

Redaktioneller Beirat:

Priv.-Doz. Dr. Jens Bahns (Kiel)
Prof. Dr. Rupprecht S. Baur (Essen)
Prof. Dr. Wilfried Bruschi (Hamburg)
Prof. Dr. Daniela Caspari (Berlin)
Prof. Dr. Rüdiger Grotjahn (Bochum)
Prof. Dr. Gabriele Kasper (Honolulu)
Prof. Dr. Friederike Klippel (München)
Prof. Dr. Frank G. Königs (Marburg)
Prof. Dr. Claire Kramsch (Berkeley)
Prof. Dr. Michael Legutke (Gießen)
Prof. Dr. Gudula List (Berlin)
Prof. Dr. Manfred Raupach (Kassel)
Priv.-Doz. Dr. Gerhard W. Schnaitmann (Heidelberg)
Prof. Dr. Helmut J. Vollmer (Osnabrück)
Prof. Dr. Günther Zimmermann (Braunschweig)

Anschrift der Redaktion (verantwortlich):

Redaktion ZFF
c/o Prof. Dr. Claudia Riemer
Universität Bielefeld
Fakultät für Linguistik und Literaturwissenschaft
Postfach 10 01 31
D-33501 Bielefeld
Fax: 0521/106-2996
E-Mail: zff@uni-bielefeld.de

Technische Redaktion:

Dipl.-Wirtsch.-Ing. Peter Leitner (Darmstadt)

Erscheinungsweise: Jährlich zwei Hefte, Umfang je Heft ca. 160 Seiten

Bezugsbedingungen: Der Preis für das Jahresabonnement ist im Mitgliedsbeitrag der DGFF enthalten. Für Nichtmitglieder beträgt das Jahresabonnement € 26,00. Einzelhefte kosten € 15,00. Alle Preise zzgl. Versandkosten.

Bestellungen: Abonnement- und Einzelheftbestellungen bei Prögel Pädagogik GmbH, Aboservice, Rosenheimer Straße 145, 81671 München, Fax 0 89/4 50 51-2 98, aboservice@proegel.de

Anzeigenverwaltung (verantwortlich): Renate Kienzler, Anzeigenagentur ConTex, Pater-Kolbe-Straße 3, 71638 Ludwigsburg, Tel. 07141/871-670, Fax -753, www.contex-lb.de. Es gilt die Anzeigenpreisliste Nr. 5, gültig ab 1.1.2005.

ZFF

Zeitschrift für Fremdsprachenforschung

Organ der Deutschen Gesellschaft für Fremdsprachenforschung
herausgegeben von

Eva Burwitz-Melzer (Gießen)
Britta Hufeisen (Darmstadt)
Karin Kleppin (Bochum)
Jürgen Kurtz (Karlsruhe)
Claudia Riemer (Bielefeld)

2006
Band 17
Heft 2

ISSN 0939-7299

ISBN 3-486-92061-8

Verlag: Oldenbourg Schulbuchverlag GmbH, Rosenheimer Straße 145, D-81671 München, www.oldenbourg-bsv.de

© 2006 Zeitschrift für Fremdsprachenforschung

Gesamtherstellung: Schöffel Druck und Verlag GmbH, Kobelweg 12 1/6, 86156 Augsburg

Individuelle Arbeitsgedächtniskapazität und fremdsprachliche Leseverstehensleistungen

Claudia Schmidt¹

Working memory refers to the ability to store and process information simultaneously. Consistent high correlation between working memory capacity and L1 reading performance has been shown, which indicates that working memory capacity is a dominant source of individual differences in reading ability. In L2 reading research only very few studies examine the sensitivity of working memory capacity to differences in L2 reading skills. This study was conducted with 16 advanced Japanese L2 German learners. The test battery included measures of reading comprehension in L2 and working memory in L1 and L2. Strong correlations between working memory capacity and L2 reading skills were found. The results will be discussed within the framework of current criticism on measuring working memory, whereby the application of strategies will be of great importance.

1. Einleitung

1.1 Arbeitsgedächtnis und fremdsprachliches Leseverstehen

Das Arbeitsgedächtnis ist eine Instanz, die für das zeitlich begrenzte Behalten und die kurzfristige Manipulation von Informationen zuständig ist. Sie spielt eine zentrale Rolle bei komplexen kognitiven Aktivitäten wie Sprachverstehen, Lernen und Problemlösen (vgl. u.a. Baddeley 1997; Just & Carpenter 1992). Beim Lesen und Verstehen von Texten werden die eingehenden graphischen Informationen dekodiert, geordnet und durch Einbeziehen des Vorwissens aus dem Langzeitgedächtnis in die bereits verarbeiteten Textinformationen integriert. An all diesen Prozessen ist das Arbeitsgedächtnis beteiligt. Die Effizienz, mit der die parallel ablaufenden Prozesse und die kurzzeitigen Bearbeitung und Speicherung von Informationen ablaufen, wird als Arbeitsgedächtniskapazität bezeichnet. Die Ressourcen des Arbeitsgedächtnisses sind begrenzt.

Ergebnisse aus der L1²-Leseforschung belegen, dass LeserInnen über sehr unterschiedliche Arbeitsgedächtniskapazitäten verfügen (Daneman & Carpenter 1980, 1983; Daneman & Green 1986). Die Leistung des Arbeitsgedächtnisses

1 Korrespondenzadresse: Dr. Claudia Schmidt, Albert-Ludwigs-Universität Freiburg, Deutsches Seminar I, 79085 Freiburg, E-Mail: claudia.schmidt@germanistik.uni-freiburg.de.

2 Ich verwende im Folgenden die Abkürzungen L1 für Erst- bzw. Muttersprache, L2 für Fremdsprache.

Informationen und Zuschriften

Informationen über die DGFF	301
Eintrittserklärung	303
Bielefelder Bibliographie Deutsch als Zweitsprache (BiDaZ)	305
Eingegangene Bücher	307
Hinweise zur Manuskripterstellung	315

wird hier als wichtigster Faktor angesehen, um zwischen guten und schlechten LeserInnen zu unterscheiden. Zahlreiche empirische Studien belegen signifikante Korrelationen zwischen Arbeitsgedächtniskapazität und unterschiedlichen Leistungen beim Leseverstehen, wie z.B. dem Erinnern von Bezugswörtern für Pronomen über größere Intervalle und dem Entdecken inhaltlicher Inkonsistenzen aufeinander folgender Sätzen (vgl. hierzu die Metaanalyse von 77 Studien von Daneman & Merikle 1996).

In der L2-Leseforschung werden die Funktionen des Arbeitsgedächtnisses seit den 1990er Jahren verstärkt thematisiert und Ergebnisse der Arbeitsgedächtnisforschung in die Beschreibung von Leseprozessen einbezogen (vgl. u.a. Brown 1998; Ehlers 1998, 1999; Lutjeharms 2006). Unter diesem Aspekt wird der Fähigkeit der automatischen Verarbeitung besondere Bedeutung zugemessen. Um die begrenzte Gedächtniskapazität für den Aufbau der semantischen Repräsentation voll nutzen zu können, müssen die Prozesse der unteren Verarbeitungsebene wie visuelle Wahrnehmung, Worterkennung und syntaktische Analyse automatisiert ablaufen. Gute LeserInnen benötigen keine Aufmerksamkeit für die Dekodierung, d.h. Fremdsprachenlernende müssen vor allem Sprachkenntnisse erwerben und diese automatisieren, um zu erfolgreichen L2-LeserInnen zu werden (vgl. Lutjeharms 2006: 146ff.).

Aktuell werden die Leistungen des Arbeitsgedächtnisses als Erklärungsansatz für individuell erreichbare L2-Lesekompetenzen diskutiert (vgl. Koda 2005: 198ff.). Enthalten ist darin die Frage nach dem Transfer von L1-Fertigkeiten auf die L2-Leseprofizienz (vgl. Ehlers 1998: 172f.). Die Ergebnisse der wenigen bisher vorliegenden Untersuchungen legen den Schluss nahe, dass eine größere Arbeitsgedächtniskapazität zu besseren Leseverstehensleistungen führt und damit als Prädiktor für individuell erreichbare L2-Lesefähigkeiten angesehen werden kann (Geva & Ryan 1993; Harrington & Sawyer 1992; Myake & Friedmann 1998; Walter 2004). Problematisch ist hierbei insbesondere die kaum hinterfragte Größe Arbeitsgedächtniskapazität. Was genau die (individuelle) Kapazität des Arbeitsgedächtnisses ausmacht, ist teilweise noch ungenügend geklärt.

1.2 Kapazität des Arbeitsgedächtnisses

Kapazitätsbestimmend ist das Zusammenwirken erhaltender, verarbeitender und regulierender Prozesse. Höchstens sieben Informationseinheiten (*chunks*) können gleichzeitig verarbeitet werden, wobei sich aber durch unterschiedliche Strategien wie z.B. semantisches Kodieren die Größe der Einheiten beeinflussen und so die Behaltensleistung erhöhen lässt. Dabei ist das Langzeitgedächtnis beteiligt (vgl. u.a. Cowan 2000; McNamara & Scott 2001; Vos, Gunter, Schriefers & Friederici 2001). Zur Kapazitätsbeschränkung tragen unterschiedliche Faktoren

bei: die Mengenbeschränkung für die zeitlich befristete Speicherung von Informationen, die zeitliche Begrenzung für die Verfügbarkeit dieser Informationen sowie die Effizienz bei der Ausführung kognitiver Operationen (vgl. Koch & Schweizer, 2001: 4f.). Weiterhin spielt die ebenfalls limitierte Aufmerksamkeitssteuerung eine wichtige Rolle (vgl. Buntig, Conway & Heitz 2004; Cowan 2000).

Die Gewichtung der unterschiedlichen Faktoren variiert je nach Arbeitsgedächtniskonzeption, auf deren theoretischer Basis die Kapazität definiert wird. Am einflussreichsten in der Sprachverstehensforschung ist das von Alan Baddeley und MitarbeiterInnen in den 1970er Jahren entwickelte Modell des Arbeitsgedächtnisses (vgl. Baddeley & Hitch 1974; Baddeley 1990)³. Die Annahmen basieren auf neurophysiologischen Daten sowie unterschiedlichen Tests, die Aufgaben zum Behalten und Verarbeiten von Informationen kombinieren (vgl. Baddeley 1992: 556ff.). Das Modell besteht in seiner ursprünglichen Version aus drei Hauptkomponenten: der Zentralen Exekutive (*central executive*) und den beiden sog. Hilfssystemen (*slave systems*), nämlich dem visuell räumlichen Skizzenblock (*sketchpad*)⁴ sowie der Phonologischen Schleife (*phonological loop*). Alle drei Systeme sind in ihrer Kapazität begrenzt. Für das Leseverstehen ist zum einen die Zentrale Exekutive und ihre Leistungsbegrenzung von Bedeutung, da sie für die Auswahl von Strategien und die Planung der arbeitsgedächtnisbezogenen Verarbeitung zuständig ist. Zum anderen spielt die Phonologische Schleife, die verbal kodierte Informationen eine gewisse Zeit aufbewahrt, eine wichtige Rolle. Hier findet die Umwandlung geschriebener Worte in eine phonologische Form statt, so dass die Informationen in den phonologischen Kurzzeitspeicher aufgenommen werden können. Die lautlichen Repräsentationen, die innerhalb von ein bis zwei Sekunden zerfallen, werden durch Rezirkulation (*rehearsal*) aufgefrischt und können so erhalten bleiben. In der aktuellen Version (vgl. Baddeley 2000a, 2000b) wurde das Modell um eine weitere Komponente, den sogenannten *episodic buffer* ergänzt. Er dient dazu, die Verbindung zwischen Langzeitgedächtnis und den Subsystemen herzustellen. Die verarbeiteten Informationen werden für begrenzte Zeit in einem multimodalen Kode gespeichert. Der Rückgriff auf die Informationen erfolgt bewusst und wird kontrolliert von der Zentralen Exekutive. Diese beeinflusst den zu speichernden Inhalt, indem sie die Aufmerksamkeit auf unterschiedliche Informationsquellen richtet, also etwa auf Komponenten des Arbeitsgedächtnisses oder auf das Langzeitgedächtnis (vgl. Baddeley 2000a: 420-423).

³ Einen kritischen Überblick geben die Sammelbände von Andrade (2001) und Miyake & Shah (1999).

⁴ Der visuell-räumliche Skizzenblock ist für die kurzzeitige Bearbeitung und Speicherung von visuellen bzw. räumlichen Informationen zuständig. Er ist für das Leseverstehen nur dann relevant, wenn Texte Illustrationen und Grafiken enthalten.

Im Gegensatz zu Baddeleys Modellvorstellung geht die häufig in der Leseforschung berücksichtigte Arbeitsgedächtniskonzeption von Daneman & Carpenter (1980, 1983) von einem einheitlichen, nur für Sprachverstehen zuständigen Arbeitsgedächtnis aus, das Verarbeitungs- und Speicherfunktion leistet⁵. Eine beschränkte Ressource dient beiden Funktionen, wobei, ähnlich wie bei Baddeley, Verarbeitung und Speicherung um begrenzte Kapazitäten konkurrieren. Individuelle Unterschiede in der Arbeitsgedächtnisleistung ergeben sich entweder durch Unterschiede in der verfügbaren Kapazität oder der Effizienz, diese Kapazität zu nutzen. Zur Messung der Arbeitsgedächtniskapazität wurden von Daneman & Carpenter verschiedene Tests entwickelt, die beide Kapazitäten beanspruchen sollen (vgl. Daneman & Carpenter 1980; Daneman & Green 1986). Als Standardtest für das Sprachverstehen hat sich der Lesespannentest (*Reading Span Test*) etabliert, der das Behalten von Informationen bei gleichzeitigem Verarbeiten misst (vgl. Daneman & Carpenter 1980, 1983).

Beschreibungen von L2-Leseprozessen und -Leseverstehensleistungen beziehen sich, mehr oder weniger explizit, sowohl auf das Arbeitsgedächtnismodell von Baddeley als auch auf die Konzeption von Daneman & Carpenter. Der Schwerpunkt liegt hierbei auf dem **Aspekt der Konkurrenz von Verarbeiten und Speichern um begrenzte Kapazitäten**. Weiterhin werden inzwischen Forschungsergebnisse zu Leistungen der Phonologischen Schleife aufgegriffen. Da diese Arbeitsgedächtniskomponente, die nach Baddeley (1992: 558) für das Verarbeiten und Behalten von unbekanntem und komplexem Wortmaterial benötigt wird, auf einem verbalen Kode beruht, ist für den Aufbau der L2-Lesekompetenz der Erwerb der fremdsprachigen Aussprache wichtig (vgl. Lutjeharms 2006: 149). Offen ist, ob neben der in der aktuellen L2-Leseforschung betonten Fähigkeit der automatisierten, ohne Kapazitätsbelastung ablaufenden Dekodierung, weitere Aspekte des Arbeitsgedächtnisses Einfluss auf die L2-Lesekompetenz haben. Von Ehlers (1998: 188) wird die kognitive Flexibilität, mit der die LeserInnen die beim Lesen sich fortwährend ändernden Verstehensaufgaben bewältigen und kontext- und aufgabenbezogene Strategien auswählen, als wichtige Einflussgröße genannt. Leseleistungen wären damit, bezogen auf Modell von Baddeley, vor allem von der individuellen Kapazität der Zentralen Exekutive abhängig, da sie für die Steuerung und Kontrolle der am Leseverstehen beteiligten Prozesse zuständig ist.

⁵ Just & Carpenter (1992: 123) beschreiben das Arbeitsgedächtnis als Teil der Zentralen Exekutive, der beim Sprachverstehen involviert ist.

2. Fragestellung

Die bisher vorliegenden Daten bestätigen den Zusammenhang von individuellen Arbeitsgedächtniskapazitäten und L2-Leseverstehensleistungen, lassen aber viele Fragen offen, und zwar nicht zuletzt aufgrund der ungenügend kontrollierten Testmaterialien. Dies betrifft insbesondere die Konzeption des Lesespannentests, der von den meisten Untersuchungen zur Bestimmung der individuellen Arbeitsgedächtnisleistung eingesetzt wird. In der Standardversion (vgl. Daneman & Carpenter 1980) werden laut einer immer größer werdende Zahl von semantisch unverbundenen Sätzen gelesen (Sets von zwei bis maximal fünf Sätzen), wobei das jeweils letzte Wort eines Satzes zu merken ist. Nach jedem Set sind die Wörter in beliebiger Reihenfolge wiederzugeben. Die Leistung wird als Lesespanne entweder auf der Basis der Gesamtzahl der richtig wiedergegebenen Wörter oder der höchsten erfolgreich bewältigten Set-Größe bestimmt. In den Studien sind zumeist weder die verwendeten Testsätze in Bezug auf ihre semantische und syntaktische Komplexität, noch die zu behaltenden Zielwörter im Hinblick auf Konkretheit, Wortlänge und Frequenz kontrolliert.

Weiterhin gibt es widersprüchliche Ergebnisse zu der Frage, ob die Kapazitätsunterschiede unabhängig von der L2-Kompetenz sind. In einer der frühesten Untersuchungen zu dieser Problemstellung belegen die Daten signifikante Korrelationen zwischen L1- und L2-Lesespannenwerten, woraus die These von der Sprachenunabhängigkeit der Arbeitsgedächtniskapazität abgeleitet wurde (Osaka & Osaka, 1992). Die für die L1-Lesefähigkeit verantwortliche Arbeitsgedächtniskapazität wäre damit übertragbar auf die L2-Leseverstehensleistungen. Allerdings wurde die L2-Kompetenz der getesteten erwachsenen japanischen LernerInnen des Englischen als nahezu bilingual eingestuft. In einer ähnlich angelegten Studie von Harrington & Sawyer (1992) mit gleicher Lernergruppe, allerdings geringerer Englischkompetenz, ergab sich kein nachweisbarer Zusammenhang zwischen japanischer und englischer Lesespanne. Meine eigenen Daten zu Arbeitsgedächtnisleistungen erwachsener japanischer Deutsch-als-Fremdsprache-LernerInnen belegen eine signifikante Korrelation zwischen japanischer und deutscher Lesespanne auf Referenzniveau C1, während auf B1-Niveau keine Korrelationen auftreten (vgl. Schmidt 2000: 94, 2003). Ein **Einfluss der Arbeitsgedächtnisleistung auf die L2-Leseperformanz erst bei hoher Sprachkompetenz lässt auf einen erst bei genügender Automatisierung der unteren Verarbeitungsebenen einsetzenden Transfer von L1-Fertigkeiten schließen**.

Zu klären ist vor allem, ob die individuelle Arbeitsgedächtniskapazität auf der Basis eines L1- oder L2-Lesespannentests zu bestimmen ist. In der Studie von Harrington & Sawyer (1992) werden signifikante Korrelation zwischen L2-Lesespanne und L2-Leseverstehensleistungen von den Autoren als Beleg für den

Einfluss der individuellen Arbeitsgedächtnisleistung auf L2-Leseverstehensleistungen interpretiert. Das ist insofern problematisch, als der Lesespannentest nicht völlig unabhängig von der Lesefähigkeit ist, damit also als Funktion der L2-Kompetenz gelten kann (vgl. Koda 2005: 200). Auch in einer aktuellen Studie von Walter (2004), in der LernerInnen (L2 Englisch und L1 Französisch) mit unterschiedlichen L2-Niveaus getestet werden, erfolgen die Korrelationsberechnungen auf der Basis der L2-Lesespannenwerte. In der häufig zitierten Studie von Geva & Ryan (1993), in der Daten zu den Lesefähigkeiten und den Arbeitsgedächtniskapazitäten von 73 Kindern an einer bilingualen Tagesschule (Englisch - Hebräisch) erhoben wurden, korrelieren die Lesefähigkeiten sowohl mit den L2- als auch mit den L1-Gedächtnisleistungen⁶. Die Autorinnen ziehen daraus den Schluss, dass eine gemeinsame zugrundeliegende Gedächtniskonstruktion angenommen werden könne und beide Leistungen gleichwertige Prädiktoren der L2-Leseverstehensleistungen seien (vgl. ebd.: 35). Zu ähnlichen Interpretationen kommen Myake & Friedmann (1998) in ihrer Untersuchung fortgeschrittener japanischer LernerInnen des Englischen.

Weitgehend unklar in diesem Zusammenhang ist, mit welchen spezifischen Lesefertigkeiten die Differenzen in den Arbeitsgedächtniskapazitäten zusammenhängen. Tests zur Prüfung der Leseverstehensleistungen wie z.B. TOEFL und *Cloze-Tests* (Geva & Ryan 1993; Harrington & Sawyer 1992) sind wenig aussagekräftig. In der Studie von Walter (2004) werden Textzusammenfassungen als Lückentexte (*summary completion*) sowie Aufgaben zum Erinnern von Bezugswörtern für Pronomina eingesetzt, bei Myake & Friedman (1998) wird das Verstehen syntaktisch komplexer Sätze geprüft. Es fehlen Aufgabenstellungen, die weitere unterschiedliche Teilleistungen des Leseverstehens mit entsprechend unterschiedlichen Belastungen des Arbeitsgedächtnisses testen.

Berücksichtigt werden muss hierbei auch das bisher kaum problematisierte Verhältnis zwischen den spezifischen aufgabenabhängigen Teilfertigkeiten des jeweils verwendeten Leseverstehentests und den beim Lesespannentest eingesetzten Teilfertigkeiten. Der Lesespannentest erfordert Verstehen und Behalten von verarbeitungseigenen Informationen sowie von Zusatzinformationen, außerdem das Steuern und Kontrollieren der unterschiedlichen Prozesse. **Aktuelle Arbeiten zum Lesespannentest zeigen auf, dass die individuellen Unterschiede zu einem großen Teil auf den mehr oder weniger effizienten Einsatz von Strategien zurückzuführen sind** (vgl. Herrmann 2003; McNamara & Scott 2001; Turley-Ames & Whithfield 2003). In der Studie von McNamara & Scott (2001) zu Lei-

⁶ Die Messung der Arbeitsgedächtnisleistung erfolgte zusätzlich zu dem an Daneman & Carpenter (1980) angelehnten Lesespannentest auf der Basis des *English Opposites Working Memory Test*, bei dem zu lautlich gegebenen Adjektiven die Gegenbegriffe genannt und behalten werden müssen (vgl. Geva & Ryan 1993: 16f.).

stungen in einem L1-Lesespannentest wird nicht nur nachgewiesen, dass Versuchspersonen mit hoher Lesespanne bessere Strategien entwickeln und diese auch effektiver einsetzen, sondern auch, dass ProbandInnen mit schlechten Lesespannenwerten durch Strategietraining das Ergebnis verbessern konnten. **Arbeitsgedächtniskapazität wird damit als eine veränderbare Größe und nicht als ein stabiles Konstrukt** aus z.B. Aktivationsressourcen und Aufmerksamkeitsspannen definiert (vgl. ebd.: 10). Die aufgeführten Kritikpunkte und Problemstellungen bilden die Grundlage für die folgenden untersuchungsleitenden Fragen:

1. Welcher Art sind die Beziehungen zwischen L2-Lesespanne und L2-Leseverstehensleistungen?
2. Welcher Art sind die Beziehungen zwischen L1-Lesespanne und L2-Leseverstehensleistungen?

3. Methodik

3.1 Versuchspersonen

An der Studie nahmen 16 erwachsene japanische Deutsch-als-Fremdsprache-LernerInnen teil (11 weiblich, 5 männlich; Altersspannweite 27-48, Mittelwert 35,31). Sie studierten nach Bestehen der DSH (Deutsche Sprachprüfung für den Hochschulzugang ausländischer Studienbewerber) in Deutschland. Die TeilnehmerInnen hatten mit durchschnittlich 19 Jahren an ihrer Heimatuniversität begonnen, Deutsch zu lernen. Englisch war die erste Fremdsprache, der Lernbeginn war hier bei 12 Jahren. Die Dauer des Aufenthaltes in Deutschland lag zwischen 1 und 10 Jahren, wobei die meisten LernerInnen sich immer wieder längere Zeit in Japan aufgehalten hatten. Die Unterrichtsjahre betragen zwischen 2 und 8 Jahren. Alle gaben an, sowohl mit Deutschen als auch Japanern Kontakt zu haben und regelmäßig in beiden Sprachen zu lesen und zu schreiben. Die Versuchspersonen waren Studierende geisteswissenschaftlicher Fächer. Die Selbsteinschätzung der Sprachkompetenz im Deutschen (auf der Skala 1-6) lag sehr hoch: Lesen 5,06 (4-6; SD=0.54), Schreiben 4,47 (3-5,5; SD=0.76) und Sprechen 4,78 (4-6; SD=0.63). Sie waren auf Referenzniveau C1 eingestuft worden.

Zur besseren Interpretation der spezifischen Leistungen der LernerInnen wurden Vergleichsdaten von deutschen MuttersprachlerInnen gewonnen. Die deutsche Kontrollgruppe setzte sich aus 16 Studierenden bzw. AbsolventInnen geisteswissenschaftlicher Fächer (11 weiblich, 5 männlich) mit einer der L2-Gruppe entsprechenden Altersstruktur zusammen (Altersspannweite 20-50; Mittelwert 28,31). Alle Versuchspersonen nahmen freiwillig und unentgeltlich an den Tests teil.

3.2 Testmaterial

Die Lesespannentests zur Messung der individuellen Arbeitsgedächtniskapazität wurden wie in vergleichbaren Studien in Anlehnung an Daneman & Carpenter (1980) entwickelt (vgl. dazu auch Schmidt 2003). Ein wesentlicher Unterschied lag in der Ergänzung des Tests durch sechs willkürlich verteilte Verständnisfragen, mit denen das Verstehen kontrolliert wurde. Die deutsche Test-Version besteht aus 70 Sätzen mit je 11 bis 12 Wörtern bzw. 21 bis 24 Silben Länge mit einfachen syntaktischen Strukturen. Die Zielitems umfassen 70 hochfrequente Konkreta (Frequenz nach CELEX zwischen 100 und 1000) von zwei Silben (62 Items) oder drei Silben (8 Items) Länge. Das kritische Item steht zumeist als Akkusativobjekt am Satzende, ist aber nicht das letzte Wort (siehe Anhang 1). Bei der Auswahl der Zielitems war darauf geachtet worden, dass keine Überschneidungen mit der japanischen Test-Version auftraten. Der japanische Lesespannen-Test wurde von Herrmann (2003) übernommen⁷. Bei den japanischen Zielitems handelt es sich ebenfalls um konkrete Nomen mit zwei oder drei Silben. Die syntaktische Struktur der Sätze, Frequenz und Stellung der Zielitems beider Test-Versionen sind aufeinander abgestimmt. Beide Test-Versionen wurden zunächst mit MuttersprachlerInnen überprüft, der deutsche Lesespannentest zusätzlich mit einer Lernergruppe auf C1-Niveau.

Zur Messung der Leseverstehensleistungen wurden zwei verschiedene Tests mit unterschiedlichen kognitiven Anforderungen eingesetzt, die jeweils unterschiedliche Teilleistungen des Leseverstehens erfassen sollten. Im ersten Test mussten die Versuchspersonen Wortbedeutungen aus dem Kontext erschließen, was einen sehr hohen kognitiven Aufwand erfordert. Die Informationen des Textes müssen verstanden, behalten und dann ausgewertet werden. Dies erfordert neben Verarbeitung und Speicherung das Gewichten und Steuern von Informationen. Hierzu wurden in Anlehnung an die Untersuchung von Daneman & Green (1986) 15 narrative Texte mit jeweils einem aus dem Russischen abgeleiteten Pseudowort erstellt (siehe Textbeispiel Anhang 3, vgl. dazu auch Schmidt 2004). Die Texte waren so geschrieben, dass erstens die Wortbedeutung eindeutig aus dem Kontext erschlossen werden konnte, zweitens genügend kontextuelle Hinweise zur Erschließung der Bedeutung enthalten waren und drittens nur durch Hinzuziehung der globalen textuellen Hinweise das erfolgreiche Erschließen möglich war. Die Texte hatten eine durchschnittliche Länge von 200,87 Wörtern. Sowohl die Assoziationsmöglichkeiten der Pseudowörter als auch die eindeutige Erschließungsmöglichkeit waren mit zwei unterschiedlichen deutschen Kontrollgruppen getestet worden (Reliabilitätsquotient nach Cronbachs Alpha 0.81 bei

⁷ Die Test-Versionen sind von Nina Herrmann und mir zusammen entwickelt worden, so dass eine optimale Abstimmung möglich war.

N=16).

Die zweite Aufgabenstellung bestand darin, kurze, aber semantisch und syntaktisch komplexe Texte unmittelbar nach dem Lesen mündlich zu reproduzieren. Im Vergleich zum ersten Test handelt es sich hierbei um eine kognitiv weniger aufwändige Aufgabe, die nur Verstehen und Behalten erfordert, aber aufgrund der Informationsdichte und syntaktischen Komplexität hohe Anforderungen an Verarbeitung und Speicherung stellt. Verwendet wurde ein standardisierter Test aus der empirischen Sprachverstehensforschung, der auf den sogenannten Wechsler-Texten basiert (vgl. Böcher 1963; Tewes 1991). Es handelte sich um zwei Zeitungsberichte, die aus jeweils drei syntaktisch komplexen Sätzen bestanden und in je 23 Inhaltseinheiten aufgeteilt waren (siehe Anhang 2). Gemessen wurde die Anzahl der Inhaltseinheiten, die nach dem Lesen wiedergegeben wurden (Reliabilitätsquotient nach Cronbachs Alpha 0.96 bei N=2000).

Da fehlender Wortschatz gerade bei fortgeschrittenen LernerInnen zumeist als größtes Problem beim L2-Lesen gilt, fand zusätzlich die Überprüfung des Wortschatzes statt. Er wurde anhand eines standardisierten Wortschatztests ermittelt, der ebenfalls aus den Wechsler-Testmaterialien stammt (vgl. Tewes, 1991: 56ff.). Bei diesem Test sind 32 nach Schwierigkeitsgrad geordnete Wörter zu definieren, beginnend mit Apfel und endend mit *Geoid* (Reliabilitätsquotient nach Cronbachs Alpha 0.90 bei N=2000).

3.3 Vorgehen

Die Versuchspersonen legten die Tests in jeweils drei Einzelsitzungen ab. In der ersten Sitzung, die durchschnittlich 30 Minuten dauerte, wurden der Wortschatztest sowie die Textreproduktionsaufgabe durchgeführt. Die japanischen LernerInnen füllten zusätzlich einen Fragebogen zur Lernsozialisation und zum Sprachgebrauch aus. Die wiederzugebenden Texte der Reproduktionsaufgabe sollten still in selbst bestimmtem Tempo gelesen werden. Die gelesenen Zeilen waren abzudecken, so dass ein nochmaliges Lesen verhindert wurde. Sofort nach dem Lesen des jeweiligen Textes musste der Inhalt mündlich wiedergegeben werden. Auch beim Wortschatztest lasen die ProbandInnen die aufgelisteten Wörter leise für sich und gaben nach jedem Wort mündlich eine Definition⁸.

In der zweiten Sitzung, die zumeist eine Woche später stattfand, legten die Versuchspersonen den Test zur Erschließung unbekannter Wörter aus dem Kontext ab. Die Dauer betrug ungefähr 60 Minuten, wobei nach acht Geschichten

⁸ Die beiden Tests wurden von mir leicht geändert; in der Originalversion werden die Texte bzw. Wörter von den Versuchsleitern/-innen vorgelesen.

eine kurze Pause eingelegt wurde. Die Versuchspersonen mussten schriftlich eine möglichst genaue Definition des zu erschließenden Wortes geben, wobei auch hier die bereits gelesenen Zeilen mit einem Blatt Papier abzudecken waren. Die Definitionen waren auf Deutsch zu formulieren, es wurde aber darauf hingewiesen, dass Rechtschreibung und Grammatik keine Rolle spielen sollten. In einem Probendurchlauf war das Verstehen der Aufgabenstellung überprüft worden. Die Korrektheit der insgesamt 480 Definitionen wurde von zwei unabhängigen Schätzern auf der Skala 0 (völlig inkorrekt) bis 4 (völlig korrekt) eingeordnet (Reliabilitätsquotient = 0.99).

Als letzter Test wurde die Messung der Lesespannen durchgeführt. Die japanischen Versuchspersonen legten zwischen der jeweils ungefähr 20 Minuten dauernden japanischen und deutschen Test-Version eine längere Pause ein. Den ProbandInnen wurden die 70 randomisierten Sätze jeder Test-Version, aufgeteilt in Blöcken von jeweils fünf Sets mit je zwei, drei, vier und fünf Sätzen visuell präsentiert (siehe Anlage 1). Jeder Satz war einzeln, quer und linksbündig auf ein weißes DIN A4-Blatt gedruckt. Die Versuchspersonen erhielten die Anweisungen, die Sätze möglichst schnell, aber in selbst bestimmtem Lesetempo laut zu lesen und nach jedem Set die unterstrichenen Zielitems wiederzugeben. Das jeweils letzte Zielitem durfte aber nicht zuerst genannt werden, da der *recency*-Effekt vermieden werden sollte: Bei freier unmittelbarer Wiedergabe mehrerer unrelatierter Items wird das zuletzt präsentierte Item immer besser wiedergegeben als Items im Mittelfeld (vgl. Baddely 1990). Die Versuchspersonen wurden außerdem angewiesen, die Sätze sorgfältig zu lesen, da zu dem Inhalt der gelesenen Sätze Fragen beantwortet werden mussten. Die Fragen wurden nach jedem Block im Anschluss an die mündliche Wiedergabe der Zielitems gestellt, bei den 4er- und 5er-Blöcken noch eine zusätzliche Frage nach dem dritten Set, so dass sich insgesamt sechs Fragen ergaben. Die Versuchspersonen bekamen die Sets mit zunehmender Satzzahl so lange präsentiert, bis sie vier Sets einer bestimmten Set-Größe nicht mehr bewältigen konnten. Die Gesamtzahl der richtig wiedergegebenen Zielitems wurde als Lesespannenmaß gewählt⁹.

4. Ergebnisse

Die deskriptiven Statistiken der Tests zum Leseverstehen, zum Wortschatz sowie zur Messung der Lesespannen sind in Tabelle 1 für die Lernergruppe und in Tabelle 2 für die deutsche Kontrollgruppe aufgeführt. Die Unterschiede zwischen den beiden Versuchsgruppen wurden in einer One-Way ANOVA

⁹ Die absolute Zahl der richtig wiedergegebenen Zielwörter hat sich als aussagekräftiger herausgestellt als die Zahl der bewältigten Set-Größe (vgl. Schmidt 2003).

gemessen. Die Leistungen der japanischen LernerInnen unterscheiden sich in der Textreproduktion und der Wortschließung aus dem Kontext nicht signifikant von der Kontrollgruppe. Ein signifikanter Unterschied besteht zwischen den Leistungen im Wortschatztest ($F [1,30] = 67.43, p < 0.001$) und den Leistungen im deutschen Lesespannentest ($F [1,30] = 15.72, p < 0.05$). Eine zusätzliche Überprüfung der Abhängigkeit des Wortschatzes von lernerbiographischen Faktoren ergab einen signifikanten Einfluss der Aufenthaltsdauer in Deutschland.

In der Lernergruppe ist die Lesespanne der japanischen Test-Version zwar höher als die der deutschen, aber sie unterscheiden sich nicht signifikant. Auffällig sind die hohen Standardabweichungen bei den Lesespannentests, der Textreproduktion und dem Wortschatztest sowie der geringe Wert bei der Wortschließung.

	Mittelwert	SD	Min	Max	Meridian	Modus
Textreproduktion	15.84	3.72	9.50	21.50	15.50	15.50
Wortschließung	1.96	0.53	1.27	2.87	1.85	-
Wortschatz	20.06	4.42	9.00	28.00	21.00	21.00
LS L2 Deutsch	50.44	10.21	31.00	68.00	53.50	52.00
LS L1 Japanisch	54.59	9.46	36.50	68.00	58.50	61.50

N=16

Tabelle 1: Deskriptive Statistiken der Tests zum Leseverstehen, zum Wortschatz und zur Lesespanne (LS) der Lernergruppe

	Mittelwert	SD	Min	Max	Meridian	Modus
Textreproduktion	15.50	3.35	10.00	21.00	16.50	15.00; 14.50; 12.50
Wortschließung	2.27	0.59	1.53	3.13	2.20	2.97; 1.57
Wortschatz	29.38	1.02	27.00	31.00	30.00	30.00
LS L1 Deutsch	62.13	5.90	49.00	69.00	61.00	69.00

N=16

Tabelle 2: Deskriptive Statistiken der Tests zum Leseverstehen, zum Wortschatz und zur Lesespanne (LS) der Kontrollgruppe

Tabelle 3 präsentiert die Ergebnisse der korrelativen Überprüfung der Leistungen in den Sprachtests mit den deutschen und den japanischen Lesespannenwerten der Lernergruppe (2-seitige Korrelation nach PEARSON). Signifikante Relationen bestehen zwischen den Lesespannenwerten der japanischen LernerInnen im deutschen Lesespannentest (Lesespanne L2 Deutsch) und den Leistungen in der Textreproduktionsaufgabe. Weiterhin korrelieren in der

Lernergruppe die Lesespannenwerte im japanischen Lesespannentest (Lesespanne L1 Japanisch) signifikant mit der Textreproduktion und der Wortschließung. In der deutschen Kontrollgruppe liegen keine signifikanten Beziehungen zwischen den Lesespannenwerten (Lesespanne L1 Deutsch) und den einzelnen Sprachtestleistungen vor.

Lesespanne	Wortschatz	Textreproduktion	Wortschließung
L2 Deutsch	-.26	0.53*	0.17
L1 Japanisch	-.23	0.74**	0.51*

* $p < 0.5$, ** $p < 0.01$, $N=16$

Tabelle 3: Korrelative Prüfung der Beziehung zwischen den Lesespannenwerten L2 Deutsch, L1 Japanisch und den Leistungen in den Leseverstehenstests und dem Wortschatztest

5. Diskussion

Die Ergebnisse der Auswertung belegen in der Lernergruppe eine signifikante Korrelation zwischen dem Lesespannentest im Japanischen und den Leistungen im Textreproduktionstest sowie der Wortschließung aus dem Kontext. Die Lesespannenwerte der deutschen Testversion korrelieren signifikant nur mit den Leistungen der Textreproduktionsaufgabe. Die Ergebnisse stehen im Widerspruch zu den Daten von Geva & Ryan (1993), die sowohl die L1- als auch die L2-Gedächtnisspanne als Prädiktor für L2-Leseverstehensleistungen belegen. Sie kontrastieren auch zu den Ergebnissen von Walter (2004), die eine signifikante Korrelation zwischen L2-Lesespannen und L2-Leseverstehensleistungen feststellt. Hier müssen der Zusammenhang zwischen den jeweils geprüften Lesefertigkeiten und den beim Lesespannentest ablaufenden kognitiven Prozesse näher erläutert werden.

Die Leistungen im Lesespannentest setzen sich aus den Teilleistungen Verstehen, Behalten und Strategieauswahl zusammen. Dass in starkem Maße die individuelle Aufmerksamkeitssteuerung und der damit zusammenhängende Strategieeinsatz verantwortlich für die Lesespannengröße ist, belegen die Studien von Herrmann (2003) und McNamara & Scott (2001). Die im Vergleich zur deutschen Kontrollgruppe signifikant geringeren L2-Lesespannenwerte der japanischen Lernergruppe weisen daraufhin, dass die Sprachverarbeitung auf den unteren Ebenen noch nicht automatisiert verläuft und damit Aufmerksamkeit erfordert, die dann für das Behalten und die Auswahl von Strategien nicht zur Verfügung steht. Ursache der Störungen können zum einen Wortschatzdefizite sein, worauf die im Vergleich zu den deutschen Versuchspersonen signifikant geringeren Leistungen im Wortschatztest hinweisen. 15 der 16 LernerInnen gaben

an, mit dem Wortschatz große Probleme zu haben, mit der deutschen Syntax dagegen keine. Auch erfolgt der lexikalische Zugriff trotz hoher L2-Kompetenz wahrscheinlich noch nicht störungsfrei (vgl. u.a. Kroll, Michael, Tokowicz & Dufour 2002). Als weiterer Störfaktor kann das laute Lesen angenommen werden, da für die japanischen Versuchspersonen die Umwandlung der Grapheminformationen in einen phonologischen Kode noch nicht genügend automatisiert ist (vgl. u.a. Segalowitz & Hébert 1990). Trotz der subjektiv als gut eingeschätzten Syntaxkompetenz könnten verlangsamte Leseprozesse aufgrund des nochmaligen Lesens von Sätzen oder Satzteilen (*re-reading*), beibehalten als Strategie aus der Anfangsphase, eine Rolle spielen (vgl. Frenck-Mestre 2002: 218-221.). Dies würde auch erklären, warum die L2-Lesespanne nur mit der Textreproduktionsaufgabe signifikant korreliert, nicht aber mit der Bedeutungserschließung aus dem Kontext. Letztere Aufgabenstellung erfordert sehr viel höheren kognitiven Aufwand als das Wiedergeben eines gelesenen Textes. Die Informationen müssen verstanden, behalten und dann zum Erschließen der Wortbedeutung eingesetzt werden. Hierbei ist ein hohes Maß an Inferieren nötig, d.h. das Gewichten, Steuern und Auswerten von Informationen. Die Aufmerksamkeit muss somit auf die semantische Verarbeitung gerichtet sein. In der Studie von Walter (2004) korreliert die L2-Lesespanne mit den Leistungen in Lückentexten, die auf Zusammenfassungen kürzerer narrativer Texte (140-150 Wörter) basieren (vgl. ebd.: 320f.). Sie wurden Englisch-als-Fremdsprache-Lehrbüchern entnommen und stellen m.E. keine großen kognitiven Anforderungen dar. Bei einer weiteren, anspruchsvolleren Aufgabe, dem Erinnern von Referenzbezügen, trat kein signifikanter Zusammenhang mit L2-Lesespannenwerten auf.

Die Leistungen der japanischen LernerInnen in den beiden Tests zum Leseverstehen entsprechen nahezu den Leistungen der deutschen Kontrollgruppe bzw. übertreffen sie sogar im Textreproduktionstest. Hier scheint die Vertrautheit mit der Aufgabenstellung eine Rolle zu spielen. Die Reproduktion von Textinhalten, die das genaue Verstehen und präzise Wiedergeben von Einzelinformationen erfordert, ist für JapanerInnen eine durch ihre Lernsozialisation bekannte Prüfungsform. Deutsche StudentInnen sind damit weniger vertraut. Auch ist für japanische MuttersprachlerInnen die Einbeziehung des Kontextes zur Bedeutungserschließung von Wörtern aufgrund ihrer L1-Lesefahrungen eine bekannte Fertigkeit. Die hohe Frequenz von Homographen im Japanischen - Kanji-Morpheme können bis zu sechs Bedeutungen und 100 Lesarten haben - erfordert ein kontinuierliches Einbeziehen des Kontextes. Zusätzlich ist zu berücksichtigen, dass L1-Fertigkeiten aufgrund des regelmäßigen Lesens japanischer Texte präsent sind.

Anzumerken ist abschließend, dass die relativ hohen Standardabweichungen

der Werte beim Wortschatztest, bei der Textreproduktion und vor allem bei den Lesespannen, sowohl bei den japanischen als auch bei den deutschen Versuchspersonen, auf individuell sehr unterschiedliche Leistungen hinweisen (vgl. Tabelle 1 und 2). Diese breite Streuung wurde insbesondere bei L2-Leseverstehensleistungen festgestellt (vgl. Chodkiewicz 2001: 43f.).

Zusammenfassend lässt sich Folgendes feststellen: Die individuelle Arbeitsgedächtniskapazität, gemessen mit einem L1-Lesepanntest, hat bei hoher L2-Sprachkompetenz und kognitiv aufwändiger Leseverstehensaufgabe einen Einfluss auf die L2-Verstehensleistungen. Da sich beim Lesepanntest unterschiedliche Aufgabenstellungen überlagern, und zwar Verstehen, Behalten und Strategieauswahl, ist der Einfluss auf die unterschiedlichen Teilleistungen beim Leseverstehen aufgrund der Komplexität der ablaufenden Prozesse allerdings nur schwer zu bestimmen. Auch ist unklar, ob die mit dem Lesepanntest erhobenen unterschiedlichen Arbeitsgedächtnisleistungen aus Differenzen der Arbeitsgedächtnisressourcen, z.B. der Speicherkapazität, resultieren oder aus der Effizienz, mit der diese genutzt werden. Folgt man aktuellen Ansätzen zur Arbeitsgedächtnisforschung, die individuelle Lesespannen auf das mehr oder weniger effiziente Einsetzen von Strategien zurückführen, bedeuten hohe Lesepanntestwerte einen der spezifischen Aufgabenstellung angepassten effektiven Strategieeinsatz. Damit hätte die "kognitive Flexibilität" (Ehlers 1998: 188) einen wichtigen Stellenwert beim Aufbau der L2-Lesekompetenz. Die von mir untersuchten japanischen Deutsch-als-Fremdsprache-Lernenden gaben weitaus häufiger als die deutschen Versuchspersonen an, Strategien wie z.B. Visualisieren oder Assoziationsbildungen einzusetzen, und zwar sowohl bei den Tests zur Messung der Lesespannen als auch bei den Leseverstehentests. Strategien sind aber offensichtlich nur dann nützlich, wenn die Prozesse der unteren Verarbeitungsebenen ungestört verlaufen, da ansonsten nicht genügend Kapazitäten zur Aufmerksamkeitssteuerung zur Verfügung stehen. Aufgrund der noch offenen Fragen in der Arbeitsgedächtnisforschung muss die in der aktuellen Forschung vertretene These, dass die (individuelle) Leistung des Arbeitsgedächtnisses Prädiktor für individuell erreichbare L2-Leseverstehensleistungen ist, weiterhin kritisch hinterfragt werden. Zukünftige Studien sollten unter Berücksichtigung der Ergebnisse aus der Arbeitsgedächtnisforschung ein breiteres Spektrum an Leseverstehensleistungen testen. Es fehlen bisher insbesondere Aufgabenstellungen, die komplexe kognitive Prozesse auf der höheren Verarbeitungsebene erfordern, wie z.B. Kohärenzherstellung. Diese Art von Aufgaben stellen besondere Anforderungen an das Arbeitsgedächtnis dar und geben möglicherweise einen besseren Einblick in die (individuellen) Leistungen des Arbeitsgedächtnisses.

Eingang des revidierten Manuskripts 20.11.2006

Literaturverzeichnis

- Andrade, Jackie (Hrsg.) (2001), *Working Memory in Perspective*. Hove: Psychology Press (Lawrence Erlbaum Association).
- Baddeley, Alan D. (1990), The development of the concept of working memory: implications and contributions of neuropsychology. In: Vallar, Guiseppe & Shallice, Tim (Hrsg.) (1990), *Neuropsychological Impairments of Short-Term Memory*. Cambridge: University Press, 54-73.
- Baddeley, Alan D. (1992), Working memory. *Science* 255, 556-559.
- Baddeley, Alan D. (1997), *Human Memory. Theory and Practice*. (überarb. Aufl.). Hove: Psychology Press (Lawrence Erlbaum Association).
- Baddeley, Alan D. (2000a), The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences* 11: 4, 417-423.
- Baddeley, Alan D. (2000b), Comment on Cowan: The magic number and the episodic buffer. *Behavioral and Brain Sciences* 24:1, 117-118.
- Baddeley, Alan D & Hitch, Graham J. (1974), Working memory. In: Bower, Gordon H. (Hrsg.) (1974), *The Psychology of Learning and Motivation*. New York: Academic Press, 47-89.
- Böcher, Walter (1963), Erfahrungen mit dem Wechsler'schen Gedächtnistest (Wechsler memory scale) bei einer deutschen Versuchsgruppe von 200 normalen Versuchspersonen. *Diagnostica* 9, 56-68.
- Brown, Carol M. (1998), L2 reading. An update on relevant L1 research. *Foreign Language Annals* 31: 2, 191-202.
- Buntig, Michael F.; Conway, Andrew R.A. & Heitz, Richard P. (2004), Individual differences in the fan effect and working memory capacity. *Journal of Memory and Language* 51, 604-622.
- Chodkiewicz, Halina (2001), The acquisition of word meaning while reading in English as a foreign language. *EUROSLA Yearbook* 1, 29-49.
- Cowan, Nelson (2000), The magical number 4 in short term memory. A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24, 78-185.
- Daneman, Meredyth & Carpenter, Patricia A. (1980), Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19, 450-466.
- Daneman, Meredyth & Carpenter, Patricia A. (1983), Individual differences in integrating information between and in sentences. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 9, 561-583.
- Daneman, Meredyth & Green, Ian (1986), Individual differences in comprehending and producing words in context. *Journal of Memory and Language* 25: 1-18.
- Daneman, Meredyth & Merikle, Philip M. (1996), Working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin & Review* 3, 422-433.
- Ehlers, Swantje (1998), *Lesetheorie und fremdsprachliche Lesepaxis aus der Perspektive des Deutschen als Fremdsprache*. Tübingen: Narr.
- Ehlers, Swantje (1999), Zum Wandel in der Lesetheorie und seine Folgen für die Fremdsprachendidaktik. *Zeitschrift für Fremdsprachenforschung* 10:2, 177-207.
- Frenck-Mestre, Chery (2002), An online look at sentence processing in the second language. In: Heredia, Roberto A. & Altarriba, Jeanette (Hrsg.) (2002), *Bilingual Sentence Processing*. Elsevier: Science B. V., 217-236.
- Geva, Esther & Ryan, Ellen B. (1993), Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning* 43: 1, 5-42.

- Harrington, Michael & Sawyer, Mark (1992), L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition* 14, 25-38.
- Herrmann, Katrin N. (2003), *Aspekte der Schriftsprachverarbeitung und der Einfluss von Strategien. Eine empirische Untersuchung*. Freiburg/Br.: Diss. Phil.
- Just, Marcel A. & Carpenter, Patricia A. (1992), A capacity theory of comprehension: individual differences in working memory. *Psychological Review* 99, 122-149.
- Koch, Wolfgang & Schweizer, Karl (2001), Kapazitätslimitierung und intellektuelle Leistungsfähigkeit. *Zeitschrift für experimentelle Psychologie* 48: 1, 1-19.
- Koda, Keida (2005), *Insights into Second Language Reading. A Cross-Linguistic Approach*. Cambridge: University Press.
- Kroll, Judith F.; Michael, Erica; Tokowicz, Natasha & Dufour, Robert (2002), The development of lexical fluency in a second language. *Second Language Research* 18: 2, 137-171.
- Lutjeharms, Madeline (2006), Zum Erwerb fremdsprachlicher Lesefertigkeiten. In: Jung, Udo O. H. (Hrsg.) (2006), *Praktische Handreichungen für Fremdsprachenlehrer* (4., vollständig neu bearbeitete Aufl.). Frankfurt a.M.: Peter Lang, 145-152.
- McNamara, Danielle & Scott, Jennifer L. (2001), Working memory capacity and strategy use. *Memory & Cognition* 29: 1, 10-17.
- Myake, Akira & Friedman, Naomi P. (1998), Individual differences in second language proficiency. Working memory as language aptitude. In: Healy, Alice F. & Bourne, Lyle E. (Hrsg.) (1998), *Foreign Language Learning. Psycholinguistic Studies on Training and Retention*. London: Lawrence Erlbaum Ass., 339-364.
- Myake, Akira & Shah, Priti (Hrsg.) (1999), *Models of Working Memory. Mechanism of Active Maintenance and Executive Control*. Cambridge: University Press.
- Osaka, Mariko & Osaka, Naoyuki (1992), Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of Psychonomic Society* 30, 287-289.
- Schmidt, Claudia (2000), Arbeitsgedächtnis und fremdsprachliches Leseverstehen. *Zeitschrift für Fremdsprachenforschung* 11: 1, 83-101.
- Schmidt, Claudia (2003), Zur Arbeitsgedächtniskapazität in L1 und L2. *Fragmentos* 24, 105-115.
- Schmidt, Claudia (2004), Wörter lernen durch Lesen: eine empirische Untersuchung zum Strategieeinsatz des indirekten Lernens bei japanischen fortgeschrittenen DaF-Lerner/innen. *Fremdsprachen Lehren und Lernen* 33, 72-82.
- Segalowitz, Norman & Hébert, Martine (1990), Phonological recoding in the first and second language reading of skilled bilinguals. *Language Learning* 40, 503-538.
- Tewes, Uwe (1991) (Hrsg.), *HAWIE-R: Hamburg-Wechsler Intelligenztest für Erwachsene. Handbuch und Testanweisung*. Bern etc.: Huber.
- Turley-Amens, Kandi J. & Whitfield, Michelle M. (2003), Strategy training and working memory task performance. *Journal of Memory and Language* 49, 446-468.
- Vos, Sandra H.; Gunter, Thomas C.; Schriefers, Herbert & Friederici, Angela D. (2001), Syntactic parsing and working memory. The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes* 16: 1, 65-103.
- Walter, Catherine (2004), Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics* 25:3, 315-339.

Anhang

Anhang 1: Testsätze aus dem Lesespannentest (2er Block mit Verständnisfrage)

In dem erst kürzlich geplanten Industrieviertel soll auch eine Fabrik errichtet werden.

Um die Leistungsfähigkeit des Rennwagens zu erhöhen, wurde ein neuer Motor eingebaut.

Bei vielen Schwarzwälder Bauernhöfen grenzt an den Garten eine bunte Wiese an. Die Ergebnisse dieser Untersuchung lassen sich gut mit einer Kurve darstellen.

Die großen Schiffe Richtung Nordsee bevorzugen den Kanal als günstigeren Verkehrsweg.

Alle Studenten hoffen am Ende des Studiums ein gutes Zeugnis zu bekommen.

In der überfüllten Straßenbahn versuchten gestern zwei Männer meine Tasche zu stehlen.

Aus ihrem Urlaubsort am Meer schickten sie eine Karte mit einem Sonnenuntergang.

Aufgrund der ungewöhnlichen Kälte mussten wir schon tagsüber ein Feuer anzünden.

Den Nonnen war vom Vatikan erlaubt worden, ein eigenes Kloster zu gründen.

Welches Motiv hatte die Urlaubskarte?

Anhang 2: Wechsler-Texte mit markierten Bewertungseinheiten (vgl. Böcher 1963)

a. Anne Schmidt / aus einem Hamburger / Vorort, / die als Reinemachefrau / in einem Geschäftshaus / arbeitet, / meldete / auf dem Polizei / Revier, / dass man sie in der vergangenen Nacht / auf der Straße / angegriffen / und ihr 100 DM / gestohlen habe. / Das Geld war bestimmt / für ihre vier / kleinen Kinder, / die seit zwei Tagen / nichts zu essen hatten. / Die Polizeibeamten, / denen die Frau leid tat, / machten für sie eine Spende. /

b. Das deutsche / Schiff / "Senator" / lief Montag / Abend / bei Cuxhafen / auf Grund. / Trotz / eines heftigen / Schneesturmes / und der Dunkelheit / wurden die 60 / Passagiere / einschließlich 18 Frauen / gerettet, / obwohl die Boote / wie Korken / auf der unruhigen See / trieben. / Am folgenden Tag / wurden sie durch einen britischen / Dampfer / zum Hafen gebracht. /

Anhang 3: Textbeispiel für die Erschließung der Wortbedeutung aus dem Kontext

Als Otto sich für den Beruf des Briefträgers entschied, hatte er von seiner zukünftigen Tätigkeit romantische Vorstellungen: Den ganzen Tag an der frischen Luft zu sein, erschien ihm immer schon erstrebenswert, und Kontakt mit Menschen suchte er, wo es eben ging. Schon in der Ausbildung merkte er aber, dass einem Briefträger mehr abverlangt wird als nur spazieren zu gehen und freundlich zu anderen Menschen zu sein. Das erste Problem waren die Hunde, die es auf seine Hosenbeine abgesehen hatten. Die zweite schmerzliche Erfahrung bezog sich auf seine Kundschaft: Viele Briefschreiber verwendeten zwar eine Schreibmaschine oder schrieben in Druckschrift, aber die Anschriften waren oft falsch. Das machte Otto beim morgendlichen Sortieren der Post große Mühe. Dann gab es aber noch Briefe, die mit unleserlichen Anschriften glänzten: Solche, bei denen nicht zu entscheiden war, ob sie in lateinischer oder kyrillischer Schrift verfasst waren, oder andere, die offensichtlich mit Tinte geschrieben, dann aber einer kurzen Wäsche unterzogen worden waren. Mindestens einen solchen *Jismeff* hatte Otto jeden Tag bei seiner Post, und es schmerzte ihn sehr, hier an die Grenzen seiner Fähigkeiten zu geraten. So wurde Otto schon sehr bald bewusst, dass die von ihm erstrebte perfekte Ausübung seines Berufes durch eine Reihe widriger Umstände behindert wurde.

Definition von *Jismeff*: ein Brief, der sehr unleserlich adressiert ist

Schriftsprachliche und mathematische Leistungen in der Erstsprache

Ein Vergleich monolingual und teilimmersiv unterrichteter Kinder der zweiten und dritten Klassenstufe

Anna Chr. M. Zaunbauer & Jens Möller¹

The present study examines the effects of early partial immersion (all school subjects including math are taught in English (L2) except learning to read and write, which are taught in German (L1)) on school attainment in grade two and grade three. $N = 81$ second and third graders educated monolingual or in a partial immersion program were tested towards the end of the school year. In addition to background variables as well as motivational variables (German interest and self-concept; math interest and self-concept) cognitive variables (concentration, nonverbal intelligence) were tested. As dependent variables German-speaking tests in the domains of reading, writing and math were administered. L1-literacy acquisition as well as mathematical output was comparable for children in the partial immersion program and monolingual educated children when controlling for cognitive and motivational differences. The results are discussed regarding selection effects as well as with regard to Cummins' Threshold and Interdependence Model.

1. Einleitung

Am herkömmlichen Fremdsprachenunterricht wird vielerlei Kritik geübt. So wird unter anderem bemängelt, dass er die Schüler nicht ausreichend auf die Begegnung mit der Fremdsprache (L2) in einer natürlichen Situation vorbereitet (Bliesener 1995) und zudem erst viel zu spät einsetzt. Als eine Alternative bietet sich daher die so genannte frühe Immersion an, eine Form bilingualen Unterrichts, bei welchem Schüler im Zuge der Einschulung in eine L2 eintauchen. Die L2 wird nicht als separates Unterrichtsfach gelehrt, sondern als Unterrichtssprache in verschiedenen Inhaltsfächern zur Vermittlung curricularer Inhalte eingesetzt (Akcan 2004). Damit geht der Fremdspracherwerb in Anlehnung an den Erstspracherwerb (L1) zumindest anfänglich ohne explizite Grammatikvermittlung vor sich.

Während Immersion in Kanada, den USA und anderen Ländern bereits

¹ Anna Chr. M. Zaunbauer & Jens Möller, Universität Kiel, Institut für Psychologie, Olshausenstr. 75a, 24118 Kiel, E-Mail: zaunbauer@psychologie.uni-kiel.de, E-Mail: jmoeller@psychologie.uni-kiel.de.

vielfältig erforscht wurde und als etablierte Unterrichtsform gilt, ist sie in Deutschland eine Ausnahme. Die von uns untersuchte Schule hat seit dem Ende der Neunzigerjahre englischsprachige Immersion als Unterrichtsprojekt erprobt. Pro Klassenstufe (Klassen eins bis vier) wird jeweils eine Klasse immersiv unterrichtet. Dabei wird eine so genannte frühe Teilimmersion (engl. *partial immersion*, Cummins & Swain 1986) realisiert, bei welcher die Schüler mit Ausnahme des Deutschunterrichts (erstsprachliche Alphabetisierung) in allen anderen Fächern, also auch in Mathematik, in der Fremdsprache Englisch (L2) unterrichtet werden. Insgesamt finden während der vier Grundschuljahre ca. 70% der Unterrichtszeit in der L2 statt. Lesen und Schreiben in der L2 sowie L2-Grammatikvermittlung werden ab der zweiten Klassenstufe eingeführt. Damit unterscheidet sich dieses Immersionsprogramm von Immersionsprogrammen im internationalen Raum. Diese sehen entweder für die ersten Schuljahre eine vollständige Immersion vor, bei welcher in den ersten Schuljahren kein erstsprachlicher Unterricht vorgesehen ist und damit die Alphabetisierung in der L2 stattfindet. Oder aber sie sehen bei Teilimmersion vor, den Anteil des erstsprachlichen Unterrichts im Laufe der Zeit zu vergrößern oder ihn von vornherein umfangreicher zu gestalten (z.B. de Courcy & Burston 2000).

Es ist wichtig zu betonen, dass durch die Immersion nicht die L1 (etwa Deutsch) verdrängt werden soll (Clyne, Wallner, Jenkins, Chen, Tsokalidou & Wallner 1995), sondern vielmehr ein additiver Bilingualismus (Lambert 1975) das Ziel ist, bei welchem L1 und L2 ebenbürtig nebeneinander bestehen können. Ein solcher additiver Bilingualismus soll mit gesteigerten kognitiven (Diaz & Klingler 1991; Johnson 1991; Lambert & Tucker 1972), linguistischen (Bialystok 2001) und metalinguistischen (Kontroll-) Fähigkeiten (Ben-Zeev 1977; Bild & Swain 1989; Cenoz & Valencia 1994) einhergehen. Damit unterscheiden sich die Bedingungen und die Situation für Schüler in einem Immersionsprogramm, wie es unserer Untersuchung zugrunde liegt, sehr deutlich von jenen Programmen (so genannte Submersionsprogramme), wie sie von Kindern mit Migrationshintergrund erlebt werden (vgl. Fazio & Lyster 1998). Denn obwohl auch Kinder mit Migrationshintergrund eine Art Immersion erleben, wenn sie etwa in Deutschland in der Unterrichtssprache Deutsch unterrichtet werden, zielt diese vornehmlich darauf ab, den Kindern ungeachtet einer Förderung ihrer L1 möglichst rasch die L2 beizubringen (subtraktiver Bilingualismus). Weitere Unterschiede zwischen der hier interessierenden Immersion und Submersionsprogrammen beziehen sich beispielsweise darauf, dass die L1 deutscher Schüler hierzulande einer Majoritätssprache mit hohem Status entspricht, während die L1 der Kinder mit Migrationshintergrund zumeist einer Minoritätssprache mit geringerem Status gleich kommt. Die L2 Englisch hat in unserer Untersuchung

einen Stellenwert als Fremdsprache, während die L2 für Kinder mit Migrationshintergrund den Stellenwert einer Zweitsprache einnimmt, die auch für den Alltag notwendig ist und vom Großteil der Bevölkerung zur alltäglichen Kommunikation verwendet wird (Oxford & Shearin 1994). Daneben finden sich zahlreiche weitere Unterschiede zwischen den beiden Populationen und den auf sie zugeschnittenen Unterrichtsprogrammen, auf die hier nicht näher eingegangen werden kann (Cummins 1979; Skutnabb-Kangas & Toukomaa 1976). Für nähere Informationen sei auf einen Überblicksartikel von Reich & Roth (2002; vgl. auch Fazio & Lyster, 1998) sowie die empirischen Ergebnisse von Stanat (2003) verwiesen.

Das Schwergewicht der internationalen Forschung zur Immersion lag bisher auf der Untersuchung früher vollständiger Immersion und zeigte, dass immersiv unterrichtete Schüler mittelfristig gleich gut oder besser in sprachlichen und mathematischen Fertigkeitsbereichen abschneiden als monolingual unterrichtete Schüler (Holobow, Genesee, Lambert, Gastright & Met 1987; Swain & Lapkin 1991). Der vermehrte L2-Input hat demnach positive Effekte auf Variablen wie das Verständnis grammatischer Strukturen der L1 (Tremaine 1975) und die figurative und metaphorische Sprachbenutzung (Neufeld 1993).

Aber auch die wenigen Studien zu Effekten früher Teilimmersion auf schulische Leistungen in der L1 und in Mathematik legen nahe, dass diese Form des Unterrichts sich insbesondere langfristig nicht negativ, sondern eher positiv auswirkt. Dies deutet auch eine erste eigene Untersuchung zum L1-Leseverständnis bei deutschen Viertklässlern an (Zaunbauer, Bonerad & Möller 2005b). In dieser wie auch in ersten längsschnittlichen Analysen (Zaunbauer & Möller 2006) zeigen sich eher Vorteile in den Schulleistungen teilimmersiv unterrichteter Schüler. Diese Vorteile sind allerdings vollständig über die nonverbale Intelligenz mediiert. Thomas, Collier & Abbott (1993) zeigen, dass teilimmersiv unterrichtete Schüler, welche Mathematik, Naturwissenschaften und Heimat- und Sachkunde in der L2, alle anderen Fächer aber in der L1 lernen, ab der zweiten Klasse im erstsprachlichen Bereich deutlich bessere Leistungen verglichen mit einer sorgfältig parallelisierten Stichprobe monolingual unterrichteter Schüler erzielen. Hinsichtlich der mathematischen Kenntnisse zeigt die Gruppe teilimmersiv unterrichteter Schüler in den Klassenstufen eins bis drei vergleichbare Leistungen wie die Kontrollgruppe (s. Studien aus dem afrikanischen Raum von Bunyi 1999; Gfeller & Robinson 1998; s. auch Johnstone 2002). Vergleichbar sind mathematische Leistungen immersiv unterrichteter Schüler, auch wenn sie mit erstsprachlichen Tests erfasst wurden. Dass teilimmersiv unterrichtete Schüler, die mathematische Konzepte in der L2 lernen, in der L1 gestellte Mathematikaufgaben erfolgreich bewältigen, sehen de Courcy & Burston (2000) als einen Hinweis darauf, dass Schüler zu Transferleistungen von der L2 zur L1 in der Lage sind (vgl. Cummins 1979).

2. Fragestellung

Im Bundesgebiet fehlen pädagogisch-psychologische Untersuchungen zur Immersion bislang fast vollständig. Die vorliegende Untersuchung analysiert, welche schulleistungsbezogenen Unterschiede zwischen monolingual und teilimmersiv unterrichteten Schülern der zweiten und dritten Klassenstufe bestehen. Insbesondere interessiert, ob und welche Unterschiede im Hinblick auf die erstsprachliche Leseleistung, die Rechtschreibleistung und (in der dritten Klasse) das Leseverständnis bestehen und ob sich Unterschiede in den erstsprachlichen mathematischen Leistungen zeigen. Diese Unterschiede werden bei Kontrolle motivationaler und kognitiver Variablen überprüft.

3. Methode

3.1 Stichprobe

Für die vorliegende Untersuchung konnten zwei Klassen der zweiten sowie zwei Klassen der dritten Klassenstufe einer Grundschule, in welcher bereits seit mehreren Jahren ein immersiver Schulzweig existiert, gewonnen werden. Bei der Schule handelt es sich um eine öffentliche Schule aus einem mittelständisch geprägten Stadtteil. Auf jeder Klassenstufe wird eine Klasse (insgesamt 48.8% der Schüler) immersiv unterrichtet, die andere Klasse wird monolingual unterrichtet. Während immersiv unterrichtete Schüler alle Fächer bis auf Deutsch (Lesen und Schreiben) in der L2 Englisch lernen und ab der zweiten Klassenstufe auch in der L2 gelesen und geschrieben wird, erhält der monolingual unterrichtete Schulzweig einmal wöchentlich Englischunterricht ab der ersten Klasse.

Die Daten von $N = 81$ Schülern (50.6% Jungen, Klassenstufe 2: 45.5%, Klassenstufe 3: 56.8%; 48.8% immersiv unterrichtet, Klassenstufe 2: 50%, Klassenstufe 3: 48.6%) konnten analysiert werden. Aufgrund von Fehltagen sind die Datensätze mancher Schüler nicht vollständig, was zu Schwankungen in den Freiheitsgraden führt. Die Schüler der zweiten Klassenstufe sind im Durchschnitt 8.15 Jahre alt ($SD = 5.4$ Monate; monolingual $M = 8.31$, $SD = 5.8$; immersiv $M = 7.98$, $SD = 4.3$), die Schüler der dritten Klassenstufe im Durchschnitt 9.06 Jahre alt ($SD = 5.9$ Monate; monolingual $M = 9.26$, $SD = 4.6$; immersiv $M = 8.85$, $SD = 6.3$). Die meisten Schüler (85.7%) haben einen Deutsch-monolingualen Sprachhintergrund (2. Klassenstufe: monolingual 88.9%, immersiv 82.4%; 3. Klassenstufe: monolingual 100%, immersiv 80%). Auf die familiäre Situation der Schüler wird im Ergebnisteil näher eingegangen.

3.2 Variablen

Die Zuweisung der Schüler zu den beiden Unterrichtsformen (monolingual vs. immersiv) erfolgte nicht zufällig, sondern aufgrund von schulischer Auswahl und Elternwünschen. Damit sind Selektionseffekte wahrscheinlich. Um zu prüfen, ob zwischen den Gruppen Unterschiede hinsichtlich relevanter Variablen bestehen, wurden in der vorliegenden Untersuchung in geringem zeitlichem Abstand vor der Untersuchung mit den Schulleistungstests gegen Ende des Schuljahres relativ viele Kontrollvariablen erhoben. Als zentrale abhängige Variablen wurden verschiedene schriftsprachliche Leistungsmerkmale sowie Leistungen in einem Mathematiktest erfasst (Maximalwerte und interne Konsistenzen (Cronbachs α) aller Variablen: vgl. für Klassenstufe 2 Tabellen 1 und 2 und resp. für Klassenstufe 3 Tabellen 3 und 4).

3.2.1 Familiäre Hintergrundvariablen

Im Elternfragebogen wurden neben dem Alter des Kindes auch sein Geburtsland und das der Eltern erfragt. Zudem wurde erfragt, ob das Kind zu Hause Deutsch gelernt hat und wie häufig Deutsch als Umgangssprache in der Familie Verwendung findet (Skala von 1 = "nie oder fast nie" bis 3 = "immer oder fast immer"). Weiterhin wurden die Eltern gebeten, ihre materielle Situation (familiärer Wohlstand) einzuschätzen ("Wie wohlhabend schätzen Sie Ihre Familie im Vergleich zu anderen Familien ein?", Skala von 1 = "überhaupt nicht wohlhabend" bis 5 = "wohlhabend") und ihren Bildungsabschluss sowie ihre berufliche Ausbildung anzugeben.

3.2.2 Häufigkeit lese- und rechenbezogener Tätigkeiten vor der Einschulung des Kindes

Anhand von insgesamt 13 Items (vierstufiges Antwortformat von 1 = "nie oder fast nie" bis 4 = "sehr oft") wurde mit der Frage "Wie oft haben Sie oder jemand anderes aus Ihrem Haushalt die folgenden Tätigkeiten mit Ihrem Kind durchgeführt, bevor es in die Schule kam?" die Häufigkeit gemeinsamer Tätigkeiten in den Bereichen Lesen/Schreiben (9 Items, z. B. "Bücher lesen", "Buchstaben oder Wörter schreiben") und Rechnen (4 Items, z. B. "Zahlen auftragen", "Größenvergleiche anstellen") vor der Einschulung des Kindes erfasst.

3.2.3 Vorschulische Kenntnisse des Kindes in den Bereichen Lesen/Schreiben und Rechnen

Mit der Frage "Wie gut konnte Ihr Kind die folgenden Dinge, als es in die Schule kam?" sollten die vorschulischen Kenntnisse des Kindes in den Bereichen Lesen/Schreiben (6 Items, z. B. "Die meisten Buchstaben des Alphabets erkennen", "Buchstaben des Alphabets schreiben") und Rechnen (4 Items, z. B.

"Die Zahlen zwischen 1 und 9 benennen", "Die Zahlen zwischen 1 und 9 schreiben") von den Eltern auf einer vierstufigen Skala (von 1 = "gar nicht" bis 4 = "sehr gut") angegeben werden.

3.2.4 Häufigkeit aktueller lese- und rechenbezogener Tätigkeiten

Anhand von insgesamt 8 Fragen ("Wie oft machen Sie oder jemand anderes derzeit zu Hause die folgenden Dinge mit Ihrem Kind?", vierstufiges Antwortformat von 1 = "nie oder fast nie" bis 4 = "sehr oft") wurde die Häufigkeit derzeit gemeinsam durchgeführter Tätigkeiten in den Bereichen Lesen/Schreiben (5 Items, z. B. "Meinem Kind beim Vorlesen zuhören") und Rechnen (3 Items, z. B. "Mit meinem Kind einfache Rechenaufgaben durchführen") erfragt.

3.2.5 Elterliches Hausaufgabenengagement

Das elterliche Hausaufgabenengagement wurde mit der Frage "Überprüfen Sie, ob Ihr Kind die Hausaufgaben macht?" anhand von zwei Items erfasst (z. B. "Sie überprüfen schriftliche Hausaufgaben auf Richtigkeit."). Das Antwortformat war dichotom (ja/nein).

3.2.6 Fachspezifisches Interesse und Selbstkonzept

Anhand eines altersgerecht gestalteten Fragebogens (vgl. Möller, Pohlmann, Streblov & Kaufmann 2002) wurden die Variablen Interesse und Selbstkonzept für die Domänen Deutschunterricht und Mathematikunterricht mit mehreren Items (Interesse, 5 Items: z. B. "Findest du Lesen und Schreiben (Rechnen) spannend?"; Selbstkonzept, 4 Items: z. B. "Findest du Lesen und Schreiben (Rechnen) leicht?") erfasst. Der Fragebogen wies ein vierstufiges Format auf (ja, eher ja, eher nein, nein). In beiden Klassenstufen wurden die einzelnen Fragen den Kindern vom Versuchsleiter vorgelesen.

3.2.7 Konzentration

Zur Bestimmung der Konzentrationsfähigkeit wurde der Aufmerksamkeits-Belastungs-Test d2 (Brickenkamp 1994) eingesetzt. Für die vorliegende Untersuchung wurde zur Quantifizierung der Konzentrations- und Aufmerksamkeitsleistung der Gesamtwert GZ-F herangezogen. Dieser ergibt sich aus der Differenz zwischen der Gesamtbearbeitungsleistung (GZ) und der Anzahl der Fehler (F).

3.2.8 Nonverbale Intelligenz

Die nonverbale Intelligenz wurde im Gruppenverfahren anhand der CPM (Coloured Progressive Matrices, Raven, Bulheller & Häcker 2002) geschätzt.

3.2.9 Leseleistung

Anhand der Würzburger Leise Leseprobe (WLLP, Küspert & Schneider 1998) wurde die Dekodiergeschwindigkeit als Maß für die erstsprachliche Leseleistung

auf Wortebene erfasst. Die Anzahl der bearbeiteten Items abzüglich der Auslassungen und Fehler ergibt die Leseleistung.

3.2.10 Leseverständnis

Der nur in Klassenstufe drei eingesetzte Test zur Erfassung des erstsprachlichen Leseverständnisses wurde der Internationalen Grundschul-Lese-Untersuchung (IGLU, Bos, Lankes, Prenzel, Schwippert, Valtin & Walther 2004) entnommen. Der Test besteht aus einem deutschsprachigen Lesetext, zu dem deutschsprachige Fragen offenen und geschlossenen Antwortformats zu beantworten sind, wobei jederzeit im Text nachgelesen werden kann. Die Leistungen wurden zu einem Gesamtscore verrechnet.

3.2.11 Rechtschreibleistung

Die erstsprachliche Rechtschreibleistung der Schüler wurde anhand der Hamburger Schreib-Probe (HSP, May 2002) erfasst. Als Maß für die Rechtschreibleistung wird die Anzahl richtig geschriebener Wörter erfasst. Die HSP gestattet zusätzlich eine Analyse der Graphemtreffer (Anzahl richtig geschriebener Buchstaben bzw. Buchstabenkombinationen, May 2002).

3.2.12 Mathematikleistung

Zur Erfassung der Mathematikleistung wurde der Deutsche Mathematiktest (DEMAT) in der Version für die zweite und für die dritte Klassenstufe eingesetzt (Krajewski, Liehm & Schneider 2004; Roick, Göllitz & Hasselhorn 2004). Der Test setzt sich aus verschiedenen, den curricularen Anforderungen der jeweiligen Klassenstufe entsprechenden Aufgabentypen zusammen (vgl. Tabellen 2 resp. 4, Untertests DEMAT). Als Maß für die Mathematikleistung wurden neben dem Gesamtscore auch die Leistungen in den einzelnen Untertests berücksichtigt, für die separate Normwerte im Handbuch berichtet werden.

3.3 Durchführung

Alle Daten wurden im Rahmen von Gruppentestungen erhoben. Dabei waren zumeist zwei Testleiter für eine Schulklasse verantwortlich. Die Erfassung motivationaler und kognitiver Variablen erfolgte wenige Tage vor der Testung mit den Schulleistungstests an einem Vormittag während zwei Schulstunden. Zu diesem Zeitpunkt wurde auch den Eltern der Elternfragebogen ausgehändigt. Die Testung mit den beiden Deutschtests WLLP (Leseleistung) und HSP (Rechtschreibleistung) erfolgte für alle Schüler in der genannten Reihenfolge vormittags in einer Schulstunde. An einem weiteren Vormittag fand während einer Schulstunde der Mathematiktest statt. Schülern der dritten Klassenstufe wurde zusätzlich der Test zur Erfassung des Leseverständnisses vorgelegt.

3.4 Statistische Analysen

In einem ersten Schritt wird für jede Klassenstufe die Vergleichbarkeit der beiden Gruppen (monolingual, immersiv) geprüft, indem familiäre Hintergrundvariablen, lese- und rechenbezogene Tätigkeiten und Vorkenntnisse, elterliches Hausaufgabenengagement sowie motivationale und kognitive Variablen verglichen werden. Erstere werden vorwiegend nonparametrisch analysiert. Für motivationale und kognitive Variablen werden getrennt multivariate Varianzanalysen mit dem Faktor Unterrichtsform gerechnet. Je nach Fragestellung sind Häufigkeiten, Mittelwerte und Standardabweichungen sowie die Ergebnisse des statistischen Vergleichs in den Tabellen 1 resp. 2 und 3 resp. 4 aufgeführt. Die Leistungen der monolingual und immersiv unterrichteten Schüler werden zunächst mit dem Mittelwert der Normstichprobe verglichen. Anschließend werden multivariate Varianzanalysen zur Prüfung von Effekten der Unterrichtsform (monolingual vs. immersiv) auf die Leistungen in erstsprachlichen Schulleistungstests durchgeführt. Danach werden diese Analysen unter Einbeziehung der relevanten motivationalen und/oder kognitiven Variablen als Kovariaten gerechnet. Die Mittelwerte und Standardabweichungen, getrennt nach Unterrichtsform und Klassenstufe, sowie die auf den Faktor Unterrichtsform bezogenen Ergebnisse der statistischen Analysen sind in Tabellen 2 und 4 dargestellt. Nachfolgend werden die Ergebnisse getrennt für die beiden Klassenstufen dargestellt.

Klassenstufe 2					
	Max. Punkte	Cronbachs α	monolingual M[SD] (N)	immersiv M[SD] (N)	Signifikanz
Familiäre Hintergrundvariablen:					
In Deutschland geboren			100% (N=9)	94.1% (N=17)	$\chi^2(1)=.55, ns$
Mütter in Deutschland geboren			100% (N=9)	94.1% (N=17)	$\chi^2(1)=.55, ns$
Väter in Deutschland geboren			100% (N=8)	93.3 (N=15)	$\chi^2(1)=.56, ns$
Deutsch zu Hause gelernt			100% (N=9)	100% (N=17)	s. Anmerkung 1
Deutsch zu Hause gesprochen	3		3.0 [1.00] (N=8)	3.0 [1.00] (N=16)	s. Anmerkung 1
Einschätzung familiärer Wohlstand	5		2.67 [1.0] (N=9)	3.13 [.92] (N=15)	$F(1,22)=1.37, ns$
Mutter: Hochschulreife			22.2 % (N=9)	82.4 % (N=17)	$\chi^2(1)=9.0, **$
Vater: Hochschulreife			14.3 % (N=7)	71.4 % (N=14)	$\chi^2(1)=7.36, **$
Mutter: Fachhochschul-/ Hochschulabschluss/Promotion			22.2 % (N=9)	68.8 % (N=16)	$\chi^2(1)=6.23, *$
Vater: Fachhochschul-/ Hochschulabschluss/Promotion			28.6 % (N=7)	78.6 % (N=14)	$\chi^2(1)=6.23, *$
Vorschulische Tätigkeiten und Kenntnisse:					
Lesbezogene Tätigkeiten	4	.68	2.62 [.26] (N=9)	2.56 [.44] (N=17)	$F(1,24)=.11, ns$
Rechenbezogene Tätigkeiten	4	.69	2.47 [.49] (N=9)	2.37 [.53] (N=17)	$F(1,24)=.24, ns$
Vorschulische Kenntnisse Lesen	4	.78	2.50 [.57] (N=9)	2.35 [.51] (N=17)	$F(1,24)=.46, ns$
Vorschulische Kenntnisse Rechnen	4	.81	3.00 [.97] (N=9)	3.03 [.61] (N=17)	$F(1,24)=.01, ns$
Aktuelle Tätigkeiten:					
Lesbezogene Tätigkeiten	4	.68	3.33 [.55] (N=9)	3.29 [.50] (N=17)	$F(1,24)=.63, ns$
Rechenbezogene Tätigkeiten	4	.67	2.52 [.56] (N=9)	2.78 [.79] (N=17)	$F(1,24)=.86, ns$
Hausaufgabenengagement	4	.55	1.78 [.67] (N=9)	1.53 [.62] (N=17)	$F(1,24)=.89, ns$

¹ Da die Variable eine Konstante ist, lässt sich keine Statistik berechnen: * $p < .05$, ** $p < .01$

Tabelle 1: Prozentualer Anteil (bezogen auf das N in Klammern) resp. Mittelwerte (M) und Standardabweichungen (SD) sowie Signifikanz zu einzelnen Fragen im Elternfragebogen, bezogen auf die Klassenstufen zwei und drei, differenziert nach der Unterrichtsform (monolingual, immersiv)

Klassenstufe 2					
	Max. Punkte	Cronbachs α	monolingual M[SD]	immersiv M[SD]	Signifikanz
Motivationale Variablen:					
Interesse Deutsch	5	.86	2.85 (1.01)	3.28 (.54)	$F(1,36)=2.88, ns$
Selbstkonzept Deutsch	4	.77	2.97 (.72)	3.49 (.48)	$F(1,36)=7.19, \eta^2=.17, *$
Interesse Mathematik	5	.85	3.05 (1.00)	3.66 (.41)	$F(1,36)=6.81, \eta^2=.16, *$
Selbstkonzept Mathematik	4	.85	2.81 (1.02)	3.24 (.65)	$F(1,36)=2.47, ns$
Kognitive Variablen:					
Konzentrationsleistung		.95 ¹	221.67 (46.17)	220.42 (30.46)	$F(1,32)=.01, ns$
Nonverbale Intelligenz	36	.82	28.20 (5.36)	32.53 (1.98)	$F(1,32)=10.61, \eta^2=.25, **$
Schulleistungsvariablen:					
Lesleistung (WILLP)	140	.98	70.58 (16.60)	70.79 (13.99)	$F(1,26)=.08, ns$
(Dekodiergeschwindigkeit)					
Rechtschreibleistung (HSP)	30	.87	20.67 (5.58)	25.42 (3.37)	$F(1,26)=2.14, ns$
Graphentreffer (HSP)	148	.80	134.42 (8.54)	142.68 (3.56)	$F(1,26)=4.10, \eta^2=.14, *$
Mathematikleistung (DEMAT)	36	.90	21.17 (5.79)	20.55 (8.82)	$F(1,26)=3.06, ns$
Untertests DEMAT:					
Zahleneigenschaften	2	.71	.64 (.84)	.58 (.84)	$F(1,28)=1.11, ns$
Längenvergleich	4	.76	3.07 (1.38)	2.84 (1.34)	$F(1,28)=3.55, ns$
Addition	4	.56	1.36 (1.15)	1.63 (1.61)	$F(1,28)=.06, ns$
Subtraktion	4	.50	2.36 (1.28)	2.16 (1.34)	$F(1,28)=.49, ns$
Verdoppeln	3	.82	2.64 (.50)	2.21 (1.27)	$F(1,28)=1.61, ns$
Division	4	.56	2.5 (1.22)	3.16 (1.12)	$F(1,28)=.01, ns$
Halbieren	3	.73	2.14 (.86)	1.79 (1.32)	$F(1,28)=2.99, ns$
Rechnen mit Geld	4	.83	2.79 (1.05)	1.63 (1.74)	$F(1,28)=9.71, \eta^2=.26, **$
Sachaufgaben	4	.50	2.71 (1.01)	2.82 (1.11)	$F(1,28)=.16, ns$
Geometrie	4	.63	1.14 (1.23)	1.74 (1.24)	$F(1,28)=.14, ns$

¹ Interne Konsistenz laut Handbuch: * $p < .05$, ** $p < .01$

Tabelle 2: Mittelwerte (M) und Standardabweichungen (SD) für motivationale, kognitive und Schulleistungsvariablen der zweiten Klassenstufe, differenziert nach der Unterrichtsform (monolingual, immersiv)

Klassenstufe 3					
	Max. Punkte	Cronbachs α	monolingual M[SD] (N)	immersiv M[SD] (N)	Signifikanz
Familiäre Hintergrundvariablen:					
In Deutschland geboren			100% (N=6)	100% (N=11)	s. Anmerkung 1
Mütter in Deutschland geboren			100% (N=6)	90.9% (N=11)	$\chi^2(1)=.58, ns$
Väter in Deutschland geboren			100% (N=4)	90.9 (N=11)	$\chi^2(1)=.39, ns$
Deutsch zu Hause gelernt			100% (N=6)	100% (N=10)	s. Anmerkung 1
Deutsch zu Hause gesprochen	3	3	3.0 [1.00] (N=5)	3.0 [1.00] (N=10)	s. Anmerkung 1
Einschätzung familiärer Wohlstand	5	5	2.33 [1.03] (N=6)	3.10 [1.32] (N=10)	$F(1,14)=4.95, \eta^2=.26, *$
Mutter: Hochschulreife			40% (N=5)	70% (N=10)	$\chi^2(1)=2.79, ns$
Vater: Hochschulreife			20% (N=5)	60% (N=10)	$\chi^2(1)=3.57, ns$
Mutter: Fachhochschul-/ Hochschulabschluss/Promotion			14.3% (N=7)	50.0% (N=10)	$\chi^2(1)=2.67, ns$
Vater: Fachhochschul-/ Hochschulabschluss/Promotion			25.0% (N=4)	60% (N=10)	$\chi^2(1)=3.57, ns$
Vorschulische Tätigkeiten und Kenntnisse:					
Lesbezogene Tätigkeiten	4	.55	2.48 [.17] (N=6)	2.43 [.43] (N=11)	$F(1,15)=.06, ns$
Rechenbezogene Tätigkeiten	4	.42	2.38 [.34] (N=6)	2.43 [.46] (N=11)	$F(1,15)=.07, ns$
Vorschulische Kenntnisse Lesen	4	.74	2.22 [.27] (N=6)	2.29 [.58] (N=11)	$F(1,15)=.07, ns$
Vorschulische Kenntnisse Rechnen	4	.76	2.54 [.64] (N=6)	2.98 [.77] (N=11)	$F(1,15)=.38, ns$
Aktuelle Tätigkeiten:					
Lesbezogene Tätigkeiten	4	.56	2.90 [.75] (N=6)	3.27 [.33] (N=11)	$F(1,15)=2.11, ns$
Rechenbezogene Tätigkeiten	4	.61	2.67 [.56] (N=6)	2.64 [.60] (N=11)	$F(1,15)=.01, ns$
Hausaufgabenengagement	4	.50	1.33 [.82] (N=6)	1.45 [.82] (N=11)	$F(1,15)=.09, ns$

¹ Da die Variable eine Konstante ist, lässt sich keine Statistik berechnen: * $p < .05$, ** $p < .01$

Tabelle 3: Prozentualer Anteil (bezogen auf das N in Klammern) resp. Mittelwerte (M) und Standardabweichungen (SD) sowie Signifikanz zu einzelnen Fragen im Elternfragebogen, bezogen auf die Klassenstufe drei, differenziert nach der Unterrichtsform (monolingual, immersiv)

Klassenstufe	Max. Punkte	Cronbachs α	monolingual <i>M(SD)</i>	immersiv <i>M(SD)</i>	Signifikanz
Motivationale Variablen:					
Interesse Deutsch	5	.85	2.91 (.72)	2.87 (.85)	$F(1,32)=.03, ns$
Selbstkonzept Deutsch	4	.56	3.57 (.37)	3.60 (.42)	$F(1,32)=.05, ns$
Interesse Mathematik	5	.75	3.07 (.74)	2.99 (.73)	$F(1,32)=.11, ns$
Selbstkonzept Mathematik	4	.86	3.07 (.71)	3.50 (.46)	$F(1,32)=4.42, \eta^2=.12, *$
Kognitive Variablen:					
Konzentrationsleistung		.95 ¹	251.00 (88.98)	275.31 (43.19)	$F(1,31)=.98, ns$
Nonverbale Intelligenz	36	.73	29.35 (3.41)	32.50 (2.58)	$F(1,31)=8.85, \eta^2=.22, **$
Schulleistungsvariablen:					
Leseleistung (WLLP) (Dekodiergeschwindigkeit)	140	.98	105.88 (24.20)	92.53 (21.04)	$F(1,28)=.35, ns$
Leseverständnis	16	.67	9.25 (2.57)	10.53 (2.23)	$F(1,28)=.16, ns$
Rechtschreibleistung (HSP)	38	.77	30.56 (4.40)	32.40 (3.85)	$F(1,28)=1.97, ns$
Graphemtreffer (HSP)	191	.78	181.44 (6.99)	186.00 (4.21)	$F(1,28)=.29, ns$
Mathematikleistung (DEMAT)	31	.73	19.88 (4.77)	22.73 (3.59)	$F(1,28)=.02, ns$
Untertests DEMAT:					
Zahlenstrahlen	3	.30	1.71 (.92)	2.31 (.48)	$F(1,30)=.63, ns$
Additionen	4	.65	2.88 (1.17)	2.19 (1.33)	$F(1,30)=6.32, \eta^2=.17, *$
Subtraktionen	4	.68	1.53 (1.42)	2.44 (1.15)	$F(1,30)=1.47, ns$
Multiplikationen	4	.55	2.71 (1.21)	2.63 (1.09)	$F(1,30)=.16, ns$
Sachrechnungen	4	-.01	3.06 (.83)	3.44 (.73)	$F(1,30)=.06, ns$
Spiegelzeichnungen	3	.57	1.94 (.90)	2.50 (.82)	$F(1,30)=.16, ns$
Formen legen	3	.36	2.41 (.71)	2.00 (.89)	$F(1,30)=2.39, ns$
Längen schätzen	2	.17	.82 (.64)	1.38 (.72)	$F(1,30)=2.99, ns$
Längen umrechnen	4	.79	2.53 (1.50)	3.50 (.73)	$F(1,30)=1.46, ns$

Tabelle 4: Mittelwerte (M) und Standardabweichungen (SD) für motivationale, kognitive und Schulleistungsvariablen der dritten Klassenstufe, differenziert nach der Unterrichtsform (monolingual, immersiv)

4. Ergebnisse

4.1 Klassenstufe 2

4.1.1 Analyse der Eingangsbedingungen

Zunächst wird die Vergleichbarkeit der beiden Gruppen monolingual und immersiv unterrichteter Schüler hinsichtlich familiärer Hintergrundvariablen, vorschulischer und aktueller Tätigkeiten und Kenntnisse in den Bereichen Lesen, Schreiben und Rechnen sowie bezüglich des elterlichen Hausaufgabenengagements eingeschätzt. Da nur für etwa die Hälfte der Schüler ($N = 25$) Elternfragebogen vorliegen, basieren die folgenden Analysen auf einer reduzierten Stichprobe.

Tabelle 1 sind getrennt für monolingual und immersiv unterrichtete Schüler Häufigkeitsangaben, Mittelwerte und Standardabweichungen zu den Elternangaben zu entnehmen. Die Ergebnisse des statistischen Vergleichs zwischen monolingual und immersiv unterrichteten Schülern erweisen sich für beinahe alle Variablen als nicht signifikant. Auffallend sind jedoch deutliche Unterschiede in den Bildungsabschlüssen der Eltern der Schüler aus Klassenstufe zwei. Während Eltern immersiv unterrichteter Schüler in mehr als der Hälfte der Fälle die Hochschulzugangsberechtigung erreichten und ein Studium (Fachhochschul-/Hochschulstudium, Promotion) abgeschlossen haben, ist dies nur bei knapp einem

Fünftel der Eltern monolingual unterrichteter Schüler der Fall. Multivariate Analysen zeigen, dass sich keine Gruppenunterschiede im Hinblick auf das Ausmaß vorschulischer Tätigkeiten (Wilks $\lambda = .99, F(2,23) = .12, ns$) oder Kenntnisse (Wilks $\lambda = .96, F(2,23) = .41, ns$) oder aktueller Tätigkeiten ergeben (Wilks $\lambda = .96, F(2,23) = .45, ns$). Auch das Hausaufgabenengagement erweist sich für beide Gruppen als vergleichbar.

4.1.2 Analyse der Gruppenunterschiede in den motivationalen und kognitiven Variablen

Eine multivariate Varianzanalyse der vier motivationalen Variablen (Interesse und Selbstkonzept in Deutsch und Mathematik) deckt signifikante Gruppenunterschiede zugunsten immersiv unterrichteter Schüler auf (Wilks $\lambda = .74, F(4,33) = 2.94, p < .05, \eta^2 = .26$), die sich auf das Selbstkonzept Deutsch und das Mathematikinteresse zurückführen lassen. Beide sind höher ausgeprägt bei den immersiv unterrichteten Schülern. Eine multivariate Analyse für die kognitiven Variablen Konzentrationsleistung und nonverbale Intelligenz entdeckt ebenfalls einen Gruppenunterschied zugunsten immersiv unterrichteter Schüler (Wilks $\lambda = .75, F(2,31) = 5.16, p < .05, \eta^2 = .25$), der sich auf die nonverbale Intelligenz bezieht. Aufgrund der beschriebenen Ergebnisse (vgl. Tabelle 2) werden für die Klassenstufe zwei das Selbstkonzept Deutsch, das Mathematikinteresse und die nonverbale Intelligenz als Kovariaten in die Analysen aufgenommen.

4.1.3 Analyse der Gruppenunterschiede im Lesen, Schreiben und Rechnen

Die Leseleistung monolingual und immersiv unterrichteter Schüler liegt beim Vergleich ihrer Prozentrangwerte mit dem mittleren Prozentrangwert der Normstichprobe im Durchschnittsbereich ($M = 50\%$; monolingual: $t(19) = -.30, ns$ resp. immersiv: $t(20) = -.84, ns$). Ihre Rechtschreibleistung hingegen liegt über dem Durchschnittsbereich der Normstichprobe ($M = 18$; monolingual: $t(19) = 2.50, p < .05$); immersiv: $t(20) = 6.83, p < .001$). Die Mathematikleistungen befinden sich im Durchschnittsbereich der Normstichprobe ($M = 20$; monolingual: $t(20) = 1.93, ns$); immersiv: $t(19) = -.04, ns$).

In einer multivariaten Varianzanalyse für die Faktoren Leseleistung, Rechtschreibleistung (Anzahl Wörter), Rechtschreibleistung (Anzahl Graphemtreffer) und Mathematikleistung zeigt sich ein signifikanter Unterschied zwischen immersiv und monolingual unterrichteten Schülern (Wilks $\lambda = .75, F(4,34) = 2.88, p < .05, \eta^2 = .25$), der sich auf beide Maße der Rechtschreibleistung bezieht (Anzahl Wörter: $F(1,37) = 5.81, p < .05, \eta^2 = .14$; Anzahl Graphemtreffer: $F(1,37) = 6.23, p < .05, \eta^2 = .14$). Die Leistung der immersiv unterrichteten Schüler ist höher als die der monolingual unterrichteten Schüler, wie die Mittelwerte in Tabelle 2 zeigen. Werden die Kovariaten

Selbstkonzept Deutsch, Mathematik-Interesse sowie nonverbale Intelligenz berücksichtigt, zeigt sich nur mehr ein tendenzieller Effekt der Unterrichtsform zugunsten der immersiv unterrichteten Kinder (Wilks $\lambda = .72$, $F(4,23) = 2.26$, $p = .09$, $\eta^2 = .28$), der auf die Graphemtreffer bezogen ist. Bezüglich der anderen Leistungen ergeben sich keine Unterschiede.

Der Mathematiktest setzt sich aus insgesamt zehn Untertests zusammen. Eine multivariate Analyse bei Berücksichtigung der drei Kovariaten Selbstkonzept Deutsch, Mathematik-Interesse sowie nonverbale Intelligenz erbringt keinen signifikanten Effekt der Unterrichtsform (Wilks $\lambda = .61$, $F(10,19) = 1.24$, *ns*).

4.2 Klassenstufe 3

4.2.1 Analyse der Eingangsbedingungen

Auch für Klassenstufe drei werden zunächst die familiären Hintergrundvariablen, die vorschulischen und aktuellen Tätigkeiten und Kenntnisse in den Bereichen Lesen/Schreiben und Rechnen sowie das elterliche Hausaufgabenengagement betrachtet. Auch hier liegt nur für etwa die Hälfte der Schüler ein Elternfragebogen vor ($N = 17$).

In Tabelle 3 sind getrennt für monolingual und immersiv unterrichtete Schüler Häufigkeitsangaben, Mittelwerte und Standardabweichungen zu den Elternangaben dargestellt. Die Ergebnisse des statistischen Vergleichs zwischen monolingual und immersiv unterrichteten Schülern erweisen sich für beinahe alle Variablen als nicht signifikant. Deutliche Unterschiede zugunsten immersiv unterrichteter Schüler bestehen jedoch hinsichtlich der Einschätzung des familiären Wohlstands.

Multivariate Analysen zeigen, dass sich auch für Klassenstufe drei keine Gruppenunterschiede im Hinblick auf das Ausmaß vorschulischer Tätigkeiten (Wilks $\lambda = .99$, $F(2,14) = .07$, *ns*) oder Kenntnisse (Wilks $\lambda = .86$, $F(2,14) = 1.19$, *ns*) oder aktueller Tätigkeiten ergeben (Wilks $\lambda = .86$, $F(2,14) = 1.16$, *ns*). Auch das Hausaufgabenengagement erweist sich für beide Gruppen als vergleichbar.

4.2.2 Analyse der Gruppenunterschiede in den motivationalen und kognitiven Variablen

Eine multivariate Varianzanalyse der vier motivationalen Variablen (Interesse und Selbstkonzept in Deutsch und Mathematik) erbringt für Klassenstufe drei keinen signifikanten Unterschied zwischen den Gruppen (Wilks $\lambda = .76$, $F(4,29) = 2.28$, *ns*). Eine multivariate Analyse für die kognitiven Variablen Konzentrationsleistung und nonverbale Intelligenz deckt einen signifikanten Unterschied zwischen monolingual und immersiv unterrichteten Schülern auf (Wilks $\lambda = .72$, $F(2,30) = 5.80$, $p < .01$, $\eta^2 = .28$). Dieser ist auf die nonverbale Intelligenz

bezogen ($F(1,31) = 8.85$, $p < .01$, $\eta^2 = .22$), die bei immersiv unterrichteten Schülern ausgeprägter ist. Aufgrund dieser Ergebnisse wird bei den Analysen der Gruppenunterschiede in den Schulleistungen die nonverbale Intelligenz als Kovariate berücksichtigt.

4.2.3 Analyse der Gruppenunterschiede im Lesen, Schreiben und Rechnen

Die Leseleistung monolingual und immersiv unterrichteter Schüler der dritten Klassenstufe liegt beim Vergleich der Prozentrangwerte mit der Norm im Durchschnittsbereich der Normstichprobe ($M = 50\%$; monolingual: $t(18) = 1.14$, *ns* resp. immersiv: $t(17) = -.16$, *ns*). Die erstsprachliche Rechtschreibleistung immersiv unterrichteter Schüler liegt oberhalb des Durchschnittsbereichs der Normstichprobe ($M = 29$; $t(17) = 3.07$, $p < .01$). Dagegen ergibt sich keine bedeutsame Abweichung der monolingual unterrichteten Schüler von den Normwerten ($M = 29$, $t(18) = 1.22$, *ns*). Immersiv unterrichtete Schüler zeigen auch bessere Mathematikleistungen als die Normstichprobe ($M = 19$; immersiv: $t(15) = 3.59$, $p < .01$). Die Mathematikleistungen monolingual unterrichteter Schüler unterscheiden sich dagegen nicht von der Normstichprobe ($M = 19$, $t(18) = .20$, *ns*).

In einer multivariaten Varianzanalyse für die Faktoren Leseleistung, Leseverständnis, Rechtschreibleistung (Anzahl Wörter), Rechtschreibleistung (Anzahl Graphemtreffer) und Mathematikleistung zeigt sich ein signifikanter Unterschied zwischen den beiden Gruppen (Wilks $\lambda = .64$, $F(5,25) = 2.85$, $p < .05$, $\eta^2 = .36$), der sich auf die Rechtschreibleistung (Anzahl Graphemtreffer: $F(1,29) = 4.76$, $p < .05$, $\eta^2 = .14$) bezieht. Die Inspektion der Mittelwerte zeigt, dass die immersiv unterrichteten Schüler bessere Leistungen aufweisen als die monolingual unterrichteten Schüler (vgl. Tabelle 4). Der Unterschied in der Mathematikleistung zugunsten der immersiv unterrichteten Schüler ist nur tendenziell ($F(1,29) = 3.51$, $p = .07$, $\eta^2 = .11$). Wird die Kovariate nonverbale Intelligenz berücksichtigt, unterscheiden sich monolingual und immersiv unterrichtete Schüler der Klassenstufe drei nicht mehr signifikant (Wilks $\lambda = .75$, $F(5,24) = 1.57$, *ns*).

Eine multivariate Analyse der neun Untertests des Mathematiktests unter Berücksichtigung der Kovariaten nonverbale Intelligenz erbringt ebenfalls keinen signifikanten Unterschied zwischen den Gruppen (Wilks $\lambda = .52$, $F(9,22) = 2.22$, *ns*). Aus der Gesamtheit der untersuchten Schüler wurde zusätzlich getrennt für jede Klassenstufe eine Gruppe monolingual und immersiv unterrichteter Schüler nach nonverbaler Intelligenz parallelisiert. Die Ergebnisse dieser Analysen ergaben keine substantiellen Veränderungen gegenüber denen der Gesamtstichprobe und bestätigen so die Annahme der Vergleichbarkeit der Leistungen beider Gruppen².

² Die Ergebnisse dieser Analysen können bei den Autoren angefordert werden.

5. Diskussion

Die vorliegende Untersuchung ging der Frage nach, ob und welche Unterschiede zwischen monolingual und immersiv unterrichteten Schülern der Klassenstufen zwei und drei im Hinblick auf erstsprachliche Fertigkeiten wie die Leseleistung, das Leseverständnis, die Rechtschreibleistung und mathematische Kompetenzen bestehen. In dieser Diskussion werden die Ergebnisse zunächst im Zusammenhang mit der unterschiedlichen Beschulung besprochen. Danach wird die Problematik der Interpretierbarkeit der Befunde unter anderem aufgrund der geringen Stichprobengröße und der Selektionseffekte diskutiert.

Wie die Analysen zeigten, erweisen sich die hier untersuchten monolingual und teilimmersiv unterrichteten Schüler der zweiten und dritten Klassenstufe hinsichtlich ihres sprachbezogenen Hintergrunds, vorschulischer Tätigkeiten und Kenntnisse in den Bereichen Lesen, Schreiben und Rechnen sowie hinsichtlich des elterlichen Hausaufgabenengagements als durchaus vergleichbar. Allerdings bestehen zwischen den beiden Gruppen deutliche Differenzen im Bildungsabschluss der Eltern und der Einschätzung des familiären Wohlstands (Klassenstufe 3) zugunsten immersiv unterrichteter Schüler. Zudem finden sich motivationale und kognitive Unterschiede. Diese Unterschiede fallen ebenfalls zugunsten immersiv unterrichteter Schüler aus, wobei anhand der vorliegenden querschnittlichen Analysen nicht geklärt werden kann, ob es sich dabei um Effekte des immersiven Unterrichts oder um Selektionseffekte handelt (Macnamara 1966; Samuels & Griffore 1979; Snow 1987; Zaunbauer, Bonerad & Möller 2005b).

Die Analysen der Schülerleistungen zeigten zunächst, dass sich die Schulleistungen beider Gruppen in beiden Klassenstufen im oder sogar oberhalb des Normbereichs bewegen. Dies deutet darauf hin, dass die Schüler im Mittel eine profunde Grundlage bezüglich des Lesens und Rechtschreibens und auch in mathematischen Konzepten entwickelt und ausgebildet haben (de Courcy & Burston 2000). Die Gruppenunterschiede zugunsten immersiv unterrichteter Schüler werden dabei durch Unterschiede in der Kovariaten nonverbale Intelligenz vermittelt und entfallen bei deren Berücksichtigung. Diese Ergebnisse decken sich weitgehend mit den Resultaten einer Untersuchung an Viertklässlern (Zaunbauer, Bonerad & Möller 2005b). Auch dort konnten die besseren Leistungen immersiv unterrichteter Schüler im L1-Leseverständnis auf Unterschiede in der kognitiven Grundfähigkeit zurückgeführt werden. Ähnlich fanden auch Harley, Hart & Lapkin (1986) bei Kontrolle des IQ keine Unterschiede zwischen monolingual und immersiv unterrichteten Schülern hinsichtlich des L1-Wortschatzes.

Somit kann festgehalten werden, dass immersiv unterrichtete Schüler der Klassenstufen zwei und drei mindestens vergleichbare erstsprachliche Lese- und

Rechtschreibleistungen sowie Mathematikleistungen zeigen, ein Befund, der der Befürchtung von Leistungsnachteilen infolge einer drastischen Reduktion der Erstsprache als Unterrichtssprache (vgl. Harley, Hart & Lapkin 1986) entgegensteht. Es zeigen sich aber auch keine deutlich besseren Leistungen immersiv unterrichteter Schüler auf den hier betrachteten Grundschulstufen zwei und drei bei Berücksichtigung motivationaler und kognitiver Gruppenunterschiede (Swain & Lapkin 1982; Turnbull, Hart & Lapkin 2000; Turnbull, Lapkin & Hart 2001). Auch dies steht in weitgehendem Einklang mit der internationalen Forschungslage (Cummins 1979). Diese berichtet bessere Leistungen immersiv unterrichteter Schüler häufig erst für höhere Klassenstufen, was zumeist im Rahmen von Cummins Schwellen- und Interdependenz-Modell (1979) erklärt wird (Turnbull, Hart & Lapkin 2000; Turnbull, Lapkin & Hart 2001). Danach ist das auf diesen frühen Klassenstufen erreichte Niveau in der L1 und L2 noch nicht hoch genug, um positive Effekte der Bilingualität mit sich zu bringen. Der Schwellenniveauhypothese zufolge lassen sich positive Effekte der Immersion auf kognitive und schulische Leistungen nur dann verzeichnen, wenn sowohl in der L1 als auch in der L2 ein hohes Niveau erreicht wird (Baker 2001; Gonzalez 1999). Dies sei erst in höheren Klassenstufen der Fall und wird mit erweitertem metalinguistischem Bewusstsein durch das Wissen um zwei Sprachen erklärt.

Wie berichtet, zeigen die hier untersuchten immersiv unterrichteten Schüler auch bezüglich mathematischer Kompetenzen vergleichbare Leistungen wie monolingual unterrichtete Schüler (vgl. de Courcy & Burston 2000). Die Tatsache, dass mathematische Konzepte in der L2 unterrichtet wurden, der Test (Instruktion, Textaufgaben) aber in der L1 geboten wurde, hatte für die Leistungen der immersiv unterrichteten Schüler keine 'Nachteile' (Bournot-Trites & Reeder 2001). Unser Ergebnis wie auch Befunde der internationalen Forschung deuten darauf hin, dass Schüler mathematische Konzepte, die sie in der einen Sprache (L2) lernen, in einer anderen Sprache (L1) abrufen können (Bruck 1982; Genesee 1992; Zaunbauer & Möller 2006). Grundlage dafür ist laut Cummins' Modell (Cummins 1979) die Existenz eines gemeinsamen, sprachunspezifischen kognitiven Systems (Dornic 1979; Verhoeven 1991; Vygotsky 1934/1962), auf das beide (mehrere) Sprachen gleichermaßen Zugriff haben (de Courcy & Burston 2000).

Einschränkungen der Interpretierbarkeit der Daten

Die Generalisierbarkeit der gefundenen Ergebnisse ist wegen der geringen Stichprobengröße nur eingeschränkt gegeben. Zudem weist die vorliegende Untersuchung ein quasiexperimentelles Design auf. Da die Zuordnung unserer Schüler zur einen oder anderen Unterrichtsform nicht zufällig erfolgte, wurden mögliche Selektionseffekte berücksichtigt.

Als Gründe für die Nicht-Zuweisung zu einem immersiven Unterrichtsprogramm gelten Schülercharakteristika, die nach Genesee (1992) risikobelastete Schüler kennzeichnen. Neben geringen kognitiven Grundfähigkeiten zählen dazu geringere erstsprachliche Fähigkeiten, ein geringerer sozioökonomischer Status und die Zugehörigkeit zu einer ethnischen Minderheit. Zudem sind Selektionsmechanismen schon bei der Anmeldung der Schüler durch die Eltern und bei der Auswahl der Lehrer für die immersiven Klassen denkbar. So könnten Schulleitungen darauf achten, dass das Unterrichtsexperiment immersiver Unterricht auch gelingt und weiter entwickelte Schüler und besonders engagierte Lehrer in dieser Unterrichtsform einsetzen. In unserer Stichprobe von Schülern der zweiten und dritten Klassenstufe zeigten sich entsprechend deutliche Unterschiede in den kognitiven Grundfähigkeiten wie auch gewisse Unterschiede in motivationalen Variablen. Zudem wurden auch Unterschiede im elterlichen Bildungshintergrund entdeckt. Eine Entscheidung darüber, ob es sich dabei um Effekte des Immersionsunterrichts und/oder um Selektionseffekte handelt, ist hier nicht möglich.

Obwohl unser Vorgehen der statistischen Berücksichtigung dieser Unterschiede insofern kritisierbar ist, als dass wir damit mögliche durch das Unterrichtsprogramm bewirkte Unterschiede zwischen den Gruppen ausgleichen (Thomas & Collier 2002), ist damit zumindest verhindert worden, dass Unterschiede zwischen immersiven und monolingualen Unterrichtsprogrammen überbewertet wurden.

Die vorliegende Arbeit zeigt somit insgesamt, dass das Niveau schriftsprachlicher und mathematischer Kompetenzen bei teilimmersiv unterrichteten Schülern der Klassenstufen zwei und drei dem der monolingual unterrichteten Schüler mindestens gleichwertig ist. Zwar ist denkbar, dass die hier teilimmersiv unterrichteten Schüler in einem monolingual geführten Unterricht noch besser abgeschnitten hätten. Dass sie aber eher besser als die Altersnorm und die Vergleichsgruppe abgeschlossen haben, deutet an, dass Teilimmersion als Unterrichtsprogramm hierzulande durchaus Erfolg versprechend ist, zumal die Schüler bereits sehr früh zusätzlich die englische Sprache auf einem recht hohen Niveau erwerben (vgl. Zaunbauer, Bonerad & Möller 2005a). Die beschriebenen Überlegungen zeigen aber auch, wie wichtig die Berücksichtigung von möglichen Selektionseffekten in querschnittlichen Analysen für die Beurteilung von Unterrichtsprogrammen ist.

Eingang des revidierten Manuskripts 27.11.2006

Literaturverzeichnis

- Akcan, Sumru (2004), Teaching methodology in a first-grade French-immersion class. *Bilingual Research Journal* 28: 2, 267-277.
- Baker, Colin (2001), *Foundations of Bilingual Education and Bilingualism* (3. Aufl.). Clevedon: Multilingual Matters.
- Ben-Zeev, Sandra (1977), The influence of bilingualism on cognitive strategy and cognitive development. *Child Development* 25, 140-147.
- Bialystok, Ellen (2001), *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge: Cambridge University Press.
- Bild, Eva-Rebecca & Swain, Merrill (1989), Minority language students in a French immersion program: Their French proficiency. *Journal of Multilingual and Multicultural Development* 10: 3, 255-276.
- Bliesener, Ulrich (1995), Weiterentwicklung des Fremdsprachenunterrichts. Eine Notwendigkeit! Aber wie? *Neusprachliche Mitteilungen* 48: 1, 18-17.
- Bos, Wilfried; Lankes, Eva-Maria; Prenzel, Manfred; Schwippert, Knut; Valtin, Renate & Walther, Gerd (Hrsg.) (2004), *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Bournot-Trites, Monique & Reeder, Kenneth (2001), Interdependence revisited: Mathematics achievement in an intensified French immersion program. *The Canadian Modern Language Review / La Revue canadienne des langues vivantes* 58, 27-43.
- Brickenkamp, Rolf (1994), *Test d2 - Aufmerksamkeits-Belastungs-Test* (8. Aufl.). Göttingen: Hogrefe.
- Bruck, Margaret (1982), Language disabled children: Performance in an additive bilingual education program. *Applied Psycholinguistics* 3, 45-60.
- Bunyi, Grace (1999), Rethinking the place of African indigenous languages in African education. *International Journal of Educational Development* 19, 337-350.
- Cenoz, Jasone & Valencia, José (1994), Additive trilingualism: Evidence from the Basque Country. *Applied Psycholinguistics* 15, 255-274.
- Clyne, Michael; Wallner, Theresa; Jenkins, Catherine; Chen, Imogen Y.; Tsokalidou, Roula & Wallner, Theresa (1995), *Developing second language from primary school: Models and outcomes*. Canberra: The National Languages and Literacy Institute of Australia.
- Cummins, James (1979), Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49: 2, 222-251.
- Cummins, James & Swain, Merrill (1986), *Bilingualism in education: Aspects of theory, research, and practice*. London: Longman.
- de Courcy, Michèle C. & Burston, Monique (2000), Learning mathematics through French in Australia. *Language and Education* 14, 75-95.
- Diaz, Rafael M. & Klingler, Cynthia (1991), Towards an explanatory model of the interaction between bilingualism and cognitive development. In: Bialystok, Ellen (Hrsg.) (1991), *Language processing in bilingual children*. Cambridge: Cambridge University Press, 167-192.
- Dornic, Stanlovitch (1979), Information processing in bilinguals - Some selected issues. *Psychological Research* 40, 329-348.
- Fazio, Lucy & Lyster, Roy (1998), Immersion and submersion classrooms: A comparison of instructional practices in language arts. *Journal of Multilingual and Multicultural Development* 19: 4, 303-317.

- Genesee, Fred (1992), Second/foreign language immersion and at-risk English-speaking children. *Foreign Language Annals* 25: 3, 199-213.
- Gfeller, Elisabeth & Robinson, Clinton (1998), Which language for teaching? The cultural messages transmitted by the languages used in education. *Language and Education* 12: 1, 18-32.
- Gonzalez, Virginia (1999), *Language and cognitive development in second language learning*. Boston, MA: Allyn and Bacon.
- Harley, Birgit; Hart, Doug & Lapkin, Sharon (1986), The effects of early bilingual schooling on first language skills. *Applied Psycholinguistics* 7, 295-322.
- Holobow, Naomi; Genesee, Fred; Lambert, Wallace E.; Gastright, Joseph & Met, Myriam (1987), Effectiveness of partial French immersion for children from different social class and ethnic backgrounds. *Applied Psycholinguistics* 8, 137-152.
- Johnson, Janice (1991), Constructive processes in bilingualism and their cognitive growth effects. In: Bialystok, Ellen (Hrsg.) (1991), *Language processing in bilingual children*. Cambridge: Cambridge University Press, 193-221.
- Johnstone, Richard (2002), *Immersion in a second or additional language at school: A review of the international research*. Stirling, Scotland: Scottish CILT, Institute of Education, University of Stirling [Online: <http://www.scilt.stir.ac.uk/pubs.htm#immersion>. 19.10.06].
- Krajewski, Kristin; Liehm, Susann & Schneider, Wolfgang (2004), *DEMAT 2+. Deutscher Mathematiktest für zweite Klassen*. Göttingen: Beltz Test GmbH.
- Küspert, Petra & Schneider, Wolfgang (1998), *Würzburger Leise Leseprobe (WLLP): Ein Gruppenlesetest für die Grundschule*. Göttingen: Hogrefe.
- Lambert, Wallace E. (1975), Culture and language as factors in learning and education. In: Wolfgang, Aaron (Hrsg.) (1975), *Education of immigrant students*. Toronto: Ontario Institute for Studies in Education, 55-83.
- Lambert, Wallace E. & Tucker, G. Richard (1972), *Bilingual education of children*. The St. Lambert experiment. Rowley, MA: Newbury House.
- Macnamara, John (1966), *Bilingualism and primary education*. Edinburgh: Edinburgh University Press.
- May, Peter (2002), *Hamburger Schreib-Probe*. 6., aktualisierte und erweiterte Auflage. Hamburg: vpm.
- Möller, Jens; Pohlmann, Britta; Streblow, Lilian & Kaufmann, Julia (2002), Spezifität von Begabungsüberzeugungen als Determinante des fachbezogenen Selbstkonzepts der Begabung. *Zeitschrift für Pädagogische Psychologie* 16, 87-97.
- Neufeld, Gerald (1993), Early French immersion and proficiency in English: Some long-range effects. *Language and Society* 43, 8-10.
- Oxford, Rebecca & Shearin, Jill (1994), Language learning motivation: Expanding the theoretical framework. *Modern Language Journal* 78, 12-28.
- Raven, John C.; Bulheller, Stephan & Häcker, Hartmut (2002), *CPM. Coloured progressive matrices* (3. Aufl.). Göttingen: Hogrefe.
- Reich, Hans H. & Roth, Hans-Joachim (2002), *Spracherwerb zweisprachig aufwachsender Kinder und Jugendlicher: Ein Überblick über den Stand der nationalen und internationalen Forschung* (in Zusammenarbeit mit Dirim, I., Jørgensen, J. N., List, G., List, G., Neumann, U., Siebert-Ott, G., Steinmüller, U., Teunissen, F., Vallen, T. & Wurnig, V.). Hamburg: Behörde für Bildung und Sport, Amt für Schule.
- Roick, Thorsten; Göllitz, Dietmar & Hasselhorn, Marcus (2004), *DEMAT 3+. Deutscher Mathematiktest für dritte Klassen*. Göttingen: Beltz Test GmbH.

- Samuels, Douglas D. & Griffore, Robert J. (1979), The Plattsburgh French language immersion program: Its influence on intelligence and self-esteem. *Language Learning* 29: 1, 45-52.
- Skutnabb-Kangas, Tove & Toukomaa, Pertti (1976), *Teaching migrant children's mother tongue and learning the language of the host country in the context of the socio-cultural situation of the migrant family*. Helsinki: The Finnish National Commission for UNESCO.
- Snow, M. Ann (1987), *Immersion teacher handbook*. Educational report series. Los Angeles: Center for Language Education and Research. University of California.
- Stanat, Petra (2003), Schulleistungen von Jugendlichen mit Migrationshintergrund: Differenzierung deskriptiver Befunde aus PISA und PISA-E. In: Baumert, Jürgen; Artelt, Cordula; Klieme, Eckhart; Neubrand, Michael; Prenzel, Manfred; Schiefele, Ulrich; Schneider, Wolfgang; Tillmann, Klaus-Jürgen & Weiß, Manfred (Hrsg.) (2003), *PISA-2000 - Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Opladen: Leske + Budrich, 243-260.
- Swain, Merrill & Lapkin, Sharon (1982), *Evaluating bilingual education: A Canadian case study*. Clevedon: Multilingual Matters.
- Swain, Merrill & Lapkin, Sharon (1991), Additive bilingualism and French immersion education: The roles of language proficiency and literacy. In: Reynolds, Allan G. (Hrsg.) (1991), *Bilingualism, multiculturalism, and second language learning: The McGill conference in honour of Wallace E. Lambert*. Hillsdale, NJ: Erlbaum, 203-216.
- Thomas, Wayne P. & Collier, Virginia P. (2002), *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz, CA: Center for Research on Education, Diversity & Excellence [Online: http://crede.berkeley.edu/research/llaa/1.1_final.html. 12.09.06].
- Thomas, Wayne P.; Collier, Virginia P. & Abbott, Martha (1993), Academic achievement through Japanese, Spanish, French: The first two years of partial immersion. *The Modern Language Journal* 77: 22, 170-179.
- Tremaine, Ruth V. (1975), *Syntax and piagetian operational thought*. Washington, D.C.: Georgetown University Press.
- Turnbull, Miles; Hart, Doug & Lapkin, Sharon (2000), *French immersion students' performance on Grade 3 and 6 provincial tests: Potential impacts on program design*. Toronto: OISE/UT Modern Language Centre.
- Turnbull, Miles; Lapkin, Sharon & Hart, Doug (2001), Grade three immersion students' performance in literacy and mathematics: Province-wide results from Ontario (1998-99). *The Canadian Modern Language Review / La Revue canadienne des langues vivantes* 58: 1, 9-26.
- Verhoeven, Lude T. (1991), Predicting minority children's bilingual proficiency: Child, family, and institutional factors. *Language Learning: A Journal of Applied Linguistics* 41: 2, 205-233.
- Vygotsky, Lev S. (1934/1962), *Thought and language*. Cambridge, MA: The MIT Press.
- Zaunbauer, Anna Chr. M.; Bonerad, Eva-Marie & Möller, Jens (2005a, 26.-28.09.2005), *Muttersprachliche und fremdsprachliche Lesekompetenz immersiv unterrichteter Kinder*. Vortrag gehalten auf der 10. Fachtagung Pädagogische Psychologie, Halle (Saale) [Online: <http://www.paeps10.uni-halle.de/>. 10.07.2006].
- Zaunbauer, Anna Chr. M.; Bonerad, Eva-Marie & Möller, Jens (2005b), Muttersprachliches Leseverständnis immersiv unterrichteter Kinder. *Zeitschrift für Pädagogische Psychologie* 19: 4, 263-265.

Zaunbauer, Anna Chr. M. & Möller, Jens (2006), Schulleistungen monolingual und immersiv unterrichteter Kinder am Ende des ersten Schuljahres. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*. Manuskript unter Vorbehalt angenommen.

Fachlichkeit und Sprachlichkeit

Zwischenbilanz eines DFG-Projekts

Helmut Johannes Vollmer¹

The research project presented here compares the basic geographical competences of bilingual and monolingual 10th graders. The central question is, whether bilingual learners, after 4 years of bilingual education, are equally competent in subject-matter terms as their monolingual counterparts, given the fact that they had used a foreign language (English) as a working language instead of their mother tongue. The comparison is done by analysing the structure of their subject-matter competence (based on an explicit model, elicited by typical subject-specific tasks and written solutions) as well as by looking at the mental processes applied (elicited through thinking-aloud-protocols, interactive conversational protocols and retrospective interviews). In particular, subject-specific discourse competence is focused upon, as identified in task solutions (products) and their editing.

Overall, the role of language as a mediator between thinking, conceptualising and communicating in subject-specific contexts is discussed. Some preliminary results are reported. The basic design of the project is hypothesis-testing, although two qualitative studies generated new hypotheses to be tested in the future.

1. Einleitung²

Bilingualer Unterricht, genauer fremdsprachlich durchgeführter Sachfachunterricht in der Sekundarstufe I an allgemein bildenden Schulen Deutschlands (besonders an Gymnasien) stellt eine Erweiterung des traditionellen Fremdsprachenunterrichts dar. Im Mittelpunkt steht allerdings nicht mehr die Vermittlung der Fremdsprache als grammatisches Regelsystem, sondern seine Nutzung als Arbeitssprache in authentischen Verwendungssituationen, wobei offensichtlich der Spracherwerb auf implizite Weise befördert wird (Vollmer 2005a). Andererseits ist bilingualer Unterricht aber vor allem *Fach*unterricht, der dem Anspruch gerecht werden muss, über das Medium der Fremdsprache dieselben fachlichen Inhalte und Kompetenzen zu vermitteln wie der Regelunterricht. Trotz jahrelanger positiver Erfahrungen mit dem bilingualen Sachfachunterricht wissen wir jedoch nur sehr wenig über die Besonderheiten und die Auswirkungen des bilingualen Lernens im Vergleich zum muttersprachlich

¹ Korrespondenzadresse: Prof. Dr. Helmut Johannes Vollmer, Universität Osnabrück, FB Sprach- und Literaturwissenschaft, D-49069 Osnabrück, Tel. 0541-969-4260, Fax 0541-969-4886, E-Mail: hvollmer@uni-osnabrueck.de.

² Ich danke Lena Heine für wertvolle Hinweise und Verbesserungsvorschläge zu einer früheren Fassung des Manuskripts.

erteilten Fachunterricht. Was sind die Stärken des jeweiligen Unterrichtsansatzes, wo liegen seine Schwächen? Wie wirkt sich die Fremdsprachlichkeit auf den Umgang mit dem Sachfachinhalt aus? Wird in beiden Fällen gleich viel gelernt? In der Vergangenheit sind vor allem die positiven Auswirkungen des bilingualen Unterrichts auf das *fremdsprachliche Können* untersucht worden, nicht aber die *Qualität des fachlichen Lernens selbst* noch die *Vergleichbarkeit der Lernergebnisse* am Ende eines mehrjährigen Bildungsganges.

Diese Fragen stehen im Mittelpunkt des Forschungsinteresses des DFG-Projekts, das z. Z. an der Universität Osnabrück unter dem Titel "Fachlernen und (Fremd-)Sprachlichkeit: Aufgabenbasierte Kognition, Kooperation, Kommunikation" unter Leitung des Autors durchgeführt wird und in das verschiedene Teilstudien als Dissertationsvorhaben eingebettet sind. Dabei handelt es sich um eine *explorative Studie*, die zugleich herausfinden will, welches die Verarbeitungsprozesse von Schülern und Schülerinnen bei der Bearbeitung und Lösung fachlicher Aufgaben und beim Aufbau inhaltlicher Konzepte mit Hilfe der L1 bzw. einer L2 sind und welche fachlichen Kompetenzen (inhaltliches Wissen, Methodenkompetenz, kommunikative Fertigkeiten, Beurteilungsfähigkeiten) am Ende von Klasse 10 nachweislich daraus resultieren. Konkret wird am Beispiel des deutsch- bzw. englischsprachigen Erdkundeunterrichts im Gymnasium beschrieben, wie sich diese *Prozesse* ebenso wie deren fachliche *Ergebnisse* differenziert darstellen lassen (vgl. Vollmer 2002).

2. Einbettung des Projekts in die forschungspolitische Landschaft

Das Projekt "Fachlernen und (Fremd-)Sprachlichkeit" war eines von vier Projekten eines Antragspakets, das in den letzten Jahren von der Deutschen Forschungsgemeinschaft (DFG) gefördert wurde. Mit diesem Paketantrag wurde der Versuch unternommen, mit intensiver Vorberatung unter den Antragstellenden Personen die Zweit- und Fremdsprachenforschung in ihrer Präsenz wie in ihrer Potenz stärker in das Bewusstsein der Wissenschaften zu heben und einen größeren Einfluss in den Einrichtungen der Drittmittelförderung zu gewinnen.

2.1 Paketantrag an die Deutsche Forschungsgemeinschaft

Diese Forschungsinitiative wurde seinerzeit durch den Autor in Kooperation mit Wolfgang Börner (Hamburg) erfolgreich auf den Weg gebracht: Von den insgesamt 10 Projektanträgen, die nach zwei internen Beratungsrunden verblieben, ist *einer* vorab außerhalb des Pakets gestellt und abgelehnt worden (vermut-

lich auf Grund einer mangelnden Kompatibilität mit den Forschungsstandards der DFG). Von den verbliebenen neun Projektanträgen wurden vier auf Anhieb als förderungswürdig eingestuft und mit gewissen Auflagen genehmigt, allerdings in keinem der Fälle in vollem Umfang. Zwei weitere Anträge wurden an die Antragssteller zur Überarbeitung zurückgegeben, mit der Aussicht auf erneute Antragsstellung und positiver Förderperspektive. Hingegen sind drei der Projektanträge auf der Basis von anonymen gutachterlichen Stellungnahmen abgelehnt worden.

Zur Begründung des Paketantrags haben Wolfgang Börner und ich die Forschungssituation innerhalb der Zweit- und Fremdsprachenforschung skizziert und die einzelnen Anträge darin eingebettet. Wir gingen davon aus, dass es innerhalb unserer Disziplin sowohl ein enormes Forschungspotential als auch einen großen Nachholbedarf gibt, insbesondere was quantifizierende Forschungsansätze anbelangt. Dasselbe gilt für eine Verknüpfung von quantifizierenden mit qualitativen Zugängen zu bestimmten Forschungsfragen; letztere erwiesen sich - zumindest bezogen auf die Qualifikationsarbeiten des wissenschaftlichen Nachwuchses - als deutlich in der Überzahl, wobei die Gründe dafür unterschiedlich sein können (vgl. Riemer 2006). Deshalb sollten sowohl Projekte zur Hypothesenüberprüfung als auch solche zur Hypothesengenerierung nach vergleichbaren methodologischen Standards ausgerichtet und zur Förderung vorgeschlagen werden. Nach bisherigen Erfahrungen mit der DFG können wir feststellen, dass deren Gutachter vermutlich eine Koppelung beider Verfahren (wo immer möglich) präferieren, zumindest aber ausschließlich qualitative Fragestellungen nur in Ausnahmefällen befürworten. Auf jeden Fall wird eine klare Ergebnisorientierung und Anwendungsperspektive der empirisch gewonnenen Einsichten eingefordert.

Die DFG hat auf unsere Bitte hin ein eintägiges Rundgespräch mit potentiellen Gutachtern arrangiert, das zu erheblichen Verbesserungen in der Formulierung einiger Anträge führte. Außerdem wurde in diesen Fachgesprächen deutlich, was die Kriterien der Begutachtung sein könnten und worauf in der spezifischen Textsorte "Drittmittelantrag" entsprechend zu achten ist - eine Erfahrung, die wir in den Folgejahren versucht haben an andere Antragsteller und Teile des wissenschaftlichen Nachwuchses weiter zu geben. Unser Antragspaket wurde im Übrigen nicht von Einzelgutachtern, sondern von einer Gutachtergruppe gesichtet und bewertet, die sich (so die Vermutung) aus Erziehungswissenschaftlern, pädagogischen Psychologen und Angewandten Sprachwissenschaftlern zusammensetzte. Federführend jedoch war innerhalb der Wissenschaftsstruktur der DFG die Fachkommission "Lehr-/Lernforschung", von der wir als Fremdsprachendidaktik offenbar ein spezifischer Anwendungsfall sind.

2.2 Rahmenbedingungen des Projekts

Während die drei anderen genehmigten DFG-Anträge von Bernt Ahrenholz, Wolfgang Börner und Una Dirks bereits zu einem früheren Zeitpunkt in der ZFF unter der Rubrik "Forschungsprojekte" in ihren zentralen Fragestellungen und ihrer Forschungsanlage vorgestellt worden sind (Ahrenholz 2003; Börner 2004; Dirks 2003), nimmt das Projekt "Fachlernen und (Fremd-)Sprachlichkeit" hiermit die Gelegenheit wahr, nicht nur die Forschungsfragen und das Design unserer Untersuchung vorzustellen, sondern bereits über gewonnene Forschungserfahrungen sowie erste Zwischenergebnisse zu berichten.

Die offizielle Laufzeit des Projekts betrug zwei Jahre mit nur einer halben wissenschaftlichen Mitarbeiterstelle - entgegen der ursprünglich beantragten Laufzeit von 3 Jahren und einer beantragten Personaldecke von drei halben Mitarbeiterstellen. Diese Politik der DFG im Sinne einer einheitlichen (reduzierten!) und strengen Mittelvergabe hat zeitweilig zu der Überlegung geführt, das Projekt gar nicht erst zu beginnen, denn angesichts der Forschungsfragestellungen (siehe unten) war eigentlich eine wohlbegründete interdisziplinäre Breite in der Personalstruktur erforderlich. Wegen dieser erheblichen personellen und finanziellen Einschränkungen konnte das Projekt weder in dem vorgesehenen Umfang noch in der beabsichtigten Tiefe durchgeführt werden; auch konnte es erst zeitverzögert beginnen, da die wissenschaftliche Mitarbeiterstelle neu besetzt werden musste, weil die dafür vorgesehene Person mit einer Fakultas in Anglistik und Geographie es vorzog, vor der Promotion erst ins Referendariat zu gehen. Die Tatsache, dass auf diese Weise keine geographische Kompetenz mehr unter den Projektmitarbeitern selbst vorhanden war, erwies sich allerdings zum Glück nur bedingt als problematisch, da wir eine überregionale Kooperation mit ausgewiesenen Geographieexperten erzielen konnten. Dasselbe galt für die psychologisch-statistische Beratung und Zuarbeit, die innerhalb der Universität Osnabrück durch punktuelle Kooperation mit einem Kollegen und einer studentischen Hilfskraft aus der Psychologie geregelt werden konnte.

Es wurde übrigens keine Verlängerung der Förderung für ein drittes Projektjahr beantragt, weil dies genau in die intensivste Phase der Auswertungsarbeit (ein halbes Jahr vor Ablauf der Regelförderung) gefallen und insofern kontraproduktiv gewesen wäre. Hier liegt ein gewisser Widerspruch, der zugleich Einsicht in die Schwierigkeiten von Forschungsförderung unter den heutigen Bedingungen der Drittmittelvergabe erlaubt. Glücklicherweise gab es andere Möglichkeiten der finanziellen Absicherung bzw. Weiterbeschäftigung der Doktorandinnen, so dass hier faktisch ein Dreijahresprojekt durchgezogen werden konnte.

3. Bilinguales Lernen und Bilingualer Unterricht als Forschungsgegenstand

Ausgehend von einer detaillierten Analyse der Forschungssituation im Bereich des bilingualen Lernens und Unterrichtens konnte im Projektantrag zusammenfassend festgestellt werden, dass diese Form des Bildungsganges immer noch erstaunlich wenig empirisch untersucht worden ist, insbesondere was die fachliche Vergleichbarkeit der Lernergebnisse zwischen bilingualen und monolingualen Lernern anbelangte. Zugleich gab es kaum vertiefte Erkenntnisse darüber, was es bedeutet, fachliche Aufgabenstellungen mit Hilfe einer Zweitsprache zu verstehen, zu bearbeiten und zu lösen. Insofern lag hier ein großes Theoriedefizit vor. Die wenigen empirischen Untersuchungen, die es damals gab, konzentrierten sich auf die Beschreibung und Erhebung des *fremdsprachlichen Zugewinns* durch bilingualen Unterricht. Hier ist vor allem die Dissertation von Bredenbröcker (2000) zu nennen, der die Standardbereiche fremdsprachlichen Könnens bereits in Klasse 7 zwischen bilingualen und nicht-bilingualen Lernern verglich und herausfand, dass die "Bilis" auch ohne einen intensivierten fremdsprachlichen Vorlauf (wie etwa in Schleswig-Holstein üblich) von vornherein den monolingualen Lernern in mancherlei Hinsicht in ihren fremdsprachlichen Fähigkeiten überlegen sind, allerdings nicht in allen Bereichen. Andere empirische Arbeiten, vor allem aus Schleswig-Holstein um Henning Wode (z.B. Claussen 1997; Cohrs 1999; Daniel 2001; Kickler 1995; Mukherjee 1999; Wode 1999; Wode et al. 1996), haben eine Reihe von Einsichten in die sprachlichen Strukturen und Besonderheiten bilingualer Äußerungen und Texte erbracht, diese aber weitgehend auf der Ebene der *Verwendung der Fremdsprache*, nicht jedoch im Hinblick auf die *sprachbasierte Fachlichkeit* der Produkte analysiert.

Eine erste Hinwendung zu stärker fachlich-inhaltlichen Fragestellungen bilingualen Lernens zeichnete sich durch zwei Arbeiten ab, die zum Zeitpunkt der Antragstellung noch nicht veröffentlicht vorlagen: da sind zum einen die Vorarbeiten zur Dissertation von Lamsfuß-Schenk (Universität Wuppertal) zu nennen, die sich auf französischsprachigen Geschichtsunterricht in Klasse 9 des Gymnasiums bezogen, von der Autorin selbst in einer bilingualen und einer monolingualen Vergleichsklasse durchgeführt. Der Vergleich der Fachleistungen dieser beiden Klassen erfolgte weitgehend auf der Ebene qualitativer Evidenz, basierend auf exemplarischen Inhaltsanalysen von Schüleräußerungen im Rahmen thematischer Gruppenarbeit. Es zeigte sich (zumindest hypothetisch), dass die bilingualen Lerner trotz der Schwierigkeiten mit Französisch als L2 inhaltlich angemessener und zielorientierter arbeiteten, dass sie möglicherweise gerade durch die ihnen besonders abverlangten Erschließungs-, Verarbeitungs- und Kommunikations-

strategien im Vergleich zu den Schülerinnen und Schülern im muttersprachlichen Geschichtsunterricht einen qualitativen Vorsprung im Fremdverstehen und damit im Geschichtsbewusstsein entwickeln, so dass sie fachlich nicht nur gleichwertig, sondern tendenziell sogar überlegen sind (vgl. Lamsfuß-Schenk 2002c).³ Zum anderen sind hier die Arbeiten von Andreas Bonnet (z.B. 1999, 2001) zu erwähnen, wobei die Dissertation (2004) zum damaligen Zeitpunkt erst im Entstehen war. Bonnet wendete sich den unterrichtlichen Interaktionsprozessen innerhalb von Gruppen bilingualer Chemielerner zu und entwickelte anhand dieser Daten ein unterrichtsnahes Modell der heuristischen Identifizierung von Kompetenzen fachlicher, personaler, sozialer und reflexiver Art durch Interaktion.⁴

Insgesamt war der Diskurs im Vorfeld der Antragsstellung darauf gerichtet, bilingualen Unterricht nicht mehr als reine Erweiterung des fremdsprachlichen Lernens und eben auch nicht nur als alternative Form des Fachlernens in einem fremdsprachlichen Kode zu verstehen, sondern als Integration von fachlichem und sprachlichem Lernen (vgl. den englischsprachigen Ausdruck *Content and Language Integrated Learning* (CLIL), der sich in Europa mehr und mehr durchsetzt). Auf der Suche nach einer eigenständigen Didaktik und Methodik des bilingualen Lernens und Lehrens war also genauer herauszuarbeiten, wie sich fachliches und sprachliches Lernen gegenseitig bedingen und beeinflussen und wie eine entsprechende sprachliche Stützarbeit (die nicht zu verwechseln ist mit Fremdsprachenunterricht im Rahmen von Fachunterricht!) zur Verbesserung und Präzisierung des fachlichen Denkens und Kommunizierens führen kann (vgl. dazu auch Bonnet, Breidbach & Hallet 2002; Vollmer 2005b).

Auf dem Zweiten Bremer Kongress zum bilingualen Sachfachunterricht unter dem Titel "Didaktiken im Dialog" (2003) ergab sich sodann unter allen beteiligten Koordinatoren und Sektionsleitern der Versuch, so etwas wie ein Forschungsprogramm zum Bilingualen Unterricht zu erstellen und entsprechende Desiderata für die Zukunft auszuweisen: Hieraus ergaben sich vielfältige Forschungsperspektiven für die Folgejahre, wobei das hier vorgestellte DFG-Projekt Teile der Fragestellungen, wie sie als notwendig erkannt worden waren, thematisiert und bearbeitet (vgl. Abendroth-Timmer et al. 2004; Vollmer 2004; Zydatiś 2002, 2004).

3 Die Dissertation von Lamsfuß-Schenk und damit stringente empirische Belege für ihre Hypothese liegen leider bis zum heutigen Tage nicht veröffentlicht vor; wohl aber sind einzelne Aufsätze dazu und Vorträge der Autorin auf Kongressen zugänglich und verfügbar (z.B. Lamsfuß-Schenk 2000, 2001, 2002a-d).

4 In den Folgejahren sind weitere Arbeiten zum bilingualen Lernen mit fachlichem Schwerpunkt entstanden, so z.B. Golay 2005 oder Timm 2006; vgl. auch die Beiträge in Wildhage & Otten 2003.

4. Zentrale Fragestellungen des Projekts

Das DFG-Projekt "Fachlernen und (Fremd-)Sprachlichkeit" hatte sich zum Ziel gesetzt, die fachlichen Kompetenzen von bilingualen und monolingualen Lernern am Beispiel des Faches Geographie in der Klasse 10 des Gymnasiums zu untersuchen und miteinander zu vergleichen. Dabei sollte insbesondere herausgefunden werden, ob die fachlichen Leistungen der bilingualen Schüler denjenigen der Regelschüler entsprechen, da bislang ungeklärt war, ob der Zuwachs an Fremdsprachenkenntnissen durch bilingualen Unterricht (der empirisch gut belegt ist) auf Kosten einer gediegenen Fachlichkeit der betroffenen Lerner geht.⁵

Laut Projektantrag vom 19.8.2002 sollten außerdem die Zusammenhänge zwischen mentaler Verarbeitung und (Fremd-)Sprachlichkeit und damit die Besonderheiten bilingualen Lernens herausgearbeitet und erklärt werden. Denn der bisherige Forschungsstand um die Grundlagen des Bilingualen Unterrichts weist, insbesondere in seiner theoretischen Fundierung, noch große Lücken auf. Schließlich sollten auch die Auswirkungen verschiedener Lernkontexte (individuelles vs. kooperatives Arbeiten) mit untersucht werden. Im Zentrum steht also die Frage, ob die Benutzung einer *Fremdsprache* als Arbeitssprache in fachlichen Lernzusammenhängen zu spezifischen kognitiven Anforderungen oder zu möglichen Erschwernissen bzw. Umwegen bei der Lösung von Fachproblemen führt. Dabei wird die Sprachlichkeit des Fachlernens sowohl mündlich (durch Lautes Denken bzw. verbale Aushandlungen beim Aufgabenlösen) als auch schriftlich (durch Verschriftlichung von Schülerantworten) erfasst. Dazu war es also nötig, einerseits Aufgaben zu entwickeln, mit denen fachspezifische Kognitionen (mit dem Ziel der Hypothesengenerierung) elizitiert werden konnten, andererseits benötigten wir als zentrales Erhebungsinstrument einen repräsentativen Geographietest, der die erreichten Fachkompetenzen der Untersuchungsteilnehmer (mit dem Ziel der Hypothesenüberprüfung) angemessen dokumentiert.

5. Zur Sprachlichkeit des Fachlernens

Obwohl das Projekt zentral auf die Erhebung und den Vergleich der *fachlichen* Kompetenzen der Probanden abhebt, lag der besondere Fokus aller Studien auf der *Beschreibung und Analyse der Rolle von Sprachlichkeit im Zusammenhang mit fachlichem Lernen*. Dieser Fokus kann als relativ neu und als Erweiterung der

5 Diese Vermutung wurde erst kürzlich wieder anlässlich der Veröffentlichung der DESI-Ergebnisse in der Frankfurter Allgemeinen Zeitung (vom 6. März 2006, Nr. 55, S.5) kritisch geäußert.

Perspektiven für die Fremdsprachenforschung gelten, weil er erstmals die *vermittelnde Rolle von Sprache für jegliche Form des Lernens*, insbesondere *für alles Fachlernen* thematisiert, und zwar nicht als separaten Untersuchungsgegenstand (wie es die Erforschung des Fremdsprachenunterrichts oder auch die Erfassung von (fremd-)sprachlichen Kompetenzen bilingualer Lerner darstellt), sondern *als integratives Moment des Fachlernens selbst*: Hier wird Sprache als kognitives Werkzeug, als funktionales Medium zur Erarbeitung und Erweiterung fachlicher Konzepte ebenso wie als Indikator fachlicher Kommunikationsfähigkeit verstehender wie mitteilender Art gesehen. Damit wird eine funktionalistische Dimension von Sprache gewonnen, die sie zwar immer schon hat (oder hatte), die jetzt aber in den Vordergrund tritt und die es erlaubt, stärker von inhaltlichen Bezügen, von konzeptuellen, wissenschaftspropädeutischen Denkansätzen und Problemlösevorgängen auszugehen und die Sprache als "Ausdrucksmittel" bzw. als Medium des Denkens, des Lernens und des Problemlösens zu sehen, so wie es nicht nur im bilingualen Unterricht, sondern *in jedem Fachunterricht* der Fall ist. Das Spezifikum des bilingualen Lernens allerdings und damit der Kern der in ihm zu erwerbenden Qualifikationen und Kompetenzen liegt in der notwendigen Entwicklung einer *fremdsprachigen* wissenschaftlichen Begriffsbildung komplexer, systemischer, theoriegeleiteter Art, basierend auf meist muttersprachlich vorliegenden Alltagsbegriffen oder "Minimalkonzepten", die es in doppelter Weise in taugliche Fachkonzepte der L2 zu überführen gilt, so wie es allgemein z.B. von Peschl (1999) oder van Looke (1999) und im Hinblick auf den bilingualen Unterricht perspektivisch u. a. von Hallet (2002, 2003) oder Niemeier (2005) skizziert wurde.

Damit sind die klassischen Fragestellungen der Fremdsprachenforschung in zentraler Weise erweitert: Die Hinwendung zum bilingualen Lernen und damit zum Fachlernen allgemein sowie die Fokussierung auf die konzeptuelle Vermittlungsfunktion von Sprache (von Fremdsprache in einem, von Erstsprache im anderen Fall) greift Stränge der Forschung auf, die in der Vergangenheit im Kontext beschränkter *fachsprachlicher* Analysen und Diskurse angesiedelt waren, mit einer deutlich terminologiekritischen und *genre*-bezogenen Ausrichtung. Diese Perspektive wird im Rahmen unseres Projekts aufgegriffen und ausgeweitet, indem Sprache als das vermittelnde Element zwischen spezifischen Herausforderungen, konzeptuellem Denken und der kommunikativ-diskursiven Aushandlung von Bedeutung in Interaktion mit anderen Personen in den Vordergrund der Untersuchung gerückt wird. Dabei geht es ebenso um Fachsprache wie um den Aufbau und die Aktivierung allgemeiner sprachlicher Repertoires im Übergang von Alltagssprache zu wissenschaftspropädeutischem Sprachgebrauch (vgl. Bonnet, Breidbach & Hallet 2002; Nation 2001). Dieser

erfordert einen zwar gegenstandsabhängigen, aber doch verallgemeinerten formalen, expliziten und rationalen Umgang mit Sprache, der von Jim Cummins (1978, 1979) als *Cognitive/Academic Language Proficiency (CALP)* charakterisiert worden ist. Sprache steuert dabei sowohl den mentalen Aufbau und das Verstehen fachlicher Zusammenhänge als auch deren Integration in vorhandene Wissensstrukturen, die dann ihrerseits versprachlicht werden können. Zugleich aber vermittelt Sprache zwischen den Bedeutungskonstruktionen verschiedener Individuen und deren Kommunikation untereinander, orientiert an den Traditionen und Bildungszielen des Faches und den dahinter agierenden fachspezifischen Diskursgemeinschaften bzw. wissenschaftlichen *Communities*. Es ist das große Verdienst von M.A.K. Halliday und der von ihm begründeten sozio-semiotischen Theorieschule, diese Zusammenhänge transparent gemacht und linguistisch aufgearbeitet zu haben (vgl. etwa Halliday & Hasan 1989; Halliday & Martin 1993; Lemke 1990; siehe auch Coetzee im Druck). Daneben schließt sich unsere Sichtweise an die Erkenntnisse der kognitiven Linguistik an, die sich momentan als Leittheorie in der Linguistik etabliert (Croft & Cruse 2004; Evans & Green 2006; Ritter 2005). Insofern tragen wir insgesamt mit unserem Projekt zu einem stärkeren Theoriefundament in der Fremdsprachenforschung bei, das zugleich Anknüpfungspunkte für *linguistische* Theoriebildung bietet.

Inhalt, Denken und Sprache hängen also auf engste miteinander zusammen. Dennoch bleibt der Fokus des Projekts die Dimension der Sprachlichkeit, weil sich in erster Linie über sie die eigene Expertise definiert. Im Laufe des Projekts hat es sich allerdings als unabdingbar erwiesen, sich auch auf den jeweiligen Gegenstand, in diesem Fall Geographie, nachhaltig einzulassen, und zwar stärker, als es ursprünglich angedacht oder vorgesehen war. Insofern liegt hier eine wesentliche Erfahrung vor, die es festzuhalten und weiterzuvermitteln gilt: Eine seriöse und weiter bringende Erforschung des bilingualen Lernens und Lehrens ist nur möglich auf der Basis interdisziplinärer Untersuchungsansätze bzw. einer stabilen Kooperation zwischen Sachfachdidaktik und Fremdsprachendidaktik. Ein Vergleich zwischen monolinguaalem und bilingualem Lernen erfordert darüber hinaus die Offenlegung bzw. das Insistieren auf der sprachlichen Dimension innerhalb der betroffenen Fachdidaktiken selbst, z.B. der Geographiedidaktik (vgl. Czapek 2000, 2004). Die meisten Fachdidaktiken sind jedoch trotz PISA-Schock und der Einsicht in die Allgegenwärtigkeit von erforderlicher Lesekompetenz sowie fachlicher Ausdrucks- und Mitteilungsfähigkeit leider erst sehr zögerlich dabei, die sprachliche Dimension ihres Faches bzw. allen Fachlernens (wieder) zu entdecken. So ist im Rahmen der Entwicklung nationaler Bildungsstandards die Fähigkeit zur "Fachkommunikation" als einer von vier zentralen Kompetenzbereichen durch die Biologie-, Chemie- und die Physik-

didaktik identifiziert sowie neuerdings auch von der Geographiedidaktik für sich übernommen worden (KMK 2004a-c; DGfG 2006; vgl. hierzu auch die neuerlichen Konzepte zu *Language Across the Curriculum*, Vollmer 2006d, im Druck). Allerdings steht man noch ganz am Anfang bei der Definition und Ausfüllung dieses Konzepts der Fachkommunikation, auf mündlicher Ebene ebenso wie auf der des fachlichen Schreibens, während die Entwicklung in der theoretischen Bestimmung von domänenspezifischer Leseverstehenskompetenz weiter fortgeschritten ist. Am Ende werden sich in diesem Bereich zwischen den Fächern viele Gemeinsamkeiten in der Verknüpfung von Sprache und Kognition zeigen ebenso wie Differenzen im Detail (siehe die Diskussion von Diskursfunktionen weiter unten im Abschnitt 6.3).

6. Geographischer Kompetenztest als zentrales Erhebungsinstrument

Um einen Vergleich fachlicher Kompetenzausprägung seriös durchführen zu können, mussten repräsentative Aufgabenstellungen entwickelt werden. Dies erforderte zunächst eine theoretisch differenzierte und im Feld akzeptierte Definition von geographischer Kompetenz, ihrer Modellierung und Dimensionierung (Vollmer 2006a, in Arbeit). Da ein solches Modell zu Beginn des Projektes nicht ausformuliert vorlag, waren wir auf die enge Kooperation mit geographischen Experten angewiesen und mussten den Fortschritt in der Entwicklung ihres Denkens immer wieder abwarten, konnten andererseits aber auch konkret bei dessen Entfaltung unterstützend tätig werden. In diesem Sinne hat der Projektleiter an der Formulierung eines geographischen Kompetenzmodells im Rahmen der Definition von Bildungsstandards für den Mittleren Schulabschluss mitgewirkt (vgl. DGfG 2006).

Während die Geographen am Ende sechs Kompetenzbereiche unterschieden, haben wir vier davon übernommen (und zwar *Fachwissen*, *Methodenkompetenz*, *Fachkommunikation* und *Beurteilungs-/Bewertungsfähigkeit*). *Fachwissen* wird verstanden als das deklarative Basiswissen eines Faches, als Verfügung über grundlegende Konzepte, Strukturen und Entwicklungsvorstellungen, um geographierelevante Phänomene, Ereignisse und Zusammenhänge benennen, beschreiben und erklären zu können. Dazu gehört auch die Fähigkeit zur Verknüpfung von Wissens-elementen zu konzeptuellen Netzwerken und zu komplexeren Wissens-einheiten, die anschluss- und ausbaufähig sind. Domänenspezifisch ist dies die "Fähigkeit, Räume auf den verschiedenen Maßstabsebenen als natur- und humangeographische Systeme zu erfassen und die Wechselbeziehungen zwischen Mensch und Umwelt analysieren zu können" (DGfG 2006: 5).

Methodenkompetenz ist laut Geographen die "Fähigkeit, geographisch/geowissenschaftlich relevante Informationen im Realraum sowie aus Medien gewinnen und auswerten sowie Schritte zur Erkenntnisgewinnung in der Geographie beschreiben zu können" (DGfG 2006: 5). Im Rahmen unseres Projekts bedeutete dies, aus dem Verständnis einer bestimmten Aufgabenstellung das darin enthaltene *Problem* zu definieren, relevante Informationen dazu zusammenzutragen und Schritte zur Lösung des Problems einzuleiten bzw. durchzuführen. Dabei wird ein Wissensraum aufgespannt, innerhalb dessen bestimmte Elemente durch Abruf aus dem Langzeitgedächtnis im Arbeitsspeicher bereit gestellt oder aber rekonstruiert werden können, andere dagegen aktiv erst neu geschaffen werden müssen. Entscheidend ist, ob die Informationen aus den verschiedenartigen Quellen, die zur Lösung einer fachlichen Aufgabe zur Verfügung stehen, so miteinander verknüpft werden, dass damit ein erfolgreicher Bearbeitungs- und Problemlöseprozess in Gang gesetzt wird.

Die dritte Kompetenzdimension *Fachkommunikation* lässt sich allgemein als diskursive Verwendung der Sprache in fachlichen Kontexten, als fachliche Verstehens- und Ausdrucksfähigkeit charakterisieren. Für die Geographen bedeutet dies die "Fähigkeit, geographische Sachverhalte zu verstehen, zu versprachlichen und präsentieren zu können sowie sich im Gespräch mit anderen darüber sachgerecht austauschen zu können" (DGfG 2006: 5). Diese *fachbasierte Diskurskompetenz* äußert sich u. a. im lesenden Verstehen größerer Mengen fachlicher Texte, in der angemessenen Wiedergabe oder Zusammenfassung des Verstandenen, in der verbalen Definition einer Problemstellung, im Festhalten von erarbeiteten Erkenntnissen (z.B. Textproduktion, non-verbale/graphische Darstellung von Ergebnissen und Zusammenhängen) und nicht zuletzt in der gezielten argumentativen Auseinandersetzung mit den fachlichen Einsichten anderer. Das interaktive Aushandeln von Bedeutung mit anderen Teilnehmern in einem "sachgerechten" Diskurs findet in der Regel auf mündlicher Ebene statt, während das Darlegen von Aufgabenlösungen ein differenziertes Maß an fachlicher Verschriftlichungsfähigkeit erfordert.

Schließlich gibt es noch die *Beurteilungs- und Bewertungskompetenz*. Diese besteht nach Definition der Geographen in der "Fähigkeit, raumbezogene Sachverhalte und Probleme, Informationen in Medien und geographische Erkenntnisse kriterienorientiert sowie vor dem Hintergrund bestehender Werte in Ansätzen beurteilen zu können" (DGfG 2006: 5). Auf der Basis der weiterführenden Diskussionen in unserem Projekt unterteilen wir dieses Konstrukt noch einmal in die zwei Teilkompetenzen: *Beurteilungskompetenz* bedeutet also die Fähigkeit zur konkreten Abwägung von Sachverhalten der genannten Art im Hinblick auf die Überprüfung bzw. Einlösung bestimmter

fachinterner Kriterien und ihrer Implikationen. Diese Beurteilungsfunktion ist unserer Meinung nach deutlich zu unterscheiden von einer *Bewertungskompetenz*, die darüber hinaus die Aktivierung, Anwendung und Explikation eigener, persönlicher (i. e. S. nicht mehr geographischer) *Werte und ethischer Normen der Lernenden* evoziert und erfordert.

Die anderen beiden Kompetenzbereiche der Geographen (*Räumliche Orientierung, Handlung*) kamen aus unterschiedlichen Gründen für uns nicht in Frage: Der erste kann als Teil des Fachwissens angesehen werden oder aber liegt quer zu allen anderen Kompetenzbereichen, wobei beispielsweise *topographisches Wissen* oder *Kartenkompetenz* hier mit eingeschlossen sind; der zweite schien uns kaum operationalisierbar zu sein, zumal er definitionsgemäß nicht nur die *Bereitschaft*, sondern eben auch die *Fähigkeit* umfassen sollte, "auf verschiedenen Handlungsfeldern natur- und sozialraumgerecht handeln zu könne" (DGfG 2006: 5; kritisch dazu Vollmer 2006a, 2006b, in Arbeit). Dennoch unterliegen wesentliche Teile des hier angesprochenen Modells unseren Definitionen und Kategorisierungen, der Bestimmung von Niveaustufen geographischer Kompetenz und der Ausformulierung entsprechender Deskriptoren und Bewertungsskalen. Durch diese enge Bindung an die theoretischen Vorarbeiten der Experten im Fach Geographie ist das Projekt zugleich gefördert wie auch bisweilen "gehemmt" worden. Insgesamt haben vier Geographiedidaktiker (allen voran Prof. Dr. Ingrid Hemmer, Eichstätt, und Prof. Dr. Michael Hemmer, Münster) sowie sechs Geographielehrer durch teilweise intensive Beratungstätigkeit zum Gelingen des Projekts beigetragen.

Nach mehrfacher Pilotierung wurden die entwickelten Aufgaben zu zwei parallelen "Kompetenztests" (deutsche und englische Version) zusammengestellt, die von insgesamt 125 bilingualen Erdkundeschülern beiderlei Geschlechts sowie von 187 deutschsprachig unterrichteten Erdkundeschülern aus insgesamt 18 Klassen (7 Schulen) auf Englisch bzw. auf Deutsch in 90 Minuten bearbeitet wurden. Die Aufgabenstellungen beziehen sich dabei auf das o.g. Strukturmodell geographischer Kompetenz, das vier Bereiche unterscheidet. Auf der Basis unserer Pilotdaten und unseres fachbasierten Theoriemodells haben wir noch eine fünfte Dimension postuliert, eine *metakognitive Reflexionskompetenz* oder *Strategiefähigkeit*. Diese zeigt sich jedoch weniger in den schriftlichen Erstbearbeitungen als vielmehr in den Prozessdaten sowie vor allem in den Ergebnissen der *Überarbeitung von Aufgabenlösungen*, die wir eine Woche nach den Ersterhebungen eingeholt haben (siehe dazu Abschnitt 8). Die genaueren Auswertungsschritte dazu stehen noch aus, deshalb soll mit dieser Kompetenzdimension erst später gearbeitet werden.

Die insgesamt 17 Testaufgaben variieren in ihrer Komplexität, in ihrem

jeweiligen Anforderungsniveau sowie in ihrer Ausrichtung auf die Erfassung bestimmter Kompetenzen und Diskursfunktionen: Dabei zeigte sich allerdings sehr schnell, dass die ersten drei Kompetenzbereiche eigentlich bei jeder Aufgabenstellung immer involviert sind, es gibt einen großen Überschneidungsbereich zwischen ihnen; insofern lassen sich diese im testtheoretischen Sinne nicht wirklich eindimensional fassen und damit getrennt messen. Wohl aber scheint es zu gelingen, trotz ganzheitlicher Performanzprodukte in den Aufgabenlösungen für die Leistungen in den definierten Kompetenzbereichen gesonderte Bewertungen vorzunehmen und so die vorfindbare geographische Gesamtkompetenz analytisch zu zerlegen. Hier tut sich allerdings eine Vielzahl methodologischer Probleme auf, die vor allem mit der Angemessenheit der theoretischen Annahmen, der Frage der Zuverlässigkeit der Erhebungsverfahren und der Generalisierbarkeit der Ergebnisse zusammenhängen; diese Probleme werden an anderer Stelle ausführlicher benannt und diskutiert (vgl. Coetzee-Lachmann 2006; Vollmer 2006b, im Druck).

Die Aufgaben selbst sind thematisch alle auf den *Tropischen Regenwald* bezogen als einem exemplarischen Gegenstandsbereich mit typischen fachspezifischen Inhalten, die typische Formen geographischen Arbeitens erfordern. Die Aufgaben variieren aber erheblich in ihrem Schwierigkeitsgrad und in der Bezugnahme auf vorhandenes Wissen sowie in der Vorgabe von Materialien verbaler und nicht-verbaler ("diskontinuierlicher") Art, so dass sich die erwarteten Rekonstruktionsleistungen der Lerner und die erforderlichen mentalen Aktivitäten bei der Neukonstruktion von Wissen bzw. bei der Erstellung von Inferenzen von Aufgabe zu Aufgabe erheblich unterscheiden. Wesentlicher Bestandteil zur Bewältigung der Aufgaben ist die Informationsentnahme aus unterschiedlichen Materialien (Karten, Diagramme, Graphen, Tabellen, Statistiken, Texte usw.). Alle Aufgaben zusammen implizieren komplexe, wissensintensive Anforderungen, die Problemcharakter haben; sie können nicht durch den Abruf von Gedächtnisinhalten allein unmittelbar gelöst werden. Vielmehr werden in ihnen unterschiedliche Aspekte des fachlich-konzeptuellen Wissenseinsatzes und der Neukonstruktion von Wissen angesprochen. Alle Aufgaben enthalten auch in unterschiedlichem Maß die Notwendigkeit zum fachbasierten Sprachverstehen und zur produktiven Versprachlichung. Die Aufgaben haben also nicht den Charakter von Sprachlernaufgaben, da der Fokus eindeutig auf fachtypischen Erdkundeansforderungen liegt; nichtsdestotrotz ist die Sprache das Medium für die Verarbeitung und Darstellung dieser Informationen. Somit kommt vorhandenes sprachliches Wissen massiv zum Einsatz. Es wird jedoch (wenn man von der Einführung einzelner Termini wie z.B. *sustainability* einmal absieht) nicht gezielt neu aufgebaut. Zum Zwecke der Illustration sollen

im Anhang zwei Aufgabenbeispiele kurz skizziert und abgebildet werden. Eine Darstellung aller Aufgaben im Überblick findet sich bei Vollmer (2006a).

Insgesamt beruhen die Testaufgaben in ihrer spezifischen Mischung auf einer Charakterisierung, die vorab auf der Basis eines eigens entwickelten Modells fachbasierter kognitiver Aufgabenanalyse erstellt wurde. Dabei wurde ermittelt, welche fachlich-konzeptuellen Wissensbestände und welche Abruf- und Neukonstruktionsprozeduren zu einer erfolgreichen Bearbeitung und Lösung der jeweiligen Aufgabe nötig sind. Dies ist hilfreich zum einen für die Entwicklung der Aufgaben selbst, zum anderen für die Analyse der Laut-Denk-Protokolle sowie der schriftlichen Lösungsprodukte.

6.1 Modell kognitiver Aufgabenanalyse

Grundlage für diese Aufgabenanalyse war wiederum ein Modell mentaler Strukturen und ihrer Veränderungen im Kontext von Problemlösehandlungen. Dies impliziert eine Vorstellung davon, was Wissen und Denken in fachspezifischen Kontexten bedeutet und wie man zu Analyse Zwecken eine begründete Typisierung vornehmen kann (siehe Heine in Arbeit). In das Kognitionsmodell fließen Erkenntnisse aus der Kognitionspsychologie ein, insbesondere aus der Problemlöseforschung sowie ihrer Adaption in der Fachdidaktik (Anderson et al. 2001; v. Aufschnaiter & v. Aufschnaiter 2003). Diese fachbasierte kognitive Aufgabenanalyse lässt sich überblicksartig wie folgt darstellen (vgl. Heine 2006):

1. *Verstehen und Aneignen der Aufgabe*: Welche Informationen liefert die Formulierung der Aufgabenstellung, welche können direkt aus den präsentierten Materialien entnommen werden? Welche Informationselemente davon sind relevant, welche vernachlässigbar (z.B. bei der Informationsentnahme aus verschiedenartigen Informationsquellen)? Wie offensichtlich wird auf relevante Informationen in der Aufgabe hingewiesen?

2. *Einsatz von Vorwissen*: Welche fachspezifischen Wissensbestandteile müssen die Lerner von sich aus aktivieren bzw. *rekonstruieren*, damit sie die Aufgabe verstehen und lösen können?

3. *Neue Erkenntnisgewinnung*: Welche fachspezifischen Wissensbestandteile müssen von den Lernern zur Lösung der Aufgabe *neu konstruiert* werden? Welche kognitiven Prozesse müssen für diese Neukonstruktion von Bedeutung ausgeführt werden?

4. Welche *inhaltlichen Wissensbestandteile* müssen am Ende *unabdingbar vorhanden* bzw. in der Aufgabenlösung *sprachlich angezeigt* sein?

Auf der Basis einer entsprechend ausdifferenzierten, fachspezifischen Kognitionsanalyse, die es in dieser Weise vorher nicht gab, konnten dann

Musterlösungen für die einzelnen Aufgaben erstellt werden, mit mehrfacher Validierung durch verschiedene Gruppen von Fachleuten: Dabei war zu entscheiden, welche fachlichen Bestandteile eine adäquate Lösung der jeweiligen Aufgabe *mindestens* aufweisen muss, um noch akzeptabel zu sein (Mindestniveau). Welche inhaltlichen Komponenten gehören dagegen zu einer *vollständigen* (guten bis sehr guten) Antwort (höchste Niveaustufe)? Wie lässt sich die Niveaustufe dazwischen (*mittleres Niveau*) inhaltlich bestimmen? Worin liegen möglicherweise besondere *Zusatzleistungen*, die gesondert bewertet werden müssen?

Eine solche Analyse macht selbstverständlich keine Annahme darüber, welche Kognitionsprozesse in welcher Reihenfolge konkret durchlaufen werden, noch darüber, wie einfach oder schwierig eine bestimmte Schrittfolge im Rahmen der Lösung einer bestimmten Fachaufgabe ist: Dies ist für unterschiedliche Lerner offensichtlich unterschiedlich, wobei auch motivationale Einflüsse eine entscheidende Rolle spielen. Die *Komplexität der Anforderungen* ist also nicht allein in der Aufgabenstellung zu suchen, sondern ebenso in den Einstellungen eines Lerners sowie in seiner Interaktion mit einer bestimmten Aufgabe, in seinen mentalen Voraussetzungen, seiner kognitiven Ausstattung und diskursiven Vorerfahrung, in seiner aktuellen Aufnahmebereitschaft, Befindlichkeit u.a.m. Wir haben es hier mit einer wichtigen Schnittstelle zwischen den Einsichten der Kognitionswissenschaften und der Filterung, Umsetzung und Anwendung dieser Theoriebasis in pädagogischen Kontexten zu tun. Die Kategorisierungen, wie sie im Diskurs vieler Fachdidaktiken zur kognitiven Charakterisierung einzelner Aufgabentypen häufig verwendet werden oder sogar üblich sind (z.B. Unterscheidung von drei Anforderungsniveaus wie Reproduktion, Analyse, Stellungnahme) und die sich primär auf die Interpretation des zentralen Operators (Verbs!) in der Aufgabenstellung selbst richten (wie z. B. Beschreibe, Vergleiche, Beurteile) - diese Kategorisierungen implizieren nach unserer Erkenntnis keineswegs eine Hierarchie im Sinne von zunehmend komplexeren Niveaustufen. Vielmehr gilt, dass jede dieser Kategorien zunächst unterschiedliche Herausforderungen und Anforderungen umfasst und in sich selbst hierarchisch zu stufen ist, und zwar je nach Aufgabenverständnis, nach Aufgabeninhalt und nach dem Niveau der geforderten oder tatsächlichen Realisierung. Mit anderen Worten: Die Stufungen liegen nicht *zwischen* den Kategorien, sondern müssen *innerhalb* einer jeden Kategorie vorgenommen werden: Beispielsweise sind Beschreibungsvorgänge sowohl auf einfachster Stufe als auch auf mittlerer Stufe möglich, ebenso aber auch auf höchster Anforderungsstufe denkbar, etwa wenn sie sich auf einen komplexen Inhalt wie die Struktur der Vegetation im Tropischen Regenwald oder seines Nährstoffkreislaufs unter Berücksichtigung von Einflussgrößen

wie Feuchtigkeit, Lichteinfall, Temperatur oder Tierleben beziehen. An dieser Stelle widersprechen wir also den in der schulischen Praxis gängigen Vereinfachungen oder Modellen von Hierarchisierung, die meistens auf der Basis der Bloomschen Taxonomien in der einen oder anderen Variante aufgebaut sind. Dagegen halten wir eine Differenzierung innerhalb ein und derselben Diskurskategorie für angemessener und wissenschaftlich haltbarer und werden an anderer Stelle diese Auseinandersetzung offensiver aufgreifen (vgl. u.a. Heine in Arbeit; Vollmer in Arbeit).

6.2 Zur Validität der Kompetenzaufgaben

Die 17 Aufgaben beanspruchen zwar nicht die Abbildung der gesamten Domäne Geographie, aber eben doch wesentlicher Aspekte davon. Die massiven theoretischen Investitionen in die Entwicklung der Aufgaben sowie ihre intensive Erprobung und partielle Verbesserung hat ihre Angemessenheit und Qualität weitgehend sichergestellt. Insofern ist ihr Gültigkeitshorizont unstrittig, wenn auch *nicht im curricularen Sinne*; denn der Regenwald wird als Thema in der Regel in Klasse 8 verbindlich durchgenommen, liegt also zum Zeitpunkt der Erhebung in der unterrichtlichen Behandlung fast zwei Jahre zurück. Andererseits wird das Thema Tropischer Regenwald in Klasse 11 des Gymnasiums in vielen Bundesländern im Sinne eines Spiralcurriculums auf höherer Abstraktionsebene wieder aufgegriffen, was seine zentrale Bedeutung innerhalb des Faches unterstreicht. Insofern lässt sich argumentieren, dass die Relevanz der verwendeten Aufgaben gegeben ist und dass ihre Auswahl (auch nach Auskunft verschiedener Expertengruppen wie Fachlehrern und Fachdidaktikern) weitgehend repräsentativ ist; sie stellen allerdings *keine Überprüfung der aktuellen Leistungen* (im Sinne der Erreichung jahrgangsspezifischer Ziele) dar, sondern zielen auf die *Erfassung von grundlegenden erdkundlichen Kompetenzen* im oben definierten Sinne, unter Einschluss von basalem Wissen im Fach, über das ein Lerner am Ende der Klasse 10 verfügen sollte (vgl. hierzu die Diskussion um Bildungsstandards, z.B. Vollmer 2005c, 2006a, fachbezogen auch DGfG 2006). Es handelt sich also um einen Aufgabentyp fachlich-sprachlicher Performanz, der nicht dazu entworfen wurde, um systematisch Lernen auszulösen oder anzuregen, sondern der gezielt vorhandenes Wissen und Können, d.h. also vorliegende fachliche Fähigkeiten, Kenntnisse und Fertigkeiten, überprüfen will, auch mit dem Ziel der Abschätzung oder Prognose zukünftiger Leistungsfähigkeit. Wir haben es hier mit einem Elizitationsinstrument zu tun, das am Ende von Klasse 10 des Gymnasiums im Fach Geographie Aufschlüsse über jeweils vorliegende Kompetenzprofile und Niveaustufungen geben soll, unabhängig davon, wie die untersuchten Kompetenzen im Einzelnen erworben worden sind.

Die genaueren Schritte zur Konstruktion und Pilotierung der einzelnen Aufgaben sowie deren Zusammenführung zu einem Kompetenztest (primär nach Gesichtspunkten der inhaltlichen Progression, der Streuung von Anforderungsniveaus und kognitiver Komplexität sowie der Abdeckung verschiedener Diskursfunktionen) werden an anderer Stelle ausführlich dargelegt (Vollmer im Druck). Auch die inhaltliche Breite der einzelnen Aufgaben sowie deren thematische Strukturierung ist anderweitig bereits beschrieben und veröffentlicht worden, so dass hier nicht näher darauf einzugehen ist (Vollmer 2006b; Coetzee im Druck; vgl. die beiden Beispielaufgaben im Anhang).⁶

6.3 Änderungen im Strukturmodell geographischer Kompetenz

Trotz des Vorhandenseins eines begründeten theoretischen Kompetenzmodells zeigte sich im Rahmen der Pilotierung sehr bald, dass die einzelnen Kompetenzbereiche nicht wirklich in sich isolierbar sind und als Einzeldimensionen kaum getrennt überprüft werden können (wie in der klassischen Testtheorie postuliert). Vielmehr haben wir nach Inspektion der ersten Daten feststellen müssen, dass die Aufgabenperformanzen der hier genannten Art immer ein Zusammenspiel aus fachlichem Wissen, methodischem Können und fachbasierter Kommunikationsfähigkeit darstellen, akzentuiert jeweils durch die eingeforderte Aufgabenstellung und die ihr zugrunde liegende Diskursfunktion wie Beschreibung, Erläuterung, Evaluation, Stellungnahme usw. (vgl. Vollmer im Druck). Die drei genannten Kompetenzbereiche Fachwissen, Prozedurale Kompetenz und Fachkommunikation interagieren beim Lösen von Fachaufgaben offenbar hochgradig miteinander, allerdings in unterschiedlicher Gewichtung bei verschiedenen Aufgaben. Insofern handelt es sich vermutlich nicht um von einander unabhängige Dimensionen, sondern um Kompetenzaspekte, die sich zwar theoretisch und analytisch trennen lassen, die jedoch im Prozess der Lösung von Aufgaben notwendigerweise immer gemeinsam aktiviert werden und sich gegenseitig bedingen.

Dies soll an einem Beispielfall deutlich gemacht werden: Wenn ein Schüler zu entscheiden hat, ob der Rückgang des Tropischen Regenwalds innerhalb eines

⁶ Es sei nur darauf verwiesen, dass wir uns im Vorfeld der Aufgabenentwicklung durch umfangreiche Hospitationen in vielen Schulen Niedersachsens mit der Praxis des bilingualen Geographieunterrichts, mit dem dort vorherrschenden intellektuellen Anspruchsniveau und den dominanten Aufgabenformaten intensiv vertraut gemacht haben. Der Autor selbst hatte in den 90er Jahren sechs Monate lang an einem Osnabrücker Gymnasium eigenständig bilingual unterrichtet, allerdings im Fach Politik/Sozialkunde in den Klassen 9 und 11. So konnten auf Grund der vorliegenden Fakultas punktuell auch im Rahmen der Pilotphase einige Unterrichtsstunden gezielt übernommen und im Hinblick auf die zu entwickelnden Kompetenzaufgaben genutzt werden, was sich als außerordentlich hilfreich erwies.

bestimmten Kontinents größer ist als in einem anderen, so wird er diese Frage nur beantworten können, wenn er zum einen über bestimmte Fachwissensbestandteile verfügt (z.B. "was ist ein Kontinent?"), zum anderen, wenn er entsprechende Informationen aus dem mitgelieferten Diagramm (Prozentangaben) sowie aus den vorgegebenen Statistiken (absolute Zahlen) richtig entnehmen und miteinander vergleichen kann. Drittens muss er diesen Vergleich in einer Weise versprachlichen können, die sein Problemverständnis auf dem Hintergrund seines Fachwissens wie seiner prozeduralen Kompetenz indiziert.

Insofern war hier eine theoretische Korrektur bzw. Neubestimmung unseres Kompetenzmodells vorzunehmen. Anstelle des ursprünglichen Kompetenzmodells musste eine Modellierung treten, welche die Eigenständigkeit der drei Komponenten ebenso zum Ausdruck bringt wie deren integrative Vernetzung. Als vorläufige Lösung des Problems bieten wir die Postulierung einer zweiten Kompetenzebene unterhalb der drei Basisdimensionen an; diese bezieht sich auf unterschiedliche *Diskursfunktionen* als integrativer Einheit von Inhalt, Denken und Sprechen (vgl. auch Zydatiß 2005). Demnach ließen sich die Aufgaben und Aufgabenarten ebenso wie die erforderlichen Kompetenzen zu ihrer angemessenen Bearbeitung verstärkt auf der Basis der jeweils dominanten Diskursfunktion(en) charakterisieren. In diesem Sinne sind *Beschreibungsaufgaben* deutlich zu unterscheiden von *Erklärungsaufgaben* und von Aufgaben, die eine *Stellungnahme* erfordern oder eben auch von solchen Aufgaben, die eine *Beurteilung* konkreter Sachverhalte im Hinblick auf die Einlösung bestimmter fachinterner Kriterien implizieren. Diese *Beurteilungsfunktion* ist noch einmal deutlich zu unterscheiden von einer *Bewertungsfunktion*, die darüber hinaus die Aktivierung und Anwendung eigener Werte und Normen durch die Lernenden erforderlich macht.

Damit ist zugleich die ursprünglich vierte Kompetenzdimension "Beurteilungs- und Bewertungsfähigkeit" mit seinen zwei Teilkompetenzen in das Raster der Diskursfunktionen aufgelöst worden und konnte mit je spezifischen Aufgaben verknüpft und überprüft werden, so dass lediglich die *drei Basisdimensionen geographischer Gesamtkompetenz* (Fachwissen, Methodenkompetenz, Kommunikationsfähigkeit) verbleiben. Andererseits haben wir mit dem Konzept der Diskursfunktionen eine Gliederung und Ausdifferenzierung geographischer Kompetenzen gewonnen, die unterhalb der drei Basisdimensionen liegen, damit aber zugleich performanznäher und langfristig auch curricular sowie unterrichtspraktisch relevanter sind. Im Rahmen der Auswertungsverfahren ließen sich die verschiedenen Kompetenzaspekte analytisch klar von einander trennen, ebenso aber auch wieder zusammenfügen, um zu einer Gesamtaussage über das Niveau vorhandener Geographiekompetenz zu einem bestimmten Messzeitpunkt zu kommen.

Der Begriff *Kompetenzaufgabe* sollte aber zunächst unter Vorbehalt benutzt werden, weil die Aufgaben (noch) nicht gezielt auf definierte oder gar empirisch abgesicherte Standards bezogen werden können, um diese in ihrer jeweiligen Ausprägung zu erfassen. Unser Projekt hat dennoch versucht, das Ensemble aller entwickelten und eingesetzten Aufgaben so zu fassen, dass damit wesentliche Aspekte geographischer Kompetenz eingefangen und dokumentiert werden konnten. Die Definition der zu erfassenden Gesamtheit und die Kriterien der Repräsentativität der Aufgaben sind eindeutig erfüllt und damit die der Validität, d.h. der tatsächlichen Erfassung jener Kompetenzen, um die es hier geht bzw. die postuliert worden sind. Allerdings bleibt, was die tatsächliche Substruktur geographischer Kompetenz anbelangt, weiter abzuwarten, ob sie sich auf Grund der Auswertung der empirischen Daten im Rahmen von faktorenanalytischen Verfahren bestätigen lässt oder nicht. Was die Zuverlässigkeit der Erhebung und des empirischen Vorgehens im Einzelnen anbelangt, so ist dies an anderer Stelle bereits ausführlich dargelegt worden (Coetzee im Druck; Vollmer im Druck). Ein gewisses Problem lag in der Länge des Tests, der in 90 Minuten nicht von allen Probanden vollständig bearbeitet wurde. Damit sind die hinteren Aufgaben nur zum Teil gelöst worden, was die Zahl der auswertbaren Antworten im Hinblick auf die Bestimmung von geographischer Gesamtkompetenz reduzierte. Eine systematische Variation in der Positionierung der einzelnen Aufgaben konnten wir jedoch aus zeitlichen und testpraktischen Gründen nicht realisieren. Andererseits war die Größe der verbleibenden Hauptstichprobe mit (nahezu) vollständiger Aufgabenbearbeitung immer noch ausreichend ($n=90$ Monos, $n=84$ Bilis), um die vorgesehenen Modellanalysen daran durchzuführen. Für eine *vertiefte* Analyse insbesondere des Fachwissens und der fachlichen Diskurskompetenz haben wir uns sogar auf sechs von 17 Aufgaben beschränken können, und zwar auf solche, zu denen am meisten Fachtext produziert wurde. Was die Objektivität der verschiedenen Auswertungsmaßnahmen angeht, so sollen diese im Folgenden genauer skizziert werden. In diesem Zusammenhang spielen die Erwartungen, die Konventionen und Bezugsnormen des Faches und der mit ihnen verbundenen wissenschaftlichen Diskursgemeinschaften eine große Rolle (vgl. dazu auch Coetzee im Druck; Vollmer im Druck).

7. Auswertungsansätze

7.1 Zum Verhältnis von geographischer Kompetenz und Performanz

Es ist hier nicht der Platz, um die Auswertungsprozeduren zur Beurteilung des demonstrierten Fachwissens bzw. *der jeweils konstruierten Bedeutung*, der vorhandenen *Methodenkompetenz* bzw. des prozeduralen Vorgehens und der *Fähigkeit zur Fachkommunikation* bzw. zur angemessenen sprachlichen Realisierung der schriftlichen Antwort genauer darzulegen und zu erörtern. Sie erfolgten zum einen auf dem Hintergrund des skizzierten fachbasierten Kognitionsmodells, in holistischer ebenso wie in analytisch kleinschrittiger Weise. Zum anderen wurde methodische Kompetenz aus der Art und Weise der Nutzung der Materialien sowie auf der Grundlage der Gestaltung der Aufgabenlösungen inferiert. Als Indikator für fachkommunikative Kompetenz wurde u.a. pauschal die Angemessenheit in der Realisierung der erforderlichen Diskursfunktion pro Aufgabe gewertet. Das Schwierigste war und ist ohne Zweifel die Entwicklung und Beschreibung von Niveaustufen innerhalb der drei Bereiche zum Zwecke der Einordnung/Zuordnung der jeweiligen Schülerantworten, die entweder im definierten Standardbereich oder aber oberhalb bzw. unterhalb der gesetzten Erwartungsnorm liegen können. Ziel war die Entwicklung möglichst klarer Deskriptoren, die semantisch relativ eindeutig sind und die einen möglichst hohen Grad an Übereinstimmung zwischen verschiedenen Bewertern bei der Beurteilung solcher Aufgabenlösungen (*inter-rater reliability*) sowie ein hohes Maß an Akzeptanz innerhalb der wissenschaftlichen bzw. didaktischen Bezugsgemeinschaften gewährleisten. Insgesamt war die Antizipation solcher differenzierten Bewertungsverfahren ein notwendiger Teil der Aufgabenkonstruktion selbst.

Für die Leser der ZFF mag es von besonderem Interesse sein, dass wir neben einer holistischen Rating-Skala zur Erfassung von Fachkommunikation insgesamt *neun weitere Maße zur Bestimmung der fachspezifischen Diskurskompetenz* (jeweils Skalen mit vier Ausprägungen) für genauere Analysezwecke entwickelt haben. Diese lassen sich in *drei* Gruppen zusammenfassen: 1. Realisierungsaspekte der erforderlichen Diskursfunktion, 2. Logik und textuelle Strukturierung der Aussage, 3. Angemessenheit der Sprachebene und des benutzten sprachlichen Inventars. Diese Skalen, die in der Forschungsliteratur bezogen auf Fachsprache so nicht vorlagen und die eine ausgeprägte fachspezifische Adaption allgemeiner Sprachskalen erforderten, sind mit hohem theoretischem Aufwand und viel Akribie von Debbie Coetzee auf Englisch entwickelt und in deutscher

Übersetzung dann entsprechend übernommen worden (vgl. Coetzee im Druck, in Arbeit; Möller in Arbeit).

Bleibt noch die Differenz zwischen Kompetenz und Performanz zu thematisieren: Der Kompetenzbegriff von Weinert (2001a, 2001b), wie er im Projekt verwendet wurde, ist bekanntlich ein relativ funktionaler, d. h. es wird nicht stringent unterschieden zwischen der Bedingung der Möglichkeit für reales symbolisches Handeln (Disposition) und der Realisierung dieser Handlungspotentiale innerhalb einer bestimmten Domäne, zur Bearbeitung einer bestimmten Aufgabenstellung oder zur Lösung eines bestimmten definierten Problems (Performanz). Die Frage, ob man von der Beobachtungsebene auf die postulierte Kompetenzebene eindeutig rückschließen kann, stellt sich also zunächst nicht, denn es wird dem Anspruch nach Kompetenz in Performanz, d.h. also in seiner performativen Erscheinungsweise erfasst. Auch die Halliday-Schule macht keinen Unterschied zwischen Kompetenz und Performanz: Bei ihr steht die Beschreibung und Erklärung von Sprache-in-Gebrauch, also der Primat der zwischenmenschlichen kommunikativen Interaktion im Vordergrund des Interesses (wer produziert Texte für wen in welcher Absicht und unter welchen Bedingungen?). Kompetenz entwickelt sich in und durch bedeutsame Interaktion (*meaningful interaction*). Da ohnehin immer nur Zugang zu dem tatsächlich demonstrierten sprachlichen *Handeln* möglich ist, lässt sich eine Generalisierung im Hinblick auf die Kompetenz (Wissensdisposition) eines Lerners nur dann vornehmen, wenn hinreichend viele Aufgaben vorliegen, wenn diese Aufgaben inhaltlich und anforderungsmäßig genügend breit gestreut sind, wenn diese Aufgaben in ausreichender Anzahl tatsächlich angegangen und gelöst worden sind und wenn sie schließlich hinreichend zwischen Lernern differenzieren. Die Aufgaben müssen sich jedoch auf basales Wissen des Faches beziehen, auf inhaltliche Basiskonzepte ebenso wie auf potentiell transferfähige Leistungen, sie müssen anschlussfähiges Wissen abfragen und nicht isolierte Wissens-elemente, sie sollten schließlich Elemente zukünftiger Anforderungen enthalten.

Theoretisch sollte man jedoch auch weiterhin zwischen Kompetenz (im basalen Sinne) und Performanz (im oberflächlichen Sinne) unterscheiden, und zwar sowohl im Bereich des inhaltlich-konzeptuellen Wissens als auch des prozeduralen Wissens ebenso wie im Bereich der Fachkommunikation, wobei klar ist, dass das gespeicherte (potentielle) Wissen immer nur in dem Maße aktiviert und zur Verfügung gestellt wird, wie es aufgabenabhängig jeweils benötigt wird. Die klassische Unterscheidung in der Testtheorie zwischen zugrunde liegenden (nicht beobachtbaren) *latent traits* und gemessener (beobachtbarer) Testleistung ist in dieser Weise nicht aufzuheben. Die Frage der Stabilität dieser als überdauernd angenommenen Merkmale menschlichen

Verhaltens bzw. das Verhältnis zwischen ihnen und den demonstrierten, tatsächlichen Verhaltensweisen zu einem bestimmten Zeitpunkt X spielen allerdings in den neueren Ansätzen zur Kompetenzdefinition keine ausgeprägte Rolle mehr - sie werden sogar vielfach gleich gesetzt.

7.2 Zur Erhebung und Auswertung von Hintergrunddaten

Um den vorgesehenen Vergleich zwischen den beiden Lernergruppen (monolingualen versus bilingualen Lernern) seriös gestalten zu können, benötigten wir Hintergrundinformationen über die jeweiligen Schülerinnen und Schüler, die mit Hilfe eines umfassenden Schülerfragebogens sowie schulbezogener Informationen erhoben wurden. Diese richteten sich zum einen auf die Notenentwicklung in ausgewählten Fächern der Sekundarstufe I, zum anderen auf Lernervariablen wie Leistungsbereitschaft, Motivationsstruktur, Fachinteresse, soziale Herkunft und familiärer Bildungshintergrund. Denn es sollten nur solche monolingualen und bilingualen Lerner miteinander verglichen werden, die von ihrem Profil und ihrer sozial-kognitiven Ausstattung her auch tatsächlich vergleichbar sind. Neben den Fragebögen wurden drei ausgewählte Subtests eines standardisierten *Kognitiven Fähigkeitstests* (KFT) eingesetzt, wie sie auch in der PISA- sowie jüngst in der DESI-Studie verwendet wurden. Die Auswahl dieser drei Subtests war insofern gerechtfertigt, als es in unserer Untersuchung nicht um Individualdiagnostik, sondern um Gruppenvergleiche im Hinblick auf kognitive Grundfähigkeiten geht. Schließlich wurde auch die globale Sprachfähigkeit in Deutsch als "Muttersprache" (zumindest für die meisten der untersuchten Schüler) sowie in Englisch als erster Fremdsprache kontrolliert, jeweils mit Hilfe eines C-Tests, bei dem die Schüler und Schülerinnen in vier kurzen Texten Lücken zu schließen hatten. Dieses Instrument hat sich vielfach als ein gültiger Indikator nicht nur für die Fähigkeit zur Textrekonstruktion, sondern für die allgemeine Sprachkompetenz in der jeweiligen Sprache erwiesen (z.B. Grotjahn 2006). Für Deutsch als Muttersprache ist es allerdings bisher kaum zum Einsatz gekommen (vgl. dazu jetzt Aguado, Grotjahn & Schlak im Druck).

Schließlich konnten auch die Lehrer der untersuchten Klassen in einem für sie entwickelten Fragebogen zu ihren Klassen, zu ihrem Unterricht, zu den einzelnen Schülern, zur Klasse als Ganzem und zu den Besonderheiten des Erdkundelernens im bilingualen bzw. monolingualen Kontext Stellung beziehen. Insgesamt kamen also je sieben Erhebungsinstrumente pro Probandengruppe (in L1 bzw. L2) zum Einsatz, wobei die Konstruktion der beiden Testhefte zur Erfassung der unabhängigen Variablen "Geographische Kompetenz" ohne Zweifel am aufwendigsten war (vgl. den Teilabdruck der Aufgaben in Vollmer 2006b). Im Rahmen dieses Überblicksartikels können die Entwicklung und der Einsatz der

erwähnten Instrumente sowie die Verfahren zur Auswertung der damit gewonnenen Daten leider nicht ausführlicher beschrieben werden; dazu sei u.a. auf Vollmer (2006c) sowie auf spätere Veröffentlichungen (z.B. Coetzee, Heine, Troschke, Vollmer, alle in Arbeit) verwiesen. Lediglich eine der Zusatzvariablen soll an dieser Stelle wegen ihrer Zentralität etwas genauer skizziert werden.

8. Editing als Prozess und Produkt der Selbstkorrektur fachlicher Mitteilung

Es war von vornherein beabsichtigt, den beteiligten Probanden die Möglichkeit der *Überarbeitung von schriftlichen Antworten* (zumindest für einen Teil der Aufgaben) einzuräumen. Diese Perspektive ergab sich u. a. aus der prozessorientierten Schreibforschung (z.B. Flower & Hayes 1981; Bereiter & Scadamalia 1987, 1993), in der deutlich wurde, dass ein zweites Hinsehen, ein "zweiter Versuch" stärker an die tatsächlich vorliegenden Fähigkeiten der Probanden heranführte. Im Projekt sollte dabei erprobt werden, ob sich dazu die eigenen Antworten oder die von anonymen Mitschülern am besten eigneten. Entsprechende Pilotierungen ergaben, dass überhaupt nur die Überarbeitung der *eigenen* Produkte zu einer ausreichenden Motivierung für solche den Schülern relativ unvertrauten Handlungen führten. Darüber hinaus experimentierten wir über mehrere Stufen in der Pilotphase mit verschiedenen Stimuli; die effektivsten Ergebnisse konnten mit einer eindeutig kommunikativen Einbettung des *Editing* erzielt werden: Die Schüler wurden aufgefordert, ihre schriftlichen Texte noch einmal daraufhin zu überprüfen, ob sie einem Mitschüler, der in der entsprechenden Unterrichtsstunde nicht anwesend sein konnte, klar und ausreichend vermitteln, worum es (fachlich) geht, was die Lösung oder die richtige Antwort zu der gestellten Frage ist bzw. was die Einsicht oder die Position des Schreibers ist.

In der Untersuchungsanlage ist sichergestellt worden, dass eine Auswahl wichtiger Aufgaben von den Lernern *mit einer Woche Abstand* und ausreichendem Zeitpolster überarbeitet werden konnte, soweit sie es denn für nötig erachteten. Dabei ist auffällig, dass die Bandbreite von Korrekturmaßnahmen außerordentlich groß ist: Sie reicht von orthographischen Fehlerbehebungen, vom Ersetzen einzelner Wörter oder Wortgruppen bis hin zum völligen Neuschreiben des Antworttextes. Mit diesem zusätzlichen Untersuchungsansatz innerhalb des Forschungsdesigns ergeben sich, wie erwartet, hochgradig interessante Einsichten in die kognitiven Metastrukturen der beteiligten Lerner. Gerade in dem so motivierten Korrekturverhalten treten so etwas wie strategische Kompetenz bzw. fachliche wie sprachliche Reflexions-

fähigkeit deutlich zu Tage, so dass man von den Überarbeitungsprodukten bis zu einem gewissen Grade zugleich auf die Präsenz, das Ausmaß und die Stärke der ursprünglich fünften Dimension von Fachkompetenz schließen kann. Die Auswertung dieser Daten verspricht hochgradig interessante Einsichten und Befunde, über die hier jedoch noch nicht im Detail berichtet werden kann (siehe aber Coetzee in Arbeit; Möller in Arbeit).

9. Zwischenbilanz: Bisherige Forschungsergebnisse

Im Folgenden sollen einige der wichtigsten Untersuchungsergebnisse präsentiert werden, wie sie sich bisher ergeben haben. Dabei muss nachdrücklich auf deren *Vorläufigkeit* hingewiesen werden, da die Auswertungen zunächst nur an einer Teilstichprobe durchgeführt wurden und noch nicht abgeschlossen sind. Die bisherigen Befunde lassen sich in zehn Punkten zusammenfassen.

1. Vergleichbare Ausprägung von Fachkompetenz bei bilingualen und monolingualen Schülerinnen und Schülern

Wichtigstes Ergebnis des Projekts ist zunächst der quantitative Befund, dass die Fachkompetenz von bilingualen Lerner in Geographie am Ende von Klasse 10 einschließlich ihrer Methodenkompetenz vergleichbar hoch (bzw. niedrig) ausgeprägt ist wie diejenige von Regelschülern, und das, obwohl sich die Bilingualen weitgehend der Fremdsprache als Arbeitssprache bedienen. Dies ist insofern kein Nachteil für ihr fachliches Lernen, wie wir weiter unten illustrieren werden. Ein endgültiges, empirisch eindeutig abgesichertes Urteil steht allerdings noch aus - ebenso wie ein Vergleich von Erstbearbeitung und Überarbeitung (Editing) bestimmter Schülerantworten, die weitere Einblicke in den vorliegenden Grad an geographischer Gesamtkompetenz einschließlich fachlicher Reflexionsfähigkeit bei den Lernern versprechen (vgl. Coetzee in Arbeit; Möller in Arbeit).

2. Einzelbefunde: Noten, Sprachfähigkeit, kognitive Ausstattung der Lerner

Im Projekt sind sämtliche Noten seit Klasse 7 für die einschlägigen Fächer erhoben worden, und zwar für alle Halbjahre. Im Hinblick auf den Zusammenhang von Sprachfähigkeitstests und Fachnoten ergibt sich das erwartete Bild: Die Note in Deutsch korreliert tatsächlich mit den Ergebnissen des C-Tests Deutsch, und zwar auf hoch signifikantem Niveau. Dasselbe gilt für die Note Englisch am Ende des 1. Halbjahres von Klasse 10 und den C-Test Englisch: Hier ist der Korrelationskoeffizient sogar noch höher, nämlich .533 auf dem Signifikanzniveau von 0,01 (2seitig). Damit kann die vorliegende Notengebung im Großen und Ganzen als valide gelten. Besonders interessant ist nun aber die Tatsache, dass die Ergebnisse des C-Tests Deutsch auch hochgradig mit der Note

in Englisch korrelieren, so wie umgekehrt die Ergebnisse des C-Tests Englisch hochgradig mit dem C-Test Deutsch und der Note Deutsch korrelieren. Am höchsten allerdings ist der Zusammenhang zwischen den beiden Sprachfähigkeitstest (*Language Proficiency Tests*) untereinander ausgeprägt. Dieser Befund spricht deutlich für eine Bestätigung der Interdependenzhypothese zwischen den Sprachen.

Was den Zusammenhang zwischen Erdkundenote und Geographiekompetenz anbelangt, so ergibt sich auch hier ein (erwarteter) Zusammenhang, jedoch nur auf dem 0,05-Signifikanzniveau und in weniger stark ausgeprägter Form. Der Korrelationskoeffizient liegt bei $r=.275$ und damit deutlich geringer als der zwischen den beiden Sprachtests.

Interessant ist nun, wie die Ergebnisse der beiden Sprachtests mit der kognitiven Ausstattung der Lerner im Hinblick auf die Geographiekompetenz interagieren. Es zeigt sich, dass die Unterschiede in der Ausprägung geographischer Gesamtkompetenz am stärksten von der Fähigkeit im Deutschen beeinflusst wird, kurz darauf gefolgt von den kognitiven Fähigkeiten und den gemessenen Englischkenntnissen. Diese Befunde bedürfen weiterer Analyse und Interpretation.

3. Weitere Lernermerkmale im Hinblick auf die Ausprägung geographischer Kompetenz

Bei der Beschreibung weiterer Lernervariablen, die möglicherweise Einfluss auf die Ausprägung der Struktur geographischer Kompetenz haben, sind wir noch nicht sehr weit vorangeschritten. Es zeichnet sich jedoch ab, dass beispielsweise der Fernsehkonsum und vor allem das Anregungspotential im elterlichen Hause (gemessen am Lesepotential (Regalmeter), am Leseverhalten und an der Kommunikationsstruktur) mit den schulischen Leistungen allgemein sowie mit der Kompetenzausprägung im Fach Geographie korreliert. Dagegen scheint der erreichte Bildungsabschluss der Eltern keinerlei signifikanten Einfluss auf die Höhe und Struktur gemessener Geographieleistungen der Lerner zu haben. Die Befragung der Schülerinnen und Schüler ergab ebenfalls, dass die Beliebtheit des Faches Erdkunde erstaunlich gering ist: Nur 10,5% aller befragten Untersuchungsteilnehmer gab an, dass Erdkunde für sie zu den beliebtesten Fächern gehört; in der Rangfolge aller Fächer rangiert es im Durchschnitt nur an 9. Stelle. Allerdings gibt es Unterschiede zwischen einzelnen Klassen, was an der Beliebtheit der Lehrperson im Einzelfall liegen mag.

4. Entwicklung eines theoretischen Modells fachspezifischer Aufgabebearbeitung

Um die Besonderheiten des Bilingualen Unterrichts genauer verstehen zu können, war eine *theoretische Erschließung dieser fremdsprachenspezifischen*

Verarbeitung von Informationen und von kognitionstheoretischen Erkenntnissen (besonders der Informationsverarbeitung, der Wissenstheorie sowie der kognitionspsychologischen Problemlösungsforschung) notwendig, die mit neueren (kognitiv-)linguistischen Theorien in Beziehung gesetzt wurden. Heine (in Arbeit) bietet dazu eine ausführliche Diskussion dieser Theoriestränge und verbindet diese in Hinblick auf ihre Relevanz für bilingualen Sachfachunterricht. Daraus wurde ein Modell fachaufgabenbasierter Aufgabenlösung erarbeitet, das eine theoretische Trennung von fachlich-inhaltlichen und sprachlichen Problemlösungsaktivitäten ermöglicht. Mit diesem Modell steht erstmals eine theoretische Basis bereit, auf dem weitere Untersuchungen über den Bilingualen Unterricht (oder andere Formen der fremdsprachlichen Verarbeitung von Fachinhalten) in Zukunft aufbauen können.

5. Empirische Validierung der Methode des Lauten Denkens in fremdsprachlichen Kontexten

Aus diesem *Modellverständnis* ergab sich eine theoretisch fundierte Sichtweise auf die empirischen Lautdenkdaten, die erhoben wurden. Sie stellen in der methodischen Diskussion um das Laute Denken als Erhebungsmethode neue Erkenntnisse bereit, die sich vor allem auf die Erhebung von Lautdenkprotokollen in fremdsprachlichen Kontexten bezieht. Die intensiv vorbereitete Datenerhebung und die tiefe kognitionstheoretische Durchdringung (Heine 2005; Heine & Schramm im Druck) gewährleisteten hoch valide Daten, anhand derer sich erstmals im Detail zeigen lässt, dass die Verwendung einer Fremdsprache als Arbeitssprache durchaus Rückwirkungen auf die fachlich-konzeptuelle Verarbeitung haben kann. Diese Auswirkungen werden von Heine (in Arbeit) im Detail beschrieben.

6. Die Rolle von Sprache bei der Verarbeitung von Fachinhalten

In Übereinstimmung mit den Ergebnissen der Schreibprozessforschung (vgl. insbesondere Bereiter & Scardamalia 1993) zeigt sich nämlich sowohl in den mutter- wie auch den fremdsprachlichen Verarbeitungsprotokollen die enge Bindung von fachlich-konzeptueller Beziehungsbildung und ihrer Versprachlichung. Insbesondere bei der Formulierung von fachlichen Aussagen übernimmt die Sprache eine Katalysatorfunktion: Weil eine Festlegung auf eine angemessene Versprachlichung eines Konzeptgefüges getroffen werden muss, werden konzeptuelle Beziehungen reflektiert, Widersprüche aufgedeckt und gelöst, zumindest aber länger beim gedachten Sachverhalt verweilt und damit konzeptuelle Verknüpfungen verstärkt. Diese positive Rückwirkung eines sprachlichen Fokus' auf die Präzisierung und Elaborierung von fachlichen Inhalten sollte unseres Erachtens in verstärktem Maß im Sachfachunterricht genutzt werden, egal, in welcher Sprache er durchgeführt wird. In der präzisen Versprachlichung von

Inhalten liegt ein bisher noch zu wenig genutztes Potential für den Wissenserwerb und die Problemlösungskompetenz von Lernern.

7. Auswirkungen der Verwendung einer Fremdsprache als Arbeitssprache

Kommt nun im Fall der bilingualen Lerner die Fremdsprache zum Einsatz, werden häufiger Grenzen der Ausdrucksfähigkeit erreicht als bei den Regelschülern, die in ihrer L1 verarbeiten. Dies führt in den Daten unserer Probandengruppe selten zu Abbrüchen, sondern eher zu einer Intensivierung in der Beschäftigung mit dem fachlichen Inhalt, für dessen Ausdruck eine Alternativlösung gefunden werden muss; der im vorangegangenen Punkt dargestellte Effekt der Fokusverschiebung auf die sprachliche Form, der in einer tieferen Reflexion der fachlich-konzeptuellen Inhalte resultiert, wird bei der fremdsprachlichen Verarbeitung daher noch verstärkt (Hypothese der semantischen Tiefenverarbeitung, vgl. Craik & Lockhart 1972, die Lesart von Wolff (1997) sowie die Kritik bei Krapp & Weidenmann 2001: 430 ff. dazu). Die sprachlichen Grenzen der bilingualen Lerner können allerdings auch dazu führen, dass begonnene Formulierungsversuche abgebrochen werden und sich weniger Wissen in den schriftlichen Texten widerspiegelt als eigentlich mental vorhanden ist (Heine in Arbeit).

Trotz dieser Fälle verfügen wir nunmehr über *empirische Befunde*, die zeigen, dass bei aller sprachlichen Begrenzungen dennoch Gewinn aus der fremdsprachlichen Verarbeitung von Fachinhalten gezogen werden kann. Lehrer sind allerdings angehalten, nicht nur aus der schriftlichen Formulierungsfähigkeit ihrer bilingualen Lerner auf ihren fachlichen Wissensstand rückzuschließen, da dieser häufig nicht adäquat in der L2 wiedergegeben werden kann.

Wie bereits erwähnt (vgl. Abschnitt 8), haben wir im Projekt einen Typ von Aufgabe erprobt und eingesetzt, bei dem eine Verschiebung der Aufmerksamkeit von der inhaltlich-semantischen Dimension auf die formale Dimension des Ausdrucks und die kommunikative Angemessenheit und Effektivität fachlicher Mitteilungen potentiell möglich ist, jedoch durch unterschiedliche Lerner unterschiedlich gut genutzt wird. Es sind dies Aufgaben der Überarbeitung, wie wir sie aus der prozessbasierten Schreibforschung kennen. Die genaueren Ergebnisse hierzu liegen allerdings z.Z. noch nicht vor, versprechen aber tiefere Einsichten in die mentalen Entscheidungsprozesse der untersuchten bilingualen Lerner (im Vergleich zu muttersprachlich agierenden Geographielernern) und deren Zusammenspiel mit sprachlichen Handlungen und Ausdrucksformen.

8. Auswirkungen verschiedener Lern- oder Arbeitskontexte

Was die *Auswirkungen verschiedener Lernkontexte (settings)* auf fachlich-sprachliches Lernen anbelangt, so gingen wir von der Annahme aus, dass eine gemeinsame Konstruktion von Wissen derjenigen überlegen ist, die ein

Individuum alleine erzeugt. Entsprechend sah das Forschungsdesign vor, dass ein Teil der Lernerpopulation einen Teil der Aufgaben (Aufgabenblock 1) in Paarbeit löst, während andere Aufgaben (Aufgabenblock 2) von eben denselben Paarteilnehmern zusätzlich in Einzelarbeit zu lösen waren. Die große Mehrzahl der monolingualen wie der bilingualen Probanden bearbeitete und löste die Geographieaufgaben lediglich schriftlich in individueller Weise.

Wenn man sich den Mittelwert der geographischen Gesamtkompetenz der drei unterschiedlichen Gruppen anschaut, so zeichnet sich ein deutlich geringerer Wert bei denjenigen Schülern ab, die zu keiner Zeit an Paarbeit beteiligt waren. Allerdings sind die Unterschiede zwischen den bilingualen und den monolingualen Paaren, die beide kooperative Aufgabenblöcke bearbeitet haben, unerheblich. Dagegen gibt es einen signifikanten Unterschied im Ergebnis der beiden Aufgabenblöcke, die entweder im Paar oder aber einzeln bearbeitet wurden. Schließlich konnten beide Paargruppen, die monolingualen und die bilingualen, zu einer einzigen Paargruppe zusammengefasst werden und mit jenen Lernern verglichen werden, die zu keinem Zeitpunkt an kooperativer Aufgabenlösung beteiligt waren. Dabei zeigt sich, dass es einen noch größeren und zwar hochsignifikanten Unterschied gibt zwischen den Paarteilnehmern insgesamt und den Nicht-Paarteilnehmern im Hinblick auf jene Aufgaben, die von allen nur individuell bearbeitet worden sind. Bei diesem Aufgabenblock 2 schneiden diejenigen Lerner, die ursprünglich zu einem Paar gehört haben, deutlich besser ab als die Gruppe derjenigen, die ausschließlich alle Aufgaben in Einzelarbeit lösten. Somit ist zu fragen, ob hier so etwas wie ein Transfereffekt vorliegt oder sogar eine Motivationsverstärkung, die aus der kooperativen Bearbeitung heraus Freude und Zuversicht an der eigenen Leistung erhöht und sich insofern auch in individueller Arbeit positiv niederschlägt. Hier ist jedoch auf Grund mangelnder Verallgemeinerungsfähigkeit interpretative Vorsicht geboten. Tendenziell jedoch scheint sich zu bestätigen, dass die von uns untersuchten Geographielerner (wie theoretisch postuliert) *durch Kollaboration* mehr Wissen erzeugen bzw. mehr Fachkompetenz demonstrieren als *individuell* (vgl. Troschke 2006, im Druck).

9. Mangelnde fachliche Angemessenheit von schriftlichen Lerneräußerungen

Was die *Ergebnisse der Hauptstudie* des Projekts anbelangt, namentlich der Vergleich der fachlichen Kompetenzen und Leistungsfähigkeit von monolingualen und bilingualen Lernern, so ist vorläufig Folgendes festzuhalten: Die *inhaltlichen* Ausführungen der Lerner aus beiden Gruppen sind meistens zwar verständlich und nachvollziehbar, aber sie entsprechen eben in aller Regel nicht den Konventionen bzw. den "thematischen Mustern" der Fachdisziplin - *die schriftlichen Äußerungen sind in den allermeisten Fällen nicht genau genug und damit fachlich nicht angemessen.*

Hier zeigen sich große *sachfachliterale Defizite* bei *beiden* Lernergruppen, die besonders in der Verschriftlichung zu Tage treten. Aber auch die konzeptuellen Modelle, mit denen die Lerner operieren, sind immer noch alltagsbasiert und vorwissenschaftlich, die wissenschaftspropädeutischen Denk- und Ausdrucksweisen am Ende der gymnasialen Klasse 10 noch nicht gelungen. Dies gilt auch für die *prozeduralen* Kompetenzen der Probanden. Sie sind im Eingebübten verhaftet und versagen häufig dann, wenn ungewohnte Herausforderungen zur Entschlüsselung von Information oder zu deren Abgleich mit anderer Information (z.B. durch Transformation von absoluten in prozentuale Zahlen oder umgekehrt) auftreten. Dabei beweisen Bilinguale allerdings eine leicht höhere Transferfähigkeit prozeduralen Wissens. Hier zeigt sich dennoch ein bedenkliches Bildungsdefizit, das wahrscheinlich auch in anderen Fächern ("Domänen") vorliegt (Vollmer 2006b, im Druck).

10. Sprachliche Kompetenzen ohne hinreichende Fachbasiertheit

Es überrascht nicht, dass die Lerner angesichts der reduzierten oder defizitären Inhaltsvorstellungen die entsprechende *Sprachlichkeit in ihren schriftlichen Produkten* relativ sicher und überzeugend handhaben - das gilt vor allem für die textuelle Organisation des Inhalts sowie für die Fähigkeit zur Herstellung von Kohäsion, nicht jedoch für die Sicherheit in der Wortwahl. Selbst bei "flachem" Inhalt gibt es große Lücken im fachspezifischen Vokabular, insbesondere von einzelwortübergreifenden Ausdrücken. Dies ist ein weiteres Indiz für die gering ausgeprägte Fachsprachlichkeit vieler Schüler (und das am Ende von Klasse 10!). Auf der grammatischen Ebene wiederum werden zwar einzelne Fehler gemacht, aber angesichts der relativ oberflächlichen inhaltlichen Aussagen können die meisten Probanden ihre Grammatikkenntnisse erfolgreich übertragen, so dass es hier zu weniger Auffälligkeiten kommt (Coetzee 2006, im Druck, in Arbeit).

Wie deutlich geworden sein dürfte, ergeben sich aus dem Projekt hoch interessante Ergebnisse und Einsichten. In mehrfacher Hinsicht glauben wir wichtige Beiträge zur Weiterentwicklung des relevanten Forschungsstandes geleistet zu haben, sowohl im Hinblick auf die Erweiterung methodologischer Erkenntnisse als auch auf die theoretischen Einsichten in die komplexen Zusammenhänge von Fachlichkeit und Sprachlichkeit einschließlich der Besonderheiten bilingualen Denkens und Problemlösens. Dabei darf auch die Aufarbeitung wichtiger theoretischer Stränge nicht unerwähnt bleiben, die bislang mehr oder minder unverknüpft miteinander waren und die nun in ihrer Verbindung zu neuen Einsichten und Beschreibungsansätzen der erzeugten fachlichen Daten in der L2, ebenso wie in der L1, führen (vgl. die Arbeiten von Coetzee, Heine, Troschke, Vollmer, alle in Arbeit).

10. Mögliche Anwendungsperspektiven und Folgeprojekte

Was die möglichen Anwendungsperspektiven betrifft, so sind wir im Moment dabei, die o. g. Ergebnisse weiter zu verallgemeinern und ihre Implikationen für die Konkretisierung zentraler Aspekte einer Didaktik und Methodik des bilingualen Unterrichtens herauszuarbeiten. Diese Ergebnisse sind zwar bis zur vollständigen Auswertung nur *Tendenzaussagen*, doch sind sie höchst alarmierend und geben fachdidaktisch wie bildungspolitisch zu denken. Es gibt allerdings auch positive Beobachtungen, die an anderer Stelle näher dazulegen sind (Coetzee in Arbeit; Möller in Arbeit). Im Prinzip lässt sich festhalten: Schüler brauchen eine anspruchsvolle fachliche Herausforderung mit genauer unterstützender Spracharbeit, um sich inhaltlich zu qualifizieren und Fachlernen positiv als Begegnung mit wissenschaftspropädeutischen Begriffen und ihrer Vernetzung zu erfahren. Ohne diese systematische Herausforderung werden sie unter den Bedingungen schulisch-institutionalisierten Lernens nicht in die Lage versetzt, erweiterte Diskursstrukturen, eine angemessene Sachfachliteraltät und fachliche Mehrsprachigkeit zu entwickeln. Damit wird ihnen die kompetente Teilhabe an zukünftigen Fachdiskursen und relevanten Diskursgemeinschaften erschwert.

Das Projekt hat einen wichtigen Beitrag zur Entwicklung und Anwendung eines domänenspezifischen Kompetenzmodells geleistet und verschiedene Zugänge holistischer wie analytischer Art für zukünftige Kompetenzmessungen erprobt, die ohnehin mit der Einführung von Bildungsstandards massiv auf die empirische Bildungsforschung sowie auf die Schulen zukommen werden. Insofern hat es im Vorgriff auf die Entwicklungen zu einer output-orientierten Steuerung des Bildungswesens und der einzelnen schulischen Bildungsgänge einen exemplarischen Beitrag zur Strukturierung von *Assessment* und Evaluation anhand eines konkreten Fachausschnitts geleistet, sowohl was die theoretische Definition von Fachkompetenz als auch deren Operationalisierung und Erfassung durch entsprechende Aufgabenstellungen anbelangt. Unsere Vorarbeiten und Befunde werden in größere Forschungszusammenhänge (Ausdifferenzierung einer domänenspezifischen Bildungsforschung, Entwicklung einer differentiellen Kompetenzdiagnostik, mögliche Weiterverfolgung im Rahmen des von der DFG soeben beschlossenen neuen Schwerpunktprogramms) einzubetten und auf eine breitere empirische Basis zu stellen sein. Andererseits könnten sie auch durch die Schulen und die Fachkollegien selbst als Ausgangspunkt für schul- und unterrichtsnahe Definitionen von Kompetenzentwicklung genutzt werden. Insofern kommt dem Projekt ein innovatives Potential, auch für die interessierte Bildungsöffentlichkeit zu.

Zur Zeit sind wir dabei, die beteiligten Schulen zu besuchen, um jeweils Rückmeldung an die beteiligten Lehrer und Lehrerinnen zu geben, ausgewählte

Ergebnisse zu präsentieren sowie einige der Implikationen der Untersuchung zu diskutieren. Diese könnten vor allem im Bereich einer Stärkung der fachbasierten Sprach- und Textarbeit in Geographie (sowohl für monolinguale als auch für bilinguale Schülerinnen und Schüler) liegen sowie in einer Neubestimmung des integrativen Verhältnisses von sprachlichem und fachlichem Lernen in *allen* Schulfächern (*language across the curriculum*). In diesem Zusammenhang werden wir dann auch im Austausch mit den beteiligten Lehrergruppen unsere didaktischen und methodischen Vorschläge zur Verbesserung und Weiterentwicklung des bilingualen Unterrichts, aber auch des Fachunterrichts insgesamt präzisieren und überprüfen können.

Mögliche Folgeuntersuchungen

In den nächsten Wochen werden die notwendigen Arbeiten abgeschlossen sein und im Laufe des kommenden Jahres dann die Projektergebnisse veröffentlicht vorliegen. Dabei gibt es absehbar einen Überhang an hochinteressanten Daten, deren Auswertung Folgeuntersuchungen überlassen bleiben muss. Auch hat das Projekt eine Reihe neuer Fragen aufgeworfen, die es auf jeden Fall lohnt weiterzuverfolgen.

1. So könnte sich eine mögliche Folgestudie auf die vielfältigen *Daten aus den Interviews* konzentrieren, die im Projekt retrospektiv (also direkt im Anschluss an die Bearbeitung und Lösung von Fachaufgaben bzw. deren Überarbeitung) mit einzelnen Schülern geführt worden sind. Sie würden einen weiteren Einblick in die fach- bzw. aufgabenbasierten mentalen Prozesse der Schüler und in ihre Ansätze zur Selbstbewertung ermöglichen. Die Aufbereitung dieser Daten ließe sich auch verknüpfen mit den Ergebnissen des Editing jener Schüler, die als Unterstichproben zu den Lautdenkern oder zu den Lernerpaaren gehören. Diese hoch interessanten Daten konnten aus Zeitgründen im Rahmen der laufenden Arbeiten nicht mehr berücksichtigt werden.

2. Eine weitere Folgestudie könnte sich z.B. auf die theoretische und empirische Durchdringung eines *C-Tests für Deutsch als Muttersprache* beziehen. Hier haben wir Neuland betreten und das Potential dieses Instruments für eine relativ schnelle und komfortable Sprachstandsbestimmung von Muttersprachensprechern erkannt (vgl. auch Aguado, Grotjahn & Schlak im Druck). Allerdings fehlte die Zeit für eine genauere Validierung dieses Forschungsinstrumentes und dessen Erprobung mit einer Vielzahl von geeigneten Texten unter systematischen Gesichtspunkten. Die Ergebnisse der von uns verwendeten zwei Varianten eines C-Tests für Deutsch als L1 liegen vor, ebenfalls deren Korrelationen mit anderen Standardmaßen wie z.B. dem C-Test Englisch oder mit drei verschiedenen Subtests des Kognitiven Fähigkeitstests (KFT).

3. Schließlich ließe sich auch eine gezielte, hochaktuelle *Analyse der spezifischen Leistungen von Risikolernern* als eine mögliche (qualitative) Folgeuntersuchung vorstellen (zumal diese im Sample vorhanden sind). Hier könnte es zu einem Vergleich von "normalen" und "benachteiligten" deutschsprachigen Gymnasiasten sowie von nicht-deutschen Migrantenkinder kommen. Wie wir allerdings aus der DESI-Studie (DESI-Konsortium 2006) wissen, müssen die Lerner mit Migrationshintergrund genauer differenziert werden, um deren unterschiedliche Sozialisations- und Sprachhintergründe angemessen interpretieren und auf ihre Profile geographischer Kompetenz beziehen zu können.

11. Zur Qualifikation des wissenschaftlichen Nachwuchses

Im Zusammenhang des Projektes entstehen drei Dissertationen (von Lena Heine, Debbie Coetzee und Randi Troschke) sowie zumindest eine weitere Monographie (vom Projektleiter), von denen nur eine durch die von der DFG bezahlte wissenschaftliche Mitarbeiterin verfasst wird. Entscheidend ist, dass das Projekt darüber hinaus den wissenschaftlichen Nachwuchskräften zu einer breiten Qualifizierung verholfen hat: Durch Hinführung zu empirischem Arbeiten, durch konkrete Einführung in verschiedenartige qualitative wie quantifizierende Forschungsmethoden, durch regelmäßige Forschungskolloquien mit der Gelegenheit, systematisch Theorieansätze, Methodenprobleme oder aber Zwischenergebnisse kritisch zu reflektieren und hochschulöffentlich zu erörtern, schließlich durch die Abfassung und gemeinsame Besprechung von wissenschaftlichen Manuskripten zur Veröffentlichung in einschlägigen Zeitschriften oder Sammelbänden.

Nicht unerwähnt bleiben sollte die Tatsache, dass zu unterschiedlichen Zeitpunkten eine ganze Reihe von Studierenden (zusammen genommen über 20) aus den Fächern Anglistik und Geographie (insbesondere solche mit bilingualem Schwerpunkt) an der Durchführung und Auswertung des Projekts als studentische Hilfskräfte beteiligt waren. Insofern hat das Projekt auch in diesem Sinne für einen erweiterten Personenkreis eine Vielzahl von Qualifizierungsmöglichkeiten geschaffen.

Im Laufe des Projekts haben wir regelmäßig öffentliche Forschungskolloquien durchgeführt. Insbesondere waren wir in der Lage, im Rahmen von ein- oder zweitägigen Forschungsworkshops mehrere Kollegen zu uns einzuladen und ihre eigenen Forschungsansätze, die sich jeweils gezielt auf Teilfragen unseres eigenen Projekts bezogen, bei uns vorzustellen. Im Einzelnen war dies u.a. Rüdiger Grotjahn (zu Fragen des C-Tests), Wolfgang Zydatis (zu Fragen des Vergleichs von bilingualen und nicht-bilingualen Lernern), Steffi Morkötter (zur

Typisierung von qualitativen Interviews und zur qualitativen Inhaltsanalyse) sowie Andreas Bonnet, der aus seinem eigenen Dissertationsprojekt über bilingualen Chemieunterricht berichtete und dabei in exemplarischer Weise interaktive Daten mit uns inspizierte. Bei diesen Gelegenheiten konnten auch die Doktorandinnen Beispiele aus ihrem jeweiligen Datenpool vortragen und mit den eingeladenen Experten analysieren. Ein Großteil der Produktivität des Projekts beruhte also auf der Tatsache der Kooperativität mit Auswärtigen sowie untereinander und damit auf Teamwork. Dessen Bedeutung kann nicht hoch genug eingeschätzt werden.

Am Ende des DFG-Projekts wurde eine wissenschaftliche Tagung organisiert, die alle laufenden DFG-Projekte im Bereich der Fremdsprachenforschung und Forschende aus anderen geplanten oder durchgeführten Drittmittelprojekten unter dem Titel "Synergieeffekte in der Fremdsprachenforschung: Empirische Zugänge, Probleme, Ergebnisse" vom 11.-12. November 2005 in Osnabrück zusammenführte. Aus dieser außerordentlich fruchtbaren Tagung entstand eine einschlägige Veröffentlichung (Vollmer (Hrsg.)), die sich z. Z. in Druck befindet und Anfang 2007 erscheinen wird.

12. Zusammenfassung und Perspektiven

Trotz der insgesamt knapp bewilligten Personal-, Finanz- und Zeitdecke konnte das ursprüngliche Forschungsprogramm nur bedingt reduziert werden, weil die Hauptfragestellung des Projekts (Messung geographischer Kompetenz von bilingualen und monolingualen Lernern in Klasse 10 des Gymnasiums) die Erhebung zusätzlicher Variablen zur Erfassung des Zusammenhangs von Sprache und fachlicher Leistung erforderte. Insofern konnte hinsichtlich der Art und des Umfangs der Instrumentenentwicklung und der Datenerhebung weniger Kompromisse gemacht werden als im Bereich der *Datenauswertung*. Dennoch hat es auch im Forschungsdesign für bestimmte Teilbereiche des Projekts eine Verschiebung weg von hypothesenprüfenden Studien zugunsten von eher qualitativen Teiluntersuchungen mit kleineren Teilstichproben gegeben, so etwa bei der *Analyse kognitiver Prozessdaten auf der Basis von Lautdenkprotokollen* (n=14 bilinguale Lerner, n=8 monolinguale Lerner) oder auch bei der *Analyse von Aushandlungsprozessen durch Lernerpaare mit Hilfe von Videographie* (n=24, d.h. also je 6 monolinguale und bilinguale Paare) - mit dem Ziel einer genaueren Durchdringung und reicheren Beschreibung von Einzelfällen. Die zuletzt genannte Teilstudie zum kollaborativen Problemlösen strebte darüber hinaus eine möglichst maximale Fallkontrastierung an. Von einer ursprünglich vorgesehenen Typenbildung bilingualer (oder monolingualer) Lerner auf der Basis inhaltsanalytischer

Auswertungsverfahren musste gänzlich abgesehen werden. Eine solche war ohnehin erst für das 3. Projektjahr eingeplant.

Der Verzicht auf quantifizierende Analyseschritte in bestimmten Teilbereichen des Projekts lässt sich im übrigen auch theoretisch legitimieren. Erst in der genauen Betrachtung einzelner Fälle waren Erkenntnisse zu gewinnen, die von höchster Bedeutung für das Verständnis der Sprachlichkeit fachlichen Denkens sind. In der Hauptstudie des Projekts wurde aus methodologischen Gründen dennoch im Wesentlichen am Ursprungsdesign festgehalten, zumal die abhängige Variable (gemessene geographische Kompetenz) zu ihrer Erklärung eine Reihe von Zusatzinformationen und Messdaten aus unabhängigen Einflussgrößen benötigte (z.B. Allgemeine Sprachfähigkeit der Lerner, Leistungsmotivation, Fachmotivation, Fachinteresse, sozioökonomischer Hintergrund, Bildungsnähe des Elternhauses usw.). Diese wurden im Rahmen von Tests (C-Tests in Deutsch und Englisch, Ausschnitte aus einem Kognitiven Fähigkeitstest), durch Fragebögen (ausführliche Schüler- und Lehrerbefragung), durch Notenerhebung in den Schulen und durch Interviews erhoben. Dadurch ergab sich zugleich die Möglichkeit, homogenere Subgruppen von Schülern zu bilden und sie nur miteinander zu vergleichen, wenn sie sich *primär* in der benutzten Arbeitssprache (L1 versus L2) unterschieden. Auf diese Weise konnte die Selbstselektivität bilingualer Lerner kontrolliert werden, denen häufig nachgesagt wird, dass sie die "besseren" Schüler aus "besserem" Hause seien.

Alle quantitativen Daten wurden entsprechend kodiert, teilweise transformiert oder aggregiert und elektronisch mit SPSS erfasst. Dieser Vorgang war äußerst zeit- und arbeitsintensiv und bedurfte einer zeitaufwändigen Kontrolle und Bereinigung, bevor die Daten weiterbearbeitet werden konnten. Die Bewertung der Schülerantworten sowie deren Überarbeitung fand nach detaillierten inhaltlichen, prozeduralen und sprachlichen Gesichtspunkten auf der Basis theoretisch begründeter Skalierungen durch mehrere Rater statt. Deren Ergebnisse wurden systematisch beschrieben, in ein Modell geographischer Gesamtkompetenz integriert und mit einer Reihe von Erklärungsvariablen in Beziehung gesetzt wurden. Aber auch hier musste der Einsatz aufwendiger multivariater Analyseverfahren notwendigerweise begrenzt bleiben. Über erste Ergebnisse ist weiter oben berichtet worden; andere werden folgen.

In der ursprünglichen Antragsstellung an die DFG war eine Serie von Tagungen vorgesehen, die alle DFG-Projekte unter dem Aspekt der Erzeugung eines gewissen Synergieeffekts in der Fremdsprachenforschung regelmäßig zusammenbringen sollte, diese Idee ist leider von der DFG nicht aufgegriffen worden. Dennoch haben wir versucht, einen Schritt in diese Richtung zu gehen - dann eben mit bescheideneren Fördermitteln -, und zwar durch die Organisation einer

Veranstaltung im November 2005 unter dem Titel "Synergieeffekte in der Fremdsprachenforschung: Empirische Zugänge, Probleme, Ergebnisse". Diese Initiative, finanziell unterstützt durch die Universität Osnabrück sowie die Deutsche Gesellschaft für Fremdsprachenforschung (DGFF), hat genau die erhoffte Funktion erfüllt, nämlich Forscher und Forscherinnen aus unterschiedlichen Projekten zusammenzubringen, die über ihre eigenen Theoriebildungsansätze und Methodenprobleme reflektieren und sich darüber in sehr offener Weise austauschen konnten. Zugleich hat diese Tagung die Funktion gehabt, jüngere und erfahrenere Forscher in einem überschaubaren Rahmen miteinander ins professionelle Gespräch zu bringen. Ein weiteres Ziel bestand darin, den Zusammenhang zwischen tatsächlich forschenden Personen und Projekten auch für die Zukunft zu stärken: In diesem Sinne wurde beschlossen, die Synergietagung regelmäßig zu wiederholen und weitere Folgekonferenzen alle zwei Jahre im November durchzuführen. Die nächste wird 2007 in Münster stattfinden, ausgerichtet von Karen Schramm (Münster/Leipzig) und Karin Aguado (Kassel).

Was wir nach den Erfahrungen unseres Projekts in Zukunft verstärkt brauchen, sind zum einen Austauschmöglichkeiten für diejenigen, die *selbst aktiv Fremdsprachenforschung betreiben*. Darüber hinaus muss es wie bisher auch Foren geben, in denen innerhalb der wissenschaftlichen *community* oder aber *in stärker auf die Öffentlichkeit ausgerichteter Weise* über laufende Forschungen berichtet wird und diese vor allem ergebnisorientiert erörtert werden. Drittens jedoch wird auch weiterhin eine systematische Plattform benötigt, um die *spezifischen Forschungs- und Praxisprobleme des bilingualen Unterrichts* voranzutreiben. In diesem Sinne sollte die Idee der Bremer Tagungen mit dem gezielten Fokus auf diese Fragestellungen fortgesetzt werden. Da dies natürlich immer auch von Personen abhängt, die als Motoren bestimmter Entwicklungen fungieren, ist mit großer Sorgfalt darauf zu achten, dass trotz des Generationenwechsels in unserer Disziplin eine Kontinuität in der Bearbeitung bestimmter ungelöster Fragen und in der Verfolgung bestimmter Forschungsansätze bestehen bleibt. Hier sind die Nachwuchsforen, die sich regional wie überregional etabliert haben, und die gegenseitigen Beratungsinitiativen zur Ausformulierung von Projektanträgen (u. a. von der DGFF) von allergrößter Bedeutung.

Eingang des revidierten Manuskripts: 13.11.2006

Literaturverzeichnis

- Abendroth-Timmer, Dagmar; Bonnet, Andreas; Breidbach, Stephan; Hoffmann, Reinhard; Kircher, Ernst; Küster, Lutz; Rymarczyk, Jutta; Vollmer, Helmut J. & Zydati, Wolfgang (2004), *Didaktiken im Dialog - Für eine integrative Didaktik des Bilingualen Unterrichts*. In: Breidbach, Stephan & Bonnet, Andreas (Hrsg.) (2004), 13-30.
- Aguado, Karin; Grotjahn, Rüdiger & Schlak, Torsten (im Druck), *Erwerbsalter und Sprachlernerfolg: zeitlimitierte C-Tests als Instrument zur Messung prozeduralen sprachlichen Wissens*. In: Vollmer, Helmut Johannes (Hrsg.) (im Druck), 123-136.
- Ahrenholz, Bernt (2003), Förderunterricht und Deutsch-als-Zweitsprache-Erwerb. Eine longitudinale Untersuchung zur mündlichen Sprachkompetenz bei Schülerinnen und Schülern nicht-deutscher Herkunft (ndH) in Berlin. *Zeitschrift für Fremdsprachenforschung* 14: 2, 291-300.
- Anderson, Lorin W.; Krathwohl, David R.; Airasian, Peter W.; Cruikshank, Kathleen A.; Mayer, Richard E.; Pintrich, Paul R.; Raths, James & Wittroch, Merlin C. (Hrsg.) (2001), *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*. Complete edition. London: Longman.
- Aufschnaiter, Claudia v. & Aufschnaiter, Stefan v. (2003), Theoretical framework and empirical evidence on students' cognitive processes in three dimensions of content, complexity, and time. *Journal of Research in Science Teaching* 40: 7, 616-648.
- Bach, Gerhard & Niemeier, Susanne (Hrsg.) (2005), *Bilingualer Unterricht. Grundlagen, Methoden, Praxis, Perspektiven*. (3., überarb. u. erweitem. Auflage). Frankfurt/M.: Lang.
- Bereiter, Carl & Scardamalia, Marlene (1987), *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Bereiter, Carl & Scardamalia, Marlene (1993), *Surpassing ourselves: An inquiry into the nature and implications of expertise*. Chicago: Open Court Press.
- Bonnet, Andreas (1999), Begriffliches Lernen im bilingualen Unterricht Chemie. *Englisch* 34: 1, 3-8.
- Bonnet, Andreas (2001), *Sprachlernen im bilingualen Chemieunterricht*. In: Breidbach, Stephan (Hrsg.) (2001), 103-118.
- Bonnet, Andreas (2004), *Bilingualer Chemieunterricht. Kompetenzerwerb durch Interaktion*. Opladen: Leske+Budrich.
- Bonnet, Andreas, Breidbach, Stephan & Hallet, Wolfgang (2002), Fremdsprachlichkeit als Spezifikum - Auf der Suche nach einer integrativen Didaktik für den Bilingualen Unterricht. In: Voss, Bernd & Stahlheber, Eva (Hrsg.) (2002), *Fremdsprachen auf dem Prüfstand. Innovation, Qualität, Evaluation*. Berlin: Pädagogischer Zeitschriftenverlag, 151-163.
- Börner, Wolfgang (2004), Bericht über das DFG-Projekt "Lernprozesse in grammatisch-kommunikativen Lernaufgaben". *Zeitschrift für Fremdsprachenforschung* 15: 1, 129-141.
- Bredenbröcker, Winfried (2000), *Die Förderung der fremdsprachlichen Kompetenz durch Bilingualen Unterricht: Eine empirische Untersuchung*. Frankfurt/M.: Lang.
- Breidbach, Stephan (Hrsg.) (2001), *Forschungsperspektiven für den Bilingualen Sachfachunterricht*. Frankfurt/M.: Lang.
- Breidbach, Stephan & Bonnet, Andreas (Hrsg.) (2004), *Didaktiken im Dialog. Konzepte des Lehrens und Wege des Lernens im bilingualen Sachfachunterricht*. Frankfurt/M.: Lang.
- Claussen, Imke (1997), *Ein Vergleich der Diskursfähigkeiten mono- und bilingual unterrichteter Schüler im Rahmen der Unterrichtserprobung in Schleswig-Holstein*. Kiel: Universität Kiel (Mimeo).

- Coetzee-Lachmann, Debbie (2006), *Eine Definition fachspezifischer Diskurskompetenz*. In: Timm, Johannes-Peter (Hrsg.) (2006), 249-265.
- Coetzee, Debbie (im Druck), *Methodological considerations in assessing task-based subject-specific writing: The case of bilingually trained geography learners of grade 10 in Germany*. In: Vollmer, Helmut Johannes (Hrsg.) (im Druck), 209-226.
- Coetzee, Debbie (in Arbeit), *Assessment of Subject-Specific Task Performance of Bilingually Trained Geography Learners: Analysing Aspects of Subject-Specific Written Discourse Competence*. Osnabrück: Universität (Dissertation).
- Cohrs, Inger (1999), *Analysen zur Syntax: Satzstrukturen bilingual und nicht bilingual unterrichteter Schüler der 10. Jahrgangsstufe*. Kiel: I & f Verlag.
- Craik, Fergus I. M. & Lockhart, Robert S. (1972), Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 671-684.
- Croft, William & Cruse, D. Alan (2004), *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Cummins, James (1978), The cognitive development of children in immersion programs. *Canadian Modern Language Review* 34, 855-883.
- Cummins, James (1979), Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49, 222-251.
- Czapek, Frank-Michael (2000), Begriffs- und Sprachbildung als Prinzip des Geographie-Unterrichts - Gedanken zum lernstrukturellen Profil des Fach-Unterrichts. *Geographie und Schule* 124, 24-30.
- Czapek, Frank-Michael (2004), Sprachliche Bildung im Geographieunterricht. In: Schallhorn, Eberhard (Hrsg.) (2004), *Erdkunde Didaktik - Praxishandbuch für die Sekundarstufe I und II*. Berlin: Cornelsen, 111-118.
- Daniel, Angelika (2001), *Lernwortschatz und Wortsatzlernen im bilingualen Unterricht*. Frankfurt/M.: Lang.
- DESI-Konsortium (2006), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch-Englisch-Schülerleistungen-International (DESI)*. Frankfurt/M.: Deutsches Institut für Internationale Pädagogische Forschung (DIPF).
- DGfG (Deutsche Gesellschaft für Geographie) (Hrsg.) (2006), *Bildungsstandards im Fach Geographie für den Mittleren Schulabschluss*. Berlin: Deutsche Gesellschaft für Geographie.
- Dirks, Una (2003), Die Professionskultur von EnglischlehrerInnen. Eine berufsbiographisch orientierte, struktur- und handlungstheoretische Untersuchung. *Zeitschrift für Fremdsprachenforschung* 14: 2, 301-312.
- Evans, Vyvyan & Green, Melanie C. (2006), *Cognitive Linguistics. An Introduction*. Edinburgh: Edinburgh University Press.
- Flower, Linda & Hayes, John R. (1981), A cognitive theory of writing. *College Composition and Communication* 32, 365-387.
- Grotjahn, Rüdiger (2006), *Der C-Test: Theorie, Empirie, Anwendungen. The C-Test: Theory, Empirical Research, Applications*. Frankfurt/M.: Lang.
- Golay, David (2005), *Das bilinguale Sachfach Geographie. Eine empirische Untersuchung zum sachfachlichen Lernzuwachs im bilingual deutsch-französischen Geographieunterricht auf der Sekundarstufe I*. Nürnberg: Selbstverlag des Hochschulverbandes für Geographie und ihre Didaktik.

- Hallet, Wolfgang (2002), Auf dem Weg zu einer bilingualen Sachfachdidaktik. Bilinguales Lernen als fremdsprachige Konstruktion wissenschaftlicher Begriffe. *Praxis des neusprachlichen Unterrichts* 49: 2, 115-126.
- Hallet, Wolfgang (2003), Bilingualer Unterricht: Qualifikationen, didaktische Konzepte und Curriculum. Bilinguales Lernen als fremdsprachige Konstruktion wissenschaftlichen Wissens. In: Hoffmann, Reinhard (Hrsg.) (2003), *Bilingualer Geographieunterricht. Konzepte - Praxis - Forschung*. Nürnberg: Selbstverlag des Hochschulverbandes für Geographie und ihre Didaktik, 45-64.
- Halliday, Michael A. K. & Hasan, Ruqaiya (1989), *Language, Context and Text: Aspects of Language in a Social Semiotic Perspective* (2nd edition). Oxford: Oxford University Press.
- Halliday, Michael A. K. & Martin, James R. (Hrsg.) (1993), *Writing Science: Literacy and Discursive Power*. London: Falmer Press.
- Heine, Lena (2005), Lautes Denken als Forschungsinstrument in der Fremdsprachenforschung. *Zeitschrift für Fremdsprachenforschung* 16: 2, 163-185.
- Heine, Lena (2006), *Wissensbasierte Lösungsprozesse von Fachaufgaben*. In: Timm, Johannes-Peter (Hrsg.) (2006), 267-282.
- Heine, Lena (in Arbeit), *Kognitive Prozesse bilingualer Lerner bei der fremdsprachlichen Bearbeitung geografiespezifischer Fachaufgaben*. Osnabrück: Universität (Dissertation).
- Heine, Lena & Schramm, Karin (im Druck), *Lautes Denken in der Fremdsprachenforschung: Eine Handreichung für die empirische Praxis*. In: Vollmer, Helmut Johannes (Hrsg.) (im Druck), 153-195.
- Kultusministerkonferenz (2004a), *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Bonn: KMK.
- Kultusministerkonferenz (2004b), *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. Bonn: KMK.
- Kultusministerkonferenz (2004c), *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. Bonn: KMK.
- Kickler, Uwe (1995), *Wortschatzerwerb im bilingualen Unterricht. Pilotstudie zur Evaluierung der lexikalischen Fähigkeiten bilingual unterrichteter Schüler anhand eines kommunikativen Tests*. Kiel: I & f Verlag.
- Krapp, Andreas & Weidenmann, Bernd (Hrsg.) (2001), *Pädagogische Psychologie. Ein Lehrbuch* (4. Aufl.). Weinheim: Beltz.
- Lamsfuß-Schenk, Stefanie (2000), Didaktik des Fremdverstehens im bilingualen Geschichtsunterricht: Eine qualitative Longitudinal-Studie im bilingualen Klassenzimmer. In: Abendroth-Timmer, Dagmar & Breidbach, Stephan (Hrsg.) (2000), *Handlungsorientierung und Mehrsprachigkeit. Fremd- und mehrsprachliches Handeln in interkulturellen Kontexten*. Frankfurt/M.: Lang, 161-175.
- Lamsfuß-Schenk, Stefanie (2001), *Geschichte und Sprache*. In: Breidbach, Stephan (Hrsg.) (2001), 167-182.
- Lamsfuß-Schenk, Stefanie (2002a), Bilingualer Geschichtsunterricht: Die Perspektive des Sachfaches. *Neusprachliche Mitteilungen* 55: 2, 87-96.
- Lamsfuß-Schenk, Stefanie (2002b), Bilingualer deutsch-französischer Geschichtsunterricht. Beobachtungen aus einer Fallstudie. *Geschichte in Wissenschaft und Unterricht* 53: 2, 109-118.
- Lamsfuß-Schenk, Stefanie (2002c), *Geschichte und Sprache. Ist der bilinguale Geschichtsunterricht der Königsweg zum Geschichtsbewusstsein?* In: Breidbach, Stephan; Bach, Gerhard & Wolff, Dieter (Hrsg.) (2002), *Bilingualer Sachfachunterricht: Didaktik, Lehrer-/Lernerforschung und Bildungspolitik zwischen Theorie und Empirie*. Frankfurt/M.: Lang, 191-206.
- Lamsfuß-Schenk, Stefanie (2002d), *Andere Sprache - andere Geschichte? Veränderung des Geschichtslernens durch den Gebrauch einer Fremdsprache*. In: Voss, Bernd & Stahlheber, Eva (Hrsg.) (2002), *Fremdsprachen auf dem Prüfstand. Innovation, Qualität, Evaluation*. Berlin: Pädagogischer Zeitschriftenverlag, 173-181.
- Lemke, Jay L. (1990), *Talking Science: Language, Learning, and Values*. Norwood, NJ: Ablex.
- Möller, Kristin (in Arbeit), *Analyse der fachbezogenen Sprachverwendung von monolingualen Geographielernern in schriftlichen Aufgabenlösungen und deren Überarbeitung*. Osnabrück: Universität (Unveröffentl. Ms).
- Mukherjee, Vera (1999), *Schriftlichkeit im bilingualen Unterricht: Kohäsive Merkmale in schriftlichen L2-Daten bilingual deutsch-englisch unterrichteter Schüler der 7. Jahrgangsstufe*. Kiel: I & f Verlag.
- Nation, I. S. Paul (2001), *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Niemeier, Susanne (2005), *Bilingualismus und "bilinguale Bildungsgänge" aus kognitiv-linguistischer Sicht*. In: Bach, Gerhard & Niemeier, Susanne (Hrsg.) (2005), 23-45.
- Peschl, Markus F. (1999), The development of scientific concepts and their embodiment in the representational activities of cognitive systems: neural representation spaces, theory spaces, and paradigmatic shifts. In: Van Looche, Philipp R. (Hrsg.) (1999), *The nature of concepts. Evolution, structure and representation*. London: Routledge, 184-214.
- Riemer, Claudia (2006), *Entwicklung in der qualitativen Fremdsprachenforschung: Quantifizierung als Chance oder Problem?* In: Timm, Johannes-Peter (Hrsg.) (2006), 451-464. (Wiederabgedruckt in: Vollmer, Helmut Johannes (Hrsg.) (im Druck), 31-42).
- Ritter, Nancy A. (2005), On the Status of Linguistics as a Cognitive Science. *The Linguistic Review* 22, 117-133.
- Timm, Hans Christoph (2006), *"You cannot make aaa. It go not." - Eine bilinguale Insel im Mathematikunterricht*. In: Timm, Johannes-Peter (Hrsg.) (2006), 213-230.
- Timm, Johannes-Peter (Hrsg.) (2006), *Fremdsprachenlernen und Fremdsprachenforschung: Kompetenzen, Standards, Lernformen, Evaluation. Festschrift für Helmut Johannes Vollmer*. Tübingen: Narr.
- Troschke, Randi Charlotte (2006), *Kollaboratives Problemlösen: inhaltliche, sprachliche und soziale Bedeutungsaushandlung bilingualer Lerner*. In: Timm, Johannes-Peter (Hrsg.) (2006), 231-248.
- Troschke, Randi Charlotte (im Druck), *Analyse interaktiver Gesprächsprotokolle: Inhaltliche, sprachliche und soziale Aushandlungsprozesse beim kollaborativen Problemlösen*. In: Vollmer, Helmut Johannes (Hrsg.), (im Druck), 227-245.
- Troschke, Randi (in Arbeit), *Kollaboratives Problemlösen: Inhaltliche, sprachliche und soziale Bedeutungsaushandlung bilingualer Lerner bei der Bearbeitung erdkundespezifischer Aufgaben*. Osnabrück: Universität (Dissertation).
- van Looche, Philipp R. (1999), The structure and representation of concepts. In: Van Looche, Philipp R. (Hrsg.) (1999), *The nature of concepts. Evolution, structure and representation*. London: Routledge, 1-7.

- Vollmer, Helmut J. (2002), *Fachlernen und Fremdsprachlichkeit im bilingualen Unterricht: Aufgabenbasierte Kognition, Interaktion, Kooperation*. Osnabrück: Universität Osnabrück (Forschungsantrag an die DFG vom 19.8.2002, unveröff. Ms. 35 S.).
- Vollmer, Helmut J. (2004), *(Fremd)Sprachlicher Kompetenzerwerb im bilingualen Sachfachunterricht*. In: Breidbach, Stephan & Bonnet, Andreas (Hrsg.) (2004), 63-64.
- Vollmer, Helmut J. (2005a), *Bilingualer Sachfachunterricht als Inhalts- und als Sprachlernen*. In: Bach, Gerhard & Niemeier, Susanne (Hrsg.) (2005), 47-70.
- Vollmer, Helmut J. (2005b), *Förderung des Spracherwerbs im bilingualen Sachfachunterricht*. In: Bach, Gerhard & Niemeier, Susanne (Hrsg.) (2005), 131-150.
- Vollmer, Helmut J. (2005c), *Das Gespenst der Standardisierung geht um oder: Lehren und Lernen fremder Sprachen auf der Basis von Bildungsstandards*. In: Bausch, Karl-Richard; Christ, Herbert; Königs, Frank G. & Krumm, Hans-Jürgen (Hrsg.) (2005), *Bildungsstandards für den Fremdsprachenunterricht auf dem Prüfstand*. Tübingen: Narr, 258-271.
- Vollmer, Helmut Johannes (2006a), *Bildungsstandards von oben - Bildungsstandards von unten*. *Der Fremdsprachliche Unterricht - Englisch* 40: H. 81, 12-16.
- Vollmer, Helmut Johannes (2006b), *Kompetenzaufgaben als Forschungs- und als Evaluationsinstrument*. In: Bausch, Karl-Richard, Burwitz-Melzer, Eva, Königs, Frank G. & Krumm, Hans-Jürgen (Hrsg.) (2006), *Aufgabenorientierung als Aufgabe*. Beiträge zur 26. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts. Tübingen: Narr, 244-255.
- Vollmer, Helmut Johannes (2006c), *Fachlernen und Sprachlichkeit: Aufgabenbasierte Kognition, Kooperation und Kommunikation bei bilingualen und monolingualen Lernern*. In: Hahn, Angela & Klippel, Friederike (Hrsg.) (2006), *Sprachen schaffen Chancen*. München: Oldenbourg, 173-187.
- Vollmer, Helmut Johannes (2006d), *Language(s) Across the Curriculum - A way towards plurilingualism*. In: Martyniuk, Waldemar (Hrsg.) (2006), *Towards a Common European Framework of Reference for Language(s) of School Education*. Kraków: Jagellonian University (erscheint). (Gleichnamige Expertise für den Europarat/Strasbourg http://www.coe.int/t/dg4/linguistic/Conference06_Docs_rev_EN.asp#TopOfPage).
- Vollmer, Helmut Johannes (im Druck), *Zur Modellierung und empirischen Erfassung von Fachkompetenz am Beispiel der Geographie*. In: Vollmer, Helmut Johannes (Hrsg.) (im Druck), 267-288.
- Vollmer, Helmut Johannes (Hrsg.) (im Druck), *Synergieeffekte in der Fremdsprachenforschung: Empirische Zugänge, Probleme, Ergebnisse*. (Dokumentation der gleichnamigen Tagung in Osnabrück, 11.-12. November 2005). Frankfurt/M.: Lang. (erscheint Anfang 2007).
- Vollmer, Helmut Johannes (in Arbeit), *Geographische Kompetenzen von monolingualen und bilingualen Lernern am Ende der Klasse 10 des Gymnasiums. Ergebnisse einer empirischen Untersuchung*. Osnabrück: Universität. (Monographie).
- Weinert, Franz E. (2001a), *Concepts of competence: a conceptual clarification*. In: Rychen, Dominique S. & Salganik, Laura H. (Hrsg.) (2001), *Defining and Selecting Key Competencies*. Seattle: Hogrefe & Huber, 45-65.
- Weinert, Franz E. (2001b), *Vergleichende Leistungsmessung in Schulen. Eine umstrittene Selbstverständlichkeit*. In: Weinert, Franz E. (Hrsg.) (2001), *Leistungsmessungen in Schulen*. Weinheim: Beltz, 17-31.

- Wildhage, Manfred & Otten, Edgar (2003), *Praxis des bilingualen Unterrichts*. Berlin: Cornelsen Scriptor.
- Wode, Henning (1999), *Incidental vocabulary learning in the foreign language classroom*. *Studies in Language Acquisition* 21, 243-258.
- Wode, Henning; Burmeister, Petra; Daniel, Angelika; Kickler, Kay-Uwe & Knust, Maike (1996), *Die Erprobung von deutsch-englisch bilingualen Unterricht in Schleswig-Holstein: Ein erster Zwischenbericht*. *Zeitschrift für Fremdsprachenforschung* 7: 1, 15-42.
- Wolff, Dieter (1997), *Bilingualer Sachfachunterricht: Versuch einer lernpsychologischen und fachdidaktischen Begründung*. In: Freie und Hansestadt Hamburg (Hrsg.): *Initiativen für den Fremdsprachenunterricht an Hamburger Gymnasien. Dokumentation der Fachtagung vom 23.-26. April 1997*. Hamburg: Behörde für Schule, Jugend und Berufsbildung - Amt für Schule, 45-53.
- Zydati, Wolfgang (2002), *Konzeptuelle Grundlagen einer eigenständigen Didaktik des bilingualen Sachfachunterrichts: Forschungsstand und Forschungsprogramm*. In: Breidbach, Stephan; Bach, Gerhard & Wolff, Dieter (Hrsg.) (2002), *Bilingualer Sachfachunterricht: Didaktik, Lehrer-/Lernerforschung und Bildungspolitik zwischen Theorie und Empirie*. Frankfurt/M.: Lang, 31-61.
- Zydati, Wolfgang (2004), *Überlegungen zur fächerübergreifenden Evaluation des bilingualen Unterrichts: Textkompetenz als Schlüsselqualifikation fremdsprachigen Sachfachlernens*. In: Bonnet, Andreas & Breidbach, Stephan (Hrsg.) (2004), 91-102.
- Zydati, Wolfgang (2005), *Diskursfunktionen in einem analytischen curricularen Zugriff auf Textvarietäten und Aufgaben des bilingualen Sachfachunterrichts*. *Fremdsprachen Lehren und Lernen* 34, 156-173.

Anhang: Aufgabenbeispiele

TASK 1. DISTRIBUTION

In this section, we would first like you to locate the tropical rain forests. Then, you are going to look at specific areas in the tropical rain forest in more detail. Please write down your answers on the separate answering sheets.

- a) Have a look at the map of the world (Figure 1) showing the distribution of the tropical rain forest. Describe as precisely as possible where exactly in the world tropical rain forests can be found. Please, do not write on the map!⁷
- b) Study the climate graph for Kisangani (Figure 2). Describe the exact location of the Kisangani weather station by also looking at the map (Figure 1).
- c) What type of climate does Kisangani have? Describe Kisangani's climate in detail with the help of the chart.
- d) Medan is a place in Indonesia and has the following coordinates: 3° N/98° E. Is the climate in Medan different from the climate in Kisangani? Motivate your answer.

⁷ Bei Figure 1 handelt es sich um eine farbige (englischsprachige) Weltkarte im DIN A1-Format, die den Schülerinnen und Schülern zur Verfügung gestellt wird, die jedoch nicht elektronisch vorliegt und deshalb hier nicht abgebildet werden kann.

e) Take a look at South America in Figure 1. Why are there no tropical rain forests in the west of South America?

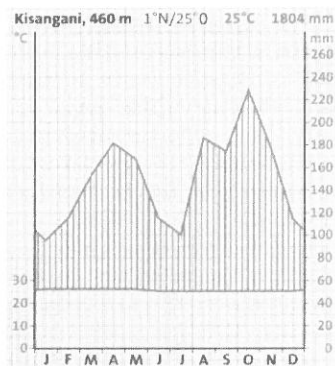


Figure 2: Climate graph for Kisangani

Zu Beispielaufgabe 1: Lage des Tropischen Regenwalds

Bei Aufgabe 1 geht es allgemein um die Lage des Tropischen Regenwaldes als auch um ausgewählte Orte und Bereiche desselben. Es ist eine sehr fachtypische Aufgabe, weil in ihr sowohl mit der Weltkarte (Figure 1) als auch mit einem Klimadiagramm (Figure 2) gearbeitet werden muss. Im Folgenden will ich mich auf die letzte der fünf Teilaufgaben (Aufgabe 1e) beschränken. Vorausgegangen ist Aufgabe 1c), die eine genaue Beschreibung des Klimas an einem Ort in Afrika namens Kisangani mit Hilfe eines vorgegebenen Klimadiagramms erforderte. Dabei waren im Detail Informationen aus diesem Diagramm zu entnehmen, zu verbalisieren und zu einer Klimabestimmung zusammenzuführen. Es folgte als Aufgabe 1d) eine Vergleichsaufgabe zwischen Kisangani und dem Ort Medan in Indonesien, mit seinen Lagedaten 3° Nord/98° Ost. Die Frage lautete: "Is the climate in Medan different from the climate in Kisangani? Motivate your answer." Es war also außer dem Vergleich und seinem Ergebnis auch eine Begründung zu geben. In der Teilaufgabe 1e) nun ist eine Besonderheit zu identifizieren und zu erklären: "1e) Take a look at South America in Figure 1. Why are there no tropical rain forests in the west of South America?" (Stamm der Aufgabe).

Die Beantwortung der Frage und damit die Erklärung der in ihr genannten Tatsache kann zu einem großen Teil mit Hilfe der Informationen geleistet werden, die auf der mitgelieferten Weltkarte im DIN A1-Format vorliegen. Damit ist diese Aufgabe im Wesentlichen eine prozedurale, die nur bedingt auf fachliches Vorwissen zurückgreift. Der Lerner muss die Weltkarte (Figure 1) mit allen Benennungen und Farbgebungen sowie sonstigen Detailinformationen (vor allem

die ausführliche Legende) richtig lesen und interpretieren können. Auf der Karte ist eindeutig die unmittelbare Nachbarschaft des Höhenzuges der Anden zu erkennen, allerdings muss der Lerner die damit verbundenen klimatischen Folgen (z.B. relativ schneller Abfall der Temperatur pro Höheneinheit) von sich aus kennen. Die zweite Information, die zunächst ebenfalls der Karte zu entnehmen ist, richtet sich darauf, dass im Westen Südamerikas eher Hochsteppen und Trockengebiete anzutreffen sind als Tropischer Regenwald. Dies bedarf einer genaueren Erklärung, die nur unter Rückgriff auf Vorwissen zu erbringen ist: es gibt hier nämlich Winde (Passatwinde), die keine tropischen Klimaverhältnisse zulassen. Damit hat der Schüler eine Reihe von inhaltlichen Möglichkeiten an der Hand, um die Frage 1e) sachkompetent zu beantworten. Entscheidend ist, wie die Versprachlichung der Antwort erfolgt bzw. am Ende ausfällt. Hier sind mehrere Argumente jeweils in Satzform zu erwarten, die miteinander verknüpft werden müssen, um damit zu einem konzeptuellen Begründungsgeflecht in Textform zu kommen - genau dies passiert allerdings in den seltensten Fällen.

TASK 6. SUSTAINABILITY

Since the world summit of Rio de Janeiro in 1992 (Agenda 21) the principle of sustainability has been globally accepted. This principle means that in terms of the use of natural resources, ecological, social and economic goals should be treated as equally important. At the same time, the rights and needs of future generations should be respected, so that they are not disadvantaged in any way through the exploitation of the resources.

Against the background of the aims of Agenda 21, how do you evaluate the following suggestion for a solution? Support your answer.

Beispielaufgabe 6. Leitbild der Nachhaltigkeit

Diese Aufgabe weicht in ihrer Diskursfunktion deutlich von der Aufgabe 1e) ab. In ihr wird weder nach einer Beschreibung noch nach einer Erklärung oder nach geographischem Grundwissen überhaupt gefragt, sondern nach der Beurteilung eines Lösungsvorschlags für die "Rettung" des tropischen Regenwalds auf dem Hintergrund eines Prinzips, das für das Selbstverständnis der Geographie zentral ist. Der Lösungsvorschlag in der Sprechblase ist also zu beziehen auf das im Aufgabenstamm beschriebene Leitbild der Nachhaltigkeit mit seinen vier aufgelisteten Elementen, nämlich der gleichrangigen Beachtung ökologischer, sozialer und ökonomischer Zielsetzungen sowie der Interessen nachfolgender Generationen, die in ihren Rechten und Bedürfnissen bei der Ausbeutung von Ressourcen heute bereits mit bedacht werden müssen. Insofern

ist die Aufgabe 6 eine recht komplexe, die es zuallererst erfordert, dass das Leitbild der Nachhaltigkeit in seinen verschiedenen Aspekten voll verstanden und dass dieses Leitbild auf den konkreten Lösungsvorschlag zum Schutz des verbleibenden Regenwaldes bezogen wird. Damit muss der gemachte Vorschlag im Hinblick auf die Berücksichtigung der vier genannten Komponenten erörtert und beurteilt werden.

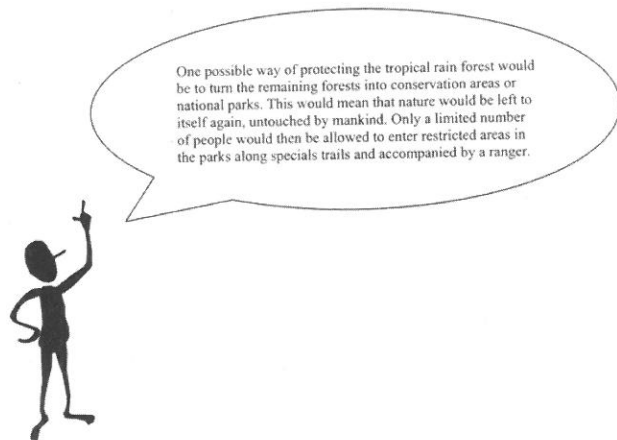


Figure 3: Suggestion for protecting the tropical rain forest

Die Begründung der Antwort erzwingt ein Abwägen der verschiedenen Kriterien des Prinzips gegenüber den Formulierungen des Vorschlags unter Einschluss der Folgen, die sich aus ihm ergeben. Insofern ist darzulegen, warum der vorliegende Lösungsvorschlag auf dem Hintergrund der Zielsetzungen, wie sie in Rio de Janeiro im Jahre 1992 ausformuliert worden sind, einseitig, nicht ausbalanciert und damit nicht haltbar ist. Dies erfordert eine differenzierte argumentative Denkweise und Verbalisierungsform. Wie bei Beurteilungsaufgaben üblich, kann der Lerner dabei seine eigene Meinung und Wertmaßstäbe gänzlich zurückhalten: es kommt lediglich auf eine rationale Überprüfung definierter Kriterien an, wie sie im Selbstverständnis und in den Zielsetzungen der Geographie selbst verankert sind. Die Aufgabe verlangt von den Lernern ein hohes Maß an sprachlichem Verstehen, an fachlich-inhaltlichem Vorstellungsvermögen und an (offenbar) relativ ungewohnter Demonstration ihrer Kompetenz. Die Relevanz und Validität der Aufgabe liegt weniger in ihrem Format begründet als in ihrem wichtigen Beitrag zur Erfassung des Konstrukts Geographische Gesamtkompetenz, hier: der Beurteilungsfähigkeit und der Diskursfunktion, die sie sprachlich realisiert. Die Aufgabe differenziert vorzüglich zwischen kompetenten und weniger kompetenten Schülerinnen Schülern.

Interkomprehension

Versuch einer Begriffsklärung

Peter Doyé¹

Of the three possible ways of defining a theoretical construct, the pragmatic one seems to be the most appropriate in the field of foreign language pedagogy. The author uses it in this article in his attempt to determine the phenomenon of intercomprehension and arrives at the following definition:

Intercomprehension is a form of communication in which each person uses his or her own language and understands that of the other.

This definition contains two critical components: the use of the mother tongue for the purpose of producing messages and the comprehension of the partner's messages when uttered in his or her language.

Two modifications of this principal determination are acceptable and are often used in the discussion of intercomprehension: the use of previously acquired foreign languages for the purpose of production and the inclusion of productive activities in the partner tongue.

The article concludes with a description of four main educational intentions concerning intercomprehension: the practical advantage, the support of language awareness and language learning awareness, the encouragement of learner autonomy, and the promotion of intercultural competence.

1. Einleitung

"Interkomprehension" ist in aller Munde, und es besteht - wie bei allen innovativen Konzepten - die Gefahr, dass durch laxen Gebrauch des neuen Terminus eine potentiell fruchtbare Idee schon in einem frühen Stadium an ihrer Realisierung gehindert wird. Wenn jeder interessierte Benutzer nach eigenem Gusto den neuen Ausdruck mit anderen, ihm wichtig erscheinenden Inhalten füllt, wird die originäre Idee verwässert und kann schon deshalb nicht zur praktischen Entfaltung kommen (vgl. Seiffert & Radnitzky 1989).

Im Falle von "Interkomprehension" ist diese Gefahr sehr akut. Gerade der Vorteil der leichten Übersetzbarkeit von einer europäischen Sprache in alle anderen bringt bei mangelnder internationaler Abstimmung den Nachteil mit sich, dass sich regional unterschiedliche Bräuche herausbilden und der Diskurs dadurch erschwert wird.

¹ Korrespondenzadresse: Prof. em. Peter Doyé, Technische Universität Braunschweig, Englisch Seminar, Bienroder Weg 80, 38106 Braunschweig, E-Mail: p.doye@freenet.de.

Intercomprehension	Intercomprensión
Intercomprensione	Intercomprehsión
Intercompreensão	Interkomprehension

Was ist der gemeinsame Nenner? Was ist die originäre Idee?

Vorsicht ist geboten hinsichtlich des Ursprungs des Konzepts, zumal der Gedanke, ein kommunikatives Gegenmodell zu dem Modell der anscheinend so nützlichen lingua franca aufzubauen, ja offenbar an mehreren europäischen Orten gleichzeitig gefasst worden ist und von Anfang an variierende Versionen auftraten. Das führte dazu, dass Rieder, Neuburg & Schindler noch 2002 meinten, dass es vielleicht zu früh sei, eine präzise und umfassende Definition des Begriffs vorzunehmen, um nicht fruchtbare Varietäten im Keim zu ersticken (Rieder et al. 2002: 6). Wir vertreten die entgegengesetzte Auffassung.

Bei allem Respekt vor einer vorläufigen Zurückhaltung (im Sinne Rieders et al.) scheint uns mehr dafür zu sprechen, bei so komplexen Phänomenen wie der Interkomprehension möglichst früh eine genaue Definition zu versuchen, um Leerlauf zu vermeiden. Wissenschaftliche Kommunikation ist auf exakte Terminologie angewiesen. Theoretiker und die ihnen folgenden Praktiker können sich nur dann über komplexe Sachverhalte und Probleme verständigen, wenn sie in ihren zentralen Konzepten übereinstimmen und mit ihren wichtigsten Bezeichnungen das Gleiche meinen.

Im Folgenden soll daher versucht werden, eine möglichst genaue inhaltliche Bestimmung des Terminus "Interkomprehension" vorzunehmen. Welche Wege sind gangbar? Es gibt hauptsächlich drei Wege zu dem genannten Ziel:

- den etymologischen,
- den konformistischen und
- den pragmatischen Weg.

Der etymologische Weg

Es liegt nahe, im Falle von Unklarheit oder Uneinigkeit über den Inhalt eines Ausdrucks zu fragen, woher denn dieser Ausdruck stammt; denn wenn man die Entwicklung des betreffenden Wortes zurückverfolge, so meinen die Anhänger dieses Vorgehens, könne man ja dessen "eigentliche" Bedeutung ermitteln. Dies ist jedoch ein kühner Schluss: Wer so vorgeht, übersieht, dass die Inhalte von Wörtern sich wandeln und dass der Rückgriff auf die frühere Bedeutung eines Wortes nicht alleiniger Maßstab für seine heutige Verwendung sein kann. Standop (1962: 129) hat daher mit Recht davor gewarnt, bei der inhaltlichen Bestimmung in Frage stehender Termini ausschließlich etymologisch zu verfahren. Der Ingenieur von heute sei eben kein "Kriegs- oder Zeugmeister" und auch nicht immer ein "Scharfsinniger". Deshalb ist auch der Hinweis von Fanzone (2005: 66)

auf den lateinischen Ursprung von "Intercomprensione" nicht gerade hilfreich: "Facciamo riferimento al etimologia latina **cum-prendere** = *prendere con sé o anche sentire con, cioè essere insieme, capire e condividere*.

Der konformistische Weg

In Ermangelung einer besseren Bezeichnung meinen wir damit den Weg, sich der von vielen Vertretern einer Disziplin verwendeten Terminologie anzupassen und den am häufigsten vorkommenden Sprachgebrauch einfach zu übernehmen. Man studiert die schriftlichen und mündlichen Äußerungen einer größeren Zahl von kompetenten Vertretern des betreffenden Bereichs und stellt fest, in welcher Bedeutung sie den fraglichen Ausdruck gebrauchen. Im Falle uneinheitlicher Verwendung ermittelt man die am meisten auftretende Bedeutung und schließt sich dieser an. Bei "Interkomprehension" ist dieses Vorgehen ausgesprochen schwierig. Die mittlerweile schier unübersehbare Vielzahl von Definitionen mit teilweise stark abweichenden Festlegungen macht den konformistischen Weg kaum gangbar.

Der pragmatische Weg

Dieser Weg besteht darin, dass man sich fragt, welche der möglichen Sprachregelungen am besten geeignet ist, Klarheit in dem betreffenden Bereich zu schaffen. Charakteristisch für dieses Vorgehen sind folgende Fragen: Welche der möglichen terminologischen Regelungen trägt am ehesten dazu bei, das betreffende Gebiet durchsichtig zu machen? Welche Definitionen der zu bezeichnenden Sachverhalte leisten die größte Denkhilfe? Ein gutes Beispiel liefert der Vorschlag von Götz & Burgschmidt (1971: 33), den Terminus "Phonologie" zum Oberbegriff für die wissenschaftliche Befassung mit der lautlichen Seite der Sprache zu erheben und ihr die Disziplinen "Phonetik" (für die physikalische Dimension) und "Phonematik" (für die kommunikative Dimension) unterzuordnen. Man erhielte dadurch - parallel zu den Bezeichnungen "Lexik" und "Grammatik" für die beiden anderen Ebenen - einen umfassenden Begriff für die Gesamtheit der Bemühungen um die lautliche Ebene der Sprache. Der Vorschlag ist klar und logisch leicht nachzuvollziehen, wurde aber leider vom Mainstream der Linguistik nicht akzeptiert, weil er so gar nicht mit der traditionellen Terminologie konform ging (s.o.).

2. Eine pragmatische Definition

Im Falle von "Interkomprehension" bieten sich folgende pragmatische Fragen als Ausgangspunkt an: Was zeichnet das neue Konzept aus? Was macht dieses Konzept nützlich? Wozu kann es dienen?

2.1 Unterordnung

Wichtig ist es zu bedenken, dass das Konzept "Interkomprehension" sowohl pädagogisch als auch sprachpolitisch motiviert ist. Es ist als Gegenmodell zu dem Modell der *lingua franca* gedacht und soll deutlich machen, dass die Menschen in Europa auch anders als über eine allgegenwärtige und deshalb dominierende Verkehrssprache kommunizieren können. Kommunikation muss also eine wesentliche Komponente des neuen Konzepts ausmachen. In genau diesem Sinne sprechen eine Reihe von Autoren und ihre Dokumente von

- "un tipo de comunicación" (Llorente Pinto 2002),
- "eine Form der Kommunikation" (Rieder et al. 2002),
- "a form of communication" (Doyé 2005),
- "um modo de comunicação" (Martins 2005),
- "un tipo di comunicazione" (Fanzone 2005).

Eine solche Festlegung auf "Kommunikation" als übergeordneten Begriff ist praktisch sehr vorteilhaft. Sie verwendet einen in mehreren Disziplinen eingeführten und klar umrissenen Oberbegriff und schafft damit eine solide Basis für die anschließend nötige Spezifizierung. Dies bedeutet aber gleichzeitig eine Absage an die Aufnahme weiterer und meist unnötiger Attribute. Deren häufigste ist die Hinzufügung von "Fähigkeit zu ...", also "intercomprehension = the ability to communicate by ..." (Capucho 2003). Diese Hinzufügung ist im pragmatischen Sinne (s.o.) unzweckmäßig, weil sie ja die Definition unnötig einengt und ihr die erforderliche Basis nimmt. Was wir zunächst brauchen, ist ein Terminus zur Bezeichnung des Phänomens selbst; daran anschließend können wir dann durch Ausdrücke wie "ability to ...", "Fähigkeit zu ...", "competência em ...", "capacité de ..." oder auch "possibilità di ..." die Möglichkeit zum Vollzug des Phänomens benennen, wenn wir eben diese Potenz artikulieren wollen.

2.2 Spezifizierung

Eine Definition im klassischen Sinne besteht aus der Unterordnung des fraglichen Begriffs unter einen Oberbegriff (*genus proximum*) und die Angabe des spezifischen Unterschieds (*differentia specifica*). Im Falle von "Interkomprehension" sind zwei unterscheidende Merkmale unstrittig. Es sind dies

- der Gebrauch der Erstsprache der Kommunikanten bei der Produktion,
- das Verstehen der Sprache der Kommunikationspartner bei der Rezeption.

Diese beiden Merkmale sind als Pfeiler, auf denen Interkomprehension ruht, generell anerkannt und stellen auch pragmatisch gesehen eine sinnvolle Spezifizierung dar. Jede Kommunikation hat nun einmal eine produktive und eine

rezeptive Seite, und deshalb ist es zweckmäßig, in dem speziellen Fall interkomprehensiver Kommunikation anzugeben, durch welche Merkmale sich diese auf den beiden Seiten auszeichnet. So gehen denn auch die meisten Theoretiker vor. Eco (1993: 19) spricht von "persone che possono incontrarsi parlando ciascuno la propria lingua e intendendo quella dell'altro". Im "Language Education Policy Guide" des Europarats heißt es parallel dazu: "Speakers of the languages in question each speak their own language and can be understood by the others" (Language Policy Division 2006: 37). Llorente Pinto (2002: 9) formuliert: "Cada interlocutor habla su lengua y entiende la del otro" (Vgl. auch Rieder 2001: 8 und Doyé 2004: 60). Es herrscht also weitgehend Übereinstimmung im Grundsätzlichen. Allerdings bezieht sich die Mehrzahl der Definitionen vornehmlich auf den mündlichen Sprachgebrauch, was nicht bedeuten muss, dass sie nicht auf den schriftlichen ausgedehnt werden können, denn dort herrschen vergleichbare Zustände. Immer mehr Menschen in Europa verkehren schriftlich miteinander – sei es brieflich oder per E-Mail – und dabei ist die interkomprehensive Form keine Seltenheit mehr. Im Folgenden sollen nun die beiden Komponenten, also die produktive und die rezeptive, näher betrachtet werden.

2.2.1. Merkmal 1: Produktion

Wenn sich ein Mensch in der interlingualen Kommunikation seiner Erstsprache bedient, tut er es in der Regel deshalb, weil er diese besser beherrscht als andere Sprachen und sich in ihr am besten ausdrücken kann. Das verschafft ihm Sicherheit und verleiht seiner Botschaft größtmögliche Klarheit – natürlich unter der Voraussetzung, dass sie der Partner versteht. Diese Sicherheit und Klarheit sind subjektiv wichtig, aber auch wegen der Schaffung gleicher Bedingungen für beide Partner. Auf diese Weise wird eine wesentliche Voraussetzung für einen "herrschaftsfreien Dialog" (Habermas) erfüllt, und die Kommunikation leidet nicht unter der Dominanz einer der beiden Seiten. Wenn einer von zwei Kommunikanten seine Erstsprache benutzen kann und der andere auf eine nicht so gut beherrschte Sprache angewiesen ist, ist ersterer immer im Vorteil. Ganz in diesem Sinne besteht die Europäische Union so beharrlich auf der Gleichberechtigung aller in der Gemeinschaft gesprochenen Sprachen – aus allgemeinen kulturellen Gründen, aber auch aus Gründen der persönlichen Befindlichkeit der Handelnden in konkreten Kommunikationssituationen (Europäisches Parlament 2000: 1).

2.2.2 Merkmal 2: Rezeption

Dass bei der Interkomprehension die Kommunikanten die – je nach Situation – mündlichen oder schriftlichen Äußerungen ihrer Partner verstehen können, auch oder gerade wenn diese in Sprachen getätigt werden, die sie nicht formal gelernt haben, ist das andere Hauptmerkmal dieser Form von Kommunikation. Hierin stimmen die meisten gängigen Definitionen überein.

Unterschiede gibt es allerdings bei der Annahme der Gründe, weshalb denn ein solches Verstehen von offenbar Unbekanntem möglich sei. Manche Autoren machen die prinzipielle Fähigkeit zur Interpretation von Zeichen, die zur geistigen Grundausstattung aller Menschen gehöre, dafür verantwortlich und stellen fest, dass diese Fähigkeit bei jeder Dekodierung sprachlicher Zeichen, auch solcher aus unbekanntem oder wenig bekannten Systemen ins Spiel komme (z. B. Pencheva & Shopov 2003). Andere führen diese Fähigkeit auf die vielfältigen Wissensbestände zurück, die jeder Mensch, wenn er die Interkomprehension aufnimmt, ja schon immer besitzt und zweckgerecht einbringen kann (z. B. Doyé 2004). Einig sind sich die Protagonisten aber wieder in der Überzeugung, dass die Verstehensfähigkeit entwickelt werden kann, und machen die Vermittlung entsprechender Strategien zum Kernstück ihrer Mehrsprachigkeitsdidaktik (z. B. Meissner 2004).

Eine Sonderstellung bei der rezeptiven Komponente nimmt das Lesen von Sachtexten und literarischen Texten ein, bei welchem ja der Kommunikationspartner, also der Autor, nicht präsent ist oder - im Gegensatz zum Briefverkehr - dem Rezipienten auch gar nicht bekannt ist. Dieser Form von Interkomprehension haben sich die Vertreter der EuroCom-Gruppe besonders zugewandt, deren Hauptwerk "Die sieben Siebe" (Klein & Stegmann 2000) deshalb ja auch den bezeichnenden Untertitel trägt: Romanische Sprachen sofort lesen können. Aber prinzipiell unterscheidet sich auch diese Art der Interkomprehension nicht von den anderen. Immer geht es um das Verstehen von Äußerungen in nicht formal gelernten Sprachen. Fassen wir die bisherigen Darstellungen zusammen, so ergibt sich folgende Definition:

Interkomprehension ist eine Form der Kommunikation, bei der sich die Kommunikanten ihrer Erstsprache bedienen und die ihrer Kommunikationspartner verstehen.

3. Modifikationen

In der gerade gegebenen Definition sind die wesentlichen Merkmale enthalten.

In den letzten Jahren sind nun aber Aspekte der Interkomprehension zur Diskussion gestellt worden, die diese grundsätzliche Inhaltsbestimmung auf interessante Weise modifizieren. Die wichtigsten von ihnen sollen im Folgenden vorgestellt werden.

3.1 Modifikation 1

Die Verfechter der ersten Modifikation stellen in Frage, ob es denn die Erstsprache der Kommunikanten sein müsse, in welcher diese sich bei interlingualen Begegnungen äußern, ob es nicht ebenso als Interkomprehension

angesehen werden könne, wenn die Partner eine andere von ihnen gut beherrschte Sprache verwenden. Die Mitglieder des spanischen ILTE-Projekts weisen mit Recht darauf hin, dass dies eine häufig vorkommende Konstellation sei und geben Beispiele.

Alguno de los interlocutores utiliza, bien lenguas distintas de la materna, que funcionarían como lenguas-pivote; por ejemplo, el italiano en el caso de una chica rusa que trabaja en Italia e qui contesta en italiano a las preguntas en español que le hace un periodista de nuestro país (Llorente Pinto 2002: 9).

3.2 Modifikation 2

Die zweite Modifikation betrifft die nach Auffassung ihrer Befürworter unnötige Beschränkung auf die Rezeption der Zielsprache bei der interkomprehensiven Kommunikation, also auf den Ausschluss der Produktion. Diese sei in der Anfangsphase von großem Vorteil: Die Lernenden brauchten sich nicht von Beginn an um beide Seiten der Kommunikation zu bemühen, sondern könnten sich auf das Verstehen konzentrieren. Dies braucht aber bei längerer Beschäftigung mit der Zielsprache nicht so zu bleiben. Protokolle des lauten Denkens beweisen, dass viele Lernende die Beschränkung auf die Rezeption der anderen Sprache als störend empfinden und sich deshalb auch in der Produktion versuchen (Meißner 2003: 98). In der Kunstpädagogik gilt der Satz "Der Eindruck strebt zum Ausdruck". Dies scheint für viele Menschen auch auf die Begegnung mit Sprachen zuzutreffen. Ausführlich behandelt hat dieses Phänomen Amoruso (2005). Sie nennt es "le due facce dell'intercomprensione" und beschreibt, wie die beiden Seiten der Kommunikation bei der Interkomprehension interagieren. Didaktisch gesehen bedeutet diese Tendenz eine große Chance: Wenn denn die Rezeption von Äußerungen in einer bisher nicht gelernten Sprache die Kommunikanten geradezu zum Erlernen der Produktion drängt, dann bedeutet dies eine ideale Motivation. Die Lernenden können nach Belieben zum Gebrauch dieser Sprache fortschreiten, es wäre töricht, diese Chance didaktisch nicht zu nutzen.

4. Pädagogische Intentionen

In einer didaktisch orientierten Publikation wie dieser bietet es sich an, das Phänomen "Interkomprehension" im Anschluss an seine pragmatische Definition und mögliche Modifikationen auf seine pädagogische Potenz hin zu durchleuchten. Dazu brauchen die an die interkomprehensiv geprägte Praxis geknüpften Erwartungen nicht in den Rang von konstitutiven Merkmalen erhoben zu werden - etwa in dem Sinne: Interkomprehension zeichnet sich durch die Intentionen A, B und C aus; wohl aber sind in der Regel bestimmte Zielvorstellungen mit der

Förderung von Interkomprehensionskompetenz verbunden. Sie sollen im Folgenden behandelt werden. Die entscheidende Frage lautet: **Weshalb ist es sinnvoll, die Fähigkeit zur Interkomprehension bei möglichst vielen Menschen zu fördern?** Folgende Gründe sind zu nennen.

4.1 Der praktische Vorteil

Wer die Fähigkeit, interkomprehensiv zu kommunizieren, besitzt, zieht daraus **praktischen Nutzen**. Die vielen Fälle gelungener Interkomprehension im skandinavischen Norden und im romanischen Süden Europas, aber auch in zweisprachigen Grenzgebieten belegen dies. Die dabei benutzten Strategien und Techniken auch möglichst vielen anderen Europäern zu vermitteln, ist eine offensichtlich sinnvolle didaktische Intention. Wie weit man dabei gehen kann und vor allem, ob ein gegenseitiges Verständnis über die Grenzen von Sprachfamilien hinweg möglich ist, muss noch erforscht werden. Jedoch ist die genannte Intention prinzipiell vernünftig. Hinzukommt, dass die Interkomprehensionsmethode durchaus als eine Alternative zum traditionellen Lehren und Lernen von Fremdsprachen in Frage kommt und dieses, wenn schon nicht ersetzen, so doch ergänzen könnte. Scherfer (2002: 83) nennt die Interkomprehension **"ein zeit- und energiesparendes Fremdsprachenlernprinzip"** und zeigt,

4.2 Die Förderung von Sprachbewusstheit

Die Befähigung zur Interkomprehension fördert auch die Sprachbewusstheit der Lernenden. Im Gefolge der Aufwertung der rationalen Erkenntnis der Welt und damit der Selbsterkenntnis des Menschen im 20. Jahrhundert ist es auch in der Sprachdidaktik zu einer Renaissance kognitiv orientierter Methoden gekommen. Language awareness und language learning awareness finden sich in fast allen Sprachprogrammen als Zielsetzungen.

Diesen Forderungen wird nun die Interkomprehensionsdidaktik in idealer Weise gerecht. Wer interkomprehensiv kommunizieren will, muss auf seine Vorerfahrungen zurückgreifen und lernen, diese Erfahrungen strategisch einzusetzen. Dabei kann ihm seine Lehrerin helfen, indem sie ihm bewusst macht, auf welchen Gebieten er Vorkenntnisse besitzt und in welchem Umfang. Deshalb ist der Interkomprehensionsunterricht geradezu prädestiniert, die Sprachbewusstheit der Lernenden zu fördern. Diese können gar nicht umhin, über Sprache nachzudenken und ihr eigenes Lernen zu reflektieren. Sie müssen auf die Vorkenntnisse in den verschiedensten Bereichen ihres sprachlichen, sachlichen und psychologischen Wissens zurückgreifen und sie bei der Entschlüsselung der Äußerungen und Texte in der fremden Sprache einsetzen (vgl. Doyé 2005).

Meißner, Meißner, Klein & Stegmann (2004: 64) betrachten die "conscientisation" als den Hauptvorteil des interkomprehensiven Lernens und weisen den Lehrenden vor allem die Aufgabe zu, diese Bewusstwerdung zu unterstützen. Rieder (2001: 34) hat die "conscientisation" als einen Stufenprozess dargestellt, der von der generellen Orientierung über die Sprachaufmerksamkeit und Sprachbewusstheit bis zum Bewusstsein der eigenen Vorkenntnisse und Vorerfahrungen reicht, die bei der Begegnung mit fremden Texten und Äußerungen ins Spiel kommen.

4.3 Die Förderung von Selbständigkeit

Die Förderung der Selbständigkeit hängt mit der gerade beschriebenen "conscientisation" eng zusammen. Mit dem Zugewinn an Bewusstheit wachsen auch die Möglichkeiten zu autonomem Lernen. Die Idee der Erziehung zur Selbständigkeit ist uralte und hat auch in Deutschland eine lange Tradition. Sie bildet ein Kernstück der pädagogischen Reformbewegung des 20. Jahrhunderts (Litt 1925; Nohl 1949). Ihre Bedeutung für das Fremdsprachenlernen und -lehren haben dann die Didaktiker des Europarats hervorgehoben (z. B. Holec & Huttunen 1997). Die Beziehung der Interkomprehension zu dieser Idee liegt auf der Hand: Um per Interkomprehension kommunizieren zu können, muss der Lernende sein eigenes Vorwissen für die Zwecke der Dekodierung aktivieren (s.o.). Er muss entscheiden, welche seiner Vorkenntnisse für die aktuelle Interpretation in Frage kommen, und diese mobilisieren. Dafür sind Strategien nötig, und diese sind lern- und lehrbar, - Strategien der Übertragung von Gelerntem auf neue Situationen.

Meißner (2004) sieht in diesen "stratégies transférentielles" das wesentliche Element der Interkomprehensionsmethode. Haben die Lernenden diese Strategien erworben, sind sie autonom in dem Sinne, dass sie neuen sprachlichen Äußerungen und Texten in bisher unbekannt Sprachen nicht mehr hilflos gegenüberstehen, sondern die sich stellenden Verstehensaufgaben selbständig lösen können. Dass diese Strategien dann auch anderen Kontexten des Spracherwerbs zugute kommen, dass sie den Lernenden auch beim konventionellen Fremdsprachenlernen helfen können, ist offensichtlich. Einen guten Überblick über das Konzept der Lernerautonomie und seine Realisierung in der Interkomprehension hat Martinez (2005) gegeben.

4.4 Der Gewinn an interkultureller Kompetenz

Zwei wesentliche Erkenntnisse haben die allgemeine Sprachdidaktik in den letzten Jahrzehnten bereichert: Erstens, dass wegen des engen Zusammenhanges von Sprache und Kultur jeder Sprachunterricht auch immer die Beziehung zu

anderen Kulturen einschließen müsse, und zweitens, dass der Sprachunterricht im Dienste interkultureller Erziehung stehen müsse. Soweit herrscht genereller Konsens. Die weitere Erkenntnis, dass die Hinführung zu interkultureller Kompetenz nicht auf die Kulturen der gerade gelernten/gelehrten Sprachen beschränkt sein dürfe, sondern die Fähigkeit zum Umgang mit Menschen anderer kultureller Herkunft überhaupt zum Ziel haben müsse, hat sich dagegen noch nicht allgemein durchgesetzt. Vorschläge, wie dies zu erreichen sei, gibt es zur Genüge (exemplarisches Lernen, interkulturelle Offenheit, Perspektivenübernahme).

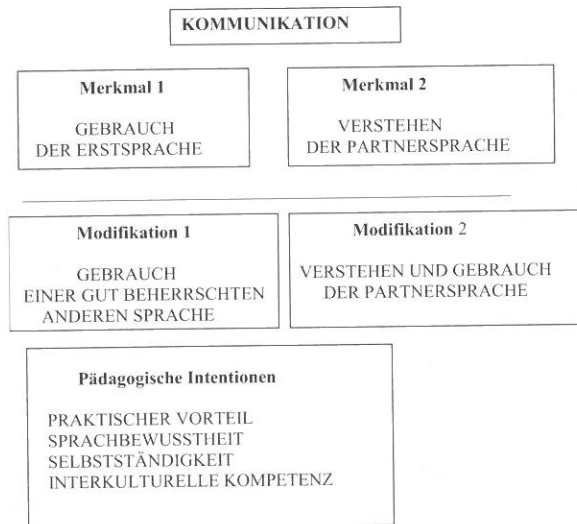


Abbildung 1: Übersicht in Stichworten

Die Interkomprehension bietet nun einen weiteren erfolgversprechenden Weg. Bei ihr lernen die Kommunikanten, sich auf fremde Sprachen überhaupt einzulassen, unabhängig davon, wie viel oder wenig sie schon von ihnen wissen oder können. Ihnen werden Strategien zur Bewältigung von Begegnungssituationen vermittelt, die nicht an bestimmte Sprachen gebunden, sondern generell gültig sind. Aber genau diese Offenheit enthält große Chancen, was für die Verfechter der Interkomprehension allein schon - im Zusammenhang mit den praktischen Vorteilen, versteht sich - ein starkes Motiv darstellt, sich für diesen neuen Weg einzusetzen. So betonen die Vertreter des spanischen ILTE-Projekts "la importancia del componente intercultural" und erklären die "competencia intercultural" zu einem übergeordneten Ziel ihrer Bemühungen (Llorente Pinto

2002). Und die Initiatorin des EU&I-Projekts nennt als Hauptziel dieses Unternehmens: "A consciência intercultural através uma focalização nas semelhanças e diferenças culturais que sustentam processos interpretativos" (Capucho 2003). Wichtige und hilfreiche Stützen für die Realisierung solcher Intentionen bilden wiederum die systematischen Darstellungen der interkulturellen Kommunikation (Donec 2002) und der interkulturellen Methodik (Müller 1995²).

Eingang des revidierten Manuskripts 23.11.2006

Literaturverzeichnis

- Amoruso, Chiara (2005), *Le due facce dell'intercomprensione*. In: Martins, Adriana (Hrsg.) (2005), 55-64.
- Capucho, Filomena (2003), *Descrição do projecto EU&I*. Viseu: Universidade Católica Portuguesa.
- Donec, Pavel (2002), *Grundzüge einer allgemeinen Theorie der interkulturellen Kommunikation*. Aachen: Shaker.
- Doyé, Peter (2004), A methodological framework for the teaching of intercomprehension. *Language Learning Journal* 30, 59-68.
- Doyé, Peter (2005), *Intercomprehension. Reference Study*. Strasbourg: Council of Europe.
- Eco, Umberto (1993), *La ricerca della lingua perfetta nella cultura europea*. Roma/Bari: Editori Laterza.
- Europäisches Parlament (2000), *Beschluss vom 17. Juli 2000 über das Europäische Jahr der Sprachen*. Brüssel: Amtsblatt der EG.
- Götz, Dieter & Burgschmidt, Ernst (1971), *Einführung in die Sprachwissenschaft für Anglisten*. München: Max Hueber Verlag.
- Holec, Henri & Huttunen, Irma (Hrsg.) (1997), *L'autonomie de l'apprenant en langues vivantes*. Strasbourg: Conseil de l'Europe.
- Fanzone, Salvatore Alberto (2005), *L'intercomprensione nella narrazione orale*. In: Martins, Adriana (Hrsg.) (2005), 65-71.
- Klein, Horst & Stegmann, Tilbert D. (2000), *EuroComRom - Die sieben Siebe*. Aachen: Shaker.
- Language Policy Division (2006), *Guide for the Development of Language Education Policies in Europe*. Strasbourg: Council of Europe.
- Llorente Pinto, Vega (2002), *Intercomprehension in Language Teacher Education 1998-2002*. Salamanca: Escuela Oficial de Idiomas.
- Litt, Theodor (1925), *Führen oder Wachsenlassen? Eine Erörterung des pädagogischen Grundproblems*. Leipzig: Quelle und Meyer.
- Martins, Adriana (Hrsg.) (2005), *Building Bridges. European Awareness and Intercomprehension*. Viseu: Universidade Católica Portuguesa.
- Martinez, Hélène (2005), Lernerautonomie - ein konzeptuelles Rahmenmodell für den Fremdsprachenunterricht ... und die Fremdsprachenlehr- und -lernforschung. *Fremdsprachen Lehren und Lernen* 34, 65-82.

2 Dass in dem vorliegenden Artikel überwiegend romanistische Quellen benutzt wurden, liegt daran, dass z.Zt. die Idee der Interkomprehension in der Romania am weitesten entwickelt ist, bedeutet aber keinesfalls, dass unser Definitionsversuch nicht auch für die germanischen, slawischen und anderen europäischen Sprachen Gültigkeit besäße.

- Meissner, Franz-Joseph; Meissner, Claude; Klein, Horst & Stegmann, Tilbert D. (2004), *EuroComRom - Les sept tamis: lire les langues romanes dès le départ*. Aachen: Shaker.
- Meissner, Franz-Joseph (2004), *Introduction à la didactique de l'eurocompréhension*. In: Meissner, Franz-Joseph; Meissner, Claude; Klein, Horst & Stegmann, Tilbert D. (Hrsg.) (2004), 7-140.
- Meißner, Franz-Joseph (2003), Grundüberlegungen zur Praxis des Mehrsprachenunterrichts. In: Meißner, Franz-Joseph & Picaper, Ilse (Hrsg.) (2003), *Mehrsprachigkeitsdidaktik zwischen Frankreich, Belgien und Deutschland*. 92-106.
- Müller, Bernd-Dietrich (1995), Steps Towards an Intercultural Methodology for Teaching Foreign Languages. In: Sercu, Lies (Hrsg.) (1995), *Intercultural Competence*. Volume I. Aalborg: Aalborg University Press, 71-116.
- Nohl, Herman (1949), *Die pädagogische Bewegung in Deutschland und ihre Theorie*. Frankfurt a. M.: Schulte-Bulmke.
- Pencheva, Maya & Shopov, Todor (2003), *Understanding Babel. An Essay in Intercomprehension Analysis*. Sofia: St. Kliment Ohridski University Press.
- Rieder, Karl (2001), *Intercomprehension. Fremdsprachige Texte entschlüsseln*. Wien: öbv & hpt.
- Rieder, Karl; Neuburg, Renate & Schindler, Ilse (2002), *Intercomprehension in Language Teacher Education 1998-2002*. Wien: Pädagogische Akademie des Bundes.
- Scherfer, Peter (2002), Was ist und wie erlernt man Interkomprehension? In: Kischel, Gerhard (Hrsg.) (2002), *EuroCom - Mehrsprachiges Europa durch Interkomprehension in Sprachfamilien*. Aachen: Shaker, 82-96.
- Seiffert, Helmut & Radnitzky, Gerard (Hrsg.) (1989), *Handlexikon zur Wissenschaftstheorie*. München: Ehrenwirth.
- Standop, Ewald (1962), Was ist "Grammatik"? *Praxis des neusprachlichen Unterrichts* 3, 129-133.

Forschungsprojekt

ADEQUA

Bericht über ein DFG-Forschungsprojekt zur Förderung von Lernstrategien im Englischunterricht

Claudia Finkbeiner, Peter H. Ludwig, Eva Wilden und
Markus Knierim¹

The ADEQUA research project, which is funded by the German Research Foundation (DFG), wants to gain empirical evidence as to how autonomous learning in the EFL classroom can effectively be supported during textbased classroom activities. For this purpose, new task formats for the EFL classroom have been developed which allow for independent work with texts in the foreign language (text comprehension) in cooperative literacy events that make cooperation among the learners compulsory. The ADEQUA project focuses on the following research questions: (1) What learning strategies do students use under what conditions? (2) Are these strategies adequate in the given learning setting? (3) What interventions do teachers identify as necessary to support the learners, and what interventions do they apply? (4) Are these teacher interventions effective? The research project is situated within the context of current efforts in education to implement high-quality student-centered, cognitively challenging classroom activities and to foster autonomous, student-active learning.

1. Zusammenfassung

Zentrales Interesse des ADEQUA-Forschungsprojekts², welches von der Deutschen Forschungsgemeinschaft (DFG) gefördert wird, ist die Gewinnung von empirisch begründeten Erkenntnissen über effektive Unterstützungsmöglichkeiten weitgehend selbstständigen Lernens in textbasierten Unterrichtsphasen des

1 Korrespondenzadressen: Prof. Dr. Claudia Finkbeiner, Universität Kassel, Sprach- & Literaturwissenschaften (FB 02), 34109 Kassel, E-Mail: cfink@uni-kassel.de; Prof. Dr. Peter H. Ludwig, Universität Koblenz-Landau, Erziehungswissenschaften (FB 05), August-Croissant-Str. 5, 76829 Landau, E-Mail: ludwig@uni-landau.de; Eva Wilden, Universität Kassel, Sprach- & Literaturwissenschaften (FB 02), 34109 Kassel, E-Mail: eva.wilden@uni-kassel.de; Markus Knierim, M.A. (USA), Universität Kassel, Sprach- & Literaturwissenschaften (FB 02), 34109 Kassel, E-Mail: mknierim@uni-kassel.de.

2 Der vollständige Titel des Forschungsprojekts lautet: "Förderung des situationsadäquaten Einsatzes von Lernstrategien in selbstständigkeitsorientierten, textbasierten Lernumgebungen im Englischunterricht". Als Kurzform wird das Kennwort ADEQUA für *adequacy of learning strategy use* verwendet.