfenen Fragen zu finden, wählten die Autoren folgendes experimentelle Untersuchungsdesign: Insgesamt 96 Kindern aus drei verschiedenen Altersgruppen (5, 7 und 9 Jahre alt) wurden verschiedene Bildergeschichten vorgelegt. Diese sollten die jungen Probanden so spannend wie möglich

nacherzählen. Im Gegensatz zu Labov und Waletzky, die für ihre Studie in eher natürlichen Situationen *selbsterlebte* Erzählungen analysierten, entschied sich die Bielefelder Forschungsgruppe für eine Laborsituation, in welcher *Nacherzählungen* evoziert wurden. Diese in der Forschung durchaus auch problematisierten Versuchsbedingungen (vgl. Bredel, 2001; Knapp, 2001; Dehn, 2006) wurden folgendermaßen gerechtfertigt:

Da Sprachgebrauch immer auch vom Kontext abhängig ist, kann man mit guten Gründen fordern, dass er da – und nur da – studiert wird, wo er sich tatsächlich ereignet, d. h. innerhalb ‚natürlicher‘ Situationen. Ein solcher im Rahmen interaktionistischer und soziolinguistischer Untersuchungen denn auch favorisierter Zugang ist für die Untersuchung der dem Sprachgebrauch ‚zugrunde liegenden‘ Kompetenzen aber weniger geeignet als ein experimentelles Setting, wie es von uns ... gewählt worden ist: Die ‚Natürlichkeit‘ der Erzählsituation und die Vergleichbarkeit der dabei gewonnen Erzähltexte schließen sich weitgehend aus (Boueke et al., 1995, S. 120).

Dieses Zitat macht zum einen die Problematik deutlich, die sich auch für standardisierte Instrumente zur Sprachstandserhebung wie das ‚Tulpenbeet‘ ergibt: Freie Textproduktionen in natürlichem Rahmen sind in der individuellen Diagnose mit Sicherheit aussagekräftiger, jedoch fordern die Testgütekriterien von einem standardisierten Verfahren einen einheitlichen Impuls sowie vergleichbare Durchführungsbedingungen. Zum anderen ist das von den Autoren entwickelte Verfahren gerade durch das experimentelle Design besonders geeignet, ein standardisiertes Verfahren wie das ‚Tulpenbeet‘ um wichtige Beobachtungskriterien zu ergänzen, da dort ebenfalls mit Bildergeschichten in homogenen Erhebungssituationen gearbeitet wurde.

Strukturtypen der Erzählkompetenz

Die Auswertung der über Bildimpulse elizitierten mündlichen Erzählungen führte zu einem Entwicklungsmodell, in welchem Boueke und andere zwischen vier verschiedenen, die Entwicklung der narrativen Kompetenz widerspiegelnden *Strukturtypen* unterscheiden (vgl. Boueke et al., 1995, S. 177 ff.):

- 1) Die fünfjährigen Kinder produzierten vor allem Erzähltexte, die von der Darstellung *isolierter Ereignisse* dominiert werden. Ebenfalls häufig ließen sich räumlich-deiktische Äußerungen beobachten: die Kinder formulieren auf dieser Stufe ‚Bildpropositionen‘, ohne inhaltliche und sprachliche Zusammenhänge herzustellen.
- 2) Ein erster wichtiger Entwicklungsschritt liegt vor, wenn die Kinder beginnen, die zuvor isoliert dargestellten Ereignisse temporal mitei-

ander zu verknüpfen. Dies geschieht häufig mittels des Konnektors ‚(und) dann‘. Die Kinder sind also in der Lage, die Bildfolge als auf einer Zeitachse angeordnet zu erkennen und dies sprachlich zu repräsentieren. Ursachen von Ereignissen oder Absichten von Personen werden auf dieser Stufe allerdings noch nicht verbalisiert.

- 3) Auf der dritten Stufe der *handlungslogischen Strukturierung* folgt die Komposition des Textes bereits einem wesentlich komplexeren Muster. Der Textaufbau orientiert sich an den von Labov und Waletzky beschriebenen Konstituenten: Einerseits werden Ort und Aktanten in einer Exposition eingeführt, zum anderen wird die Diskontinuität des Handlungsablaufs als Komplikation herausgestellt. Konkret wird dies durch *Ereignisstruktur-Markierungen* geleistet. Als zentral sind hier adversative Konnektoren aufzufassen, mit denen Brüche in der Handlung und damit besagte Diskontinuität im Handlungsablauf verdeutlicht werden. Somit folgt die Textgestaltung auf dieser Stufe nicht mehr nur einer simplen Chronologisierung der Ereignisse, sondern vielmehr einem komplexeren, auf die Bedeutung der Ereignisse bezogenen Muster. An dieser Stelle lassen sich Parallelen zum Textmodell Roland Harwegs identifizieren. Er unterscheidet mit „Kette“ und „Durchschuss“ zwischen zwei Erzeugungsrichtungen eines Textes (vgl. Harweg, 1973, zitiert nach Gülich & Raible, 1977, S. 119 ff.): Die ‚Kette‘ als horizontale Dimension meint die Verknüpfung benachbarter Sätze mittels pronominaler Substitution. Der ‚Durchschuss‘ dagegen ist als eine vertikal zur ‚Kette‘ stehende Dimension zu verstehen, welche die spezifische Makrostruktur von Textsorten determiniert. Auf die vorliegende Kompetenzstufe des Erzählens gemünzt bedeutet dies, dass die Kinder nicht nur in der Lage sind, einen Text mit einer lückenlosen Referenzprogression im Sinne der ‚Kette‘ zu produzieren, sondern mit der Segmentierung des Textes in die genannten Makrokonstituenten nunmehr auch den für Geschichten charakteristischen ‚Durchschuss‘ beherrschen.
- 4) Auf der vierten und höchsten Stufe der Erzählkompetenz schließlich strukturieren die Kinder ihre Texte nicht nur entsprechend der zugrunde liegenden Handlungslogik – als zusätzliches Kriterium erscheint hier ein spezifischer Adressatenbezug in Form von inhaltlichen und sprachlichen Elementen, die den Rezipienten in die Handlung involvieren und die Geschichte spannender und interessanter werden lassen. Diese *affektiven Markierungen* dienen also dazu, die Ereignisfolge emotional zu qualifizieren. Boueke und andere nehmen hier eine theoriegeleitete Klassifikation vor, indem sie die affektiven Markierungen den drei Bereichen ‚Valenz‘, ‚Plötzlichkeit‘ und ‚psychologische Nähe‘ zuordnen: Mittel, die den Hörer in das Geschehen hineinversetzen, konstituieren demnach die psychologische Nähe.

Dies kann durch direkte und indirekte Rede, eine Thematisierung der Gedanken von Aktanten oder die Wiedergabe von akustischen Sinneseindrücken geschehen. Plötzlichkeit dagegen wird vor allem durch entsprechende Temporaladverbien wie ‚auf einmal‘ oder eben ‚plötzlich‘ markiert. Unter Valenz schließlich subsumieren die Autoren solche Mittel, die Positives oder Negatives hervorheben, was durch expressive Lexeme oder Intensifikatoren wie ‚sehr‘ geleistet werden kann.

2.2 Die Indikatoren in unserer Stichprobe

Wie kann das oben dargestellte Modell nun für die Auswertung der in FÖRMIG erhobenen Schreibproben nutzbar gemacht werden? Zunächst muss auf einige grundlegende Unterschiede bezüglich der Untersuchungsbedingungen hingewiesen werden: Bei unserer Stichprobe handelt es sich mit Schülerinnen und Schülern am Beginn der Sekundarstufe I nicht nur um ältere Kinder, die Texte wurden zudem im Medium der Schriftlichkeit verfasst; es galten also andere Produktionsbedingungen als in der Untersuchung von Boueke und anderen. Dies bedeutet für die Kinder eine Zerdehnung der Kommunikationssituation über Raum und Zeit hinweg (vgl. Ehlich, 1983). Somit richten sich die Texte an einen abstrakten Adressaten, und die kommunikative Aufgabe wird insgesamt anspruchsvoller, da ein direktes Rückmeldeverhalten nicht möglich ist. Andererseits sind im Schriftlichen differenziertere Planungs- und Revisionsprozesse möglich, aber auch notwendig, die gerade für Kinder im Zweitspracherwerb je nach Entwicklungsstand eine eigene Herausforderung darstellen.

Aufgrund des Alters der Kinder handelt es sich bei den Schreibproben in den meisten Fällen um Texte, die den Strukturtypen III oder IV zugeordnet werden können. Für die Abbildung der Entwicklung der Erzählkompetenz kann eine bloße Bestimmung des Strukturtyps nicht ausreichend sein. Daher wählten wir die *Häufigkeit der Vorkommen der Ereignisstruktur- und affektiven Markierungen* als empirische Grundlage, um eine angemessene ‚Körnigkeit‘ des Entwicklungsmodells zu gewährleisten.

2.2.1 Ereignisstruktur-Markierungen

Die von Boueke und anderen (1995, S. 90 f.) zugrunde gelegten Markierungen der einzelnen Konstituenten wurden von uns modifiziert: Sie unterscheiden zwischen Markierungen des Settings (Exposition), der Episode (Komplikation) und des Abschlusses (Auflösung). In unserer

Untersuchung fokussieren wir hinsichtlich bestimmter Konstituenten nur die Differenziertheit der Exposition, da diese am ehesten impulsunabhängig ist. Dagegen sind die Markierungen, die von Boueke und anderen bezüglich der Episode und des Abschlusses genannt werden², meist direkt der Bildsequenz entnehmbar und lassen damit weniger auf die individuelle Ausprägung der Textkompetenz schließen. Adversative Konnektoren, die ‚makrosematische Schnittstellen‘ bilden, wurden von uns verzeichnet, ohne zu berücksichtigen, welche Konstituenten damit im Einzelnen voneinander abgegrenzt werden.

Im Fokus unserer Auswertungen stand das ‚temporale Gerüst‘ der einzelnen Texte, d.h. inwiefern die Schülerinnen und Schüler über einfache ‚und dann‘-Verknüpfungen, welche den Strukturtyp II kennzeichnen, hinausgehen. Zu diesem Zweck wurden die Texte daraufhin untersucht, ob und wie oft eine (explizite) Markierung von Vorzeitigkeit, Gleichzeitigkeit oder Nachzeitigkeit von Ereignissen vorliegt. Vorzeitigkeit kann neben Konjunktionen auch durch Adverbien oder ein kontrastives Tempus (Plusquamperfekt nach Präteritum oder Perfekt) gekennzeichnet werden. Gleichzeitigkeit von Ereignissen wird in den meisten Fällen durch die subordinierende Konjunktion ‚als‘ markiert, ebenso aber auch koordinativ durch Adverbien wie ‚dabei‘. Nachzeitigkeit wird von uns als vom aktuellen Handlungszeitpunkt aus auf die Zukunft bezogen verstanden, so dass sie zumeist in Form des Futurs realisiert wird. Einfache ‚und dann‘-Verknüpfungen wurden an dieser Stelle explizit nicht berücksichtigt. Weiterhin wurde die Häufigkeit von strukturell komplexen Zeitangaben wie z.B. in Präpositionalphrasen nach dem Muster ‚nach einem kurzen Moment ...‘ verzeichnet, die – funktional betrachtet – ebenfalls Ereignisabfolgen gliedern und auf semantischer Ebene eine höhere Präzision und Differenzierungstiefe bewirken.

Zur Ereignisstruktur einer Geschichte gehört es auch, Begründungen von Ereignissen bzw. Handlungen und mögliche Folgen dieser zum Ausdruck zu bringen. Dies kann durch einfache kausale Konnektoren wie *weil* oder Partikel wie *also* geleistet werden. Im Folgenden werden alle der Kategorie Ereignisstruktur-Markierungen zugeordneten Phänomene noch einmal zusammenfassend aufgeführt:

- Differenzierung der Exposition:³
 - Einführung der Zeit,
 - Einführung des Handlungsortes,

2 Z.B. die Einführung neuer Aktanten oder die Einführung eines neuen Handlungsortes.

3 Zur Berechnung der Differenziertheit der Exposition wurde zunächst für jede der im Folgenden genannte Unterkategorie eine dichotome Variable gebildet (Wert 1 = Kategorie vorhanden; Wert 0 = Kategorie nicht vorhanden). Der individuelle Wert der Differenziertheit der Exposition errechnet sich aus der Summe dieser dichotomen Variablen und liegt demnach zwischen 0 und 4.

- Einführung der Aktanten,
- Einführung von laufenden Handlungen.
- Markierung der Temporalstruktur durch:
 - Vorzeitigkeit,
 - Gleichzeitigkeit oder
 - Nachzeitigkeit von Ereignissen,
 - (nichtadverbiale) Zeitangaben.
- Markierung der Ursache-Wirkungs-Struktur durch:
 - Begründungen von Ereignissen und Handlungen,
 - Darstellung von Folgen von Ereignissen oder Handlungen.

2.2.2 Affektive Markierungen

Der triadischen Aufteilung in Valenz, psychologischer Nähe und Plötzlichkeit wird im Weiteren nicht mehr gefolgt – sinnvoller für eine überblicksweises Darstellung erscheint zunächst eine Aufteilung der emotional qualifizierenden Phänomene in einen inhaltlichen und einen sprachlichen Bereich, wie Becker ihn in Anlehnung an Boueke und andere vorgenommen hat (vgl. Becker, 2005, S. 93). Nachfolgend werden die Phänomene aufgelistet, die in unserer Stichprobe erfasst und statistisch ausgewertet wurden:

- Inhaltlich:
 - indirekte Rede,
 - direkte Rede,
 - Verbalisierung von Empfindungen und Sinneseindrücken der Aktanten,
 - Verbalisierung von Wünschen und Absichten der Aktanten,
 - Verbalisierung von Handlungszwängen der Aktanten,
 - Verbalisierung von Gedanken der Aktanten,
 - Verbalisierung von inneren Zuständen der Aktanten.
- Sprachlich:
 - Iteration als Gestaltungselement (z.B. ‚Der Vater ging weiter und weiter zurück‘),
 - Lautmalerei,
 - Plötzlichkeit markierende Temporaladverbien,
 - expressive Verben,
 - evaluierende Adjektive,
 - Intensifikatoren (z.B. ‚sehr‘ oder ‚ganz‘).

Becker kritisiert die von Boueke und anderen vorgenommene Typologisierung dahingehend, dass sie der indirekten und direkten Rede eine af-

fektive Qualität abspricht. Dies erscheint auf den ersten Blick plausibel, wird doch allein durch das Formalium ‚direkte Rede‘ nichts über emotionale Zustände der Aktanten ausgesagt. Auf der Ebene des Rezipienten einer Geschichte kann allerdings durchaus von einer affektiven Wirkung ausgegangen werden, da diesem die Aktanten darüber als sprachlich handelnde und interagierende Personen im Sinne eines ‚dramatischen‘ Effekts näher gebracht werden. Somit erscheint es in theoretischer Hinsicht sinnvoll, die affektiven Markierungen auf der inhaltlichen Ebene als handlungsinterne sowie adressatenbezogene und als rein adressatenbezogene zu unterscheiden.

2.3 Ergebnisse

Zunächst soll hier die Differenziertheit der Exposition beleuchtet werden. Bereits in den ursprünglichen Auswertungen⁴ hatte sich gezeigt, dass dies einer der wenigen Teilbereiche ist, in denen eine Entwicklung zu beobachten war. Auch in der Stichprobe wuchs der Anteil der Schülerinnen und Schüler, die überhaupt eine Exposition in den Text integrierten, von 78 Prozent in der ersten auf 86 Prozent in der zweiten Welle. Auch hinsichtlich der Differenziertheit dieser dem Rezipienten Orientierung bietenden Makrokonstituente hat ein Zuwachs stattgefunden.

In der zweiten Welle führten die Lernenden einen potentiellen Leser der Geschichte wesentlich ausführlicher und vollständiger in den Handlungskontext ein als noch in der ersten; der Mittelwert stieg von 2,11 auf 2,64⁵. Diese Entwicklung ist statistisch signifikant.

Weitere Ereignisstruktur-Markierungen

Eine ausführliche Darstellung sämtlicher Einzelphänomene würde den Rahmen dieses Beitrags sprengen, daher seien nur diejenigen genannt, die die quantitativen Pole bilden: Am häufigsten wurden kausale Zusammenhänge und eine Gleichzeitigkeit von Ereignissen markiert, wohingegen Folgen von Handlungen oder Ereignissen oder ihre Vorzeitigkeit am seltensten zu beobachten waren.

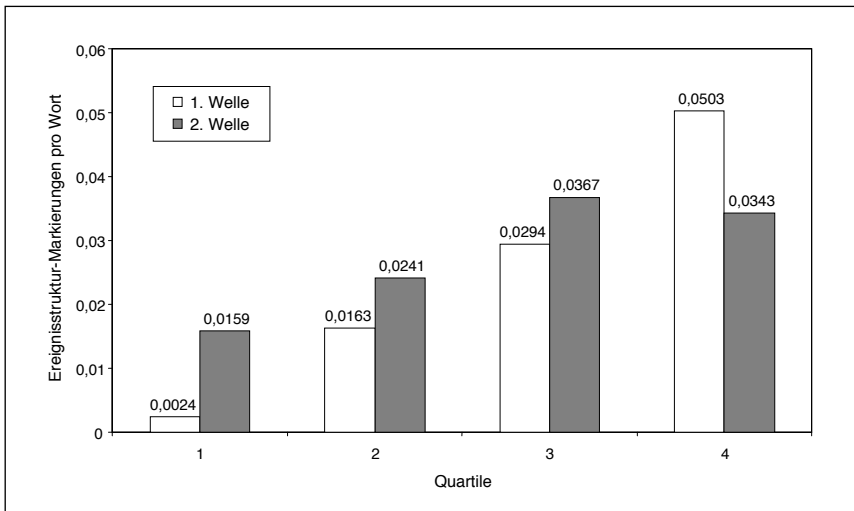
Da wir deduktiv von dem bestehenden Theoriemodell von Boueke und anderen und damit von der Eindimensionalität der erfassten Phänomene ausgehen, bietet es sich an, die Summe der Ereignisstruktur-

4 Dort wurde unter dem Punkt ‚Gestaltung des Anfangs‘ erfasst, ob die Kinder eine Überschrift, einen formelhaften Einstieg verwenden bzw. ob der eigentlichen Erzählsequenz überhaupt eine Exposition vorangestellt wird.

5 Bei der hier verwendeten Einheit handelt es sich um den Mittelwert der in der Exposition eingeführten Kategorien (vgl. Abschnitt 2.2.1); das mögliche Maximum liegt demnach bei 4,0.

Markierungen zu bilden. Der darüber errechnete Mittelwert steigt mäßig von 2,64 auf 2,80 solcher Markierungen pro Text, dies aber wohlgermerkt bei teilweise rapide sinkenden Werten in der Aufgabenbewältigung und dem verbalen Wortschatz. Eine deutlichere Zunahme zeigt sich hingegen, wenn man die Anzahl der Ereignisstruktur-Markierungen zur Anzahl der im Text geschriebenen Wörter in Beziehung setzt: Dieser Quotient steigert sich von 0,025 auf 0,028. Demnach ist der zuvor berichtete augenscheinliche Eindruck der verbesserten Textkompetenz durch den professionellen Auswerter belegt. Es liegt den Texten der zweiten Welle trotz stagnierender bzw. zum Teil abnehmender Textmenge eine qualitativ komplexere Ereignisstruktur zugrunde. Ein Element dessen ist beispielsweise darin zu sehen, dass viele Kinder, die ihre Texte in der ersten Welle dem Strukturtyp II entsprechend mit vielen ‚und dann‘-Verknüpfungen gestalteten, zum zweiten Erhebungszeitpunkt der ‚Normalform‘ von Erzählungen gerechter wurden.

Abbildung 3: Quartile des Quotienten der Ereignisstruktur-Markierungen nach Erhebungszeitpunkt



Diese Vermutung lässt sich anhand von Abbildung 3 erhärten. Sie stellt die Entwicklung hinsichtlich der relativen Häufigkeit der Ereignisstruktur-Markierungen für die Quartile der ersten Welle dar: Darin zeigt sich deutlich, dass die Kinder des ersten Quartils die stärksten Zuwächse zu verzeichnen haben, während bei der ‚Spitzengruppe‘ ein regelrechter Einbruch festzustellen ist. Dieser lässt sich auch in zum Teil noch deutlicherer Ausprägung bei einer Reihe anderer Indikatoren der ersten Aus-

wertung beobachten (siehe unten).⁶ Es lässt sich die Annahme formulieren, dass die in der ersten Welle stärksten Schülerinnen und Schüler in der zweiten Welle aufgrund einer zu geringen Herausforderung oder Motivation weniger von ihrem Können realisierten. Für das ‚obere Feld‘ der Untersuchungsgruppe des vierten Quartils manifestiert sich dieses Phänomen demnach auch in der Erzählkompetenz.

Eine den Befund tiefer gehend betrachtende Frage lautet: Drückt sich der oben festgestellte Anstieg der relativen Häufigkeit der Phänomene auch in einer höheren Differenziertheit der Verwendung dieser Mittel aus? Zur Beantwortung dieser Frage wurden die einzelnen Variablen dichotomisiert und im Anschluss addiert. Der Dichtomisierung, sprich der Umwandlung der einzelnen Summen in eine vorhanden/nicht vorhanden-Kategorisierung liegt die theoretische Annahme zugrunde, dass das Vorhandensein qualitativ verschiedener Phänomene der Abbildung sprachlicher Kompetenz näher kommt als eine Berechnung nach Vorkommenshäufigkeiten. Am Ergebnis ändert sich allerdings nichts: Steigerungen waren auch mit dieser Berechnungsmethode nicht festzustellen.⁷

Abgesehen von der Differenziertheit der Exposition erweisen sich die Ereignisstruktur-Markierungen damit für die vorliegende Stichprobe insgesamt als wenig aussagekräftig zur Erfassung der Entwicklung von Erzählkompetenz. Es kann nicht ausgeschlossen werden, dass diese unter optimalen Erhebungsbedingungen auch über den Zeitraum eines knappen Jahres sprachlicher Förderung durchaus Fortschritte ausweisen lassen – sie erweisen sich aber nicht als robust genug, um ein Instrument auch für einen eben nicht immer unter optimalen Bedingungen ablaufenden Einsatz tauglich zu machen.

Affektive Markierungen

In der Verwendung affektiver Markierungen offenbart sich dem hier zur Anwendung gebrachten Modell gemäß die höchste Erzählkompetenz. Um mit den Worten von Boueke und anderen zu sprechen verfolgen die Kinder eine ‚Involvierungsstrategie‘ in der ‚zerdehnten‘ Kommunikationssituation (vgl. Abschnitt 2.2).

Die Häufigkeitsverteilung ergibt zu beiden Erhebungszeitpunkten in etwa das gleiche Bild: Mit großem Abstand werden Wünsche der Aktanten, direkte Rede und evaluative Adjektive am häufigsten realisiert. Dagegen wurden Gedanken und Zwangslagen von Aktanten selten verbalisiert, fast ebenso selten bezogen die Kinder onomatopoetische Mittel

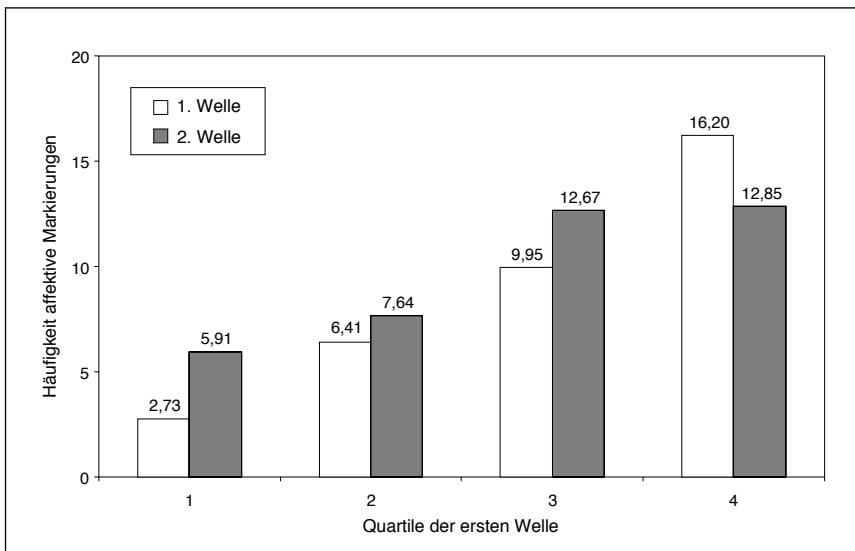
6 Nach unserer Einschätzung handelt es sich nicht um eine ‚Regression zur Mitte‘, sondern um beobachtbare Schreibphänomene, die gerade bei den weniger weit entwickelten Schreibern reale Leistungszuwächse ausweisen.

7 In der ersten Welle gebrauchten die Kinder im Schnitt 2,01 verschiedene Ereignisstruktur-Markierungen, in der zweiten dann 2,02.

in ihre Textgestaltung mit ein. Im Gegensatz zu den Ereignisstruktur-Markierungen wurden in der zweiten Welle trotz geringerer Textmenge sogar absolut mehr affektive Markierungen verwendet als in der ersten, der Mittelwert stieg von 8,63 auf 9,66. Dieser Zuwachs fällt hinsichtlich der relativen Häufigkeit noch deutlicher aus, bezogen auf die Anzahl der Wörter wird er statistisch signifikant ($p = .009$).

Auch die 13 Variablen, welche die Häufigkeit der affektiven Markierungen wiedergeben, wurden aufgrund der oben genannten Überlegung dichotomisiert, um zu prüfen, ob ihre Verwendung in der zweiten Welle differenzierter ausgefallen ist. Es lässt sich ein Zuwachs nachweisen. Im Mittel wurden zum zweiten Erhebungszeitpunkt 4,82 verschiedene affektive Markierungen verwendet, während es zum ersten 4,29 waren. Diese Entwicklung ist ebenfalls signifikant ($p = .01$) und markiert noch deutlicher als die Summe in absoluten Zahlen eine ausdifferenzierte Textkompetenz der Lernenden. Diese Beobachtung gilt nicht für alle Schülerinnen und Schüler: Hinsichtlich der Häufigkeit der affektiven Markierungen ist erneut festzustellen, dass die Spitzengruppe der ersten Welle ihre Werte in der zweiten Erhebung nicht reproduzieren konnte. Ablesen lässt sich dies an Abbildung 4, in der die Entwicklung der affektiven Markierungen nach den Quartilen der ersten Welle dargestellt ist.

Abbildung 4: Mittelwert der affektiven Markierungen zu beiden Erhebungszeitpunkten nach den Quartilen der ersten Welle



Somit wird die Tendenz aus der bisherigen Auswertung nochmals bestätigt, wonach sich vor allem die Schülerinnen und Schüler mit in der ersten Welle eher schwachen Ergebnissen weiterentwickeln konnten. Der oben beschriebene ‚Einbruch‘ betrifft allerdings nur das oberste Quartil. Darüber hinaus ist eine Entkopplung des in der ersten Welle feststellbaren Zusammenhangs zwischen Aufgabenbewältigung und affektiven Markierungen zu verzeichnen (1. Welle: $r = .38$; 2. Welle: $r = .22$.); dies wird auch bei einem Vergleich der zwei Diagramme in Abbildung 5 und 6 ersichtlich, in denen nach den Quartilen der Aufgabenbewältigung zum einen der Mittelwert der Aufgabenbewältigung, zum anderen die affektiven Markierungen dargestellt sind.

Weiterhin wird deutlich, dass bezüglich der Aufgabenbewältigung in fast allen Gruppen ein Rückgang der Werte auf ein fast einheitliches Niveau festzustellen ist, während dieselben Gruppen – mit Ausnahme des obersten Quartils – in der zweiten Welle wesentlich mehr affektive Markierungen verwenden. Mit anderen Worten: Trotz der geringeren Vollständigkeit in der sprachlichen Darstellung der Bildsequenz haben die Schülerinnen und Schüler die spezifischen Gestaltungsmittel der Textsorte Erzählung in der zweiten Welle häufiger und differenzierter angewendet. Die Aufgabenbewältigung dagegen erweist sich für die vor-

Abbildung 5: Mittelwert der Aufgabenbewältigung zu beiden Erhebungszeitpunkten nach den Quartilen der ersten Welle

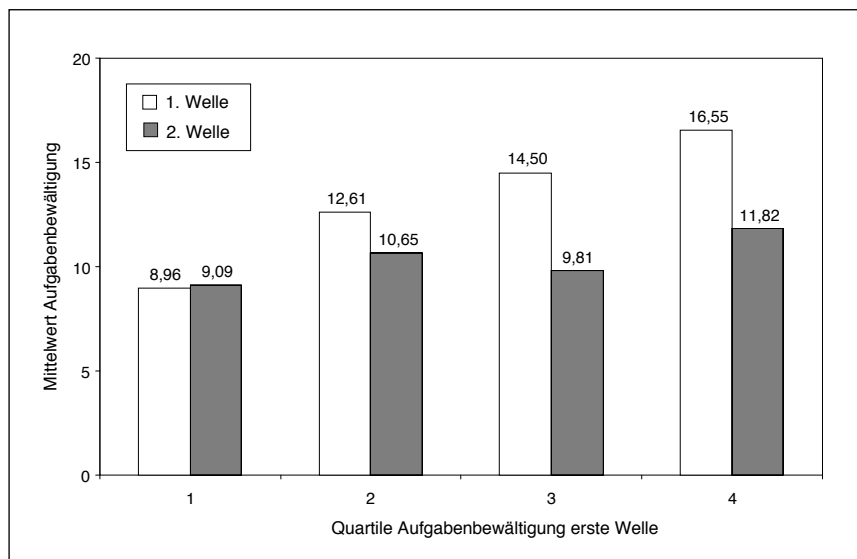
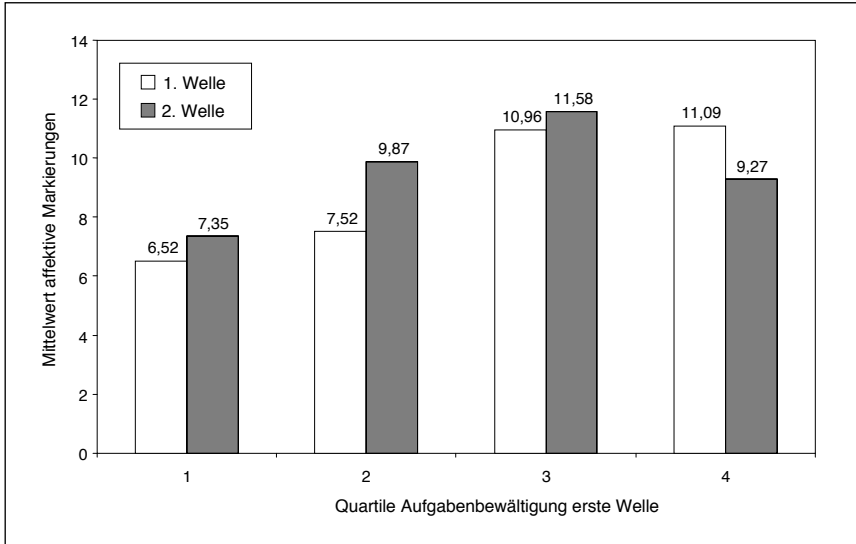


Abbildung 6: Mittelwert der affektiven Markierungen zu beiden Erhebungszeitpunkten nach den Quartilen der Aufgabenbewältigung der ersten Welle



liegende Stichprobe nur hinsichtlich der Einstiegsdiagnostik als stabiler Indikator. Zusammenfassend kann somit durchaus von einer Steigerung der Textkompetenz gesprochen werden, auch wenn offensichtlich motivationale Einbrüche – aufgrund nicht vergleichbarer Erhebungssituationen oder des der Wiederholung desselben Verfahrens geschuldeten Gewöhnungseffekts – vorliegen. Für die Instrumentenentwicklung ist von Bedeutung, dass die affektiven Markierungen anscheinend so weit robust sind, auch unter suboptimalen Einsatzbedingungen qualitative Aussagen über die Entwicklung der Textkompetenz von Schülerinnen und Schülern zu ermöglichen.

Wie lässt sich nun diese Diskrepanz zwischen den Indikatoren erklären? Eine Möglichkeit besteht darin, auf das Konzept von implizitem und explizitem Wissen zurückzugreifen: In diesem Fall würde sich sprachliche Kompetenz auf der Textebene ähnlich wie bei Muttersprachlern auf der Satzebene als ein weitgehend unbewusstes Wissen darstellen. Analog zu der Tatsache, dass Muttersprachler in den seltensten Fällen die grammatische Wohlgeformtheit einzelner Sätze anhand von Regeln erklären können, ginge man unter dieser Prämisse davon aus, dass auch spezifische, kollektiv vereinbarte Muster auf der höheren Hierarchieebene des Textes in ihrer konkreten Realisierung nicht nur auf bewussten Planungsprozessen beruhen. Auf den hier vorliegenden Fall übertragen bedeutet dies, dass die Lernenden im Zeitraum zwischen den Erhebun-

gen solch implizites Wissen über die Struktur der Textsorte Erzählung erworben haben. Diese Fertigkeit wurde trotz – dies ist an den Werten der Aufgabenbewältigung ablesbar – geringerer Motivation bzw. geringerer bewusster Auseinandersetzung mit der Schreibaufgabe in den Texten realisiert.

Abschließend bleibt die Frage zu beantworten, welche der neu eingeführten analytischen Kategorien sich als hinreichend aussagekräftig erweisen, um sie als Indikatoren in den Auswertungsbogen des Instruments aufzunehmen.

- Hinsichtlich der Differenziertheit der Exposition war eine hochsignifikante Steigerung zu beobachten, daher sollte der bestehende Auswertungsblock um die Kategorien *Einführung der Aktanten*, *Einführung des Handlungsortes*, *Einführung der Handlungszeit* sowie *Einführung laufender Handlungen* erweitert werden.
- Der Zuwachs der weiteren Ereignisstruktur-Markierungen fiel dagegen nicht signifikant aus, daher erscheint ihre Implementierung unter Berücksichtigung der Auswertungsökonomie zweifelhaft. Ohnehin werden im bestehenden Block ‚Satzverbindungen‘ viele dieser Phänomene bereits verzeichnet, so dass ein genauerer Blick auf die verwendeten Konjunktionen und Adverbien bereits aufschlussreich sein kann.
- Die affektiven Markierungen schließlich erwiesen sich in der Stichprobe als stabile Indikatoren und können zu einer Erhöhung der Validität des Verfahrens beitragen. Tabelle 1 zeigt die Entwicklung indizierende Einzelphänomene.

Tabelle 1: Zuwachs indizierende affektive Markierungen

	Phänomen	Mittelwert der zweiten Welle	Entwicklung (in % der 1. Welle)
Affektive Markierungen	a) Indirekte Rede	0,31	163 %
	b) Verbalisierung von Empfindungen	0,51	119 %
	c) Verbalisierung von Wünschen	2,53	116 %
	d) Verbalisierung von inneren Zuständen	0,46	139 %
	e) Iteration als Gestaltungselement	0,22	158 %
	f) Lautmalerei	0,12	200 %
	g) Plötzlichkeit markierende Temporaladverbien	0,28	104 %
	h) expressive Verben	0,87	119 %
	i) evaluierende Adjektive	2,05	122 %

Somit erweisen sich neun von 13 Indikatoren als aussagekräftig, die folgendermaßen in die Auswertung einbezogen werden können: Die Phänomene a) bis g) können als vertiefende Beobachtungen dem bestehenden Block ‚Literarische Elemente‘ zugeordnet werden. Anhand einer größeren Stichprobe gilt es dann faktorenanalytisch zu prüfen, ob sich auch auf empirisch-induktivem Weg ein Zusammenhang zwischen den Variablen nachweisen lässt. Expressive Verben und evaluierende Adjektive dagegen können im Rahmen der Wortschatzuntersuchung gesondert erfasst werden. Damit würde der Wortschatz nicht nur in seiner globalen Differenziertheit, sondern auch spezifisch hinsichtlich der Textsorte ‚Erzählung‘ zugänglich.

3. Fehleranalyse

Im Zuge der vertiefenden Auswertungen wurde neben den in den vorangegangenen Abschnitten vorgestellten Beobachtungen zusätzlich eine Fehleranalyse durchgeführt. An dieser Stelle wird nur das für den Zusammenhang mit der Textkompetenz relevante Ergebnis zur Entwicklung von Orthografie und Grammatik in geraffter Form vorgestellt.⁸

Insgesamt zeigt sich eine ebenfalls erfreuliche Entwicklung: Der Mittelwert des Fehlerquotienten sank von 0,233 auf 0,206. Diese Entwicklung ist signifikant ($p = .002$). Es wiederholt sich allerdings das bereits zuvor beobachtete Phänomen: Die Schülerinnen und Schüler mit den wenigsten Abweichungen in der ersten Welle (1. Quartil) konnten sich zur zweiten nicht verbessern, während dies den weiteren drei Gruppen umso stärker gelang, je mehr Abweichungen sich in den Texten der ersten Welle finden ließen.

Angesichts dieser Parallelität bezüglich der Entwicklung der untersuchten sprachlichen Teilfähigkeiten stellt sich die Frage nach dem generellen Zusammenhang zwischen Textkompetenz und sprachsystematischen Fähigkeiten. In diesem Kontext wurde eine neue Variable berechnet, die sich aus der Differenz von 1 und dem Fehlerquotienten ergibt und somit als Prozentwert der abweichungsfreien Textmenge gelesen werden kann. Eine Korrelation dieser Variablen mit dem Quotienten der affektiven Markierungen zeigt zu beiden Erhebungszeitpunkten einen positiven Zusammenhang, der allerdings schwach und nicht signifikant ausfällt (1. Welle: $r = .17$; 2. Welle: $r = .14$). Somit kann ein systematischer Zusammenhang ausgeschlossen werden, d.h., grammatische und

8 Die untersuchten Phänomene lassen sich den folgenden übergeordneten Kategorien zuordnen: 1) Phonem-Grafem-Korrespondenz, 2) Groß-Kleinschreibung und sonstige Orthografie, 3) Grammatik, 4) Lexikalische und semantische Abweichungen, 5) Interferenzen und Sprachmischungen.

orthografische Korrektheit besitzt keinen Vorhersagewert für textuelle Fähigkeiten im Sinne der Verwendung involvierender Gestaltungselemente.

4. Resümee

Zusammenfassend lässt sich festhalten, dass sich die Schülerinnen und Schüler der vorliegenden Stichprobe in beiden zusätzlich untersuchten sprachlichen Teilfähigkeiten teilweise signifikant verbessern konnten. Es muss jedoch einschränkend festgehalten werden, dass diese positive Entwicklung vor allem für die Lernenden gilt, die zum ersten Erhebungszeitpunkt die tendenziell schwächeren Ergebnisse erzielten. Die jeweils besten 25 Prozent der ersten Welle konnten sich weder in der Erzählfähigkeit noch hinsichtlich Orthografie und Grammatik verbessern, für diese Gruppe waren vielmehr leichte bis mäßige negative Entwicklungen festzustellen. In diesem Zusammenhang lässt sich die Frage, welchem Faktor dieses Phänomen geschuldet ist, nicht abschließend beantworten. In Betracht kommen sowohl personeninterne als auch instrumenten- sowie durchführungsspezifische Kriterien.

Jedoch können die hier angewendeten Indikatoren dazu beitragen, die Validität des Instruments insofern zu erhöhen, als ein breiteres Spektrum sprachlicher Fähigkeiten in den Blick genommen wird. Durch die Einbeziehung textstruktureller Merkmale wird die in den ursprünglichen Auswertungen beobachtete Abhängigkeit der Ergebnisse von der Textmenge reduziert. Daher werden die Merkmale ‚Differenziertheit der Exposition‘ sowie eine Reihe der affektiven Markierungen in das Auswertungsverfahren integriert. Im Anhang dieses Bandes findet sich dazu sowohl der modifizierte Auswertungsbogen als auch ein Manual für die Hand der Lehrkräfte.

Literatur

- Becker, T. (2005). *Kinder lernen erzählen. Zur Entwicklung der narrativen Fähigkeiten von Kindern unter Berücksichtigung der Erzählform* (2., korrigierte Aufl.). Baltmannsweiler: Schneider Verlag Hohengehren.
- Boueke, D., Schüle, F., Büscher, H., Terhorst, E. & Wolf, D. (1995). *Wie Kinder erzählen. Untersuchungen zur Erzähltheorie und zur Entwicklung narrativer Fähigkeiten*. München: Fink.
- Bredel, U. (2001). Ohne Worte – Zum Verhältnis von Grammatik und Textproduktion am Beispiel des Erzählens von Bildergeschichten. *Didaktik Deutsch*, 11, 4–21.

- Dehn, M. (2006). Anforderungen an Aufgaben zum Texteschreiben. *Grundschule*, 7/8, 39–43.
- Ehlich, K. (1983). Text und sprachliches Handeln. Die Entstehung von Texten aus dem Bedürfnis nach Überlieferung. In A. Assmann, J. Assmann & C. Hardtmeier (Hrsg.), *Schrift und Gedächtnis* (S. 24–43). München: Fink.
- Gülich, E. & Raible, W. (1977). *Linguistische Textmodelle*. München: Fink.
- Hayes, J. R. & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale: Erlbaum.
- Knapp, W. (2001). Erzähltheorie und Erzählerwerb. Zur Diskussion neuerer Forschungsergebnisse. *Didaktik Deutsch*, 10, 26–48.
- Labov, W. (1980). Der Niederschlag von Erfahrungen in der Syntax von Erzählungen. In N. Dittmar & B.-O. Rieck (Hrsg.), *William Labov: Sprache im sozialen Kontext* (S. 287–336). Königstein: Athenäum.
- Labov, W. & Waletzky, J. (1973). Erzählanalyse: Mündliche Versionen persönlicher Erfahrung. In J. Ihwe (Hrsg.), *Literaturwissenschaft und Linguistik*, Bd. 2 (S. 78–126). Frankfurt a.M.: Fischer Athenäum.
- Roth, H.-J., Gogolin, I. & Neumann, U. (2007). Abschlussbericht über die italienisch-deutschen, portugiesisch-deutschen und spanisch-deutschen Modellklassen. Hamburg: Arbeitsstelle Interkulturelle Bildung. Verfügbar unter: http://www2.erzwiss.uni-hamburg.de/institute/interkultur/Bericht_2007.pdf [06.05.2008].