

Jude, Nina; Klieme, Eckhard; Eichler, Wolfgang; Lehmann, Rainer H.; Nold, Günter;
Schröder, Konrad; Thomé, Günther; Willenberg, Heiner

Strukturen sprachlicher Kompetenzen

Klieme, Eckhard [Hrsg.]: Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie. Weinheim u.a. : Beltz 2008, S. 191-201

urn:nbn:de:0111-opus-31595

in Kooperation mit:

BELTZ

<http://www.beltz.de>

Nutzungsbedingungen

pedocs gewährt ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit dem Gebrauch von pedocs und der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Kontakt:

peDOCS

Deutsches Institut für Internationale Pädagogische Forschung (DIPF)

Informationszentrum (IZ) Bildung

Schloßstr. 29, D-60486 Frankfurt am Main

eMail: pedocs@dipf.de

Internet: www.pedocs.de

*Nina Jude / Eckhard Klieme / Wolfgang Eichler / Rainer H. Lehmann /
Günter Nold / Konrad Schröder / Günther Thomé / Heiner Willenberg*

18 Strukturen sprachlicher Kompetenzen

Sprachliche Kompetenzen sind ein wesentlicher Bestandteil von Bildung und werden durch schulischen Unterricht in unterschiedlicher Weise beeinflusst, so dass sie in empirischen Studien zu Schülerleistungen immer stärkere Beachtung finden (Bos u.a. 2003; Deutsches PISA Konsortium 2001). Neben der Feststellung von Ausprägungen der sprachlichen Kompetenzen, also der Einordnung von Schülerinnen und Schülern auf Kompetenzniveaus, stellt sich vermehrt auch die Frage nach der Differenzierung von Sprachkompetenzen, also danach, welche spezifischen sprachlichen Phänomene wie gut beherrscht werden. Aus sprachwissenschaftlicher und fachdidaktischer Sicht ist danach zu fragen, welche Zusammenhänge zwischen einzelnen sprachlichen Teilbereichen bestehen: Gibt es beispielsweise systematische Gemeinsamkeiten im Lesen und Schreiben, so dass ein übergeordneter, schriftsprachlicher Kompetenzfaktor angenommen werden kann? Oder sind die Zusammenhänge zwischen Leseverstehen und Hörverstehen so hoch, dass sich eine rezeptive Sprachkomponente damit beschreiben lässt? Wie lassen sich also Strukturen von Sprachkompetenz abbilden, wenn dazu Daten aus Schulleistungstudien analysiert werden?

Aus entwicklungspsychologischer Perspektive stellt der Spracherwerb einen sehr komplexen Prozess dar, der sich über einen langen Zeitraum erstreckt. Dabei werden nicht alle Teilfähigkeiten gleichzeitig und gleich gut erlernt (Szagun 2000; Haberzettl/Wegener 2003). Hörverstehen und Sprechen zählen zu den früh in der menschlichen Entwicklung ausgebildeten Kompetenzen. Der Erwerb von schriftsprachlichen Fähigkeiten beginnt meist mit der Einschulung. Auch die institutionelle Vermittlung weiterer spezifischer Teilkompetenzen wie bspw. von Wortschatz und Grammatik erfolgt dann in schulischen Lehr-Lernkontexten (Detering 2000; Eckerth 2000). Es wird angenommen, dass sich mit zunehmender gesamtsprachlicher Kompetenz die Ausprägungen der unterschiedlichen Facetten einander annähern, d.h. dass die schriftsprachlichen Leistungen in der Erstsprache den auditiven entsprechen und ihre Entwicklung mit dem Eintritt ins Erwachsenenalter größtenteils abgeschlossen ist (Oerter/Montada 2002). Die Strukturen von Sprachkompetenz bei Erwachsenen würden sich in diesem Fall in einem homogenen, globalen Faktor abbilden lassen.

Entsprechende Modelle finden sich in linguistischen Ansätzen der 1970er Jahre, die von einer singulären übergreifenden Sprachkompetenz ausgehen und alle spezifischen sprachlichen Fähigkeiten hierauf beziehen (Oller 1976; Kunnan 1992). Inzwischen legen jedoch aktuelle Forschungsergebnisse nahe, dass die Struktur von Sprachkompetenz als mehrdimensional betrachtet werden muss, und in Abhängigkeit von Alter oder Sprachhintergrund der untersuchten Personengruppen durchaus variieren kann (North 2000). Neue kommunikative Ansätze in der Leistungserhebung,

die stärker offene und mündliche Testformate verwenden, unterstützen die facettenreiche Messung von Sprachkompetenz ebenso wie statistische Analysemethoden, die die Überprüfung solcher mehrdimensionaler Modelle ermöglichen (McNamara 1995; Kuhlemeier/van den Bergh/Rijlaarsdam 2002). So können methodische Artefakte reduziert und differenzierte Modelle von Sprachkompetenz konfirmatorisch überprüft werden (Gräfe-Bentzien 2000).

Die Mehrzahl der psycholinguistischen Modelle in diesem Forschungsfeld postulieren vier Basiskomponenten – Lesen, Schreiben, Hörverstehen und Sprachproduktion – die sich wiederum den Prozessen der Sprachproduktion und -rezeption bzw. den Modalitäten der schriftlichen und auditiven Sprachverwendung zuordnen lassen. Ein differenziertes Modell der Struktur von Sprachkompetenz mit exemplarisch ausgewählten sprachlichen Teilbereichen ist in Tabelle 18.1 dargestellt.

Tabelle 18.1: Prozess-Modalitätsmodell mit differenzierten Facetten sprachlicher Kompetenz.

Modalität Prozess	Auditiv	Schriftsprachlich
Produktion	Sprechen <i>Grammatik</i> <i>Wortschatz</i> Aussprache	Schreiben <i>Grammatik</i> <i>Wortschatz</i> <i>Rechtschreiben</i>
Reflektion Rezeption	Hören <i>Grammatik</i> <i>Wortschatz</i> <i>Wahrnehmungstempo</i>	Lesen <i>Grammatik</i> <i>Wortschatz</i> <i>Lesegeschwindigkeit</i>

Aus Perspektive der Lehr-Lernforschung interessiert, wie diese Teildimensionen zusammenhängen und welche Struktur der Sprachkompetenz sie, ggf. in Faktoren höherer Ordnung, bilden. So könnten sich die Leistungen im Lesen und Hörverstehen auf die gemeinsame Dimension „Rezeption“ zurückführen lassen (Rost/Hartmann 1992). Alternativ könnte der Zusammenhang zwischen Sprechen und Hören stärker ausgeprägt sein und sich somit die Dimension „auditive Sprachkompetenz“ abbilden lassen. Solche Zusammenhänge können Rückschlüsse auf gemeinsame kognitive Vorgänge oder spezifische Lehr-Lernansätze ermöglichen: Ein stark kommunikativ ausgerichteter Sprachunterricht könnte dazu führen, dass sich die auditiven sprachlichen Teilkompetenzen Sprechen und Hörverstehen ähnlich stark ausprägen. Entsprechend könnte ein eher schriftbasierter Unterricht die schriftsprachlichen Subkomponenten parallel fördern.

18.1 Theoretische Modelle sprachlicher Kompetenzen

Die Leistungen im Deutschen und Englischen wurden in DESI in jeweils sechs unterschiedlichen Bereichen erfasst (vgl. Kapitel 2), dabei kamen verschiedene Testformate zum Einsatz, wobei der Schwerpunkt auf der schriftgestützten Erfassung von Sprachkompetenz lag (Beck/Klieme 2007). Der Test zur mündlichen Sprachproduktion im Englischen (Set-10, vgl. Kapitel 16) wurde lediglich von einer

kleinen Stichprobe von Schülern bearbeitet und wird in die Gesamtanalysen nicht einbezogen.

Die Auswahl der sprachlichen Kompetenzen wurde durch die Zielsetzung der DESI-Studie bestimmt. Anhand der erhobenen Daten ist es möglich, eine Vielzahl unterschiedlicher, theoretisch fundierter Modelle zu analysieren. Besonders interessant für DESI ist dabei die Frage, wie differenziert die Kompetenzstruktur auf Individualebene abgebildet werden kann. Je besser dies gelingt, umso deutlicher stellen sich die unterschiedlichen Testkonstrukte dar, welche die vielfältigen sprachlichen Teilbereiche in DESI abbilden sollen. Im Folgenden werden die in Abb. 18.1 dargestellten Modelle vergleichend analysiert:

- Modelle, die einen *Gesamtfaktor* der Sprachkompetenz annehmen (G-Faktor-Modell),
- Modelle, die eine *mehrdimensionale* Differenzierung korrelierender sprachlicher Teilkompetenzen postulieren und
- Modelle, die spezifische *produktive und rezeptive* Faktoren annehmen.

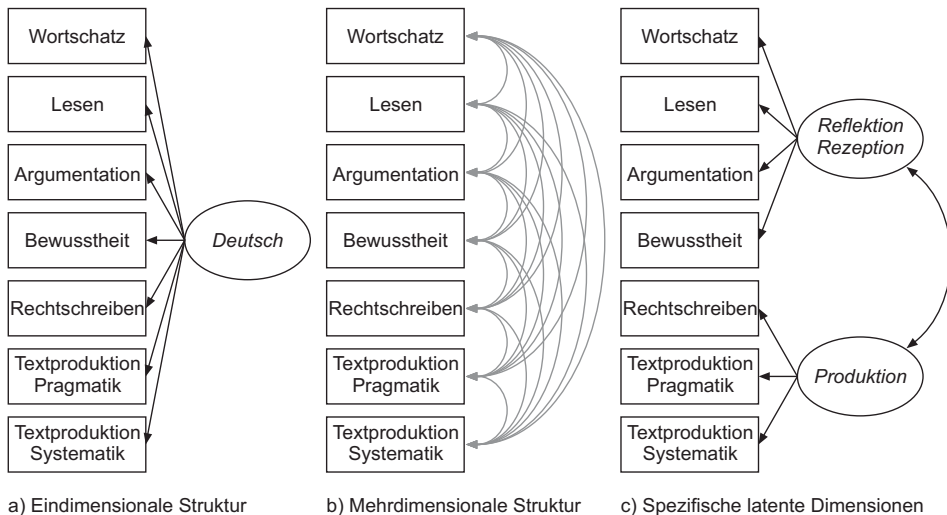


Abbildung 18.1. Unterschiedliche Modelle der Struktur von Sprachkompetenz am Beispiel Deutsch.

Die Analyse dieser Modelle von Sprachkompetenz erfolgt mittels Strukturmodellierungen, die auf Kovariationen der unterschiedlichen Teiltests miteinander basieren (Kunnan 1998; Muthén 2004). Diese Kovariationen können sich in ihrem Ausmaß zwischen der Individualebene und der Klassenebene unterscheiden, d.h. innerhalb einer Klasse können die Leistungen der Schüler homogener oder heterogener sein als die von Schülern in einer anderen Klasse (Van den Noortgate/Opendakker/Onghena 2005). Um dieses Phänomen auch statistisch zu berücksichtigen wird daher die Struktur von Sprachkompetenz mittels Mehrebenen-Faktorenanalysen auf zwei Ebenen modelliert: Auf Individualebene und auf Klassenebene (Hox 1993, 2002; vgl. auch Kapitel 3). Um zu beurteilen, welches theoretische Modell den Daten der

DESI-Studie am besten gerecht wird, werden die drei oben genannten Modelle (G-Faktor-Modell, mehrdimensionales Modell und hierarchisches Modell) jeweils gegenübergestellt und verglichen (Jöreskog 1993). Die Vergleichskriterien sind dabei die bei statistischen Modellvergleichen gängigen Modellgütekriterien (Tanaka 1993; Schermelleh-Engel/Moosbrugger 2002).

18.2 Allgemeines Modell der Deutschkompetenz

Die DESI-Testmodule erfassen mit unterschiedlichen Testformaten verschiedene Aspekte der Deutschkompetenz. Nicht nur dieses methodische Vorgehen, sondern auch bereits vorhandene empirische Befunde lassen erwarten, dass sich eine differenzierte Kompetenzstruktur abbilden lässt. In der Tat weist die Modellierung eines rein eindimensionalen Modells (G-Faktor), das auf Individual- und auf Klassenebene nur einen gemeinsamen latenten Faktor als Ursache der Leistungsvariationen in den Teilkompetenzen annimmt, lediglich einen befriedigenden Fit auf. Besonders auf Klassenebene weisen die Ergebnisse darauf hin, dass sich nicht alle Varianz der Testmodule durch eine gemeinsame Dimension erklären lässt. Eine Ursache hierfür könnte die testmethodische Umsetzung sein: Die Testmodule Rechtschreiben und Textproduktion stellen Testformate dar, in denen die (freie) Produktion geschriebener Sprache gefordert wird, und unterscheiden sich schon dadurch von den eher geschlossenen Formaten der vier anderen Testmodule. Der Vergleich der drei oben beschriebenen, theoretischen Modelle von Sprachkompetenz weist auf eine Struktur hin, die der Datenlage angemessener erscheint: Das Modell, das auf Individualebene eine mehrdimensionale Struktur sowie auf Klassenebene die hierarchischen Faktoren Sprachrezeption und Sprachproduktion annimmt, weist eine gute Passung auf (vgl. Abbildung 18.2).

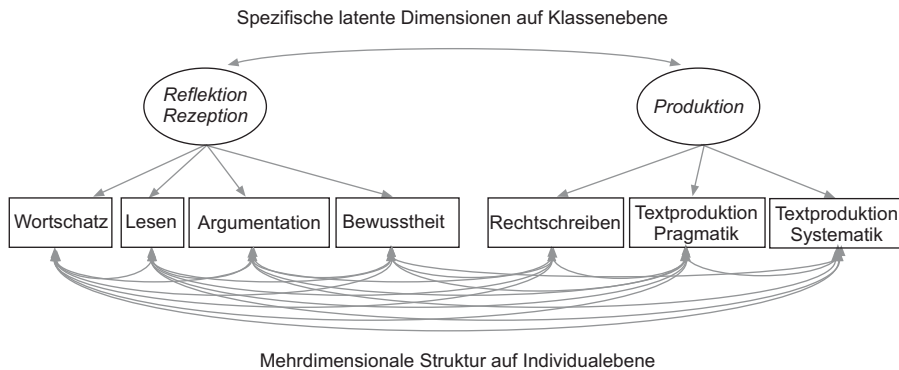


Abbildung 18.2: Mehrebenen-Strukturmodell der Deutschkompetenz in DESI. Modelpassung: $CFI=.99$, $RMSEA=.02$; $SRMR_{within} = .00$, $SRMR_{between} = .01$.

Die Kompetenzstruktur auf Individualebene stellt sich durchaus differenziert dar und verdeutlicht die Heterogenität der sprachlichen Fähigkeiten. Diese korrelieren zwar untereinander, jedoch nicht so hoch, dass von einer eindimensionalen Struktur ge-

sprochen werden kann. Die DESI-Testverfahren erfassen also durchaus differenzierbare Teildimensionen von Sprachkompetenz. Die Korrelationen auf Individualebene sind in Tabelle 18.2 dargestellt.

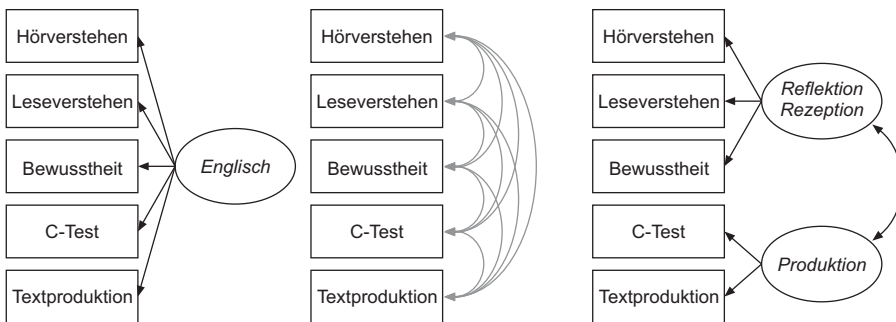
Tabelle 18.2: Korrelationen der Testmodule Deutsch auf Individualebene.

	Lesen	RS	KA	TP2	TP3	WS
Rechtschreiben (RS)	.15					
Argumentation (KA)	.28	.18				
Textproduktion Pragmatik (TP 2)	.25	.22	.29			
Textproduktion Systematik (TP 3)	.17	.36	.17	.37		
Wortschatz (WS)	.35	.18	.26	.20	.17	
Bewusstheit (BD)	.34	.29	.29	.31	.29	.36

Die Korrelationen der sprachlichen Teilkompetenzen auf Individualebene variieren deutlich: Zwischen Rechtschreiben und Lesen betragen sie lediglich $r = .15$, zwischen Wortschatz und Sprachbewusstheit hingegen $r = .36$, ebenso zwischen Rechtschreiben und Textproduktion Systematik. Die beiden letzteren Testmodule lassen sich auf Klassenebene auf die gemeinsame Dimension *Produktion* zurückführen. Auf Klassenebene kann darüber hinaus eine zweite differenzierbare latente Dimension, *Rezeption und Reflektion*, abgebildet werden. Diese Dimension bezieht die Indikatoren des Lesens, der Sprachbewusstheit, der Argumentation und des Wortschatzes ein. Die latenten Dimensionen auf Klassenebene korrelieren zu $r = .98$ miteinander. Trotz dieser sehr hohen Korrelation erreicht das Modell mit einer eindimensionalen Struktur auf Klassenebene eine schlechtere Anpassung an die Daten.

18.3 Allgemeines Modell der Englischkompetenz

Die Indikatoren für die Sprachkompetenz im Englischen beinhalten im Unterschied zu den Deutschtests ein Testmodul zum Hörverstehen sowie C-Tests als Indikatoren für die produktiven Fähigkeiten im Englischen. Zur Analyse der Struktur der Englischkompetenz wurden entsprechend den Überlegungen zur Deutschkompetenz drei Modelle miteinander verglichen (vgl. Abbildung 18.3).



a) Eindimensionale Struktur

b) Mehrdimensionale Struktur

c) Spezifische latente Dimensionen

Abbildung 18.3: Unterschiedliche Modelle der Struktur der Englischkompetenz.

Auch hier zeigt sich analog zur Kompetenz im Deutschen, dass nicht pauschal von einer Eindimensionalität der Fremdsprachenkompetenz ausgegangen werden kann. Im Vergleich der verschiedenen Modelle für die Englischkompetenz der DESI Schülerinnen und Schüler lässt sich das in **Abbildung 18.4** dargestellte Mehrebenenmodell abbilden: Auf Individualebene besteht eine mehrdimensionale Struktur mit korrelierenden, differenzierten Kompetenzen, auf Klassenebene hingegen können die latenten Dimensionen *Sprachreflektion/ Sprachrezeption* und *Produktion* gebildet werden (vgl. **Abbildung 18.4**).

Spezifische latente Dimensionen auf Klassenebene

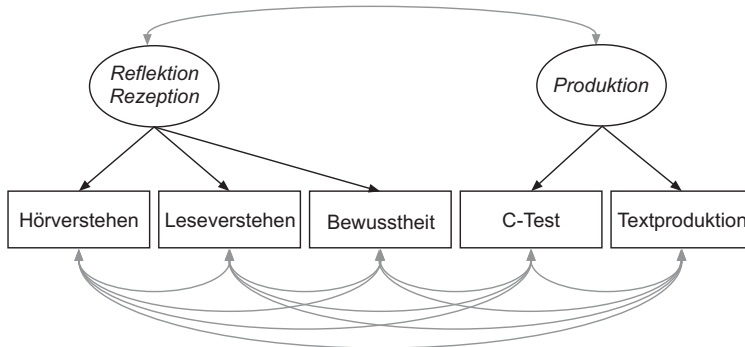


Abbildung 18.4: Mehrebenen-Strukturmodell der Englischkompetenz in DESI. Modellpassung: $\chi^2/df = 2.8$; $p = .02$; $CFI = .99$, $RMSEA = .01$; $SRMR_{within} = .00$, $SRMR_{between} = .00$.

Die Zusammenhänge zwischen den sprachlichen Teilbereichen im Englischen lassen sich parallel zur Struktur im Deutschen beschreiben: Auf Individualebene zeigt sich ein differenziertes, mehrdimensionales Bild, auf Klassenebene stellen sich die latenten Dimensionen der Sprachproduktion sowie der Reflektion-Rezeption dar, letztere korrelieren zu $r = .99$. Die Korrelationen der Teilmodule auf Individualebene sind in **Tabelle 18.3** dargestellt.

Tabelle 18.3: Zusammenhänge der Indikatoren der Englischkompetenz auf Individualebene

	Lesen	HV	CT	BE
Hörverstehen (HV)	.43			
C-Test	.49	.39		
Bewusstheit (BE)	.35	.30	.41	
Textproduktion (ST)	.42	.28	.52	.35

Die Korrelationsmuster auf Individualebene zeigen etwas höhere Zusammenhänge zwischen den Teilmodulen im Englischen, als dies bei den Deutschtests zu beobachten ist. Die höchste gemeinsame Variation findet sich für die Formate Textproduktion und C-Test. Hier werden ähnliche Kompetenzen, nämlich das (freie) Schreiben von Text unter Rückgriff auf Vokabular und Rechtschreiben, gefordert. Die insgesamt höheren Korrelationen könnten sich dadurch erklären lassen, dass der insti-

im Deutschen und im Englischen korrelieren zu $r = .97$, die reflektiv-rezeptiven Bereiche in beiden Sprachen zu $r = .94$. Auf Individualebene zeigen sich stärker differenzierte Strukturen, die Korrelationen sind in Tabelle 18.4 dargestellt.

Tabelle 18.4: Zusammenhänge der Indikatoren der Englisch- und Deutschkompetenz.

		Englisch Individualebene				
		Hören	Lesen	Bewusst-heit	C-Test	Textpro- duktion
Deutsch Individualebene	Wortschatz	.23	.29	.16	.30	.24
	Lesekompetenz	.24	.31	.18	.32	.27
	Argumentation	.15	.24	.13	.26	.28
	Bewusstheit	.23	.29	.23	.42	.36
	Rechtschreiben	.17	.26	.26	.43	.34
	Textproduktion Pragmatik	.13	.13	.13	.28	.37
	Textproduktion Systematik	.13	.21	.18	.33	.27

Besonders jene Testmodule im Deutschen und im Englischen, die auf ähnliche Bereiche rekurren, weisen auf Individualebene deutliche Zusammenhänge auf. Dazu zählen besonders die Lesekompetenz ($r_{DE} = .31$) und die Module zur Textproduktion ($r_{DE} = .37$) in beiden Sprachen. Hier wurden im Deutschen wie auch im Englischen vergleichbare Aufgaben- und Antwortformate gewählt, so dass Gemeinsamkeiten in den spezifischen Kompetenzen über die Sprachen hinweg auch aus testmethodischer Sicht angenommen werden können (vgl. Beck/Klieme 2007). Hohe Korrelationen mit den Deutschmodulen weist auch der englische C-Test auf. Dieses Verfahren, das als Indikator der allgemeinen Sprachkompetenz verwendet wird (vgl. Kapitel 14), scheint neben den spezifischen Iteminhalten in der Tat auch mit einer allgemeinen Sprachkompetenz in Zusammenhang zu stehen, die vor allem Sprachbewusstheit und Rechtschreiben, sowie Lesekompetenz und Wortschatz einbezieht.

18.5 Zur Struktur von Sprachkompetenz

Mit den Instrumenten der DESI Studie werden die sprachlichen Kompetenzen wesentlich differenzierter und facettenreicher erfasst, als dies in vergleichbaren Schulleistungsstudien bisher der Fall war. Dies ermöglicht die Analyse der Struktur von Sprachkompetenz, die über eine Modellierung einfacher eindimensionaler Modelle hinausgeht. Durch die Verwendung unterschiedlicher Testformate lassen sich sowohl produktive als auch rezeptive Dimensionen der sprachlichen Kompetenzen im Deutschen wie im Englisch abbilden. Es verwundert daher nicht, dass die vorliegenden Ergebnisse über jene theoretisch postulierten und in Studien der 1970er Jahre auch empirisch gefundenen, eindimensionalen Modelle von Sprachkompetenz hinausgehen. Dies zeigt sich auf der Ebene der individuellen Schülerinnen und Schüler

darin, dass die verschiedenen sprachlichen Kompetenzen zwar miteinander korrelieren, diese Zusammenhänge jedoch durchaus starke individuelle Variationen aufweisen, so dass mehrdimensionale Modelle angemessen abgebildet werden können.

Für das Beispiel der Sprachproduktion bedeutet dies: Schülerinnen und Schüler, deren Rechtschreibkompetenz hoch ausgeprägt ist, müssen nicht zwangsläufig auch gute Aufsätze schreiben. Auch ein großer Wortschatz begünstigt zwar das Leseverstehen, aber alle diese Kompetenzen sind zumindest auf Individualebene aber nicht auf einen gemeinsamen, übergeordneten Faktor zurückzuführen. Die Annahme, dass beispielsweise ein Leseverstehentest als ausschließlicher und ausreichender Indikator für die gesamtsprachliche Kompetenz dienen könnte, ist also nicht haltbar. Vielmehr sollten für Aussagen über sprachliche Leistungen von Schülerinnen und Schülern, die sich auf die individuellen Kompetenzen beziehen, immer facettenreiche Erhebungsinstrumente zum Einsatz kommen. Auch auf Klassenebene zeigt sich ein stärker strukturiertes Kompetenzmodell als in einigen bisherigen Studien.

In DESI lassen sich für das Deutsche und das Englische jeweils zwei latente Dimensionen unterscheiden: Sprachproduktion und Sprachreflektion-Rezeption. Diese Ergebnisse unterstützen die Validität der verwendeten Testinstrumente, die Variationen auch auf Klassenebene noch abbilden lassen. Besonders die Struktur von Sprachkompetenz auf Klassenebene bietet Ansatzpunkte für die Wirkungsforschung des Sprachunterrichts, da die Variationen durchaus auf Unterschiede in den Unterrichtskulturen zurückgeführt werden könnten (De Fraine u.a. 2003). Erst die in DESI realisierte differenzierte Erfassung von Sprachkompetenz als mehrdimensionales Phänomen ermöglicht Rückschlüsse über spezifische Kompetenzen, die auch im Rahmen eines *large scale assessments* durch den Einsatz curriculumsspezifischer, differenzierter Testinstrumente durchaus möglich sind. Zukünftige Studien sollten daher neben der Lesekompetenz, die allgemein als wesentliche Komponente von Sprachkompetenz angesehen wird, auch die Fähigkeiten in weiteren sprachlichen Bereichen im Blick behalten.

Auf Basis dieser differenzierten Testdaten lassen sich in weiteren Analysen zusätzliche Modelle abbilden, die spezifischen theoretischen Ansätzen zur Struktur von Sprachkompetenz nachgehen. Hierzu gehören Zusammenhangsanalysen in Mehrebenenregressionsmodellen und Mehrebenenstrukturmodellen, welche auch die Analyse differenzieller Unterschiede zwischen Personengruppen mit einbeziehen (vgl. Jude im Druck).

Literatur

- Beck, B./Klieme, K. (Hrsg.) (2007): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Bos, W./Lankes, E.-M./Prenzel, M./Schwippert, K./Walther, G./Valtin, R. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.

- De Fraine, B./Damme, J.v./Landeghem, G.v./Opdenakker, M.-C. (2003): The effect of schools and classes on language achievement. In: *British Educational Research Journal* 29, H. 6, S. 841-859.
- Detering, K (Hrsg.) (2000): *Wortschatz und Wortschatzvermittlung. Linguistische und didaktische Aspekte*. Frankfurt am Main: Lang.
- Deutsches PISA Konsortium (Hrsg.) (2001): *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Eckerth, J. (2000): Zielsprachliche Kommunikation über Grammatik im Fremdsprachenunterricht. In: *Zeitschrift für Fremdsprachenforschung* 11, H. 1, S. 9-30.
- Gräfe-Bentzien, S. (2000): Evaluierung bilingualer Kompetenz. Eine Pilotstudie zur Entwicklung der deutschen und italienischen Sprachfähigkeiten in der Primarstufe beim Schulversuch der Staatlichen Europa-Schule Berlin (SESB). Berlin: Freie Universität.
- Haberzettl, S./Wegener, H. (Hrsg.) (2003): *Spracherwerb und Konzeptualisierung*. Frankfurt am Main: Lang.
- Hox, J.J. (1993): Factor analysis of multilevel data. In: Oud, J.H.L./Blokland-Vogeleesang, R.A.W. (Hrsg.): *Gauging the Muthén Model. Advances in longitudinal and multivariate analysis in the behavioral sciences*. Nijmegen, NL: ITS, S.141-156.
- Hox, J.J. (2002): *Multilevel Analysis. Techniques and Applications*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Jöreskog, K.G. (1993): Testing structural equation models. In: Bollen, K.A./Long, J.S. (Hrsg): *Testing structural equation models*. Newbury Park, California: Sage Publications, Inc., S. 294-316.
- Jude, N. (im Druck): *Strukturen sprachlicher Kompetenzen im Deutschen und Englischen*.
- Köller, O./Trautwein, U. (2004): Englischleistungen von Schülerinnen und Schülern an allgemein bildenden und beruflichen Gymnasien. In: Köller, O./Watermann, R./Trautwein, U./Lüdtke, O. (Hrsg): *Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien*, S. 285 – 326.
- Kuhlemeier, H./ van den Bergh, H./Rijlaarsdam, G. (2002): The dimensionality of speaking and writing: A multilevel factor analysis of situational, task and school effects. In: *British Journal of Educational Psychology* 72, S. 467-482.
- Kunnan, A. J. (1992): An investigation of a criterion-referenced test using G-theory, and factor and cluster analyses. In: *Language Testing* 9, H. 1, S. 30-44.
- Kunnan, A. J. (1998): An introduction to structural equation modelling for language assessment research. In: *Language Testing* 15, H. 3, S. 295-332.
- McNamara, T. F. (1995): Modelling performance: opening Pandora's box. In: *Applied linguistics* 16, H. 2, S. 159-175.
- Mehta, P.D./Foorman, B.R./Branum-Martin, L./Taxlor, W.P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. In: *Scientific Studies of Reading* 9, H. 2, S. 85-116.
- Muthén, B. (2004): Latent variable analysis. In: Kaplan, D. (Hrsg): *Handbook of quantitative methodology for the social science*. Newbury Park, CA, Sage Publications.
- North, B. (2000): *The development of a common framework scale of language proficiency*. Frankfurt am Main: Lang.
- Oerter, R./Montada, L. (2002): *Entwicklungspsychologie*. Weinheim: Psychologie Verlags Union.
- Oller, J. W. (1976): Evidence for a general language proficiency factor: An expectancy grammar. In: *Die Neueren Sprachen* 76, S. 165-174.
- Rost, D. H./ Hartmann, A. (1992): Lesen, Hören, Verstehen. In: *Zeitschrift für Psychologie* 200, S. 345-361.
- Schermelleh-Engel, K./ Moosbrugger, H. (2002). Beurteilung der Modellgüte von Strukturgleichungsmodellen. Frankfurt: Arbeiten aus dem Institut für Psychologie der J. W. Goethe-Universität, Heft 4/2002.

Szagun, G. (2000): Sprachentwicklung beim Kind. Weinheim: Beltz

Tanaka, J. S. (1993): Multifaceted conceptions of fit in structural equation models. In: Bollen, K. A./Long, J. S. (Hrsg): Testing structural equation models. Newbury Park, California: Sage Publications, S. 10-39.

Van den Noortgate, W. M./Opendakker, C./Onghena, O. (2005): The effects of ignoring a level in multilevel analysis. In: School effectiveness and school improvement 16, H. 3, S. 281-303.