



Otfried Börner
Gaby Engel
Bernd Groot-Wilken
(Hrsg.)

Hörverstehen Leseverstehen Sprechen

Diagnose und Förderung von
sprachlichen Kompetenzen
im Englischunterricht
der Primarstufe

WAXMANN

Hörverstehen
Leseverstehen
Sprechen

Otfried Börner, Gaby Engel,
Bernd Groot-Wilken
(Hrsg.)

Hörverstehen Leseverstehen Sprechen

Diagnose und Förderung von
sprachlichen Kompetenzen
im Englischunterricht der Primarstufe



Waxmann 2013
Münster / New York / München / Berlin

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Print-ISBN 978-3-8309-2937-6
E-Book-ISBN 978-3-8309-7937-1 (PDF)

© Waxmann Verlag GmbH, 2013
Postfach 8603, 48046 Münster
Waxmann Publishing Co.
P.O. Box 1318, New York, NY 10028, USA

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Inna Ponomareva, Münster
Titelfoto: © moodboard/fotolia.com
Satz: Stoddart Satz- und Layoutservice, Münster
Druck: Hubert & Co., Göttingen
Illustrationen auf den Seiten 107, 108, 130, 132, 206
und 208-210 von Hartmut Bieber

Die Inhalte der im Print beiliegenden CD befinden sich
im E-Book im Anhang.

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706



Printed in Germany
Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des
Verlages in irgendeiner Form reproduziert oder unter Verwendung
elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Vorwort	7
---------------	---

Eike Thürmann

Sprachtests im Englischunterricht der Grundschule:

Orientierungen auf dem Weg von Behauptungen zu Gewissheiten..... 11

1. Vorbemerkungen.....	11
2. Sprachstandstests und die systemische Evaluation auf der Makro-Ebene	13
3. Sprach(stands)tests auf der Meso-Ebene.....	29
4. Sprach(stands)tests auf der Mikro-Ebene	32
Literatur	38

Gaby Engel und Gisela Ehlers

Hören – Zuhören – Verstehen

Möglichkeiten der Analyse, Diagnose und gezielten

Förderung des Hörverstehens..... 43

1. Wissenschaftliche Erkenntnisse und Grundlagen zum Hörverstehen im fremdsprachlichen Lernprozess	43
2. Schulung des Hörverstehens im Fremdsprachenunterricht	51
3. Überprüfung des Hörverstehens	61
4. Fazit	67
Literatur	68

Otfried Börner und Stefanie Frisch

Förderung und Erhebung des Lesens im Englischunterricht

der Grundschule..... 71

Einleitung: Zur Rolle des Lesens im Englischunterricht der Grundschule.....	71
1. Stand der Forschung: Lesen tut dem Fremdsprachenlernen gut.....	72
2. Der Leseprozess in der Fremdsprache Englisch.....	73
3. Die englische und deutsche Orthografie im Vergleich und Konsequenzen für den Unterricht.....	77
4. Praktische Anregungen zur Förderung der Lesemotivation und der Lesefertigkeiten	79
5. Lesefertigkeiten überprüfen	85
6. Fazit.....	90
Literatur	90

Bärbel Diehr und Markus Kötter

Englisch sprechen lernen:

Kompetenzaufbau im Licht der EVENING-Studie..... 95

1. Alltagstheorien und wissenschaftliche Theorien zum Englischsprechen in der Grundschule.....	95
2. Das Konstrukt „Sprechen“	97
3. Grundsätze der Leistungserhebung für Sprechttests im Grundschulenglisch	104
4. Der Sprechttest in EVENING	105
5. Konsequenzen für den Kompetenzaufbau	115
6. Rückschlüsse aus den Ergebnissen des EVENING-Sprechttests	118
Literatur	119

Bernd Groot-Wilken und Vera Husfeldt

Die Testinstrumente und -verfahren des EVENING-Projekts –

Eine empirische Betrachtungsweise 121

1. Zur Genese des EVENING-Projekts	121
2. Der EVENING-Test – Erhebung des Lernstands nach zwei Jahren	122
3. Ergebnisse.....	133
4. Einsatz des EVENING-Tests in der eigenen Klasse bzw. an der eigenen Schule.....	138
Literatur	140

Jörg-U. Keßler

Gelingsbedingungen für guten Englischunterricht ab der Grundschule..... 141

1. Acht Thesen zum Englischwerb in der Grundschule	142
2. Qualitätskriterien für guten Englischunterricht in der Grundschule	148
3. Bilanz und Ausblick.....	155
Literatur	156

- Anleitungen zur Testdurchführung (Hör- und Leseverstehen, Sprechen)
- Testheft Hör- und Leseverstehen
- Aufnahmeskript Hörverstehenstexte
- Soundtrack Hörverstehenstexte
- Auswertungstabelle
- Sprechttest mit Bildmaterial
- Diagnosebogen zur Auswertung des Sprechttests

Vorwort

*„If we think of our children as plants ...
summativ assessment of the plants
is the process of simply measuring them.
It might be interesting to compare and analyse
measurements but, in themselves,
these do not affect the growth of the plants.
Formative assessment, on the other hand,
is the equivalent of feeding and watering the plants
appropriate to their needs – directly affecting their growth.“¹*

Kinder und auch Jugendliche wachsen heute in einem Europa auf, das durch sprachliche und kulturelle Vielfalt geprägt ist. Mehrsprachigkeit gehört daher zu den wichtigen Schlüsselqualifikationen. Im Zeitalter der Globalisierung ist Englisch internationales Verständigungsmittel, Arbeitssprache und Weltsprache zugleich. Die Grundschule legt die Basis für die Erziehung zur Mehrsprachigkeit und bereitet damit auf eine Lebenswirklichkeit vor, die in Europa durch sprachliche und kulturelle Pluralität geprägt ist. Sie begreift Mehrsprachigkeit als Chance und tut dies, indem sie die Kinder durch die konkrete und bewusste Arbeit mit der Sprache Englisch für diese Pluralität sensibilisiert. Damit verändern sich auch Funktion und Aufgaben des Englischunterrichts in der Sekundarstufe I. Der systematische Erwerb der Fremdsprache in der Grundschule führt zu einer neuen Auseinandersetzung des Kindes mit Sprache. Entwicklungspsychologisch ist dieses Stadium gut für den Erwerb einer weiteren Sprache geeignet.

Inzwischen ist der Englischunterricht fester Bestandteil der meisten Grundschulcurricula im In- und Ausland. In den vergangenen Jahren hat es eine Reihe von Untersuchungen gegeben zu Voraussetzungen und Bedingungen, Unterrichtspraktiken sowie den möglichen Fähigkeiten und Fertigkeiten, die Kinder dieses Alters im Fremdsprachunterrichts der Grundschule erreichen können. Die umfangreiche Studie EVENING (**E**valuation **E**nglish **i**n der **G**rundschule) wurde 2006 vom Ministerium für Schule in Nordrhein-Westfalen in Auftrag gegeben. Deren Ergebnisse wurden in dem 2009 erschienenen Buch „Englisch in der Primarstufe – Chancen und Herausforderungen“² vorgestellt. Darin geht es vor allem um eine Positionsbestimmung des Englischunterrichts in der Grundschule auf nationaler

1 Clarke, S. (2001): *Unlocking Formative Assessment – Practical Strategies for Enhancing Pupils' Learning in the Primary Classroom*. London: Hodder and Stoughton, S. 2.

2 Engel, G. et al. (Hrsg.). Berlin: Cornelsen.

wie internationaler Ebene. Im Fokus stehen dabei auch Unterrichtspraktiken sowie Erfahrungen mit dem frühen Englischunterricht in verschiedenen Ländern, mit den unterschiedlichen Untersuchungen bzw. Ergebnissen sowie deren Auswirkungen auf die Entwicklung von Unterricht und Lehrplänen.

In Ermangelung geeigneter Testmaterialien für diese Lernstufe entwickelte Nordrhein-Westfalen im Rahmen der Studie in Kooperation mit nordrhein-westfälischen Universitäten einen Test mit verschiedenen Aufgabenschwerpunkten für die Kompetenzbereiche Hören, Lesen und Sprechen. Im Zuge der Entwicklung bundesweiter Bildungsstandards und dem damit einhergehenden Paradigmenwechsel hin zur Kompetenzorientierung – d.h. auch zu kompetenzorientierten Lehrplänen – wachsen die Anforderungen an den Umgang mit der Leistungsbewertung. Der nordrhein-westfälische Englischlehrplan für die Grundschule setzt hier klare Maßstäbe. Im Vordergrund steht eindeutig die kriteriengeleitete prozessbegleitende oder formative Leistungsfeststellung, d.h. die Beobachtung des Lernprozesses des einzelnen Kindes, wobei sich die Kriterien an den ausgewiesenen Kompetenzerwartungen orientieren. Dies gilt im gleichen Maße für die punktuelle oder summativ Leistungsfeststellung, die ebenfalls zur Überprüfung des Hör- und Leseverstehens, aber auch des Sprechens dosiert eingesetzt werden kann.

Was aber genau geht eigentlich vor beim Prozess des Hörverstehens, Leseverstehens, Sprechens? Wie lässt sich etwas derart Komplexes valide und effizient erfassen bzw. messen? Welche Parameter müssen beim Erfassen berücksichtigt werden, welche Aufgabenformate stehen zur Verfügung, wie passt das alles in den normalen Unterrichtsalltag? Diesen Fragen widmet sich der hier vorliegende Band, der sich vor allem mit den Möglichkeiten der Erfassung von Fähigkeiten und Fertigkeiten in den genannten Bereichen des Englischunterrichts in der Grundschule befasst. Gleichzeitig werden praktische Vorschläge für die Förderung der genannten rezeptiven und produktiven Kompetenzen gemacht.

Der das gesamte Themenfeld überspannende Beitrag von Eike Thürmann, der am Anfang des Buches steht, gibt eine Übersicht über Arten und Funktionen von Sprachstandstests und Lernstandserhebungen für den Englischunterricht auf der Primarstufe. Unterschieden werden drei Ebenen: (a) die Ebene des Bildungssystems, (b) die der einzelnen Schule und (c) die der individuellen Leistungserfassung der Lernenden. Angesichts der in der Öffentlichkeit anzutreffenden (interessengeleiteten) kontroversen Positionierungen zum Frühbeginn des Fremdsprachenunterrichts, die oft genug durch Behauptungen und Mythen begründet werden, plädiert Thürmann für die Herstellung von Gewissheiten durch eine methodensichere Evaluation, die vor allem der Steuerung von Entwicklungsprozessen auf den unterschiedlichen Ebenen dient.

Der Beitrag von Gaby Engel & Gisela Ehlers befasst sich mit dem Hörverstehen. Dieses nimmt im Fremdsprachenunterricht der Grundschule eine Schlüsselstellung ein. Primär über das Hören knüpfen die Kinder ihre ersten Kontakte mit der Fremdsprache zu Beginn der Lernzeit. Das Hörverstehen gehört zu den rezeptiven Fertigkeiten, was jedoch keinesfalls Passivität auf Seiten des Hörers bedeutet. Der Prozess des Hörverstehens ist – im Gegenteil – ein aktiver und komplexer Prozess, bei dem es um weitaus mehr geht als um die ‚Wahrnehmung‘ und Entschlüsselung akustischer Signale. Die verschiedenen Prozessabläufe, die Voraussetzungen und Gelingensbedingungen für einen erfolgreichen Kompetenzaufbau sowie Möglichkeiten der Schulung und Unterstützung durch entsprechende Techniken und Methoden stehen im Mittelpunkt dieses Beitrags. Das Wissen um diese Zusammenhänge ist eine wesentliche Voraussetzung für die Entwicklung adäquater Materialien sowohl für die Schulung des Hörverstehens durch Lernaufgaben als auch zur validen Kompetenzüberprüfung durch Testaufgaben.

Aus den Ergebnissen der EVENING Studie geht hervor, dass Grundschul Kinder im Englischunterricht gerade im Bereich der mündlichen Kompetenz beachtliche Lernfortschritte machen. Der Beitrag von Bärbel Diehr und Markus Kötter greift diese Mut machende Erkenntnis auf und beleuchtet die Bestandteile und Einflussfaktoren des Sprechvorgangs sowie Möglichkeiten, den Kompetenzaufbau zu fördern. Dabei beziehen die Autoren auch die Chancen ein, die das Testinstrument aus EVENING für die Diagnose und Beratung bereithält.

Im Anschluss an einen Überblick über den Forschungsstand beschreiben Otfried Börner und Stefanie Frisch in ihrem Beitrag die komplexen Vorgänge beim fremdsprachlichen Leseprozess. Daraus ergeben sich praktische Anregungen einerseits zur Lesemotivierung, andererseits Übungen und Aufgaben zur Förderung der Lesekompetenz. Die zentrale Frage dieses Buches nach den verschiedenen Möglichkeiten der Überprüfung der Lesefertigkeiten, auch im Licht der EVENING-Studie, steht am Ende dieses Beitrags. Auch wenn es noch an fundierten wissenschaftlichen Befunden mangelt, plädieren die Autoren ausdrücklich für die Förderung des Lesens im Englischunterricht der Grundschule.

Bernd Groot-Wilken und Vera Husfeldt betrachten in ihrem Beitrag den EVENING-Test unter psychometrischen Aspekten. Dabei stehen die Fragen nach Reliabilität und Validität des Tests in seinen beiden rezeptiven Domänen Hör- und Leseverstehen sowie in der produktiven Domäne Sprechen im Vordergrund. Grundlage dieser Betrachtungen sind Ergebnisse, die in unterschiedlichen Untersuchungen gewonnen wurden und eine Analyse über die in der ersten Phase des Projektes in Nordrhein-Westfalen gewonnenen Daten ermöglichen. Darüber hinaus werden Genese und Entwicklung des Tests in aller Kürze dargestellt.

Der Beitrag von Jörg-U. Keßler analysiert, wie guter Englischunterricht in der Grundschule gelingen kann. Dabei wird die aktuelle Diskussion um den Grundschulenglischunterricht ab Klasse 1 aufgenommen und gängige Vorurteile gegen diese Unterrichtsart werden konstruktiv widerlegt. Ausgehend von acht Thesen zum Englischerwerb in der Grundschule werden in dem Beitrag Qualitätskriterien für guten frühen Englischunterricht entwickelt und begründet. Auf diese Weise wird gezeigt, dass Englischunterricht in der Grundschule in Europa nicht nur etabliert ist und beachtliche Erfolge erzielen kann, sondern dass inzwischen eine umfangreiche wissenschaftliche Basis zu diesem Aspekt des schulischen Fremdsprachenlernens entstanden ist. Auf dieser Basis gilt es die Qualität des Unterrichts zu sichern und weiter auszubauen.

1. Dezember 2012

Gaby Engel, Soest und Dieter Wolff, Essen

Eike Thürmann

Sprachtests im Englischunterricht der Grundschule: Orientierungen auf dem Weg von Behauptungen zu Gewissheiten

1. Vorbemerkungen

Die Verständigung über das frühe schulische Lernen und Lehren von Fremdsprachen ist auf den unterschiedlichen Ebenen von Evaluation (Makro-Ebene = Innovation und Steuerung des Bildungssystems, Meso-Ebene = Schul- und Unterrichtsentwicklung, Mikro-Ebene = individuelle Leistungserfassung im Unterricht) recht heikel. Sie ist geprägt von einer weitgehenden Abwesenheit von Gewissheiten und kontroversen Positionierungen, die in der öffentlichen und an Schule interessierten Diskussion oft durch Behauptungen und Mythen kompensiert wird.

Auch von Pädagogen und Experten werden Schlüsselfragen wie die folgenden widersprüchlich beantwortet:

- Was ist von „Je früher, desto besser“ zu halten?
- Ab welcher Altersstufe ist Fremdsprachenunterricht sinnvoll?
- Bringt das Fremdsprachenlernen in der Grundschule überhaupt nachhaltige Effekte?
- Sollte im frühen Fremdsprachenunterricht (mehr) geschrieben werden?
- Reichen Sprachbad und verständlicher Input aus, und kann dadurch das bewusstmachende Lernen ersetzt werden?
- Sollen Fehler sofort korrigiert werden, damit sie sich nicht einschleifen? Sollen Lernleistungen bewertet werden?
- Soll im Unterricht nur die Zielsprache verwendet werden?
- Ist ein Lehrwerk das beste Mittel für die planvolle Anbahnung von Sprachkompetenzen?
- Werden Kinder mit Migrationshintergrund durch den Fremdsprachenunterricht überfordert, wenn sie Deutsch als Schulsprache noch nicht oder nur unvollkommen beherrschen?
- Müssen Lehrkräfte zu Beginn der Sekundarstufe alles wiederholen, damit der Spracherwerb auf der gesicherter Basis von Aussprache, Wortschatz, Rechtschreibung und Grammatik erfolgen kann?

Wer sich die Mühe macht, im Internet in den einschlägigen Blogs und Foren zum frühen Fremdsprachenunterricht nachzuschauen, wird über das Ausmaß der Mythen, Mutmaßungen und Kontroversen erstaunt sein. Allein in dem Forum, das sich an den Beitrag („Effekt gleich Null“) von Lena Greiner (2009) in

DER SPIEGEL vom 21.01.2009 anschließt, kann das Spektrum der Meinungen aus mehr als 530 Stellungnahmen zu dem Thema „Wie sinnvoll ist früher Fremdsprachenunterricht?“ eingesehen werden.

Problematischer ist, dass die Verständigung über das frühe Lehren und Lernen von Fremdsprachen auch innerhalb des Bildungssystems nicht störungsfrei verläuft, vor allem weil Lehrkräfte für den Fremdsprachenunterricht der Primarstufe und der Sekundarstufe I in der Regel nicht in einen institutionellen Dialog eingebunden sind und folglich wenig von den Zielen, Methoden und Kontextfaktoren des vorausgehenden bzw. folgenden Bildungsabschnitts wissen, manchmal auch nicht wissen wollen. Keßler (2009, S. 58) reduziert in Anlehnung an Börner (2001) und Sommerschuh (2003) übliche wechselseitige Vorbehalte und Mythen an der Schnittstelle zwischen den beiden Schulstufen auf folgende Darstellung:

Grundschullehrer	Sekundarstufenlehrer
„In der Sekundarstufe gibt es nur noch Grammatik und Vokabeln, sonst nichts, da wird nur gepaukt.“	„Mit Spielerei und ein paar Liedchen kann man keine Fremdsprache lernen.“
„Die nehmen überhaupt keine Rücksicht mehr auf die Kinder, die nehmen sie auch nicht mehr richtig ernst.“	„Nach ein paar Wochen sind alle Schüler gleich, egal ob sie vorher Frühenglisch hatten oder nicht.“

In der überwiegend qualitativen empirischen Studie von Böttger (2009a) wurden mithilfe von je 45-minütigen Leitfadenterviews im Zeitraum von Oktober 2007 bis Oktober 2008 67 Lehrkräfte der fünften Klassen aus bayerischen Realschulen und Gymnasien zu den Themen „Übergang auf weiterführende Schulen“ und „Individualisierung/Differenzierung“ befragt. Bezogen auf den Englischunterricht in den 5. Klassen fasst Böttger (2009b, S. 4) die Ergebnisse folgendermaßen zusammen:

- Etwa 61 Prozent der Lehrkräfte der weiterführenden Schulen kannten die Ziele und Inhalte des Englischunterrichts an Grundschulen nicht oder nur sehr eingeschränkt.
- Knapp die Hälfte gab an, in punkto Aussprache, Grammatik und Schrift höhere Erwartungen zu Beginn des 5. Schuljahres an die ehemaligen Grundschüler zu haben. Oft wurde genannt, die Nachhaltigkeit des Englischunterrichts an Grundschulen fehle, er sei nicht ernsthaft genug, viel zu spielerisch.
- Etwa 95 Prozent der Befragten sagten aus, es sei kein Unterschied am Ende der 5. Klasse mehr zu bemerken, ob den Kindern in der Grundschule Englischunterricht erteilt worden sei oder nicht.
- Daraus folgerten etwa 65 Prozent, dass Englisch an Grundschulen überflüssig sei, bezogen den Befund aber nicht auf ihren eigenen Englischunterricht in der 5. Klasse.

Die Diskussion um den Fremdsprachenunterricht im Primarbereich bedarf also auf allen Ebenen der Versachlichung und der datengestützten Argumentation. Sprachtests bzw. Sprachstandserhebungen (*language assessment*) können Orientierungen auf dem Weg von Behauptungen zu Gewissheiten geben und Daten liefern, mit deren Hilfe Missverständnisse und interessengeleitete Polarisierungen überwunden werden. Allerdings werden auf diesem Weg längst nicht alle Kontroversen aufzulösen sein, da die Kontextfaktoren für das frühe Lehren und Lernen von Fremdsprachen vielfältig und selbst innerhalb eines Bundeslandes keinesfalls einheitlich oder wenigstens vergleichbar ausgeprägt sind.

Welche Funktionen Sprachtests bzw. Sprachstandserhebungen für die Weiterentwicklung des frühen Fremdsprachenunterrichts erfüllen können und wie sich der Test EVENING in die inzwischen angewachsene Fülle empirisch-quantitativer Evaluationsvorhaben einordnen lässt, davon handelt der folgende Beitrag.

2. Sprachstandstests und die systemische Evaluation auf der Makro-Ebene

Auf der Makro-Ebene kann im Zusammenhang mit dem frühen Fremdsprachenunterricht im Primarbereich die Feststellung fremdsprachlicher (Teil-) Kompetenzen drei grundlegenden Zielen dienen:

- Sie können zur Überprüfung herangezogen werden, ob unter den gegebenen Bedingungen des Regelschulwesens innovative Unterrichtsangebote zu nachhaltigen und für den weiteren Bildungsweg der Schüler relevanten und vorteilhaften Lernergebnissen führen. In der Regel sollen dabei Sprachstandstests verlässliche Daten für die im politischen bzw. administrativen Raum zu treffenden Entscheidungen liefern, ob zum Beispiel ein Schul- bzw. Modellversuch aus dem Versuchsvorbehalt entlassen und flächendeckend als Regelangebot implementiert werden kann. Auch in dem Fall, dass der Nutzen früher Fremdsprachenangebote angezweifelt wird, wird der empirische Nachweis des Nutzens vermittels Sprachstandserhebungen angemahnt, wie unten mit einem Beispiel belegt wird. Insbesondere in ideologisierten und politisierten öffentlichen Diskursen zur Thematik des frühen Fremdsprachenunterrichts setzen Befürworter und Gegner auf Sprachstandstests zur datengestützten Begründung von Entscheidungen. Sprachstandstests haben also in solchen Evaluationsvorhaben *Begründungsfunktion* in bildungsökonomisch und fachpädagogisch strittigen Entscheidungssituationen.
- Auf der Makro-Ebene spielen Sprachstandstests auch eine wichtige Rolle im Rahmen der wissenschaftlichen Begleitung von innovativen fremdsprachlichen Unterrichtsangeboten. Hier sollen Antworten auf Fragen gefunden werden wie zum Beispiel „Welche Voraussetzungen und welche Unterstützungsleistungen

müssen vom Bildungssystem erbracht werden, damit die Innovation erfolgreich im Regelbetrieb des Bildungssystems weiterentwickelt und optimiert werden kann?“ In diesem Sinne kann Sprachstandstests eine *Steuerungsfunktion* für die Implementation und Weiterentwicklung des Fremdsprachenfrühbeginns zukommen, der von den relevanten politischen Institutionen bereits beschlossen wurde und von den schuladministrativen Institutionen mit entsprechenden Ressourcen (zum Beispiel Lernzeitkontingente, Lehreraus- und Lehrerfortbildung, curriculare Vorgaben zu Bildungsstandards und didaktisch-methodischen Prinzipien, Lehr- und Lernmittel) auszustatten ist.

- Wenn der frühe Fremdsprachenunterricht Bestandteil des Regelschulwesens ist, kann mit zentral gestellten Sprachstandstests am Ende der Primarstufe überprüft werden, ob und in welchem Maße Schulen bzw. Lerngruppen vorgegebene Ziele bzw. curriculare Standards erreichen. Solchen Sprachstandstests, die auf der Makro-Ebene Aussagen über den Zustand des Bildungssystems erlauben, kommt eine *Kontrollfunktion* bzw. mit Blick auf Parlament und Öffentlichkeit eine *Rechtfertigungsfunktion* zu hinsichtlich der eingesetzten finanziellen Mittel.

2.1 Evaluationsstudien mit Begründungsfunktion

Ein Beispiel aus der jüngeren Vergangenheit des Englischunterrichts an Grundschulen in Deutschland mag zur Konkretisierung der (naiven) Hoffnungen dienen, die sich mit dem Einsatz von Sprachstandstests im Rahmen von systemischer Evaluation auf der Makro-Ebene verbinden. In einem Schreiben des Landeselternbeirates eines deutschen Bundeslandes fordert die Vorsitzende das Schulministerium auf, „den Nutzen des Fremdsprachenunterrichts in der Grundschule zu evaluieren. Dies könnte durch einen Vergleich der Sprachkompetenz der jetzigen Siebtklässler (mit Grundschulfremdsprache ab Klasse 1) mit den Leistungen früherer 7. Klassen (ohne Grundschulfremdsprache) an den weiterführenden Schulen erfolgen“ (Landeselternbeirat Baden-Württemberg, 2010).

Als Begründung wird unter anderem formuliert: „Den Landeselternbeirat erreichen aus den weiterführenden Schulen immer wieder Klagen von Lehrkräften über die Fremdsprachenkenntnisse der Grundschüler. Die häufigsten Klagen betreffen das unterschiedliche Niveau der Schüler in der Fremdsprache und eine falsch vermittelte Aussprache. Die Kinder kämen mit einem völlig unterschiedlichen Kenntnisstand an die weiterführenden Schulen. Während manche ihren Fremdsprachenunterricht als ‚Sprachbad‘ spielerisch erlebten, hätten andere schon Vokabeln gelernt und Tests geschrieben.“

Schließlich stellt sich dem Landeselternbeirat die Frage, ob „angesichts des großen Aufwandes für den Fremdsprachenunterricht in der Grundschule der Nutzen nicht zu bescheiden ist und die Unterrichtsstunden für die Fremdsprache nicht in anderen Fächern wie zum Beispiel Ethik besser angelegt wären“.

Ob die Sehnsucht nach Gewissheit bezüglich des Nutzens frühen Fremdsprachenlernens die oben zitierten Forderungen veranlasst hat oder ob die organisierte Elternschaft auf diesem Wege die Interessen der Lehrkräfte für den Fremdsprachenunterricht in der Sekundarstufe zu vertreten hofft, mag dahingestellt sein.

Die Strategie, den „Nutzen“ des Fremdsprachenbeginns in der Primarstufe durch Sprachstandserhebungen nachzuweisen, war und ist in Deutschland und darüber hinaus im europäischen Kontext weit verbreitet. In Deutschland wurden solche summativen Sprachstandstests seit dem Braunschweiger Forschungsprojekt FEU, das in den 1970er Jahren von Doyé & Lüttge (1975, 1977) durchgeführt wurde, vor allem in der Evaluation von Schul- und Modellversuchen mit der Hoffnung eingesetzt, politische Entscheidungsinstanzen über empirisch gewonnene Fakten dazu bewegen zu können, den Englischunterricht zum Regelfach für alle Kinder in der Grundschule zu machen. Später folgten in den Grundstrukturen ähnliche, in Details der Sprachstandstests und Durchführungs- und Auswertungsmodalitäten durchweg differente Evaluationsvorhaben unter anderem in Hamburg (Kahl & Knebler, 1996), in Rheinland-Pfalz (Helfrich, 1995), in Bayern (Andreas & Hillen, 1997, Andreas, 1998), in Hessen („*Merry-Go-Round-Projekt*“ Gladenbach, unter anderem Legutke, 2002). Auch außerhalb Deutschlands sind Schul- und Modellversuche unter Einsatz von Sprachstandstests evaluiert worden, so zum Beispiel in Schottland (Low, Brown, Johnstone & Pirrie, 1995), in Österreich (zum Beispiel „*Lollipop-Projekt*“, Peltzer-Karpf et al., 1996, Hasiba et al., 1997), in der Schweiz (zum Beispiel das Schulprojekt 21 im Kanton Zürich, mit dem der CLIL-Ansatz in der Grundschule erprobt wurde, Stotz & Meuter, 2001). Zieht man die von der Europäischen Kommission beauftragten Übersichtsstudien von Blondin et al. (1998), Edelenbos et al. (2006) heran, in denen die sozial-empirische Forschung unter anderem zu den Ergebnissen des frühen Fremdsprachenunterrichts dokumentiert und ausgewertet wird, steigt die Zahl einschlägiger und umfassender Kompetenzerhebungen auf der Makro-Ebene noch erheblich.

Generalisierbare Gewissheiten über den „Nutzen“ des in den Primarbereich vorverlegten Fremdsprachebeginns lassen sich meines Erachtens aus diesen Studien kaum ziehen,

- weil die evaluierten Schul- und Modellversuche immer auch anfällig für Pygmalion-Effekte sind und bei der Rekrutierung der Versuchsgruppen die Wirklichkeit des Regelschulwesens (zum Beispiel Motivation und Qualifikation der Lehrkräfte, Bildungsehrgeiz und Unterstützungsverhalten der Schülerfamilien, Interesse und Unterstützungsverhalten der Schulleitungen) nicht zuverlässig abgebildet werden kann;
- weil sich die eingesetzten Testinstrumente trotz mancherlei Gemeinsamkeiten (summativ und standardisiert, überwiegend performanzorientiert, hoher Anteil geschlossener Aufgabenformate für Überprüfung von Verstehensleistungen) in vielerlei Hinsicht unterscheiden (zum Beispiel Lernfortschritts- bzw. *attainment* Tests vs. Sprachstands- bzw. *proficiency* Tests, den Probanden bekannte vs.

unbekannte Itemformate, umfassende vs. eingeschränkte Berücksichtigung der kommunikativen Fertigkeiten, zufällige vs. gezielte Erfassung von Sprachmittelbeständen zum Beispiel Wortschatz, Morpho-Syntax, Unterschiede im Design der Sprechtestanteile, unterschiedliche Testdauer, fremde vs. vertraute Testadministratoren, zentrale vs. dezentrale Auswertung);

- weil mit „Primarbereich“ international unterschiedliche Altersgruppen gemeint sind (oft einschließlich der Jahrgangsstufen 5 und 6);
- schließlich weil es für die oben erwähnten Studien keine einheitliche Kalibrierung gibt, mit der der „Nutzen“ des Frühbeginns gemessen werden kann.

Gerade der letzte Punkt ist hoch problematisch. Inzwischen gibt es international einen gewissen Konsens, dass Schüler im Fremdsprachenunterricht in der Primarstufe das Niveau A1 des „Gemeinsamen europäischen Referenzrahmen Sprachen“, GeR (Europarat, 2001) erreichen sollten. Damit lässt sich die elementare Sprachverwendung hinlänglich konkret für die Konstruktion von Sprachstandstests beschreiben. Dennoch sind damit die Schwierigkeiten für die Erfassung des Lernfortschritts im Anfangsstadium des Fremdsprachenunterrichts nicht ausgeräumt, da es für die Stufen auf dem Weg nach A1 keine konsensfähigen Deskriptoren gibt, die als Grundlagen für die Konstruktion von Sprachstandserhebungen dienen können.

Für die Mehrzahl der oben aufgeführten Evaluationsstudien war der GeR noch nicht verfügbar. Um einen möglichen „Nutzen“ des Fremdsprachenfrühbeginns nachzuweisen, wurde überwiegend nicht die Frage gestellt, welche Fähigkeiten Schüler am Ende der Primarstufe erworben haben, sondern wie lange der Vorsprung Bestand hat im Vergleich zu Schülern, die am Anfang der Sekundarstufe mit der ersten Fremdsprache begonnen haben, wie es auch in der oben zitierten Forderung eines Landeselternbeirats der Fall ist. Das sollte dann eigentlich die Frage aufwerfen, wie pädagogisch professionell und pfleglich die weiterführenden Schulen mit den fremdsprachlichen Fähigkeiten und Kenntnissen umgehen, die von den Kindern nachweislich in der Grundschule erworben wurden.

Die Tatsache, dass in der Regel die Vorverlegung des Fremdsprachenunterrichts dem bildungspolitischen Primat und dem erwarteten Bildungswert für die Primarstufe zu verdanken ist und nicht den Messwerten von Leistungsdifferenzen zwischen „Früh-“ und „Spätbeginnern“, muss nicht unbedingt bedauert werden angesichts einer Unterrichtspraxis und -kultur, die sich erst noch entwickeln und verstetigen muss sowie angesichts der mangelnden Konvergenzen bei Instrumenten und Verfahren, mit denen Lernergebnisse bzw. Interimskompetenzen verlässlich gemessen, ausgewertet und gedeutet werden.

2.2 Evaluationsstudien mit Steuerungsfunktion

Unter dieser Überschrift werden Studien angesprochen, die Lernstandserhebungen im Zusammenhang mit der Erkundung der innovativen Praxis des Englischunterrichts in Grundschulen im Sinne eines Regelangebots für alle Schüler einsetzen.

Der in diesem Band thematisierte Test EVENING ist als Sprachstandstest im Zusammenhang einer umfassenden punktuellen Zustandsbeschreibung zur Einführung des obligatorischen Englischunterrichts ab Klasse 3 der Grundschulen in Nordrhein-Westfalen entwickelt und in 88 Klassen an 60 Schulen eingesetzt worden. Diese auf der Makro-Ebene angesiedelte systemische Evaluationsstudie stand unter der zentralen Frage: In welchem Maße gelingt es den Grundschulen des Landes zwei Jahre nach Einführung des verbindlichen Englischunterrichts flächendeckend ein verlässliches und ergebnisorientiertes Sprachlernangebot zu realisieren und im jeweiligen Schulprogramm zu verankern? Dieser Erkundungsauftrag war mit vielerlei Detailspekten ausgestattet, zum Beispiel:

- In welchem zeitlichen Umfang wird Englischunterricht angeboten?
- Über welche sprachlichen und didaktisch-methodischen Qualifikationen verfügen die Lehrkräfte?
- Sind altersangemessene Medien und Lehr- und Lernmittel verfügbar und wie werden sie eingesetzt?
- Welche didaktisch-methodischen Prinzipien bestimmen das Unterrichtsgeschehen?
- Über welche fremdsprachlichen Fähigkeiten und Kenntnisse verfügen die Schüler?
- Wie wirkt sich der Frühbeginn des Englischunterrichts auf das Lernverhalten von Kindern mit Migrationshintergrund aus und welche Konsequenzen sind in Hinblick auf den Förderschwerpunkt Lernen im Gemeinsamen Unterricht (GU) zu ziehen?

So waren vier unterschiedliche Module erforderlich, um Antworten auf solche und weitere Fragen zu ermitteln (nach Groot-Wilken, 2009, S. 125):

Grundgesamtheit aller Grundschulen in NW (N=3461)	
Flächendeckende Schulleiterbefragung (N=2355, entspricht einem Rücklauf von 68%)	
Befragung von Fachlehrkräften (Ca. 430 Lehrkräfte von 250 ausgewählten Schulen)	
Kriteriengestützte Unterrichtsbeobachtung + anschließende strukturierte Lehrerinterviews (Expertenteams in 88 Klassen an 60 Schulen)	
Sprachstandstests (88 Klassen an 60 Schulen, identisch mit den Klassen, in denen Unterricht beobachtet wurde)	
Hör- und Leseverstehenstest (mit ca. 1500 Schülern)	Sprechttests (mit ca. 400 Schülern)

Mit dem Design der Evaluationsstudie wurde ermöglicht, Befunde aus den einzelnen Erhebungsmodulen aufeinander zu beziehen und damit auch zu deuten (in Form von Triangulationen: zum Beispiel Lehrkraftbefragung – Unterrichtsbeobachtung – Sprachstandserhebung). Die Ergebnisse der Evaluationsstudie sind ausführlich in dem Sammelband Engel, Groot-Wilken & Thürmann (2009) dokumentiert. Hier soll es im Wesentlichen nur um den Sprachstandstest gehen, der im Sinne einer summativen Evaluation fremdsprachliche Teilkompetenzen in den Bereichen

- Hörverstehen
- Leseverstehen
- Sprechen (an Gesprächen teilnehmen, zusammenhängend sprechen)

in direkter und (quasi-) authentischer Weise als Performanztest erfasst und in erweiterter Form auch in anderen Kontexten eingesetzt wurde, zum Beispiel in der Schweiz im Kanton Aargau (vgl. Husfeldt & Bader-Lehmann, 2009). Das bedeutet, dass die Verfügbarkeit sprachlicher Mittel (zum Beispiel Wortschatz, Morpho-Syntax) nicht gesondert mit eigenen Testitems ermittelt wurde. Wohl aber ist diese Dimension bei der Testkonstruktion und den Auswertungsverfahren berücksichtigt worden: Einmal durch die Kontrolle und Steuerung der verwendeten Wörter, grammatischen Formen, syntaktischen Strukturen und Redemittel (*chunks*) in den

Verstehenstests sowie durch die stichprobenartige lexikalische und sprachstrukturelle Auswertung der aufgezeichneten Sprechleistungen. Der im Evaluationsvorhaben verwendete Sprachstandstest ist wiederholt pilotiert und in den Items und Durchführungsmodalitäten entsprechend optimiert worden, sodass er den üblichen Gütekriterien entspricht. Die Auswertung der Ergebnisse erfolgte auf der Basis von Klassen/Lerngruppen.

Meines Erachtens hat die Evaluationsstudie EVENING die erwartete Steuerungsfunktion erfüllt, da die Ergebnisse (vgl. Engel, 2009) von den Instanzen der Lehrerbildung rezipiert und umgesetzt sowie bei der Weiterentwicklung curricularer Vorgaben für Nordrhein-Westfalen (Ministerium für Schule und Weiterbildung NRW, 2008) berücksichtigt wurden.

Zwei Erfahrungen aus dem Evaluationsvorhaben EVENING waren ausschlaggebend für die Entscheidung, den Sprachstandstest in der Weise weiterzuentwickeln, dass er mit entsprechenden Anleitungen von Schulen bzw. einzelnen Lehrkräften – auch ohne Beteiligung von Experten – eingesetzt werden kann:

- Die Befragung der Lehrkräfte sowie die Interviews im Anschluss an die beobachtete Unterrichtsstunde zeigten, dass viele Lehrkräfte an Instrumenten interessiert sind, die sich für eine datengestützte Standortbestimmung bezüglich des in ihren Lerngruppen erzielten Sprachentwicklungsstandes eignen, um sich – ausgehend von den erzielten Ergebnissen – allein oder gemeinsam mit Kollegen der eigenen bzw. von befreundeten Schulen – auf Vorhaben zur Unterrichtsentwicklung einzulassen.
- Das in dem Evaluationsvorhaben eingesetzte Instrument erwies sich auch in der Durchführung und in der Auswertung der Ergebnisse als praktikabel, zeit-/ personalökonomisch und robust und fand durchaus Akzeptanz bei Lehrkräften und Schulleitungen.

Damit waren die Grundlagen für eine weitere psychometrische Optimierung gegeben.

Als Beispiel für den Einsatz von Sprachstandstests in wiederholten punktuellen Evaluationsstudien (quasi Längsschnittstudien) sei hier auf die Berichte von Schaer & Bader (2003) und Bader & Schaer (2006) verwiesen. Seit 2001/02 wird an allen Appenzell Innerrhoder Primarschulen ab der 3. Klasse Englisch unterrichtet. Nachdem 2003 bereits nach zwei Jahren der Englischunterricht in Klasse 4 evaluiert wurde, wurde zwei Jahre später in allen sechsten Klassen eine zweite Evaluation durchgeführt. „Ziel der Evaluation war, Stärken und Entwicklungsmöglichkeiten des Projekts aufzuzeigen sowie das Material, den Unterricht, den Kenntnisstand der Schülerinnen und Schüler und den Fortbildungsbedarf der Lehrkräfte zu hinterfragen“ (Schaer & Bader, 2003, S. 3). Der Sprachstand wurde in der vierten Klasse mit folgenden Testkomponenten erhoben:

Testserie rezeptive Fertigkeiten • Rezeptiver Wortschatz • Hörverstehen • Leseverstehen	von den Evaluatorinnen ausgearbeitet, von den Klassenlehrkräften mit gleichen Arbeitsanweisungen von CD durchgeführt und von einem unabhängigen Team korrigiert
Beurteilungsbogen • Hörverstehen und Sprechen	Beurteilung der mündlichen Fertigkeiten jedes einzelnen Lernenden durch die Lehrkraft
Interviews • Sprechtest • Interview	durchgeführt von den Evaluatorinnen

Die nach zwei Jahren Englischunterricht erzielten Ergebnisse waren exzeptionell positiv – auch im Vergleich zu der oben skizzierten Evaluationsstudie EVENING:

„In den Vokabel-, Hör- und Leseverstehenstests haben die Kinder im Durchschnitt 90% der Aufgaben richtig gelöst. Alle Aufgaben wurden zu mehr als 80% richtig gelöst.

Dabei waren die Aufgaben keineswegs zu einfach. Ein Vergleich der Anforderungen in den Tests mit Lehrplänen für die gleiche Stufe in Deutschland erwies, dass die Kinder im Bereich Hör- und Leseverstehen über dem dort formulierten Niveau liegen. Im Bereich Sprechen scheinen sie dem dort geforderten Ansprüchen gut zu entsprechen.“
(Schaer & Bader, 2003, S. 3)

Natürlich lassen sich die eindrucksvollen Ergebnisse auch mit der festgestellten hohen Motivation der Lehrkräfte und Schüler des ersten fremdsprachlichen (Pilot-) Jahrgangs erklären. Wichtiger jedoch für die Erklärung der Leistungsdifferenzen unter anderem zu EVENING ist die Tatsache, dass in Appenzell Innerhoden summative Lernfortschrittstests (*attainment*) eingesetzt wurden, wohingegen die Lernstandserhebung in Nordrhein-Westfalen eher den Charakter eines *proficiency* Tests hatte. Schaer & Bader (2003, S. 14) kommentieren ihre Entscheidung so:

„Ähnlich wie in Schottland, wo Anfang der 90er Jahre der Leistungsstand der Primarschüler in den Fremdsprachenprojekten erhoben wurde, wurden summative Tests entwickelt, die der Situation der Kinder und des Unterrichts gerecht werden, das heißt es wurden nicht Tests verwendet, um ihren allgemeinen Sprachstand (*proficiency*) festzustellen, sondern, was auf der Primarstufe eher angebracht ist, ihren Leistungsstand in Bezug auf das Programm, das sie durchlaufen haben (*linguistic attainment*). Die Testinstrumente sollten den Ansatz, der im Unterricht verwendet wurde, widerspiegeln (...).“

Im Übrigen basierte der Englischunterricht in allen an der Evaluation beteiligten 15 Lerngruppen auf einer vergleichbaren Materialbasis, dem Lehrwerksystem *Here comes Super Bus* (1-3) von Lobo & Subirá (2000) bestehend aus *Pupil's Book*, *Activity Book*, *Class CDs*, Video und *Teacher's Book*. Gleichzeitig erhielten die Lehrkräfte eine CD-ROM mit Bildermaterial aus dem Lehrwerk, Spielen und zusätzlichen Arbeitsblättern und expliziten Pensenvorgaben für die Bearbeitung des Lehrwerks. In anderen Kontexten des frühen Fremdsprachenunterrichts, so zum Beispiel in Nordrhein-Westfalen, lassen (Rahmen-) Lehrpläne große Spielräume für die inhaltliche Unterrichtsgestaltung, und es gibt auch keine zentralen verbindlichen Vorgaben für den Einsatz bestimmter Lehr- und Lernmittel. So lässt sich die Homogenität der guten Testergebnisse, die sich im Übrigen auch in der zweiten Testserie in Klasse 6 fortgesetzt haben, plausibel erklären.

Eine weitere Strategie der Evaluation mit Steuerungsfunktion ist die der wissenschaftlichen Begleitung über einen längeren Zeitraum. Beispielhaft dafür ist die „Wissenschaftlichen Begleitung der Pilotphase Fremdsprache in der Grundschule – Zielsprache Englisch und Zielsprache Französisch (WiBe)“, die ein Team unter der Leitung von Erika Werlen über einen Zeitraum von viereinhalb Jahren in Baden-Württemberg durchgeführt hat. In diesem Bundesland wurde der Fremdsprachenunterricht ab Klasse 1 2001 zunächst an 464 Grundschulen begonnen, im Schuljahr 2003/4 wurde dann das Modell zum Regelangebot an allen 2500 Grundschulen des Landes. Aufgabe des WiBe-Projekts war es, pädagogisch-didaktische Entscheidungsgrundlagen für die Weiterentwicklung des Grundschul-Fremdsprachenunterrichts (Zielsprachen Englisch und Französisch) bereitzustellen. Es handelt sich dabei um eine im Wesentlichen qualitative Longitudinalstudie, die als linguistisch-didaktische Untersuchung vor allem empirisch begründete Ziele des Grundschul-Fremdsprachenunterrichts zu formulieren und didaktische und infrastrukturelle Wege zu diesen Zielen aufzuzeigen beabsichtigt (Werlen et al., 2005). Die Herangehensweise an Sprachentwicklungsstände der Schüler – auch in ihrer Progression über vier bzw. fünf Jahrgangsstufen – unterscheidet sich testmethodisch deutlich von der weitaus überwiegenden Zahl sonstiger Sprachstandserhebungen im Kontext des frühen Fremdsprachenunterrichts. WiBe geht von der damals aktuellen Diskussion um die Ableitung und Festlegung von Bildungsstandards aus und hat zum Ziel, „auf die linguistisch-didaktische Longitudinalstudie abgestützte Bildungsstandards in ihren drei Ebenen: *Leitgedanken*, *Kompetenzen* und *Inhalte, Niveaunkonkretisierungen* und Kompetenzmodelle zu liefern“ (ebd., S. 4). Für die Bildungsstandards des Fremdsprachenunterrichts der Grundschule wird das Primat der interaktiven und sprachlernstrategischen Kompetenzen postuliert – in Abgrenzung zu sonst üblichen Teilkompetenz-Modellen der *communicative skills*: „Diese Kompetenzen rangieren hierarchisch und erwerbsmäßig vor den Teilkompetenzen *Hören, Sprechen, Lesen, Schreiben*“ (ebd., S. 5). Konsequenterweise werden dann auch die Verfahren und Instrumente der Sprachstandsbeschreibung auf dieses Axiom abgestimmt. Vermittels so genannter „Beschreibungen der