

---

Liebe Leserinnen und Leser,

das aktuelle ELiSe-Heft beginnt auf der folgenden Seite mit dem Titelblatt. Diese Leerseite, die dem Heft von ELiSe vorgeschaltet ist, soll Ihnen ermöglichen, auf einer Druckseite im DIN-A4-Format zwei ELiSe-Seiten mit der korrekten Paginierung auszu-  
drucken. Gesetzt ist die Zeitschrift in A5.

Bei weiteren Fragen wenden Sie sich bitte an die Redaktion.

[elise@uni-essen.de](mailto:elise@uni-essen.de)

# ELiS\_e

[e'li:zə]

<Essener Linguistische Skripte – elektronisch>

## E-Papiere zu Sprachwissenschaft und Sprachdidaktik

Jahrgang 2, Heft 2 (Dezember 2002)

ISSN 1617-5425

[elise@uni-essen.de](mailto:elise@uni-essen.de)

<http://www.elise.uni-essen.de>

Liebe Leserinnen und Leser,

wie im Editorial angekündigt, möchten wir mit ELiSe besonders die wissenschaftliche Diskussion anregen und einen schnellen Gedankenaustausch ermöglichen. Deshalb eröffnen wir zu jedem Heft ein Diskussionsforum, in dem Anregungen, Kritik, Hinweise und Stellungnahmen von Lesern zu den einzelnen Beiträgen veröffentlicht werden. Die Autorinnen und Autoren sind von uns gebeten worden, an diesen Diskussionen teilzunehmen und haben das ausdrückliche Recht zur Kommentierung.

Bitte senden Sie Ihre Anregungen an die Redaktion, damit wir sie in das Diskussionsforum stellen können und die Autorinnen und Autoren die Möglichkeit haben, Stellung zu beziehen.

[elise@uni-essen.de](mailto:elise@uni-essen.de)

Bitte gestalten Sie die Betreff-Zeile nach dem Schema: **Beitrag: Autor, Heft**

---

# **ELiS\_e**

**<Essener Linguistische Skripte – elektronisch>**

## **E-Papiere zu Sprachwissenschaft und Sprachdidaktik**

**Jahrgang 2**

**Heft 2**

**Dezember 2002**

---

## Impressum

ELiSe wird herausgegeben von:

Christoph Chlosta • Hermann Cölfen • Joachim Raith  
Werner Schöneck • Christoph Schroeder • Wilfried Timmler

Kontakt: Dr. Christoph Schroeder, Universität Essen, Fachbereich 3, 45117 Essen

E-Mail: [elise@uni-essen.de](mailto:elise@uni-essen.de)

© ELiSe-Redaktion. Die Zeitschrift und alle in ihr enthaltenen Beiträge sind urheberrechtlich geschützt. Die Nutzung der Zeitschrift und der in ihr enthaltenen Beiträge ist insoweit frei, als nichtkommerziell handelnden Personen, Firmen, Einrichtungen etc. ein begrenztes Recht auf nichtkommerzielle Nutzung und Vervielfältigung in analoger und digitaler Form eingeräumt wird. Das betrifft das Laden und Speichern auf binäre Datenträger sowie das Ausdrucken und Kopieren auf Papier. Dabei obliegt dem Nutzer stets die vollständige Angabe der Herkunft, bei elektronischer Nutzung auch die Sicherung dieser Bestimmungen.

Es besteht – außer im Rahmen wissenschaftlicher und schulischer Veranstaltungen öffentlicher Träger – kein Recht auf Verbreitung. Es besteht kein Recht zur öffentlichen Wiedergabe. Das Verbot schließt das Bereithalten zum Abruf im Internet, die Verbreitung über Newsgroups und per Mailinglisten ein, soweit dies durch die ELiSe-Redaktion – oder durch den/die Urheber des betreffenden Beitrags – nicht ausdrücklich genehmigt wurde. Darüber hinausgehende Nutzungen und Verwertungen sind ohne Zustimmung des Urhebers unzulässig und strafbar.

ISSN: 1617-5425

---

## **Inhalt**

*Michael Fritsche*

Kollektivsymbolik in den Nationalbewegungen  
des 19. und 20. Jahrhunderts.

Ein Vergleich zwischen Deutschland und Südosteuropa ..... 7

*Christian Krekeler*

TestDaF und DSH – ungleiche Sprachtests im Vergleich ..... 19

*Thomas Niehr*

International vergleichende Diskurs- und Argumentations-  
analyse. Vorstellung eines Forschungsprogramms ..... 51



**Michael Fritsche**  
**Kollektivsymbolik in den Nationalbewegungen**  
**des 19. und 20. Jahrhunderts.**  
**Ein Vergleich zwischen Deutschland und Südosteuropa**

**1. Nation und Nationalcharakter im 19. Jahrhundert**

**1.1. Nation als Produkt der Nationalideologie**

Fragt man sich nach dem Ausgangspunkt von Nation und Nationalbewegungen, so ist er nach einem materialistischen Geschichtsverständnis darin zu sehen, daß – etwas schematisch ausgedrückt – kapitalistische Produktionsformen, bei den sie vorantreibenden Menschen zu einem Klassenbewußtsein führten, das nach Selbstbestimmung und Teilhabe an der Herrschaft strebte. Die erwünschte Herrschaftsform war der bürgerliche Nationalstaat. Die von bürgerlichen Eliten ausgelösten Prozesse, die diesen erreichen wollten, werden als Nationalbewegungen oder als nationale Emanzipationsbewegungen bezeichnet. Ihre ideologische Basis hatten diese Bewegungen zunächst in der Aufklärung mit ihrem Anspruch der Befreiung aus politischer und geistiger Bevormundung.<sup>1</sup>

Wenn jedoch die ökonomischen und politischen Machtverhältnisse diesem Anspruch widersprachen, d. h. wenn die bürgerlichen Eliten zu schwach oder zu wenig ausgeprägt und die reale Basis für bürgerliche Nationalstaaten fehlte, wenn die durch den beginnenden Kapitalismus am Ende des 18. Jahrhunderts ausgelöste Massenmobilität zur allgemeinen Verunsicherung, Leere und Entwurzelung führten<sup>2</sup>, entwickelte sich im nationalen Diskurs mit seiner ausgeprägten Kollektivsymbolik ein System von Metaphern, die – bestimmt durch binäre Oppositionen – nur noch die eigene Nation als lebenswert fokus-

---

<sup>1</sup> De Vincenz, Andrzej: Völker, Nationen und Nationalsprachen, in: Hentschel, Gerd (Hg.): Über Muttersprache und Vaterländer. Frankfurt a. M. u. a. 1997, S. 8.

<sup>2</sup> Vgl. Baba, Homi: DissemiNation: Zeit, Narrative und die Ränder der modernen Nation, in: Bronfen, Elisabeth/Marius, Benjamin (Hgg.): Hybride Kulturen: Beiträge zur anglo-amerikanischen Multikulturalismusdebatte. Tübingen 1997, S. 150.



sierten und die anderen Nationen bis hin zur Vernichtung abwerteten. Hier wurde also das ursprünglich aufklärerische Ziel in sein Gegenteil verkehrt – ein Prozeß, den schon Franz Grillparzer prägnant charakterisierte: „Von Humanität durch Nationalität zur Bestialität.“<sup>3</sup>

## 1.2. Kollektivsymbolik als Konstituente der Nationalideologie

Die Kollektivsymbolik, insbesondere die Metaphorik, die sich mit z. T. verheerenden Folgen verselbständigte, die für den nationalen Diskurs besonders im zentralen und süd-östlichen Europa des 19. und 20. Jahrhundert typisch ist, möchte ich im folgenden zumindest in Grundzügen versuchen zu rekonstruieren. Eine detaillierte Rekonstruktion erfordert eine detaillierte Bearbeitung der politischen Literatur (insbesondere Reden) und der politisch motivierten Belletristik (besonders Lyrik) sämtlicher Balkanstaaten einschließlich der Türkei, die ich hier nicht bewältigen kann. Bei der skizzenhaften Rekonstruktion des nationalen Diskurses bediene ich mich weitgehend der Terminologie von Gerhard und Link.<sup>4</sup>

## 1.3. Besonderheit des deutschen Nationalismus

In Westeuropa gelang der Prozeß der Nationsbildung insbesondere der Bildung von Nationalstaaten relativ gradlinig, dadurch daß vorhandene politische Strukturen durch die Beteiligung der Bourgeoisie an der politischen Macht gewissermaßen uminterpretiert werden konnten, die Machtgebilde jedoch erhalten blieben. So entwickelten sich z. B. England zur konstitutionellen Monarchie und Frankreich von einer zentralistischen absoluten Monarchie durch die Revolution zu einem zentralistischen bürgerlichen Nationalstaat, der sich ideologisch auf den *contrat social* gründete.

In Deutschland oder sagen wir besser in Mitteleuropa konnte sich aufgrund der 700 Jahre andauernden Feudalkrise ein Machtzentrum wie in Frankreich nicht entwickeln. Städtisches Bürgertum und Bauern lebten nach den katastrophalen Niederlagen in den auf die Reformation folgenden Kriegen in den zahlreichen Feudalstaaten im Bewußtsein absoluter Ohnmacht und der Vergeblichkeit jeglichen Widerstandes. Auch die bürgerlich revolutionären Bewegungen im Gefolge der Französischen Revolution führten nicht zum Erfolg, was eine massenhafte Resignation zur Folge hatte, die auch die spezifische Variante des deutschen Nationalismus prägte.

Der deutsche Nationalstaat wie er dann 1871 insbesondere auf das Betreiben Bismarcks hin gegründet wurde, beruhte auf einem Klassenbündnis zwischen Feudaladel und Großbourgeoisie und entsprach nicht den Zielsetzungen der bürgerlich-revolutionä-

---

<sup>3</sup> Grillparzer, Franz: Sprachenkampf, in: Sämtliche Werke. Bd 1. München 1960, S.500.

<sup>4</sup> Gerhard, Ute / Link, Jürgen: Zum Anteil der Kollektivsymbolik an den Nationalstereotypen, in: Link, Jürgen/Wülfing, Wulf (Hgg.): Nationale Mythen und Symbole in der zweiten Hälfte des 19. Jahrhunderts. Stuttgart 1991, S. 16-52.

ren Elite von 1848. Große Teile dieser Elite jedoch identifizierten sich mit dieser Gründung von oben<sup>5</sup>

#### **1.4. Eine Metapher als Ausgangspunkt für nationale Kollektivsymbolik**

Diese Identifikation war möglich auf der Basis der Nation. Dieser Begriff (natio) bedeutet etymologisch das Geborenwerden oder die durch Geburt bedingte Beschaffenheit, also Art oder Gattung. Dieser Ausdruck wurde schon in römischer Zeit auf menschliche Gesellschaften übertragen, und zwar auf Gruppen, deren Existenz man eine gemeinsame Abstammung unterstellt. „Nation“ wie sie heute verstanden wird beruht also auf einer Metaphorisierung. Dabei werden genetische Eigenschaften auf politisch motivierte Gruppen übertragen. Diese zugeschriebenen Eigenschaften erscheinen dann den Mitgliedern aufgrund der nicht durchschauten Metaphorisierung tatsächlich als genetisch und damit natürlich.<sup>6</sup>

Dies führte dann im 19. Jahrhundert zu der massenhaften Solidarität der Bürger mit ihrem Nationalstaat. Da diese nationale Solidarität, dieses nationale Zugehörigkeitsgefühl, die nationale Identität höchst unsicher und unbestimmt ist – alles andere als die Geborgenheit einer Großfamilie vermittelt –, mußte der Nationalstaat für eine Kontinuität und Verfestigung dieser Solidarität sorgen. Dies geschah durch ein System von Metaphern und Mythen, die in nationalen Kollektivsymbolen (diskursiver Art, z. B. Nationalsprache, Sprachgeschichte, oder materieller Art, z. B. Denkmäler, Feiern) vergegenständlicht wurden. Diese Kollektivsymbole wurden in Institutionen durch z. T. neu geschaffene akademische Fächer (insbesondere die nationalen Philologien und Geschichtswissenschaften) entwickelt und in einem möglichst flächendeckenden Schulsystem oder in den nationalen Armeen und Kirchen verbreitet und zum Massenbewußtsein oder zum nationalen Habitus.<sup>7</sup>

#### **1.5. Der Nationalcharakter – Basis für die Biologisierung gesellschaftspolitischer Prozesse**

Basis für diesen nationalen Habitus war der von Herder erfundene Nationalcharakter. Entsprechend der etymologischen Bedeutung von Nation liegt diese im Ursprung jeglichen Menschseins und bringt die Menschen mit ihren besonderen Nationalcharakter erst hervor. Der Nationalcharakter bestimmt dann die Kultur, deren wichtigster Ausdruck die Sprache ist. Waren für Herder – ganz im Sinne der Aufklärung – die verschiedenen Nationalcharaktere noch gleichwertig – die Welt verglich er mit einem Garten, in dem die „Nationalpflanzen“ blühen sollten<sup>8</sup> –, so deuteten seine Nachfolger und Vulgarisatoren diese organistische Metapher aus dem Bildbereich der Biologie um, indem sie die

---

<sup>5</sup> Mosse, George L.: Die Nationalisierung des Massen. Frankfurt a. M. u. a. 1993, S. 28 ff.

<sup>6</sup> Vgl. Klotz, Reinhold: Handwörterbuch der lateinischen Sprache. Bd. 2. Braunschweig 1879, S. 478. De Vincenz a.a.O., S. 7.

<sup>7</sup> Finkelkraut, Alain: Die Niederlage des Denkens. Reinbeck 1989, S. 13, 33.

<sup>8</sup> Reiter, Norbert: Gruppe, Sprache, Nation. Wiesbaden 1984, S. 398.

Überlegenheit einzelner Nationen aufgrund ihrer Sprachen betonten z. B. Grimm.<sup>9</sup> Damit wurde – von Herder wahrscheinlich nicht beabsichtigt – eine Metapher zur Basis für die spätere Biologisierung der genealogischen Sprachwissenschaft, indem nämlich Schleicher den Darwinismus zu deren Modell erklärte.<sup>10</sup>

Dieser Nationalcharakter, der die Basis für das Konzept der Volksnation bildete, ist zu sehen als das Ergebnis einer beschädigten Entwicklung des mitteleuropäischen Bürgertums, das sich in Westeuropa – allerdings gestützt auf koloniale Eroberung und Ausbeutung – gradliniger entwickelte und ein stärkeres Selbstbewußtsein behaupten konnte. Für dieses Bürgertum trifft das Konzept der Staatsnation zu.<sup>11</sup>

Zur Stabilisierung des labilen Volkscharakters wurde in den mittel- und südosteuropäischen Nationalbewegungen, die dem Konzept der Volksnation folgten, wie oben erwähnt, ein System nationaler Kollektivsymbole diskursiv entwickelt. Diese beziehen sich auf den Bereich der Sprache, der (weitgehend mythologisierten) Geschichte und der daraus konstruierten kollektiven Mentalitäten.

## **2. Nationale Kollektivsymbolik der Sprache in Deutschland. Das Modell der Sprachfamilie**

### **2.1. Die mörderische Macht einer Metapher**

In Deutschland konnte die Nationalbewegung immerhin auf eine bereits polyfunktionale Standardsprache zurückgreifen, die sich in einem ungefähr vierhundert Jahre andauernden komplizierten Prozeß durch schriftsprachliche Tätigkeit und die genialen Entwürfe einzelner Literaten bis zum Ende des 18. Jahrhunderts entwickelt hatte. Es war also damit (im Gegensatz zu Südosteuropa) eine kollektivsymbolische Basis für die nationale Identität bereits vorhanden.<sup>12</sup>

Die Entdeckung der strukturellen und lexikalischen Ähnlichkeiten von ungefähr 55 lebenden und toten Sprachen seit dem späten 18. Jahrhundert, die sich auf einem Gebiet zwischen Westeuropa und Indien verteilten, führte zu einer Klassifizierung dieser Sprachen, die in Deutschland aufgrund der sie im Osten und Westen begrenzenden Unterklassen – nämlich die indischen und germanischen Sprachen – als indogermanisch (außerhalb des deutschen Sprachgebiets als indoeuropäisch) bezeichnet wurden. Diese indogermanischen Sprachen wurden mit der Metapher „Sprachfamilie“ gekennzeichnet.<sup>13</sup> Diese Metapher produzierte nun die Vorstellung, daß diese Sprachen miteinander verwandt seien und es ein Urvolk (die Indogermanen) gegeben haben müsse. Diese Bezeichnung wurde umgedeutet, indem die Germanen als ältester integraler Bestandteil dieses angeblichen Urvolks gesehen wurden. Da sich die deutschen intellektuellen Eliten

---

<sup>9</sup> Arens, Hans: Sprachwissenschaft. München 1969, S. 196.

<sup>10</sup> Schleicher, August: Die Darwinsche Theorie und die Sprachwissenschaft. Weimar 1873, S. 67f.

<sup>11</sup> De Vincenz a.a.O., S. 9

<sup>12</sup> Polenz, Peter von: Geschichte der deutschen Sprache. 9. überarb. Aufl. Berlin u. a. 1978, S. 108ff.

<sup>13</sup> Crystal, David: Die Cambridge Enzyklopädie der Sprache. Frankfurt a. M. u. a. 1995, S. 292ff.

im 19. und 20. Jahrhundert als die Nachkommen der Germanen sahen, war damit die Kontinuität der deutschen Nation von den frühesten Anfängen der Menschheit an scheinbar wissenschaftlich erwiesen. Diese Verselbständigung der Metapher „Indogermanische Sprachfamilie“ führte dazu, daß „indogermanisch“ sich zum Inbegriff der nazistischen Rassentheorie entwickelte: Durch die Verknüpfung von sprachlichen Ähnlichkeiten mit Spekulationen über archäologische Funde und angebliche psychische Eigenschaften, die man den Nachfahren der Indogermanen zuordnete, kam man zur indogermanischen Rasse, einem Synonym für die arische Rasse, die den Nazis als die allein lebenswerte galt.<sup>14</sup> Hier wird die mögliche mörderische Macht der Metaphern deutlich, und zu spät (1939) erschien der warnende Artikel eines der Begründer der modernen Linguistik, Nikolaj S. Trubetzkoy, der die indogermanische Sprachfamilie als rein linguistische Kategorie bezeichnete.<sup>15</sup> Das Modell der Sprachfamilie hatte sich als unbestrittenes System nationaler Kollektivsymbolik insbesondere in Südosteuropa unter Bezeichnungen wie Panslavismus, Jugoslawismus und Panturkismus durchgesetzt.

## **2.2. Kontinuität durch metaphorisierte Sprachgeschichte**

Doch zunächst galt es im 19. Jahrhundert, die würdige Position und die Stärke der Nationalsprache innerhalb der indogermanischen Sprachfamilie zu zeigen. Dies tat die deutsche Nationalphilologie, die Germanistik, indem sie mit den Methoden der genealogischen Sprachwissenschaft Alter und Kontinuität der deutschen Sprache bewies. Allerdings unterstellt die Einteilung in Althochdeutsch, Mittelhochdeutsch und Neuhochdeutsch eine Kontinuität, die es in Wirklichkeit so nicht gegeben hat<sup>16</sup>, denn die beiden erstgenannten Sprachperioden sind keine Standardsprachen und damit mit dem Neuhochdeutschen nicht ohne weiteres vergleichbar. Jedenfalls wurde es der deutschen Elite seit dem 19. Jahrhundert damit selbstverständlich, zu einer alten und damit würdigen und wertvollen Kulturnation zu gehören – mit Anspruch auf einen gleichberechtigten oder sogar höherrangigen Platz neben den anderen europäischen Nationen, die in der Entwicklung des Imperialismus schon weiter fortgeschritten waren. Auch ließ die Mitgliedschaft in einer das Territorium überschreitenden Sprachfamilie Gebietsansprüche zu, die die Kommunikationsbasis erweitert hätten.

## **3. Sprache als Kollektivsymbol in Südosteuropa**

### **3.1. Diglossie und unzureichende Standardisierung im 19. Jahrhundert**

In Südosteuropa waren im 19. Jahrhundert Standardsprachen im Sinne national verwendbarer kollektivsymbolischer Systeme noch nicht vorhanden. Es herrschte dagegen Diglossie oder Polyglossie. D. h. verschiedene Sprachen waren auf Verwendungsberei-

---

<sup>14</sup> Vgl. Kattmann, Ulrich: Rasse und Volk als Inhalt kollektiver Identität. Kritik des biologischen Volksbegriffs. Oldenburger Vordrucke. Heft 62/1989, S.2.

<sup>15</sup> Trubetzkoy, Nikolaj S.: Gedanken über das Indogermanenproblem, in: Acta Linguistica 1. 1939, S. 81-89.

<sup>16</sup> Vgl. Polenz, Peter von: a.a.O., S. 40, 53 ff., 75ff.

che mehr oder weniger festgelegt und die von den nach Emanzipation strebenden nationalen Eliten bevorzugten Sprachen waren aufgrund ihrer bisherigen funktionalen Einschränkungen nicht verwendbar. Vertreter der südosteuropäischen Nationalbewegungen, die sich aus neu entstandenen gesellschaftlichen Zwischenschichten rekrutierten, verfolgten deshalb das Ziel, die Diglossie zu beseitigen und durch Standardisierung der Sprache, die sie als Symbol für ihre nationale Identität ansahen, eine polyfunktionale Sprache zu schaffen, die die nationale Kommunikationsbasis ausmachen sollte. Ich möchte nun diesen Prozeß bei den einzelnen Nationen Südosteuropas – notwendigerweise vereinfachend – in m. E. wesentlichen Grundzügen darstellen.

### 3.2. Griechen

Griechen hatten im Osmanischen Reich eine Position, die manche Südosteuropa-Historiker als osmanisch-griechisches Condominium bezeichnen. Sie hatten zumindest bis zum Beginn des 19. Jahrhunderts einen hohen Anteil an der finanziellen Oberschicht, und die griechische orthodoxe Geistlichkeit war gewissermaßen Herr über alle orthodoxen Christen in den europäischen Teilen des Osmanischen Reiches. Und so richtete sich die nationale Emanzipation der nichtgriechischen orthodoxen Christen (Rumänen, Bulgaren, Serben und z. T. Albaner) insbesondere, was die Standardisierung und Polyfunktionalisierung der Sprachen angeht, v. a. gegen den Druck griechischsprachiger Institutionen. Im Griechischen selber herrschte eine Diglossie, die bis in die 70er Jahre des 20. Jahrhunderts andauerte. Konservative Kräfte bestanden auf einer Varietät, die am klassischen, attischen Griechisch orientiert war und die Identität der zeitgenössischen mit den antiken Griechen symbolisieren sollte. Diese Varietät (Katharevousa) dominierte das Bildungswesen und große Teile des öffentlichen Lebens. Daneben hatte sich längst eine schriftsprachliche Varietät der lebenden Volkssprache entwickelt (Dhimotiki), die schon seit Jahrhunderten die Belletristik beherrschte. Die Kontinuität zum alten Griechenland wurde auch in der schönen Literatur betont. Möglicherweise war die Durchsetzung eines modernen Standards besonders schwierig aufgrund der nachgewiesenen schriftsprachlichen Tradition und der negativen Sichtweise der Europäer auf das sogenannte Neugriechische, was ja schon eine diskriminierende Bezeichnung darstellt.<sup>17</sup>

### 3.3. Rumänen

Das Rumänische im 18. und z. T. auch 19. Jahrhundert war als Schriftsprache nur sehr eingeschränkt gebräuchlich. Gebildete Rumänen mußten Griechisch, Lateinisch, Deutsch oder Ungarisch verwenden. Anhänger der rumänischen Nationalbewegung versuchten nun, das Rumänische zu einer einheitlichen Schriftsprache zu standardisieren. Sie beriefen sich dabei auch auf Alter und Würde des Rumänischen, indem sie dessen Romanität betonten. Für den tatsächlichen Sprachgebrauch wurde die Sprache der rumänischen Volksdichtung bevorzugt. Andererseits war diese Volkssprache für den mo-

---

<sup>17</sup> Apostolidis-Kusserow, Karin: Die griechische Nationalbewegung, in: Reiter, Norbert (Hg.): Nationalbewegungen auf dem Balkan. Berlin 1983, S. 85 ff.

dernen schriftsprachlichen Gebrauch nicht ohne weiteres verwendbar. Es gab deshalb vor und nach der Gründung eines unabhängigen rumänischen Staates einen jahrzehntelang andauernden Streit über die Beschaffenheit des Rumänischen, der sich mit puristischen und eher pragmatischen Positionen beschreiben läßt. Mit der raschen Entwicklung des Bildungswesens und der sich daraus ergebenden schriftsprachlichen und übersetzerischen Tätigkeit rumänischer Intellektueller ergab sich eine Standardsprache, die sich gegen die Sprachpfleger durchsetzte und deren theoretische Erwägungen, die z. T. in einem radikalen Purismus gipfelten, überholte.<sup>18</sup>

### **3.4. Serben**

Bezogen sich die Rumänen bei der Begründung ihrer Nationalsprache auf die Romantät, so betonten Serben und Kroaten im 19. Jahrhundert im Sinne des Panslavismus den südslavischen Charakter des Serbischen und Kroatischen und arbeiteten gemeinsam an der Standardisierung einer sowohl bei Serben und Kroaten verbreiteten Varietät. Dabei waren die Kroaten, die ihre sprachlichen und politischen Emanzipationsbestrebungen Illyrizam (Illyrismus) nannten, eher um historische Argumente bemüht, um Alter und Würde der Sprache zu betonen, während Vuk Karadžić, der geniale serbische Sprachreformer, sich auf die lebende Volkssprache bezog, die er bewußt als Rinder- und Schweinehirtensprache bezeichnete. Er setzte damit einen wichtigen Akzent gegen das zu seiner Zeit sehr mächtige Kirchenslavisch. Auch bei Serben und Kroaten spielte die Übersetzerstätigkeit eine wichtige Rolle bei der Herausbildung der modernen Standardsprache. Die Ideologie des Jugoslavismus, die zunächst von den Südslaven in der Donaumonarchie – besonders von Kroaten – instrumentiert wurde, zeigte allerdings nach der politischen Neuaufteilung nach dem Ersten Weltkrieg verheerende Folgen.<sup>19</sup>

### **3.5. Albaner**

Die Entwicklung des Albanischen zur nationalen Standardsprache wurde auch mit Alter und Würde begründet – schließlich wird Albanisch als einzige Fortsetzung einer vorromanischen Balkansprache (des Illyrischen) betrachtet. Daß sich eine albanische Standardsprache erst im 20. Jahrhundert etablieren konnte, liegt wohl besonders an der religiösen Diversität der Albaner (sie gehören mehrheitlich dem Islam aber zu kleinen Teilen auch der römisch-katholischen und der orthodoxen Kirche an), wobei die Muslime als integraler und in mehrfacher Hinsicht bedeutender Bestandteil der Osmanen angesehen wurden und es wohl auch waren. Bezeichnend dafür sind die Brüder Abdyl, Naim und Sami Frashëri, die als die Väter der albanischen Nation gelten. Naim und Sami waren gleichzeitig bedeutende osmanische Schriftsteller und Gelehrte. Insbesondere der Letztgenannte, der den ersten Roman in türkischer Sprache verfaßte, dürfte unter dem

---

<sup>18</sup> Fritsche, Michael: Sprachkonzeption und Sprachpolitik am Beispiel des rumänischen Sprachnationalismus, In: Osnabrücker Beiträge zur Sprachtheorie, 12, 1979, S. 90-109.

<sup>19</sup> Friesel-Kopecki, Dubravka: Die serbische Nationalbewegung, in: Reiter, Norbert (Hg.): Nationalbewegungen auf dem Balkan. Berlin 1983, S. 245-252.

Namen Şemsettin Sami Bey jedem gebildeten Türken von heute ein Begriff sein. Die griechische Orthodoxie (vgl. Abschnitt 3.2.) leistete gegenüber der Standardisierung des Albanischen beträchtlichen Widerstand.<sup>20</sup>

### 3.6. Bulgaren

Ähnlich verhielt es sich mit der bulgarischen Nationalbewegung, die – ähnlich der Konzeption des Vuk Karadžić – hinsichtlich der anzustrebenden Standardisierung des Bulgarischen, gegen griechischen Druck eine populistische Richtung, also Standardisierung der Volkssprache anstrebte.<sup>21</sup>

Es ist auffällig, wie südosteuropäische Nationalapologeten konkurrierenden Nationalbewegungen Würde und Alter der Sprache streitig machten und sich gegenseitig sprachliche Korruptiertheit vorwarfen. Diese Korruptiertheit wurde dann auch auf andere angeblich kollektive Mentalitäten übertragen. Damit wurde die Überlegenheit der eigenen Nation symbolisiert.

### 3.7. Türken

Die türkische Nationalbewegung zeigt sich als besonders kompliziert, weil ihre Promotoren sich notwendigerweise aus Teilen der Eliten des Osmanischen Reiches zusammensetzen mußten, die zunächst ihr Establishment verteidigen wollten. Diesen Eliten boten sich verschiedene Lösungsmöglichkeiten an. Zunächst die Modernisierung des Osmanischen Reiches durch die Verbesserung des Bildungswesens und die Erzeugung eines osmanischen Patriotismus. Durch die Modernisierungsbestrebungen im militärischen Bereich entstand hier eine Zwischenschicht: Offiziere und Bürokraten, die – ähnlich wie in anderen Balkanstaaten – die Nationalbewegung vorantreiben sollten. Nachdem sich der osmanische Staat unter der Regierung des Sultans Abdülhamit II als reformunfähig erwiesen hatte, bot sich mit der Organisation *İttihat ve Terakki* (Einheit und Fortschritt) der Jungtürken eine Lösung in Richtung Türkisierung des Osmanischen Reiches an. Dabei spielten die Beispiele der balkanischen Nationalbewegung eine wichtige Rolle. Die kollektivsymbolische Umdeutung des Türkischen, seine Geschichtlichkeit, sein würdiges Alter und die Stärke der türkischen Nation wurde der Ideologie des Panturkismus entnommen, der wie alle anderen Pan-Bewegungen sprachlich motiviert ist. Er ist dem Panславismus nachgebildet und gehört wie dieser und auch der oben erwähnte Indogermanismus oder Pangermanismus zur Basis des Rassismus des 20. Jahrhunderts. Erst nach dem endgültigen Versagen und Zusammenbruch des Osmanischen Reichs versuchte Mustafa Kemal, die türkische Republik auf eine andere ideologische Basis zu stellen, nämlich auf die der Staatsnation, und sich von dem in Mittel- und Südosteuropa dominierenden Konzept der Volksnation zu lösen. Deswegen wurde der Panturkismus auch von ihm bekämpft. Dieser war jedoch offensichtlich in Teilen der türkischen Intel-

<sup>20</sup> Faensen, Johannes: Die albanische Nationalbewegung. Berlin 1980, S. 19 ff, 41, 63, 99-112.

<sup>21</sup> Springborn, Radina: Die bulgarische Nationalbewegung, in: Reiter, Norbert (Hg.): Nationalbewegungen auf dem Balkan. Berlin 1983, S. 307 ff.

lizenzen schon derart verwurzelt, daß er im rassistischen Diskurs in der Türkei bis heute eine wichtige Rolle spielt.<sup>22</sup>

Im Sinne der Modernisierung wurde in der türkischen Republik – wie in den anderen südosteuropäischen Staaten – der Sprachenfrage ein besonderes Gewicht beigemessen. Sie beherrscht bis in unsere Tage den nationalen Diskurs. Es geht auch hier um die Beseitigung der Diglossie und um die Etablierung einer polyfunktionalen Standardsprache. Dies wäre vermutlich schneller gelungen, wenn sich die Vorgänger und Nachfolger der *Türk Dil Kurumu* (d. i. die Türkische Sprachgesellschaft, die für die Modernisierung und Standardisierung des Türkischen zuständige staatliche Institution, die mit der höchsten sprachpflegerischen Autorität ausgestattet ist) nicht immer wieder in eine Entwicklung eingemischt hätten, die durch erfolgreiche schriftsprachliche und literarische v. a. übersetzerische Tätigkeit längst Eigendynamik entwickelt hatte.

### **3.8. Sprachnationalismus in den südosteuropäischen Nationalbewegungen**

Es bleibt aber festzuhalten, daß es auch in der Türkei wie bei allen anderen südosteuropäischen Nationalbewegungen darum ging, auf der Basis einer standardisierten Volkssprache nationale Identität und Solidarität zu symbolisieren. Dieser Solidarisierungsprozeß gelang in einem Diskurs, in dem die Erkenntnisse der genealogischen Sprachwissenschaft (Indogermanistik) in der Weise metaphorisiert und instrumentalisiert wurden, daß der vorhandenen oder projektierten Nationalsprache eine würdige und überlegene Position in einer alten und vitalen Sprach- und Rassegemeinschaft nachgewiesen wurde. Dieser Nachweis geschah mit wissenschaftlichen Methoden und Argumenten. Der so gewonnene Anspruch auf Überlegenheit wurde dann durch weitere Kollektivsymbole verabsolutiert. Vergegenständlicht wird er v. a. in Hymnen, feierlichen Inszenierungen und in monumentaler Bau- und bildender Kunst.

## **4. Nationale Kollektivsymbolik im historischen Diskurs**

### **4.1. Anachronistische Sichtweisen**

Ein weiteres System der nationalen Kollektivsymbolik, welches die Vorstellung von Einheit, Kontinuität und Überlegenheit der Nation verfestigen sollte, fanden die Nationalbewegungen in dem von ihnen in Gang gesetzten historischen Diskurs. Dieser zeichnet sich dadurch aus, daß nationales Denken und Handeln in Zeiten zurückprojiziert wird, in denen es überhaupt noch nicht vorhanden war.

So sehen deutsche Geschichtsbücher ein deutsches Nationalbewußtsein schon mit dem ersten Auftreten des Ethnonyms also im frühen Mittelalter; in Griechenland und Rumänien wurden und werden antike Ortsnamen wiederbelebt; in der Türkei werden vortürkische Ortsnamen volksetymologisch umgedeutet und entsprechend angepaßt. Mittelalterliche oder antike Stammesallianzen werden nationalisiert: so werden die ver-

---

<sup>22</sup> Vgl. Georgeon, François: *Aux Origines du Nationalisme Turc*. Yusuf Ak-çura (1876-1935). Paris 1980. Oba, Ali Engin: *Türk milliyetçiliğinin doğusu*. Ankara 1995.



schiedenen zentralasiatischen tribal organisierten Machtgebilde (gekennzeichnet durch Attila, Dschingis Khan u. a.) zu türkischen Staaten ernannt, wobei eine Kontinuität zur heutigen türkischen Republik behauptet wird. Der Aufstand des germanischen Stammesführers Arminius gegen die Römer wurde als erste deutsch-nationale kollektive Handlung zur Behauptung der deutschen Nation interpretiert, der mit der römischen Kaiserwürde versehene mittelalterliche König Friedrich I (Barbarossa) wird als mittelalterliche Vergegenständlichung der deutschen Nation gesehen, der römische Kaiser Traian, Mihai Viteazul, Fürst der Wallachei und der als Pfahl-Vojevode auch in der Türkei bekannte und berühmte Vlad Dracul (türk. kazıklı voyvoda) gelten als Väter der rumänischen Nation, der Spartaner Leonidas und andere Helden des antiken Griechenlandes gelten als direkte Vorgänger der griechischen Nation, ganz Kosova ist offensichtlich für beachtliche Teile der serbischen Elite ein Kollektivsymbol für den heldenhaften Kampf der mittelalterlichen Serben um nationale Unabhängigkeit.

Diese anachronistische Sichtweise, die verschleiert, daß es sich bei Nationalbewegungen und Nationalstaaten grundsätzlich um Phänomene des 19. Jahrhunderts handelt, wurde und wird im nationalen Diskurs auch heute noch eher selten in Frage gestellt. Dies zeigen u. a. die geläufigen Bezeichnungen für Nationalbewegungen als nationale Wiedergeburten (dieser Terminus impliziert, daß etwas früher Vorhandenes erneut auflebt), so in Rumänien, Bulgarien und Albanien. Gefördert und emotional aufgeladen wurde diese Sichtweise auch durch verschiedene Hymnen, die mit „Erwache...“ beginnen. Mir sind allein drei Hymnen dieser Art aus Rumänien, Ungarn und der Türkei bekannt. Auch das früher in Deutschland sehr populäre Lied mit dem Text von Friedrich Rückert „Der alte Barbarossa...“ folgt diesem Muster nach einem metonymischen Prinzip: Stellvertretend für alle Deutschen wird Barbarossa erwachen.

## **4.2. Nationale Mythen als Basis kollektiver Mentalitäten**

### **4.2.1. Entstehungsmythen**

Zu der im historischen Diskurs entwickelten nationalen Kollektivsymbolik gehören auch die verschiedenen nationalen Mythen, insbesondere die Entstehungsmythen. Ein solcher ist u. a. in Frankreich der Sturm auf die Bastille, in der Schweiz der Rütlichschwur, in England die Schlacht von Hastings, in Ungarn die Krönung Stefans des Heiligen – es sind immer Mythen, die einen Beginn, häufig einen befreienden Akt symbolisieren. Im Gegensatz dazu orientierte sich der deutsche Nationalmythos am Nibelungenlied, das den Untergang eines Volkes zum Thema hat. Dabei wurde der Held des Nibelungenliedes, Siegfried, gleichgesetzt mit Arminius und Kaiser Wilhelm I. Daraus wurden als angebliche kollektive Eigenschaften und Mentalitäten der Deutschen männliche Kraft, Todesbereitschaft und unbedingte Treue gefiltert, aus denen sich dann leicht die bürgerlichen Sekundärtugenden wie Gründlichkeit, Fleiß und Sauberkeit ableiten ließen.<sup>23</sup>

<sup>23</sup> Münkler, Herfried / Storch, Wolfgang: Siegfrieden. Politik mit einem deutschen Mythos. Berlin 1988, S. 54ff.

Diese Eigenschaften rechtfertigen im nachhinein die bereits erwähnten Niederlagen der deutschen bürgerlichen Eliten gegen die feudale Unterdrückung; sie setzten das Bewußtsein der Vergänglichkeit jeglichen Widerstandes, den deutschen Untertanengeist, in ein positives Licht.

#### 4.2.2. Mythen der lustvollen Unterwerfung und des lustvollen Todes

Dieser Untertanengeist äußerte sich dann in der lustvollen Unterwerfung, die sich bis zur Begeisterung für die Massenabschlachtungen des Ersten Weltkrieges steigerte und im Langemark-Mythos vergegenständlicht und dann später von den Nazis instrumentalisiert wurde. Diese durch den Nibelungenmythos gerechtfertigte Haltung fand ihren literarischen Niederschlag in den Todesphantasien eines Walter Flex oder Ernst Jünger, die von einer erotischen Lust am Sterben und Töten bestimmt sind, die heutige Leserinnen und Leser pervers anmutet.<sup>24</sup>

Dahinter stand die Leere und Bodenlosigkeit der deutschen nationalen Identität, die sich nur negativ über erfundene Feinde definieren ließ. Deutschsein bedeutete Nicht-Franzose-Sein und Nicht-Jude-Sein. Diese mit Feindbildern aufgefüllte Leere und Bodenlosigkeit des deutschen Nationalbewußtseins ist m. E. die Ursache für die mörderische und perfektionistische Konsequenz, mit der deutsche Antisemiten ihre jüdischen Opfer, die doch selber Deutsche waren, bis hin zum Holocaust verfolgten.<sup>25</sup>

#### 4.2.3. Vergegenständlichung dieser Mythen in Sprache und Kunst

Nationale Mythen, die Zerstörung und Tod beinhalten, finden sich in einer auffälligen Dichte auch in der Kollektivsymbolik südosteuropäischer Nationen. Die Schlacht von Thermopylai, die Schlacht auf dem Kosovo Polje und die Dardanellenschlacht im Ersten Weltkrieg werden im nationalen Diskurs zu Geburtsstunden der Nationen (der griechischen, serbischen und türkischen) erklärt. Aus diesen Mythen werden nationale Eigenschaften wie Mut, Tapferkeit, soldatische Ehre, Treue und Liebe zum Vaterland, Todesbereitschaft herausgefiltert.

Diese durch Kollektivsymbole erzeugten Mentalitäten äußern sich z. B. in Redensarten wie „*Türk asker doğar*“ (Der Türke wird zum Soldaten geboren) oder „*Türk korkmaz*“ (Der Türke fürchtet sich nicht) oder „*Samo sloga Srbina spašava*“ (Nur die Einheit rettet den Serben). Sie wurden und werden in entsprechenden nationalen Liedern verherrlicht. Eine Vergegenständlichung fand auch in der bildenden Kunst statt. Dies zeigen v. a. die Nationalgalerien in den südosteuropäischen Hauptstädten, in denen monumentale Schlachtengemälde (meist Schlachten gegen die Türken) dominieren. Eine positive Ausnahme bildet hier übrigens die Nationalgalerie in Ankara, wo friedliche Genrebilder und Interieurs vorherrschen.

---

<sup>24</sup> Vgl. Krockow, Christian Graf von: Von deutschen Mythen. Rückblick und Ausblick. Stuttgart 1995, S. 25, 61.

<sup>25</sup> ebd., S. 38ff.

### 4.3. Die Rolle von Linguistik und Sozialwissenschaften im nationalistischen Diskurs

Diese kollektivsymbolischen Mythen zeigen – wie erwähnt – die eigene Leere und Bodenlosigkeit nationaler Identitäten. Dieser Leere entspricht auch heute noch die kollektive Labilität von Teilen der nationalen Eliten in Südosteuropa (und natürlich nicht nur dort), die sich ständig bedroht fühlen oder sich durch ein konstruiertes Bedrohungspotential überhaupt erst legitimieren, nach dem Prinzip: „Wir gehen unter, wenn wir uns nicht mit allen – auch den aggressivsten – Mitteln wehren“. Sie müssen ständig die eigene Nation überhöhen und dabei die anderen abwerten. Die politischen Folgen sind verheerend, wenn diese kollektive Labilität die politische Macht erlangt. Es ist zu erwarten, daß mit den massenhaften Verunsicherungen, die die Globalisierung mit sich bringt, das Angebot der mittels der nationalen Kollektivsymbolik konstruierten Volksnation eher an Attraktivität gewinnt. Um so mehr ist es die Aufgabe der Sozialwissenschaften aufzudecken, daß es sich bei der nationalen Identität um das Ergebnis eines Diskurses, in welchem Metaphern und Mythen in nationalen Kollektivsymbolen vergegenständlicht werden, und eben nicht um natürliche Eigenschaften handelt. Dieser nationalistische Diskurs aber geht über die nationalen Grenzen hinaus, ist also eigentlich international und hebt die Einzigartigkeit der Nationen auf.

Ich habe – notgedrungen sehr vereinfacht – versucht zu zeigen, daß zwischen der deutschen und der südosteuropäischen nationalen Kollektivsymbolik auffällige Parallelitäten bestehen, und habe auch deutlich gemacht, daß die türkische Nationalbewegung sich der gleichen Kollektivsymbolik bedient hat und somit als ein Bestandteil der südosteuropäischen Nationalbewegungen zu sehen ist. Die enge Verflechtung wird bereits deutlich, wenn man die Biographien einzelner Nationalapologeten betrachtet – wie z. B. die der Brüder Frashëri. In Zukunft sollte das Forschungsinteresse sich also wesentlich mehr auf die Beziehungen **zwischen** den südosteuropäischen Nationalbewegungen und ihrer Kollektivsymbolik richten als auf die einzelnen Nationalbewegungen.

**Christian Krekeler**

## **TestDaF und DSH – ungleiche Sprachtests im Vergleich**

### **Zusammenfassung**

Anlass für die Studie waren Unsicherheiten von Zulassungsstellen im Umgang mit dem TestDaF-Zeugnis. Mit einer quantitativen Studie (Durchführung der DSH und des TestDaF mit C-Test), welche durch eine qualitative Studie ergänzt wurde (Befragung einzelner Teilnehmer), wurde die Frage verfolgt, welches Ergebnis im TestDaF einer bestandenen DSH entspricht bzw. bei welchem Ergebnis die sprachlichen Voraussetzungen für ein Fachstudium erfüllt sind.

Ein Vergleich der Testergebnisse wird durch die unterschiedlichen Formate und die fehlende Standardisierung der DSH erschwert. Da die Tests der Studie unter unterschiedlichen Voraussetzungen (mit/ohne Konsequenzen für die Kandidaten) und in einem zeitlichen Abstand von drei Monaten stattfanden, waren weitere Einschränkungen zu beachten. Die Studie kommt zu dem Ergebnis, dass ein Recht auf Befreiung von der DSH ausschließlich bei einem TestDaF-Ergebnis „TestDaF-Niveaustufe (TDN) 5 in allen vier Subtests“ unangemessen ist. Studienbewerber sollten mit einem TestDaF-Ergebnis von „durchschnittlich TDN 4“ von der DSH befreit werden. Ob das Ergebnis „DSH bestanden“ auch einem niedrigeren TestDaF-Ergebnis entspricht, kann mit dieser Studie weder bestätigt noch ausgeschlossen werden. Möglicherweise sollte die Frage nach der Äquivalenz jedoch umgekehrt gestellt werden: Sollte sich die DSH nicht vielmehr am standardisierten TestDaF ausrichten? Voraussetzung ist eine Festlegung von angemessenen TestDaF-Ergebnissen für die Studienzulassung.

### **1. Vergleichbarkeit der Tests**

Mit dem „Test Deutsch als Fremdsprache für Studienbewerber“ (TestDaF) existiert neben der „Deutschen Sprachprüfung für den Hochschulzugang“ (DSH) seit dem Jahr 2000 ein weiterer Test, der sich auf die Prüfung der Deutschkenntnisse ausländischer Studienbewerber konzentriert. Der TestDaF, ein zentral gestellter und zentral ausgewerteter Sprachtest für den Hochschulzugang, bietet eine Alternative zur DSH. Studienbewerber oder Studieninteressierte können den TestDaF schon im Heimatland ablegen, sich vor der Einreise nach Deutschland ein Bild über ihre Deutschkenntnisse machen

und den erforderlichen Nachweis der Deutschkenntnisse erbringen. Die Verbreitung des TestDaF wird in den nächsten Jahren zunehmen, denn die Möglichkeit, den Test bereits im Heimatland abzulegen, ist für viele Studienbewerber trotz der Prüfungsgebühren in Höhe von € 90 bis 110 sicherlich attraktiv. Damit wird die Zahl derjenigen zunehmen, die sich mit einem TestDaF-Zeugnis um einen Studienplatz in Deutschland bewerben.

Aus dem Nebeneinander zweier Sprachtests mit unterschiedlichen Vorgehensweisen ergibt sich die Frage nach der Vergleichbarkeit der Ergebnisse. Bislang wird die Zulassung von Studienbewerbern mit TestDaF-Zeugnis unterschiedlich gehandhabt. Die Studie soll Hinweise geben, mit welchem TestDaF-Zeugnis ausländische Studienbewerber zum Studium zugelassen bzw. mit welchem TestDaF-Ergebnis sie von der DSH befreit werden sollten. Grundlage des Beitrags ist eine qualitative und eine quantitative Studie mit Studierenden des Studienkollegs der Fachhochschule Konstanz<sup>1</sup>. Zunächst möchte ich einige Unterschiede zwischen beiden Tests vorstellen, die im Rahmen der Studie von Belang sind.

*TestDaF: differenzierte Beschreibung der sprachlichen Fertigkeiten*

*DSH: „bestanden – nicht bestanden“*

Ein wichtiger Unterschied zwischen beiden Tests liegt in der unterschiedlichen Rückmeldung über die Ergebnisse. Zunächst zur DSH: Das DSH-Zeugnis enthält in der Regel folgenden Satz:

„Frau/Herr ... hat die ‘Deutsche Sprachprüfung für den Hochschulzugang’ (DSH) bestanden und damit die für ein Studium an einer deutschen wissenschaftlichen Hochschule erforderlichen Kenntnisse der deutschen Sprache nachgewiesen.“  
(FaDaF 2001: 10/3)

Wer das Zeugnis vorlegen kann, erfüllt die sprachlichen Voraussetzungen für ein Studium und kann zugelassen werden. Wer ein derartiges Zeugnis nicht vorlegen kann, kann nicht zugelassen werden. Damit erübrigt sich eine Interpretation. In Ausnahmefällen kann ein „Zeugnis über die bestandene DSH mit Auflagen“ ausgestellt werden (vgl. FaDaF 2001: 3/11), differenzierte Aussagen über die Leistungen in den Subtests sind im Unterschied zum TestDaF nicht vorgesehen. Wenn die DSH Teil der Feststellungsprüfung für Studierende von Studienkollegs ist, erhalten die Kandidaten eine Zensur, welche in die Gesamtnote der Feststellungsprüfung eingerechnet wird. Aber auch hier gibt es den Schwellenwert, so dass das wichtigste Resultat „bestanden“ oder „nicht bestanden“ bleibt. Im Studienkolleg der Fachhochschule Konstanz müssen die Kandidaten in der DSH als Teil der Feststellungsprüfung mindestens die Gesamtnote 4,0 auf einer Notenskala von 1,0 bis 5,0 erreichen.

---

<sup>1</sup> Ich spreche bei den Teilnehmerinnen und Teilnehmern des Studienkollegs von „Studierenden“. Sie haben zwar ihr Fachstudium noch nicht aufgenommen, sind aber im Besitz eines „Studierendenausweises“.

Der TestDaF stellt die Ergebnisse nach sprachlichen Fertigkeiten getrennt dar. In den Subtests Leseverstehen (LV), Hörverstehen (HV), Schriftlicher Ausdruck (SA) und Mündlicher Ausdruck (MA) erfolgt jeweils eine Rückmeldung in TestDaF-Niveaustufen (TDN): TDN 3, TDN 4, TDN 5 (vgl. Tabelle 1). Liegt das Ergebnis unter der TDN 3, wird nicht weiter differenziert, sondern das Ergebnis „unter TDN 3“ vermerkt.

Tabelle 1: TestDaF-Niveaustufen am Beispiel des Subtests Schriftlicher Ausdruck

<b>TDN 5</b>	Kann sich in studienbezogenen Alltagssituationen (z. B. Bericht für Stipendiengeber) sowie im fächerübergreifenden wissenschaftlichen Kontext (z. B. Protokoll, Thesenpapier) zusammenhängend und strukturiert sowie der jeweiligen Textsorte angemessen äußern.
<b>TDN 4</b>	Kann sich in studienbezogenen Alltagssituationen (z. B. Bericht für Stipendiengeber) sowie im fächerübergreifenden wissenschaftlichen Kontext (z. B. Protokoll, Thesenpapier) weitgehend zusammenhängend und strukturiert sowie der jeweiligen Textsorte weitgehend angemessen äußern, allerdings mit sprachlichen Mängeln, die jedoch das Textverständnis nicht behindern.
<b>TDN 3</b>	Kann sich in studienbezogenen Alltagssituationen (z. B. Bericht für Stipendiengeber) nur eingeschränkt zusammenhängend und strukturiert schriftlich äußern; kann sich im fächerübergreifenden wissenschaftlichen Kontext (z. B. Protokoll, Thesenpapier) schriftlich nicht so äußern, dass das Textverständnis hinreichend gesichert ist.

(TestDaF-Institut 2001)

Ein derartig differenziertes Ergebnis ermöglicht eine differenzierte Zulassungspraxis. Es stellt gleichzeitig die Abnehmer der Zeugnisse, ausländische Studienbewerber und Zulassungsstellen, vor die Aufgabe, das Zeugnis zu interpretieren. Es muss festgelegt werden, mit welchem TestDaF-Ergebnis eine Zulassung ausländischer Studierender sinnvoll ist. Bislang gelten folgende Regelungen:

„Wird TDN 5 in allen vier Subtests erreicht, dann bedeutet dies – aufgrund eines Plenarbeschlusses der HRK [Hochschulrektorenkonferenz] vom Februar 2000 – die uneingeschränkte Zulassung zu einem Fachstudium. Hat der Kandidat TDN 5 in zwei Subtests und TDN 4 in den beiden übrigen Subtests erreicht, dann wird den Hochschulen die Zulassung empfohlen. Bei einem Ergebnis von TDN 4 in zwei und mindestens TDN 3 in den beiden anderen Subtests kann die Hochschule unter der Auflage studienbegleitender Sprachkurse eine Zulassung zum Fachstudium aussprechen. Ein Ergebnis von TDN 3 in drei oder vier Subtests bedeutet, dass die sprachliche Leistung des Kandidaten noch nicht für die Zulassung zu einem Fachstudium ausreicht. Die Hochschulen können unter bestimmten Voraussetzungen – wie z. B. im Fall bilingualer Studiengänge – Ausnahmen zulassen.“  
(Grotjahn/Kleppin 2001: 427)

Das TestDaF-Institut hat eine Handreichung mit Empfehlungen ausgegeben, aus der das Bestreben deutlich wird, Hochschulen in Bezug auf die Sprachkenntnisse ausländischer Studienbewerber differenzierte Aufnahmeregelungen zu ermöglichen.

Tabelle 2: Vorschlag des TestDaF-Instituts zur Zulassung ausländischer Studierender

<b>Subtest</b> Fächergruppe	<b>Leseverstehen</b>	<b>Hörverstehen</b>	<b>Schriftlicher Ausdruck</b>	<b>Mündlicher Ausdruck</b>
Geistes-, Sprach-, Kulturwissenschaften	<b>TDN 5</b>	<b>TDN 4</b>	<b>TDN 4</b>	<b>TDN 4</b>
Jura, Sozial-, Wirtschaftswissenschaften	<b>TDN 4</b>	<b>TDN 4</b>	<b>TDN 4</b>	<b>TDN 4</b>
Informatik, Ingenieurwissenschaften, Kunst und Musik, Mathematik, Naturwissenschaften	<b>TDN 3</b>	<b>TDN 3</b>	<b>TDN 3</b>	<b>TDN 3</b>
	<b>unter TDN 3</b> keine Zulassung in deutschsprachigen Studiengängen			

(TestDaF-Institut 2002)

Der differenzierte Ergebnisausweis des TestDaF hat zu Diskussionen geführt. Zwei Aspekte sind hervorzuheben: Erstens die Möglichkeit einer nach Studienfach differenzierter Zulassung und zweitens die Zulassung mit Auflagen.

Die Frage, mit welchem TestDaF-Ergebnis Studienbewerber zu einem bestimmten Studiengang aufgenommen werden können, scheint aktueller als die Äquivalenz mit einer bestandenen DSH. Die folgenden Gründe haben mich bewogen, die Vergleichbarkeit der Prüfungen dennoch in den Mittelpunkt der vorliegenden Untersuchung zu stellen: Erstens sind die sprachlichen Anforderungen einzelner Studiengänge meiner Kenntnis nach nicht systematisch erhoben und vor allem nicht verglichen worden. Ob Maschinenbau-Studierende weniger Deutschkenntnisse benötigen als Geisteswissenschaftler, ob die sprachlichen Anforderungen für Kunststudierende deutlich unter denen für Studierende der Kulturwissenschaften liegen, sind offene Fragen, die einer eigenen Untersuchung bedürfen. Zweitens ist die Frage nach der Äquivalenz für die Zulassungsstellen von großer Dringlichkeit. Der Beschluss der Hochschulrektorenkonferenz wird von einigen Zulassungsstellen so ausgelegt, dass nur Studienbewerber von der DSH befreit werden, welche ein TestDaF-Zeugnis mit der TDN 5 in allen vier Subtests vorlegen. Schließlich sollte die Frage nach der Äquivalenz beider Prüfungen – bei allen Schwierigkeiten auf Grund der Uneinheitlichkeit der DSH – zufriedenstellend beantwortet sein, bevor gesonderte Regelungen für einzelne Studienfächer aufgestellt werden. Verzichtet man auf diesen Schritt, bedeutet das auch, auf die langjährigen Erfahrungen mit der DSH zu verzichten.

Zum zweiten Aspekt: Durch den differenzierten Ergebnisausweis des TestDaF ist häufig von einer Zulassung mit Auflagen die Rede. Bei einem mittleren TestDaF-Ergebnis könnte beispielsweise eine bedingte Zulassung ausgesprochen werden mit der Auflage, am studienbegleitenden Sprachunterricht teilzunehmen. Der Sinn studienbegleitender Sprachförderung steht außer Zweifel. Ich habe aber Zweifel, ob man im Sinne ausländischer Studienbewerber handeln würde. Im Normalfall ist es sicherlich günstiger,

wenn ausländische Studierende bereits zu Beginn des Studiums über fortgeschrittene Deutschkenntnisse verfügen. Im Rahmen dieser Studie werden Überlegungen zu einer Zulassung mit Auflagen daher nicht weiter berücksichtigt. Im Mittelpunkt stehen die Fragen, bei welchem TestDaF-Ergebnis ausländische Studienbewerber von der DSH befreit werden sollten bzw. bei welchem TestDaF-Ergebnis ausreichende Sprachkenntnisse für die Aufnahme eines Fachstudiums in Deutschland vorliegen.

*TestDaF: standardisiert*

*DSH: nicht standardisiert*

Der TestDaF ist testmethodisch „auf der Höhe der Zeit“. Er ist ein standardisierter Test, d. h. er wird an einer repräsentativen Stichprobe geeicht. Die Entwicklung der Testsätze verläuft in mehreren Stufen, die Aufgaben werden vorerprobt, erprobt, gegebenenfalls revidiert und geeicht. Das Ergebnis ist ein Sprachtest, der die Gütekriterien von Sprachtests in hohem Maße erfüllt. Man kann davon ausgehen, dass die Ergebnisse von zwei Kandidaten unabhängig von der Testversion vergleichbar sind. Ein Kandidat, welcher an unterschiedlichen Testversionen teilnimmt, würde bei gleicher Leistung gleiche Ergebnisse erzielen.

Die DSH ist ein Beispiel für eine Prüfungstradition, welche vom Unterricht, von den Unterrichtenden und von der Institution ausgeht. Die Rahmenordnung für die DSH wurde von der Hochschulrektorenkonferenz 1995 in Kraft gesetzt und inzwischen wohl von allen Hochschulen in Deutschland umgesetzt. Im Auftrag des „Fachverbandes Deutsch als Fremdsprache“ (FaDaF) wurde 2001 ein Handbuch für Prüferinnen und Prüfer herausgegeben, das die Rahmenordnung durch Anwendungsbeispiele und Hinweise zur Durchführung konkretisieren soll. Auch über eine „Zertifizierung“ von Prüfungen soll eine Vergleichbarkeit gewährleistet werden. Dennoch: Die DSH ist keine standardisierte Prüfung, die Gütekriterien von Sprachtests dürften kaum zufriedenstellend erfüllt werden. Angesichts der fehlenden Standardisierung der DSH bezweifelt Alderson (2002), dass die DSH überhaupt als Test anzusehen ist. Seiner Meinung nach ist es angemessener, von einem Gerüst („framework“) zu sprechen. Vergleicht man DSH-Prüfungen miteinander, wird deutlich, dass die DSH-Rahmenordnung ein breites Spektrum an Tests ermöglicht bzw. dass sie unterschiedlich ausgelegt wird. Deutliche Abweichungen im Subtest Hörverstehen offenbarte beispielsweise die Studie von Koreik/Schimmel (2002). Kandidaten, welche am DSH-Subtest Hörverstehen verschiedener Institute teilnahmen, erzielten sehr unterschiedliche Ergebnisse. Die Prüfungsordnung erlaubt außerdem die Kombination von Testteilen, so dass eine DSH aus zwei bis vier Subtests bestehen kann. Für Untersuchungen über die DSH hat diese Problematik Folgen: Aussagen über den Schwierigkeitsgrad einer DSH oder einer ihrer Subtests sind immer unter der Einschränkung zu betrachten, dass die Tests möglicherweise stark voneinander abweichen. Auch für Kandidaten hat die eingeschränkte Vergleichbarkeit Folgen: Sie können nicht wissen, ob sie bei einer anderen DSH das gleiche Ergebnis erzielt hätten.



*TestDaF: integriertes Testen der sprachlichen Fertigkeiten**DSH: zusätzlicher Grammatiktest*

Der TestDaF besteht aus vier Subtests, in denen die sprachlichen Fertigkeiten integriert geprüft werden (vgl. Tabelle 3). Ein getrennter Test sprachlicher Elemente (Grammatik, Phonetik, Lexik) findet nicht statt. Bei den Subtests Schriftlicher Ausdruck und Mündlicher Ausdruck handelt es sich um direkte Performanztests. Sie stellen die Frage, ob die Kandidaten ihre Sprachfähigkeit in einer spezifischen Situation anwenden können. Auch die Subtests Leseverstehen und Hörverstehen sind als Performanztest anzusehen, auch wenn sie eher indirekt vorgehen müssen, denn das Verstehen eines Textes wird anhand von Fragen über den Text geprüft.

Tabelle 3: Subtests des TestDaF und der DSH im Vergleich

<b>TestDaF</b>	<b>DSH</b>
<b>Leseverstehen</b> Der Subtest LV besteht aus drei Texten mit unterschiedlichem Schwierigkeitsgrad. Insgesamt sind 30 Items zu bearbeiten (Zuordnung, Mehrfachauswahlaufgaben, Auswahlitems). Die TestDaF-Niveaustufen werden aus den Rohwerten ermittelt (max. 30 Punkte). Dauer: 60 Minuten.	<b>Verstehen und Bearbeiten eines Lesetextes</b> (DSH-LV) Aufgaben zu einem Lesetext, der authentisch, studienbezogen und wissenschaftsorientiert ist.
<b>Hörverstehen</b> Der Subtest HV besteht aus drei Hörtexten mit unterschiedlichem Schwierigkeitsgrad. Insgesamt sind 25 Items zu bearbeiten (gesteuerte Notizen, Auswahlitems). Die TestDaF-Niveaustufen werden aus den Rohwerten ermittelt (max. 25 Punkte). Dauer: 40 Minuten.	<b>Verstehen und Verarbeiten eines Hörtextes</b> (DSH-HV) Aufgaben zu einem Hörtext, der max. zweimal präsentiert wird. Der Text soll „der Kommunikationssituation Vorlesung/Übung angemessen Rechnung tragen“.
<b>Schriftlicher Ausdruck</b> Beim Subtest SA soll ein Text zu einer Aufgabenstellung, welche eine Grafik umfasst und eine begründete Stellungnahme verlangt, verfasst werden. Die Bewertung (TestDaF-Niveaustufen) erfolgt durch zwei, im Zweifel durch drei Prüfer. Dauer: 60 Minuten.	<b>Vorgabenorientierte Textproduktion</b> (DSH-TP) Texterstellungsaufgabe. Sie kann „erklärender, vergleichender oder kommentierender Art sein, sie kann auch die sprachliche Umsetzung von Grafiken, Schaubildern, Diagrammen zum Gegenstand haben.“
	<b>Verstehen und Bearbeiten wissenschaftssprachlicher Strukturen</b> (DSH-Grammatiktest) Textgebundene Aufgabenstellung. Die Leistung wird nach sprachlicher Richtigkeit bewertet.
<b>Mündlicher Ausdruck</b> Der MA wird als „Simulated Oral Proficiency Interview“ (SOPI) durchgeführt. In diesem kassettengesteuerten Subtest müssen die Kandidaten auf zehn kurze Sprechanlässe reagieren. Bewertung erfolgt in TestDaF-Niveaustufen. Dauer: 30 Minuten.	<b>Mündliche Prüfung</b> Prüfungsgespräch, Dauer: max. 20 Minuten. Sie kann entfallen, wenn „für die Beurteilung der mündlichen Kommunikationsfähigkeit hinreichende Erkenntnisse vorliegen.“

(vgl. TestDaF-Institut 2001, FaDaF 2001. Prüfungsordnung der DSH, zit. nach FaDaF 2001)

Eine Besonderheit der Subtests MA, LV und HV sind Aufgabengruppen auf unterschiedlichen Anforderungsniveaus. Beim Leseverstehen müssen beispielsweise drei unterschiedliche Textsorten mit unterschiedlichen Aufgabentypen bearbeitet werden (vgl. Grotjahn 2000). Der Subtest Mündlicher Ausdruck wird als kassettengesteuertes Interview durchgeführt. Die Kandidaten sprechen nicht mit einem Prüfer, sondern erhalten einen Sprech Anlass vom Band. Dieses Format bewährte sich als zuverlässig und valide (vgl. Kenyon 2000, Kniffka/Üstünsöz-Beurer 2001).

Die Subtests der DSH sind zwar in der Prüfungsordnung beschrieben und im „DSH-Handbuch für Prüferinnen und Prüfer“ erläutert, doch ob sie von den durchführenden Institutionen in dieser Weise umgesetzt werden, wird nicht geprüft. Einige Eckpunkte sind hervorzuheben: Die DSH-Subtests Textproduktion und Mündliche Prüfung sind ebenfalls den direkten Performanztests zuzurechnen. Allein beim DSH-Grammatiktest handelt es sich um einen Kompetenztest. Sprachliche Kompetenztests haben die allgemeine Beherrschung von Sprache unabhängig von einer konkreten Anwendungssituation zum Gegenstand. Sie bestehen eher aus dem Abfragen von Einzelfertigkeiten. Die übrigen Subtests sind durchaus mit denen des TestDaF vergleichbar, allerdings verzichten sie auf abgestufte Aufgabenstellungen. Verallgemeinernde Aussagen zum Format der einzelnen Subtests sind jedoch kaum zu treffen. Eine Besonderheit der DSH, die im Rahmen dieser Untersuchung durchgeführt wurde, ist ein Fachbezug der Subtests Textproduktion und Leseverstehen. Die Kandidaten konnten aus mehreren Aufgaben/Texten wählen. Bei der Textproduktion gab es beispielsweise Aufgaben aus den Bereichen Naturwissenschaften, Technik, Wirtschaft, Informatik, Design und Bauingenieurwesen.

Ich fasse zusammen: Ein Vergleich der Testergebnisse aus der DSH und dem TestDaF ist wünschenswert. In dieser Studie geht es allein um die Äquivalenz der Ergebnisse. Die Vergleichbarkeit beider Tests wird durch die mangelnde Standardisierung der DSH und die unterschiedlichen Ergebniskategorien erschwert. Die Stärke der DSH lässt sich nicht an den klassischen Kriterien der Testgüte festmachen. Sie liegt in der Nähe der Prüfer zu den Kandidaten, zum Unterricht und zum Studium. Beim TestDaF sind geringe Abweichungen zwischen einzelnen Testjahrgängen zwar unvermeidbar (vgl. Grotjahn 2000), doch die aufwändige Eichung gewährleistet eine hohe Vergleichbarkeit der Tests. Aussagen zum Anforderungsniveau des TestDaF können daher verallgemeinert werden.

## **2. Beschreibung der Studie**

Grundlage dieses Beitrags ist eine quantitative Studie mit 67 Studierenden des Studiengangs an der Fachhochschule Konstanz. Sie wurde durch eine qualitative Studie mit ausgewählten Teilnehmern der Testreihe erweitert (vgl. Tabelle 4).

Tabelle 4: Ablauf der Untersuchung

<b>Quantitative Studie</b>	
<b>Phase 1</b> TestDaF-Erprobung, C-Test, DSH-Grammatiktest 4 Lerngruppen des Studienkollegs (58-67 Probanden)	<b>Phase 2</b> (3 Monate später) DSH 2 Lerngruppen des Studienkollegs (39) und Externe (192 Probanden)
<b>Qualitative Studie</b>	
Befragung und Beobachtung ausgewählter Kandidaten	

*Quantitative Studie*

Im Rahmen der quantitativen Studie wurden mehrere Sprachtests mit einer Stichprobe von 67 ausländischen Studienbewerbern durchgeführt, die zum Zeitpunkt der Erhebung als Studierende am Studienkolleg der Fachhochschule Konstanz eingeschrieben waren. Zunächst nahmen die Studentinnen und Studenten an einer Erprobungsfassung des TestDaF teil, welche auch einen C-Test umfasste. Zeitnah (zwei Wochen später) machten sie den ersten DSH-Grammatiktest. Die DSH wurde etwa drei Monate später durchgeführt (vgl. Tabelle 4). Auch 192 externe Kandidaten nahmen an der DSH teil. Der TestDaF, welcher von den Studierenden bearbeitet wurde, war eine Erprobungsfassung in der zweiten Stufe. In dieser Phase werden die vorerprobten Aufgaben nochmals auf ihre Güte überprüft und die Ergebnisse der Subtests Leseverstehen sowie Hörverstehen werden den TestDaF-Niveaustufen zugeordnet (vgl. Arras/Eckes/Grotjahn 2002, TestDaF-Institut 2001). Die eingesetzte Erprobungsfassung unterschied sich nur geringfügig von der Endversion. Der erste DSH-Grammatiktest wurde dem DSH-Handbuch entnommen, wo er beispielhaft angeführt wird. Die DSH wurde von einem Team erstellt und korrigiert. Eine Besonderheit der DSH an der Fachhochschule Konstanz ist der Fachbezug der Testteile Textproduktion und Leseverstehen. Bei beiden Subtests werden den Kandidaten mehrere Texte bzw. Aufgaben zur Auswahl angeboten, welche sich auf unterschiedliche Studienfächer beziehen. Da bei den Kandidaten aus dem Studienkolleg insgesamt nur vier Mündliche Prüfungen durchgeführt wurden, wurde dieser Prüfungsteil nicht in die Untersuchung einbezogen.

Mit Ausnahme der TestDaF-Subtests Mündlicher Ausdruck und Schriftlicher Ausdruck lagen intervallskalierte Daten vor, das heißt die Abstände zwischen den einzelnen Skalenwerten sind konstant. Bei den TestDaF-Niveaustufen könnte das Vorliegen einer Intervallskala bestritten werden, was an dem Ergebnis „unter 3“ liegt. Beim TestDaF werden alle Ergebnisse unter der TDN 3 dem Ergebnis „unter 3“ zugeordnet. Die Erfahrung zeigt aber, dass sich auch das Ergebnis „unter 3“ wie eine weitere Ergebnisklasse auf einer quantifizierenden Skala verhält. Die bivariaten Korrelationen wurden mit Hilfe von Rangkorrelationskoeffizienten bestimmt, welche von der „niedrigeren“ Rang- oder Ordinalskala ausgehen.

Bei der Interpretation sind folgende Einschränkungen zu beachten:

- Auch in der zweiten Phase der TestDaF-Erprobung kann es zu Änderungen in der Testfassung kommen. Die Ergebnisse im Subtest LV könnten beispielsweise zu einer Überarbeitung der Items geführt haben. Das Schwierigkeitsniveau der Endfassung kann über oder unter der Erprobungsfassung liegen.
- Der TestDaF und der erste DSH-Grammatiktest traf alle Kandidaten völlig unvorbereitet. Die Tests wurden ohne Ankündigung am Beginn eines neuen Semesters durchgeführt. Man kann davon ausgehen, dass Teilnehmer bei einer Sprachprüfung besser abschneiden, wenn sie sich vorher mit dem Format vertraut gemacht haben. Wie groß die Auswirkungen einer Vorbereitung auf die Ergebnisse der einzelnen Testteile sind, ist jedoch unklar.
- Die Ergebnisse aus dem TestDaF und dem ersten DSH-Grammatiktest hatten für die Teilnehmer keine Auswirkungen. Es ist möglich, dass sich einige Kandidaten nicht in gleichem Maße angestrengt haben, bei anspruchsvollen Aufgaben beispielsweise eher aufgegeben haben oder erst gar keinen Antwortversuch unternommen haben als bei einer Prüfung mit gewichtigen Konsequenzen. Die ersten beiden Aspekte sprechen dafür, dass die Ergebnisse aus dem TestDaF und dem ersten Grammatiktests möglicherweise unter den Ergebnissen liegen, die sich in einer Testsituation ohne Probecharakter ergeben hätten.
- Zwischen der ersten (TestDaF, C-Test und DSH-Grammatiktest 1) und der zweiten Phase (DSH mit DSH-Grammatiktest 2) lag ein Semester Unterricht im Studienkolleg, welcher zehn Wochenstunden Deutschunterricht einschloss. In diesem Unterricht wurden die Studierenden auf die spezifischen Anforderungen der DSH vorbereitet und mit dem Prüfungsformat vertraut gemacht. Im Vergleich zu den externen Kandidaten haben die Studierenden aus dem Studienkolleg signifikant bessere Ergebnisse erzielt.
- Die DSH am Ende des Semesters hatte gewichtige Konsequenzen für die Kandidaten. Nur mit einer bestandenen DSH erhalten die Studierenden aus dem Studienkolleg das Feststellungszeugnis, welches sie für ein Fachstudium in Deutschland benötigen. Man muss davon ausgehen, dass sie ihr ganzes sprachliches Können aufgeboren haben, um ein gutes Ergebnis zu erzielen.
- Die mangelnde Standardisierung stellt eine Übertragbarkeit von Studien mit nur einer DSH grundsätzlich in Frage.

Für die Interpretation der Daten haben diese Umstände Folgen: Die Stichprobenergebnisse im TestDaF sind im Vergleich zur Population eher unterbewertet, die Ergebnisse in der DSH eher überbewertet.

### *Qualitative Studie*

Bei der Analyse der Ergebnisse aus der quantitativen Studie traten zwei Schwierigkeiten auf: Zum einen war die Größe der Stichprobe gering. Lagen in der ersten Phase der quantitativen Studie mit 57 bis 65 Probanden noch ausreichend große Stichproben vor,

so waren statistisch signifikante Aussagen auf Basis der Ergebnisse von 29 bis 32 Probanden in der zweiten Phase nicht immer möglich. Zum anderen gaben die statistischen Daten wenig Hinweise auf die Ursachen auffälliger Erscheinungen. Die Interpretation der Ergebnisse stieß auf Hindernisse, weil Informationen über die individuellen Voraussetzungen der Kandidatinnen und Kandidaten fehlten. Vor allem die letztgenannte Schwierigkeit veranlasste mich dazu, im Anschluss an die statistische Auswertung der Tests aus der ersten Phase eine qualitative Studie durchzuführen. Die ausgesprochene Heterogenität als durchgehendes Merkmal der Population „ausländische Studienbewerber“ war ein weiterer Grund für die Berücksichtigung individueller Lernwege und Erfahrungen bei der Interpretation der Testergebnisse. Mit einer Kombination von quantitativen und qualitativen Methoden sollte ein umfassenderes Bild entstehen als dies allein mit statistischen Untersuchungen an der Stichprobe möglich war.

Leitgedanke der Auswahl war nicht, wie repräsentativ die Kandidaten waren; die Auswahl war vielmehr ein Ergebnis der Datenanalyse. Die Gespräche, welche in Form eines Leitfaden-Interviews geführt wurden, waren unterschiedlich formell und von unterschiedlicher Dauer: Manche Gespräche fanden nach der Lehrveranstaltung im Unterrichtsraum oder auf dem Flur statt, andere etwas formaler im Büro. Manchmal waren auch andere Studierende anwesend. In der Regel fertigte ich während der Gespräche Notizen an, in Ausnahmefällen danach. Die Kandidaten wurden nach biografischen Angaben und vor allem nach ihrer Sprachlernbiografie befragt. Außerdem wurden sie um Kommentare zu den Tests und den Testergebnissen gebeten.

### *Beschreibung der Stichprobe*

Die Stichprobe entspricht in wesentlichen Merkmalen der Population „ausländische Studienbewerber“, etwa was die Verteilung der Herkunftsländer und Ausgangssprachen betrifft. Hinsichtlich der Herkunft und der Muttersprache(n) waren die Gruppen sehr heterogen. Die größten Gruppen stellten Studierende aus China, Marokko und dem Libanon. Die Stichprobe ist mit einer Einschränkung als spezifisch repräsentativ anzusehen. Die Einschränkung bezieht sich auf den Besuch des Studienkollegs. Nicht alle ausländischen Studienbewerber haben die Möglichkeit (bzw. die Pflicht), an einem einjährigen studienvorbereitenden Kurs teilzunehmen. Die Studierenden werden in zwei Semestern auf die DSH, die sie als Teil der Feststellungsprüfung ablegen, vorbereitet. An den Ergebnissen der DSH, an der 39 Studierende aus dem zweiten Semester des Studienkollegs und 192 externen Kandidaten teilnahmen, wird deutlich, dass sich die Studierenden aus dem Studienkolleg von externen Bewerbern unterscheiden, die sich – häufig im Heimatland – auf die sprachlichen Anforderungen des Studiums vorbereiten (vgl. Tabelle 5).

Die DSH gilt als bestanden, wenn sowohl die mündlichen als auch die schriftlichen Prüfungsleistungen mindestens mit 4,0 bewertet werden. Da sich nicht alle Kandidaten einer Mündlichen Prüfung unterziehen müssen, werden nur die Ergebnisse der schriftlichen Prüfung untersucht. 38 Studierende aus dem Studienkolleg haben den schriftlichen

Teil der DSH mit der Note 4,0 oder besser bestanden. (Dieses Resultat veränderte sich durch die Mündlichen Prüfungen nicht.) Nur eine Studentin aus dem Studienkolleg bestand nicht (Ergebnis der Schriftlichen Prüfung: 4,2). Zum Vergleich: Bei den 192 externen Kandidaten lautete das Ergebnis der DSH-Gesamtprüfung (Mündliche und Schriftliche Prüfung) 56 Mal „nicht bestanden“, was etwa 30 Prozent entspricht. Im Mittel erzielten die internen Kandidaten in der Schriftlichen Prüfung ein deutlich besseres Ergebnis als die externen Kandidaten (Interne:  $M = 2,67$ ; Externe:  $M = 3,42$ ). Mit einem  $t$ -Test für unabhängige Stichproben konnte gezeigt werden, dass der Leistungsunterschied signifikant ist ( $t(229) = 4,672$ ,  $p < 0,01$  2-seitig). In der Textproduktion war der Unterschied zwischen den externen und internen Kandidaten am größten.

Tabelle5: Statistische Kennwerte, DSH (Schriftliche Prüfung)

	Gruppe	Mittelwert (M)	Standardabweich. (s)	Min	Max	Anzahl (n)
<b>DSH-Ergebnis der Schriftlichen Prüfung</b> (Note 1,0-5,0)	Studienkolleg	2,67	0,83	1,1	4,2	39
	Externe	3,42	0,93	1,1	5,0	192
<b>DSH-Leseverstehen</b> (Note 1,0-5,0)	Studienkolleg	2,95	0,96	1,2	4,5	39
	Externe	3,53	0,96	1,2	5,0	192
<b>DSH-Hörverstehen</b> (Note 1,0-5,0)	Studienkolleg	2,26	1,04	1,0	4,7	39
	Externe	3,15	1,30	1,0	5,0	192
<b>DSH-Textproduktion</b> (Note 1,0-5,0)	Studienkolleg	2,35	0,92	1,0	4,0	39
	Externe	3,52	1,22	1,0	5,0	192
<b>DSH-Grammatik 2</b> (Note 1,0-5,0)	Studienkolleg	3,14	1,08	1,1	5,0	39
	Externe	3,50	1,38	1,1	5,0	192

Welche Ursachen gibt es für die besseren Ergebnisse der internen Kandidaten? Zunächst dürfte ein nahe liegender Grund zutreffen, eine bessere Sprachkompetenz der Studierenden aus dem Studienkolleg. Sie halten sich seit mindestens einem Jahr in Deutschland auf, wurden anhand eines Aufnahmetests für das Studienkolleg aus einer großen Gruppe von Bewerbern ausgewählt und nahmen anschließend zwei Semester an einem intensiven Deutschunterricht teil. Sie sind außerdem mit dem Format der DSH gut vertraut, es werden aber keine Texte oder Aufgaben in die DSH aufgenommen, die bereits im Unterricht behandelt wurden. An den Leistungen im Subtest Textproduktion zeigt sich die Wirkung der wissenschaftspropädeutischen Arbeit im Studienkolleg. Nach der zweisemestrigen Ausbildung im Studienkolleg sind die Studierenden in der Lage, den Leseerwartungen deutscher Prüfer eher zu entsprechen als die externen Kandidaten. Außerdem sind einige Kandidaten nicht zur Prüfung angetreten, wenn sich aus den Leistungen im Semester abzeichnete, dass die Chancen auf Bestehen der DSH/Feststellungsprüfung gering sein würden. Bei den externen Kandidaten sind die Voraussetzungen sehr unterschiedlich, was sich auch in einer größeren Streuung der Ergebnisse zeigt (Externe:  $s = 0,93$ , Interne:  $s = 0,83$ ; vgl. Tabelle 5). Gemeinsames Merkmal der externen Kandi-

daten ist ein Sprachzeugnis der Mittelstufe, das als Voraussetzung für die Teilnahme an der DSH verlangt wird.

Tabelle 6: Teilnehmer der quantitativen Studie

	<b>Sem. 1</b> W-Kurs	<b>Sem. 1</b> T-Kurs	<b>Sem. 2</b> W-Kurs	<b>Sem. 2</b> T-Kurs	<b>externe</b> <b>Kandi-</b> <b>daten</b>
<b>Phase 1:</b> TestDaF Erprobungsfassung mit C-Test, DSH-Grammatik 1	✓	✓	✓	✓	
<b>Phase 2:</b> DSH Gesamtprüfung (einschl. DSH-Grammatik 2)			✓	✓	✓

Anmerkung: Der „W-Kurs“ (Wirtschaftskurs) bereitet auf wirtschaftswissenschaftliche Studiengänge vor. Der „T-Kurs“ (Technikkurs) bereitet auf technische Studiengänge vor (auch Informatik, Architektur).

Die Studierenden werden im Studienkolleg in vier Gruppen mit 18 bis 24 Teilnehmern unterrichtet. Es gibt zwei Lerngruppen für Teilnehmer, die ein technisches Studienfach anstreben (auch Informatik oder Architektur), und zwei Lerngruppen für Teilnehmer, die wirtschaftswissenschaftliche Studiengänge belegen möchten. Dass sich die sprachlichen Leistungen der beiden Gruppen unterscheiden, ging aus den Ergebnissen des vom TestDaF-Institut entwickelten C-Tests hervor, welcher zur Erprobungsfassung des TestDaF gehört. Der C-Test gilt als valider und ökonomischer Test der „allgemeinen Sprachkompetenz“ (vgl. Grotjahn 1995, 1997). Die Leistungen der Studierenden aus den W-Kursen lagen deutlich über denjenigen aus T-Kursen (W-Kurs:  $M = 60\%$ , T-Kurs:  $M = 51\%$ ). Wenn man davon ausgeht, dass der C-Test ein gutes Maß für die allgemeine Sprachkompetenz darstellt, liegt die Sprachkompetenz der Studierenden aus dem W-Zweig signifikant über derjenigen von Studenten aus dem T-Zweig ( $t$ -Test für unabhängige Stichproben:  $t(62) = 2,487$ ,  $p < 0,05$  2-seitig). Die Deutschkenntnisse der Studienbewerber mit den Studienzielen Informatik, Maschinenbau, Elektrotechnik usw. liegen demnach deutlich unter derjenigen von Studienbewerbern, welche wirtschaftswissenschaftliche Studiengänge anstreben. Dies ist mit Blick auf eine differenzierte Zulassungspraxis, welche der TestDaF ermöglicht, zwar interessant, aber kein Hinweis darauf, ob Studierende technischer Studienfächer weniger Deutschkenntnisse benötigen.

An der Studie nahmen Studierende aus dem ersten und dem zweiten Semester des Studienkollegs teil (vgl. Tabelle 6). Es war zu erwarten, dass die Ergebnisse der Studierenden aus dem zweiten Semester im C-Test im Mittel über denen aus dem ersten Semester lagen (1. Sem.:  $M = 51,4\%$ , 2. Sem.:  $M = 59,6\%$ ). Auch dieser Unterschied ist signifikant ( $t(62) = 2,17$ ,  $p < 0,05$  2-seitig). Die Studierenden aus dem ersten Semester hatten einige Wochen vorher an einem von uns konzipierten Aufnahmetest teilgenommen, der auch einen Deutschteil enthält. Die andere Gruppe der Teilnehmer befand sich am Beginn des zweiten Semesters. Sie absolvierten bereits ein Semester des Studienkollegs und wurden in das zweite Semester versetzt. Voraussetzung dafür sind u. a. mindestens ausreichende Leistungen im Fach Deutsch.

### 3. Ergebnisse

#### 3.1 Ergebnisse der quantitativen Studie

##### *C-Test als „Screening Test“ für den TestDaF geeignet*

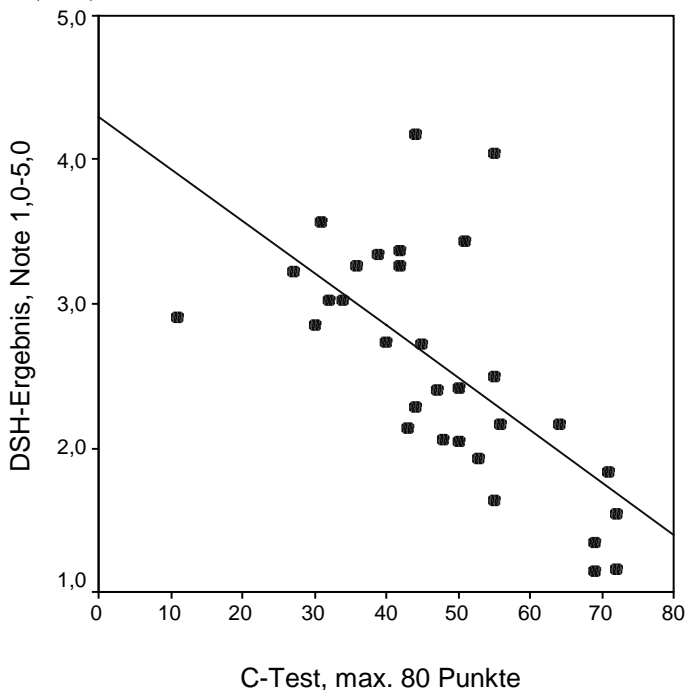
Vor dem Vergleich der Ergebnisse aus dem TestDaF und der DSH möchte ich kurz auf den C-Test eingehen, der mit jeder TestDaF-Erprobungsfassung zur Verankerung der Subtests durchgeführt wird. In einer Studie von Arras, Eckes und Grotjahn (2002) zeigte der C-Test hohe Korrelationen mit allen vier Subtests des TestDaF. Auf Grund der hohen Korrelationen wird außerdem ein Einsatz des C-Tests als „Screening-Test“ für den TestDaF angeregt. Kandidaten könnten von den Ergebnissen eines frei zugänglichen C-Tests ihre Teilnahme am TestDaF abhängig machen. In der vorliegenden Untersuchung liegen die Rangkorrelationskoeffizienten (Spearman) der TestDaF-Subtests mit dem C-Test im Mittelbereich zwischen 0,526 und 0,589 (vgl. Tabelle 7). Sie sind etwas niedriger als die von Arras, Eckes und Grotjahn beobachteten, was eine Folge unterschiedlicher Auswertungsmethoden des C-Tests sein könnte. Sie weisen auf einen mittleren bis guten Zusammenhang zwischen dem C-Test und den TestDaF-Subtests hin, was den Einsatz des C-Tests als Prädiktor für den TestDaF als sinnvoll bestätigt.

##### *C-Test möglicherweise auch als „Screening Test“ für die DSH geeignet*

Wäre ein Einsatz des C-Tests als „Screening-Test“ auch für die DSH sinnvoll? Trotz des zeitlichen Abstandes findet man beträchtliche Korrelationen zwischen dem C-Test und den Subtests der DSH und nicht zuletzt mit dem DSH-Gesamtergebnis ( $r_s = -0,676$ ; vgl. Tabelle 7). Das Gesamtergebnis der DSH wird als ungewichteter additiver Index aus den Ergebnissen der vier Subtests gebildet (Schriftliche Prüfung). Die enge Korrelation wird durch das Streudiagramm bestätigt (vgl. Abbildung 1). Mit Ausnahme der Ergebnisse von drei oder vier Kandidaten liegen die Ergebnisse dicht an der linearen Regressionsgeraden. Vor allem im oberen Leistungsspektrum besteht ein enger Zusammenhang. Kandidaten, welche zuvor im C-Test ein hohes Ergebnis erzielt haben, können bei der DSH ebenfalls mit einem guten Ergebnis rechnen. Einschränkend muss darauf hingewiesen werden, dass drei Monate zwischen den beiden Tests lagen und dass die Studierenden des Studienkollegs besonders gut auf die DSH vorbereitet waren. Außerdem haben 30 der 31 Kandidaten, welche sowohl an der DSH als auch am C-Test teilnahmen, die DSH bestanden, so dass die Frage nicht beantwortet werden kann, bei welchem C-Test-Ergebnis das Bestehen der DSH gefährdet wäre.



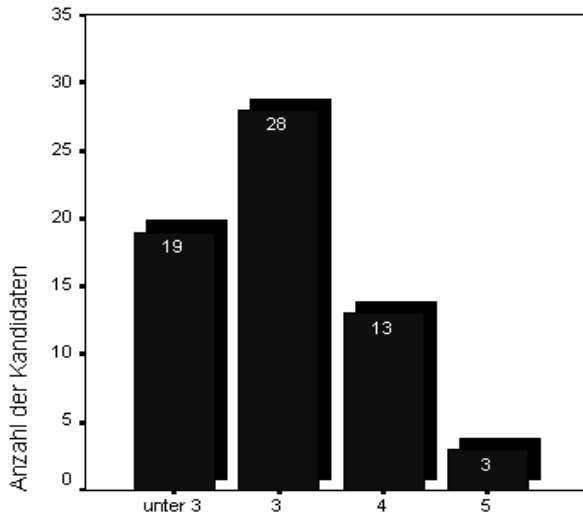
Abbildung 1: Ergebnisse aus der DSH und dem C-Tests im einem Streudiagramm mit linearer Regressionsgeraden (n=31)



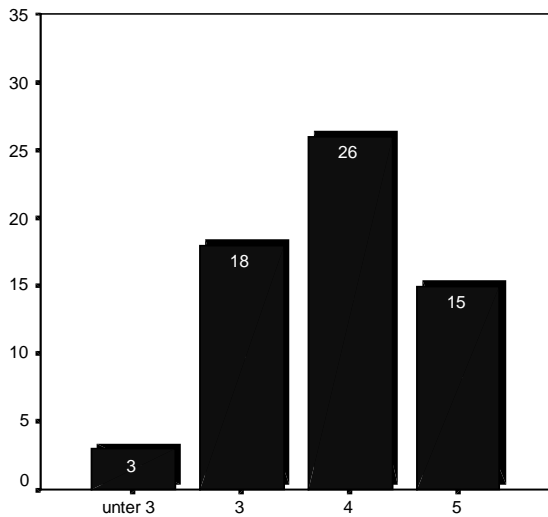
#### *TestDaF: TDN 5 selten*

Wie zu erwarten war, sind die mittleren Niveaustufen häufiger vertreten als die Extremwerte (vgl. Abbildung 2). Die TDN 5 ist unterschiedlich oft erzielt worden: Beim Mündlichen Ausdruck und beim Hörverstehen erzielte fast jeder vierte Kandidat die TDN 5 (24 %). Anders beim Leseverstehen. Dort erreichten nur 5 Prozent die höchste Niveaustufe. Das häufigste Ergebnis (Modus) war die TDN 3. Nur beim Hörverstehen wurde die TDN 4 am häufigsten erreicht. Da sich viele Kandidaten zum Testzeitpunkt bereits seit mehreren Monaten in Deutschland aufhielten, verwundern die relativ guten Leistungen in den Prüfungsteilen HV und MA nicht. Überraschend sind aber die Schwierigkeiten, welche die Kandidaten mit dem LV hatten.

Abbildung 2: Ergebnisse der einzelnen TestDaF-Subtests (n = Summe der Säulenwerte)



Leseverstehen, TDN



TestDaF Hörverstehen TDN

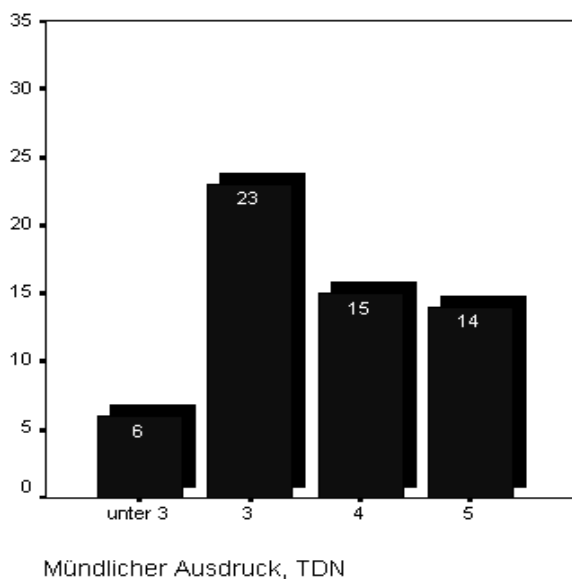
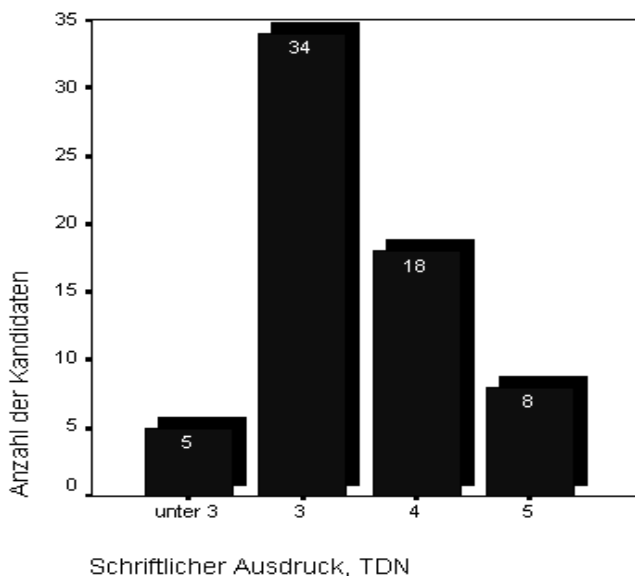


Tabelle 7 zeigt die Korrelationen zwischen den Subtests der DSH und des TestDaF. Die Korrelationen wurden mit den Rangkorrelationskoeffizienten von Spearman (untere

Dreiecksmatrix) bzw. Kendalls-Tau-b (obere Dreiecksmatrix) berechnet. Korrelationen sind ein Maß für die Übereinstimmung von Messwerten. Der Kendalls-Tau-Koeffizient gibt Auskunft über die Wahrscheinlichkeit, mit der die Reihenfolge zweier zufällig ausgewählter Kandidaten wie beobachtet eintrifft (vgl. Bortz/Döring 2002: 508-509). Die Werte liegen jeweils unter den Korrelationen von Spearman, sie folgen aber dem gleichen Muster.

Tabelle 7: Korrelationen zwischen den Tests der Studie (gesamte Stichprobe)

	DSH-LV	DSH-HV	DSH-TP	DSH-Grammatik 2	DSH-Grammatik 1	C-Test	Test-DaF LV	Test-DaF HV	Test-DaF SA	Test-DaF MA
<b>Kendall-Tau-b (<math>\tau</math>)</b>										
<b>DSH LV</b>		,456** n=231	,394** n=231	,395** n=231	,490** n=34	,433** n=31	,408** n=31	,492** n=31	,430** n=31	,339* n=29
<b>HV</b>	,627** n=231		,371** n=231	,267** n=231	,420** n=34	,317** n=31	,537** n=31	,439** n=31	,390** n=31	,568** n=29
<b>TP</b>	,539** n=231	,518** n=231		,397** n=231	-,411** n=34	-,571** n=31	-,561** n=31	-,378** n=31	-,360* n=31	-,320* n=30
<b>Grammatik 2</b>	,554** n=231	,384** n=231	,552** n=231		,534** n=34	,398** n=31	,283* n=31	,227 n=31	,356* n=31	,164 n=29
<b>Grammatik 1</b>	,644** n=34	,549** n=34	-,564** n=34	,697** n=34		,340** n=59	,295** n=58	,324** n=57	,449** n=60	,164 n=55
<b>C-Test</b>	,613** n=31	,437* n=31	-,713** n=31	,555** n=31	,477** n=59		,395** n=62	,439** n=61	,431** n=63	,415** n=57
<b>Test DaF-LV</b>	,582** n=31	,733** n=31	-,714** n=31	,403* n=31	,411** n=58	,543** n=62		,386** n=61	,357** n=63	,400** n=57
<b>HV</b>	,610** n=31	,609** n=31	-,513** n=31	,316 n=31	,459** n=57	,589** n=61	,523** n=61		,470** n=61	,594** n=56
<b>SA</b>	,525** n=31	,480** n=31	-,472** n=31	,436* n=31	,537** n=60	,533** n=63	,440** n=63	,571** n=61		,437** n=57
<b>MA</b>	,423* n=29	,683** n=29	-,414* n=29	,213 n=29	,228 n=55	,526** n=57	,494** n=57	,718** n=56	,500** n=57	
<b>Spearman's rho (<math>r_s</math>)</b>										

\*\* Korrelation ist auf dem Niveau von 0,01 signifikant (2-seitig).

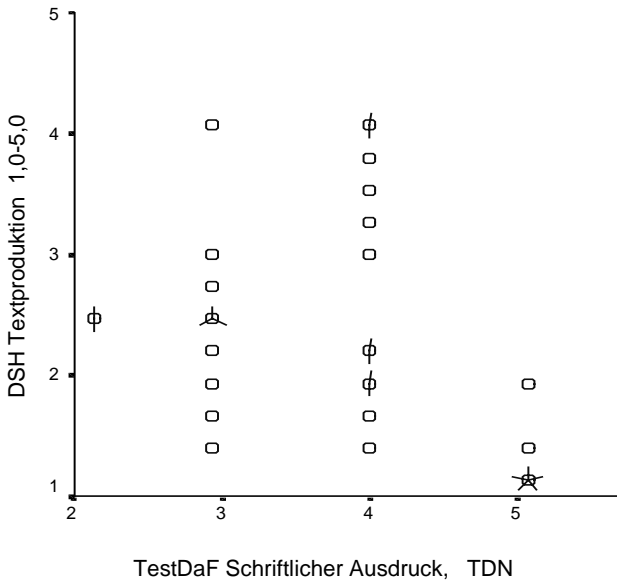
\* Korrelation ist auf dem Niveau von 0,05 signifikant (2-seitig).

### Subtests zum Schreiben messen ähnliche Fertigkeiten

In beiden Subtests zum Schreiben erzielten die Kandidaten gute Ergebnisse: Im TestDaF blieben nur fünf Kandidaten unter der TDN 3 und sieben der 31 Kandidaten, die auch an der DSH teilnahmen, erreichten die TDN 5. Das schlechteste Ergebnis der DSH war die Note 4,0, im Mittel wurde von den Kandidaten aus dem Studienkolleg 2,67 erzielt. Wie eng war der Zusammenhang zwischen den Ergebnissen in der DSH-TP und im TestDaF-SA? Die Korrelation zwischen den Subtests zum Schreiben liegt auf einem mittleren Niveau

( $r_s = -0,472$ ; vgl. Tabelle 7). Eine hohe TDN im TestDaF-SA war in vielen Fällen auch ein Hinweis auf ein gutes Ergebnis in der DSH-Textproduktion. Besonders bei den Ergebnissen der Spitzengruppe ist ein enger Zusammenhang zu erkennen (vgl. Abbildung 3). Die besten fünf Kandidatinnen erzielten im TestDaF-SA die TDN 5 und in der DSH-Textproduktion eine 1,0. Mehrere Kandidaten haben in der DSH-TP ein Ergebnis erzielt, das über demjenigen im TestDaF liegen dürfte: Es gibt sieben Studierende, die am Anfang des Semesters im TestDaF-SA ein niedriges Ergebnis (TDN 3), in der DSH-Textproduktion am Ende des Semesters aber ein durchschnittliches oder gutes Ergebnis erzielten (2,5 oder besser). Allerdings konnten sich nicht alle Kandidaten verbessern: Zwei Studentinnen erreichten am Beginn des Semesters im Schriftlichen Ausdruck die TDN 4, in der DSH aber nur eine 4,0.

Abbildung 3: Ergebnisse aus der DSH Textproduktion und dem TestDaF Schriftlicher Ausdruck in einem Sonnenblumen-Streudiagramm (n=31)



(Anm.: In einem Sonnenblumen-Streudiagramm werden Wertepaare zu Gruppen zusammengefasst. Ein Kreis steht für ein Ergebnis, zusätzliche Striche jeweils für Kandidaten, die das gleiche Ergebnis erzielten.)

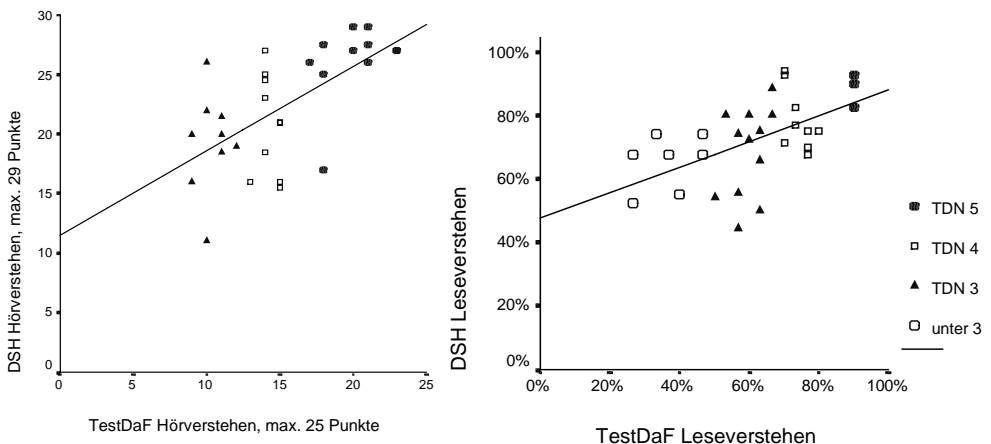
Die Aufgabenstellungen beider Subtests ähnelten sich, daher war ein Zusammenhang zwischen den Ergebnissen zu erwarten. Die Verbesserungen sind sicherlich mit dem Unterricht zu erklären. Wie konnte es zum Einbruch der zwei Studentinnen kommen? Möglicherweise ist der Fachbezug eine Ursache dafür. Im TestDaF wird ein Fachbezug vermieden, für die Textproduktion werden Themen gewählt, bei denen man erwartet,

dass alle Kandidaten sie unabhängig vom Studienziel und fachlichen Vorkenntnissen bearbeiten können. Thema des Subtests SA waren die Kosten für ein Studium, welche vor dem Hintergrund eines Schaubildes zu kommentieren waren. Im Subtest Textproduktion der vorliegenden DSH wird ein Fachbezug jedoch hergestellt. Die Kandidaten konnten aus mehreren Themen auswählen, z. B. ein Schaubild über Unternehmensfusionen kommentieren. Denkbar ist, dass die fachgebundene Aufgabe die beiden Kandidatinnen vor besondere Schwierigkeiten stellte.

### *Leseverstehen und Hörverstehen: Hohe Korrelationen zwischen TestDaF und DSH*

Der Subtest Leseverstehen des TestDaF besteht aus drei Texten mit insgesamt 30 Items (Zuordnungs-, Mehrfachauswahl- und Auswahlaufgaben). Der TestDaF-Subtest LV ist als schwierig empfunden worden, nur drei Kandidatinnen erzielten die TDN 5. Beim DSH-Leseverstehen konnten die Kandidaten einen aus drei Texten auswählen. Die Texte wurden von unterschiedlichen Prüfern erstellt und korrigiert. Die Ergebnisse liegen im Mittel knapp über dem Mittel des Gesamtergebnisses, wurden also nicht als besonders schwierig empfunden (LV:  $M = 2,84$ , DSH-Gesamtergebnis  $M = 2,57$ ). Wie aus dem Streudiagramm (Abbildung 4) hervorgeht, differenziert der Subtest DSH-LV vor allem im unteren Bereich nur wenig, der TestDaF differenziert deutlicher. Es gibt dennoch einen engen Zusammenhang zwischen den Ergebnissen der Subtests TestDaF-LV und DSH-LV ( $r_s = 0,582$ , vgl. Tabelle 7). Die Kandidaten konnten im DSH-LV mehr Items lösen als im TestDaF-LV, was z. B. an der geringen Steigung der Regressionsgeraden zu erkennen ist. Sie schneidet die y-Achse bei 45 % und nicht im Nullpunkt.

Abbildung 4: Ergebnisse DSH-HV und TestDaF-HV sowie DSH-LV und TestDaF-LV in einem Streudiagramm mit linearer Regressionsgeraden ( $n = 31$ )



Auf einen ähnlichen Sachverhalt trifft man beim Vergleich des DSH-Hörverstehens mit dem TestDaF HV: Die Kandidaten konnten in dem leichtesten Subtest der DSH drei Monate nach dem TestDaF mehr Items beantworten als im TestDaF (vgl. Tabelle 4). Die Korrelationskoeffizienten deuten auf einen engen Zusammenhang ( $r_s = 0,609$ ; vgl. Tabelle 7).

*TestDaF Mündlicher Ausdruck und Hörverstehen korrelieren hoch.*

Auffällig sind hohe Korrelationen zwischen dem TestDaF-Subtest Mündlicher Ausdruck und den Subtests Hörverstehen der DSH und des TestDaF (DSH-HV:  $r_s = 0,683$ ; TestDaF-HV:  $r_s = 0,718$ ). Möglicherweise ist dies ein Hinweis auf Ähnlichkeiten bei der Verarbeitung von gesprochener Sprache. Erwähnenswert sind außerdem die hohen Interkorrelationen zwischen den jeweiligen Subtests Hörverstehen und Leseverstehen, welche methodenspezifische Ursachen haben könnte (z. B. TestDaF-LV und DSH-HV:  $r_s = 0,733$ , vgl. Tabelle 7).

*DSH-Grammatiktest korreliert mit Subtests zum Schreiben und zum Lesen*

Die DSH-Grammatiktests weisen hohe Korrelationen mit Subtests zum Schreiben auf (TestDaF-SA:  $r_s = 0,537$ ; DSH-TP:  $r_s = 0,552$ ; vgl. Tabelle 7). Bei der Aufgabe, Sätze bzw. einen Text zu produzieren, werden vergleichbare Fertigkeiten geprüft wie bei der Aufgabe, kontextualisierte Sätze zu transformieren. Hier spiegeln sich auch die Leitlinien für die Bewertung des Schriftlichen Ausdrucks, bei der die sprachliche Korrektheit ein wichtiges Kriterium darstellt. Ein hoher Zusammenhang zwischen Grammatiktests und Tests zum Leseverstehen ist in mehreren Studien beobachtet worden (vgl. Alderson 1993, Bernhardt 1999, Clapham 1996). In der vorliegenden Untersuchung konnten die Beobachtungen nur teilweise bestätigt werden. Der DSH-Grammatiktest 2 korrelierte hoch mit dem DSH-Leseverstehen ( $r_s = 0,554$ ), die Korrelation des DSH-Grammatiktests 1 mit dem TestDaF-Leseverstehen war jedoch geringer ( $r_s = 0,411$ , vgl. Tabelle 7). Dieser Unterschied könnte seine Ursache auch in unterschiedlichen Formaten der Subtests zum Leseverstehen haben: Im TestDaF besteht er aus 30 Zuordnungs-, Mehrfachwahl- und Auswahlitems, in der DSH aus Fragen, die mit einem Stichwort oder in einem Satz beantwortet werden mussten. Es ist nicht auszuschließen, dass im Falle der DSH nicht nur das Textverständnis, sondern auch die Sprachproduktion in die Bewertung eingeflossen sind. Eine signifikante Korrelation des DSH-Grammatiktests mit dem TestDaF-Subtest Mündlicher Ausdruck konnte nicht beobachtet werden. Niedrig ist der Zusammenhang mit dem Subtest Hörverstehen ( $r_s = 0,384$ ; vgl. Tabelle 7). Das Konstrukt des Grammatiktests hat demnach mehr mit geschriebener als mit gesprochener Sprache zu tun. Dies war auch eine mögliche Interpretation von zwei Faktoranalysen, von denen eine mit den Subtests der DSH und eine mit dem TestDaF und dem DSH-Grammatiktest durchgeführt wurde. Bei einer Extraktion von zwei Faktoren erwies sich der Grammatiktest jeweils als bestimmende Größe für den zweiten, durch den Umgang mit Schriftsprache geprägten Faktor. Der erste Faktor, welcher über einen deutlich höheren Eigenwert verfügt, wird eher durch den Umgang mit gesprochener Sprache bestimmt.

### *Zusammenfassung*

Der C-Test korreliert mittel bis hoch mit den Subtests des TestDaF und hoch mit dem Gesamtergebnis der DSH, er dürfte sich daher als „Screening-Test“ sowohl für den TestDaF als auch für die DSH eignen. Er ist vor allem im oberen Leistungsspektrum sehr aussagekräftig. Obwohl TestDaF und DSH auf unterschiedlichen Skalen messen, wurde deutlich, dass den vergleichbaren Subtests (z. B. Textproduktion und Schriftlicher Ausdruck) vergleichbare Konstrukte zugrunde liegen. Dabei wurden der TestDaF-Subtest Mündliche Ausdruck und die Mündliche Prüfung der DSH nicht berücksichtigt. Deutliche Zusammenhänge konnten auch zwischen Subtests zum Hören und zum Lesen beobachtet werden, was methodenspezifische Ursachen haben könnte. Der DSH-Grammatiktest korreliert mit den Subtests zum Lesen und zum Schreiben. Gering bzw. nicht nachzuweisen waren Zusammenhänge des DSH-Grammatiktests mit Subtests zum Hören und zum Sprechen. Weitere Untersuchungen zur Austauschbarkeit vergleichbarer Subtests wurden nicht unternommen, da nicht die Verschiedenheit der Subtests, sondern die Verschiedenheit der Auswertungsmethoden die Vergleichbarkeit der Testergebnisse erschweren.

### **3.2. Ergebnisse der qualitativen Studie**

#### *TestDaF-Format erfreut sich hoher Akzeptanz.*

Die Kandidatinnen und Kandidaten haben sich insgesamt positiv über den TestDaF geäußert. Der Test und seine Teile schienen den Kandidaten angemessen und fair zu sein. Zwei Aspekte wurden häufiger angesprochen: die Bearbeitungszeit sowie die Auswahlaufgaben des Subtests Leseverstehen. Die Bearbeitungszeit für die Testteile LV und HV wurde als zu knapp empfunden. Dieses Problem kann durch eine Vorbereitung auf das Testformat möglicherweise verringert werden. Die Auswahlaufgaben beim LV, welche nicht eindeutig genug gewesen seien, waren ein weiterer Kritikpunkt. Der kassettengesteuerte Testteil MA wurde zwar thematisiert und als ungewöhnlich empfunden, aber nicht negativ bewertet. Ein Grammatiktest wurde nicht vermisst.

#### *Kandidaten empfanden DSH-Grammatiktest als schwierig.*

In den Gesprächen über die DSH wurde der Grammatiktest häufig angesprochen. Er wurde als verwirrend und als zu schwierig empfunden. Die Aufgaben in der Prüfung waren nach Meinung einiger Studierender komplexer als diejenigen im Unterricht. Einige Kandidaten bemerkten zu spät, dass sich auch auf der Rückseite des Aufgabenblattes Aufgaben befanden und gerieten daher in Zeitnot. Deutlich wurde: Der Grammatiktest ist ein Testteil, über den die Kandidaten sprechen, denn es geht um richtig oder falsch; man möchte die Ergebnisse vergleichen. Er spielt für das Ansehen der DSH bei den Kandidaten eine wichtige Rolle; der Grammatiktest lässt die DSH „schwierig“, aber im Gegensatz zum Leseverstehen oder Hörverstehen als berechenbar erscheinen. Das bedeutet meiner Ansicht nach allerdings nicht, dass ein Verzicht notwendigerweise zu einem Verlust an Augenscheinvalidität führen muss.



### *Auswirkungen unterschiedlicher Zulassungskriterien.*

Welche Auswirkungen das TestDaF-Ergebnis auf die Zulassung zum Studium bzw. auf die Befreiung von der DSH hat, hängt von den Maßstäben ab, die Anwendung finden. Die Auswirkungen der Ergebnisse bei unterschiedlichen Zulassungshürden sind in der Tabelle 8 dargestellt. Dabei sind auch Kandidaten berücksichtigt worden, bei denen ein Prüfungsteil fehlt. Sie wurden anhand der Tendenz der Ergebnisse aus den anderen Subtests eingeordnet.

Nur eine Kandidatin, welche sowohl am TestDaF als auch an der DSH teilnahm, bestand die DSH nicht. Die Libanesin erzielte in der DSH nur in der Textproduktion ein gutes Ergebnis (2,5), in den anderen Subtests fiel sie durch. Ihre Ergebnisse im TestDaF drei Monate vorher waren unterdurchschnittlich: Im Leseverstehen erzielte sie 17 von 30 Punkten (TDN 3), im Hörverstehen 10 von 25 (TDN 3), im Schriftlichen Ausdruck die TDN „unter 3“, im Mündlichen Ausdruck die TDN 3. Auffällig ist das Ergebnis im ersten DSH-Grammatiktest (0 Punkte), den sie wohl abgebrochen hat. Weitere Informationen über ihre Lernbiografie liegen mir nicht vor. Aus dem Bestehen der übrigen 30 Kandidaten zu folgern, ihre Ergebnisse im TestDaF seien mindestens äquivalent zum Bestehen der DSH, wäre unzutreffend. In Kapitel 2 wurde bereits auf Besonderheiten bei der Durchführung der Tests hingewiesen, welche zur Folge haben, dass die Ergebnisse der Stichprobe (im Vergleich zur Population) im TestDaF eher niedriger und in der DSH eher höher ausgefallen sein dürften.

### *Zulassung bei TDN 5-5-5-5?*

Wenn Studienbewerber mit einem TestDaF-Ergebnis von TDN 5 in allen Subtests von der DSH befreit werden, würde das auf drei der 66 Kandidatinnen und Kandidaten zutreffen, deren Sprachlernbiografie ich kurz vorstellen möchte:

- Die Studentin aus Bulgarien wurde wegen ihrer außergewöhnlich guten Leistungen im Aufnahmetest zum Studienkolleg direkt in das zweite Semester des Studienkollegs aufgenommen. Sie lernte in Bulgarien vier Jahre Deutsch, spricht auch Englisch und verfügt bereits über Arbeitserfahrungen. In Deutschland besuchte sie einen Intensivkurs, in dem sie ein Zeugnis über die Mittelstufe II erhielt. Ihre guten Deutschkenntnisse schreibt sie einer engagierten Lehrerin in Bulgarien zu, welche sie zum Lernen motiviert habe. Dies allein erklärt allerdings nicht, warum sie auch in allen anderen Fächern sehr gute Noten erzielte.
- Die zweite Studentin, welche das Ergebnis TDN 5 in allen Testteilen erzielte, kommt aus der Slowakei. Sie war zum Testzeitpunkt 21 Jahre alt, lernte Deutsch in der Schule, arbeitete vor dem Besuch des Studienkollegs ein Jahr als Aupairmädchen in einer deutschen Familie und nahm gleichzeitig an einem Oberstufenkurs an einer Sprachschule am Ort teil. Sie besuchte das zweite Semester des Studienkollegs, das sie mit sehr guten Noten abschloss.
- Auch eine Studentin aus Ungarn erfüllt das Kriterium. Sie besuchte in Ungarn eine zweisprachige Schule, ist Fremdenführerin und legte bereits die Zentrale Oberstu-

fenprüfung des Goethe-Instituts ab, könnte also bereits auf dieser Grundlage von der DSH befreit werden.

Tabelle 8: TestDaF – Auswirkungen unterschiedlicher Zulassungskriterien

<b>TDN 5 in allen Subtests (TDN 5-5-5-5)</b>	
<ul style="list-style-type: none"> <li>Hochschulrektorenkonferenz: „müssen zugelassen werden“ (Befreiung von DSH)</li> </ul>	
3 Kandidatinnen erfüllen das Kriterium:	5555, W-Kurs, 2. Sem., Slowakei, DSH 1,3
5555, W-Kurs, 2. Sem., Bulgarien, DSH 1,1	5555, W-Kurs, 2. Sem., Ungarn, DSH 1,5
<b>TDN 5 im LV und TDN 4 in allen anderen Subtests (TDN 5 LV 4-4-4)</b>	
<ul style="list-style-type: none"> <li>TestDaF-Institut: Zulassung für Fächergruppe Geistes-, Sprach-, Kulturwissenschaften (Befreiung von DSH)</li> </ul>	
Zusätzlich keine weiteren Kandidaten, insgesamt 3.	
<b>TDN 4 in allen Subtests (TDN 4-4-4-4)</b>	
<ul style="list-style-type: none"> <li>TestDaF-Institut: Zulassung für Fächergruppe Jura, Sozial-, Wirtschaftswissenschaften (Befreiung von DSH)</li> <li>Hochschulrektorenkonferenz: „sollen zugelassen werden“</li> </ul>	
Zusätzlich 7 Kandidaten, insgesamt 10:	4555, W-Kurs, 2. Sem., Tschechische Rep., DSH 1,6
4555, T-Kurs, 2. Sem., Slowakei, DSH 1,1	4555, W-Kurs, 2. Sem., Slowakei, DSH 1,8
4555, W-Kurs, 2. Sem., Ungarn, DSH 1,8	4545, W-Kurs, 1. Sem., Russland
4555, W-Kurs, 2. Sem., Slowakei, DSH 2,4	4545, W-Kurs, 1. Sem., Peru
<b>Durchschnitt TDN 4 (TDN 4-4-4-4, TDN 3-3-5-5, TDN 3-4-4-5, nicht „unter 3“)</b>	
<ul style="list-style-type: none"> <li>Vorschlag, der im Text diskutiert wird (Befreiung von der DSH)</li> </ul>	
Zusätzlich 3 Kandidaten, insgesamt 13:	4543, W-Kurs, 2. Sem., Indonesien, DSH 2,0
3535, T-Kurs, 2. Sem., Irak, DSH 1,9	3545, W-Kurs, 1. Sem., Slowakei
<b>Durchschnitt 3,5 (TDN 3-3-4-4; TDN 3-3-3-5, nicht „unter 3“)</b>	
<ul style="list-style-type: none"> <li>Vorschlag, der im Text diskutiert wird.</li> </ul>	
Zusätzlich 12 Kandidaten, insgesamt 25:	
4434, T-Kurs, 2. Sem., Iran, DSH 3,2	3434, T-Kurs, 1. Sem., Kenia
3443, T-Kurs, 2. Sem., Mauretanien, DSH 2,5	3434, W-Kurs, 1. Sem., Kenia
3443, T-Kurs, 2. Sem., Korea, DSH 2,1	3434, W-Kurs, 1. Sem., Mexiko
4433, T-Kurs, 2. Sem., Vietnam, DSH 2,4	3444, T-Kurs, 1. Sem., Kenia
3344, T-Kurs, 2. Sem., Indonesien, DSH 2,1	3543, W-Kurs, 1. Sem., VR China
3435, W-Kurs, 1. Sem., Russland	4434, W-Kurs, 1. Sem., Südafrika
<b>TDN 3 in allen Subtests (TDN 3-3-3-3)</b>	
<ul style="list-style-type: none"> <li>TestDaF-Institut: Zulassung für Fächergruppe Informatik, Ingenieurwissenschaften, Mathematik, Kunst und Musik, Naturwissenschaften (Befreiung von DSH)</li> <li>Hochschulrektorenkonferenz: „können zugelassen werden“</li> </ul>	
Zusätzlich 12 Kandidaten, insgesamt 37:	
3433, T-Kurs, 2. Sem., Indonesien, DSH 2,3	343-, T-Kurs, 1. Sem., Ghana
344-, W-Kurs, 2. Sem., Peru, DSH 2,7	3433, W-Kurs, 1. Sem., Kenia
344-, W-Kurs, 2. Sem., Litauen, DSH 3,4	3433, T-Kurs, 1. Sem., Marokko
4333, T-Kurs, 2. Sem., VR China, DSH 2,0	3334, T-Kurs, 1. Sem., Marokko
3334, T-Kurs, 2. Sem., Syrien, DSH 2,9	3333, T-Kurs, 1. Sem., VR China
3333, T-Kurs, 2. Sem., Marokko, DSH 3,3	435-, T-Kurs, 1. Sem., Iran

**nicht TDN 3 in allen Subtests**

- TestDaF-Institut: keine Zulassung zum Studium

29 Kandidaten:	4332, W-Kurs, 1. Sem., VR China
3323, T-Kurs, 2. Sem., Libanon, DSH 4,2	333-, T-Kurs, 1. Sem., VR China
2443, T-Kurs, 2. Sem., Mauretanien, DSH 2,9	2434, W-Kurs, 1. Sem., Litauen
3424, T-Kurs, 2. Sem., Libanon, DSH 3,2	2423, T-Kurs, 1. Sem., Libanon
2545, W-Kurs, 2. Sem., Slowakei, DSH 4,0	2-33, T-Kurs, 1. Sem., VR China
2444, W-Kurs, 2. Sem., Senegal, DSH 3,3	2233, W-Kurs, 1. Sem., Rep. Kongo
2342, W-Kurs, 2. Sem., VR China, DSH 3,0	3232, T-Kurs, 1. Sem., Vietnam
3332, W-Kurs, 2. Sem., VR China, DSH 3,3	-3-4, T-Kurs, 1. Sem., Mongolei
2343, W-Kurs, 2. Sem., Vietnam, DSH 3,6	2-34, T-Kurs, 1. Sem., Marokko
2333, T-Kurs, 2. Sem., VR China, DSH 2,7	2444, T-Kurs, 1. Sem., Indonesien
2433, W-Kurs, 2. Sem., VR China, DSH 3,0	2433, T-Kurs, 1. Sem., Mauretanien
3232, W-Kurs, 1. Sem., VR China	2332, W-Kurs, 1. Sem., VR China
233-, T-Kurs, 1. Sem., VR China	--3-, T-Kurs, 1. Sem., Libanon
2323, T-Kurs, 1. Sem., Libanon	2433, T-Kurs, 1. Sem., Indonesien
2433, W-Kurs, 1. Sem., Libanon	--2-, W-Kurs, 1. Sem., Marokko

**Erläuterungen:**

vierstellige Nummer: TestDaF-Ergebnis als TDN (LV, HV, SA, MA), „unter 3“ = „2“, „-“ = Prüfungsteil fehlt;

T-Kurs = Vorbereitungskurs für technische Studiengänge; W-Kurs = Vorbereitungskurs für wirtschaftliche Studiengänge.

*Zulassung bei TDN 5 (LV) 4-4-4?*

Die differenzierte Rückmeldung des TestDaF ermöglicht auch unterschiedliche Interpretationen des Ergebnisses für unterschiedliche Zielgruppen. Das TestDaF-Institut regt an, für unterschiedliche Studienfächer unterschiedliche TestDaF-Ergebnisse vorauszusetzen (vgl. Tabellen 2 und 8). Für die Studiengänge Geistes-, Sprach-, und Kulturwissenschaften sollen Kandidaten von der DSH befreit werden, wenn sie im Leseverstehen die TDN 5 und in allen anderen Testteilen mindestens die TDN 4 erzielen. Für die vorliegende Stichprobe ergibt sich keine Veränderung zum Kriterium „TDN 5 in allen Subtests“, denn im Leseverstehen haben nur drei Kandidatinnen die TDN 5 erzielt. Diese erfüllen gleichzeitig aber auch das höhere Kriterium (vgl. Tabelle 8).

*Zulassung bei TDN 4-4-4-4?*

Kandidaten, die ein Studium der Wirtschaftswissenschaften oder der Rechtswissenschaften anstreben, müssten nach dem Vorschlag des TestDaF-Instituts in allen Testteilen mindestens die TDN 4 erzielen. Zehn der 66 Teilnehmer aus dem Studienkolleg erfüllen dieses Kriterium. Die Stichprobengröße der Untersuchung reicht nicht aus, um signifikante Aussagen über die Verteilung der Herkunftsländer machen zu können. Die Verteilung der Herkunftsländer ist jedoch so auffällig, dass sie erwähnt werden soll: Mit Bulgarien, der Slowakei (4), Russland, der Tschechischen Republik, Ungarn (2) und Peru sind fast nur Kandidaten aus Mittel- und Osteuropa vertreten. Nur der Student aus

der Slowakei besucht den Technikkurs, die acht Studentinnen besuchen den Wirtschaftskurs. Bei diesen Kandidaten stand das Bestehen der DSH drei Monate nach dem TestDaF nicht in Frage, sie gehörten zur sprachlichen Spitzengruppe des Studienkollegs, erzielten gute bis sehr gute Ergebnisse in der DSH. Auch zwei Studentinnen aus dem ersten Semester erfüllten das Kriterium:

- Die Studentin aus Peru erzielte zwei Mal die TDN 4 und zwei Mal die TDN 5. Nur im DSH-Grammatiktest ist ihr Ergebnis unterdurchschnittlich. In ihrem Heimatland besuchte sie in den zwei Jahren vor ihrer Ausreise nach Deutschland eine zweisprachige Schule. Ebenfalls in Peru erhielt sie das Kleine Deutsche Sprachdiplom (KDS). Sie absolvierte den TestDaF zu Beginn des ersten Semesters des Studienkollegs. Zu dem Zeitpunkt konnte sie bereits flüssig und verständlich sprechen. Auch in den anderen TestDaF-Subtests und im C-Test stellte sie ihre fortgeschrittenen Deutschkenntnisse unter Beweis. Mit Ausnahme des DSH-Grammatiktests lagen ihre Ergebnisse in den übrigen Testteilen durchweg deutlich über dem Durchschnitt. Im Gespräch gab sie an, dass sie das Testformat des DSH-Grammatiktests verwirrt habe und sie nicht gewusst habe, was von ihr verlangt wird. Die grammatischen Phänomene, die für die geforderten Transformationen notwendig sind, seien ihr eigentlich nicht fremd gewesen, was sie bei der Besprechung des Tests bemerkt habe. In ihrem Fall liefert ein unvorbereiteter DSH-Grammatiktest die zusätzliche Information, dass sie Schwierigkeiten mit dem Testformat hat. Schlussfolgerungen für die Studierfähigkeit lassen sich daraus nicht ziehen, denn ihrer Interpretation des Testergebnisses ist wahrscheinlich zuzustimmen. Sie müsste sich mit dem Testformat vertraut machen, dann würde sie auch in diesem Testteil ein besseres Ergebnis erzielen.
- Die Russin absolvierte in ihrem Heimatland bereits eine Ausbildung zur Deutschlehrerin und arbeitete vor dem Besuch des Studienkollegs ein Jahr als Aupair bei einer deutschen Familie. Das Studienkolleg besucht sie, um in Deutschland wirtschaftswissenschaftliche Fächer studieren zu können.

### *Zulassung bei TDN 3-3-3-3?*

Kandidaten, die Informatik, Ingenieurwissenschaften, Mathematik, Kunst, Musik oder Naturwissenschaften studieren möchten, sollen nach dem Vorschlag des TestDaF-Instituts zugelassen werden, wenn sie in allen Testteilen mindestens die TDN 3 erzielten. Dieses Kriterium wird von insgesamt 37 der 66 Kandidaten erfüllt, auch viele Studierende aus dem ersten Semester gehören dazu. Bei dieser Gruppe ist eine regionale Häufung nicht mehr zu beobachten.

Einige Lernbiografien sollen kurz skizziert werden:

- Der 21-jährige Chinese bringt sehr gute Leistungen aus dem Heimatland mit. Nach der Schule studierte er in China zwei Semester Chemie – was ihn allerdings nicht interessierte. In der Schule und während des Studiums lernte er Englisch, aber noch kein Deutsch. Seinen Wunsch, ein Studium in Großbritannien oder in den USA,

konnte er aus finanziellen Gründen nicht verwirklichen. Ein Jahr vor der Teilnahme an der Prüfung kam er ohne Deutschkenntnisse nach Deutschland und nahm an einem einjährigen Intensivkurs an einer privaten Sprachschule teil. Der Intensivkurs schloss mit einer Mittelstufenprüfung ab. Wegen seiner sehr guten Leistungen im (schriftlichen) Aufnahmetest zum Studienkolleg wurde er gleich in das zweite Semester des Studienkollegs aufgenommen. Der intelligente und sprachbegabte junge Mann kann auf einen strukturierten Deutscherwerb in Deutschland zurückblicken. Durch seine Englischkenntnisse ist er sprachlernerfahren. Dies scheint ihm allerdings beim Mündlichen Ausdruck und beim Hörverstehen nicht zu helfen (TDN 4-3-3-3). Beim DSH-Grammatiktest gehörte er zur Spitzengruppe (90 % richtige Lösungen,  $M = 64 \%$ ).

- Die Kandidatin aus Kenia belegte nach ihrem Schulabschluss Sprachkurse in Deutsch und Französisch in ihrem Heimatland. Ein halbes Jahr vor dem TestDaF hat sie am Goethe-Institut in Nairobi das Zertifikat Deutsch erworben. Vom TestDaF fühlte sie sich überrascht, sie hätte gerne mehr Zeit für die Bearbeitung gehabt. Mit dem TestDaF-Ergebnis von TDN 3-4-3-3 könnte sie nach den Vorstellungen des TestDaF-Instituts nur zu einem technischen Fachstudium zugelassen werden.
- Der Kandidat aus Indonesien hätte das Kriterium TDN 3-4-3-3 ebenfalls erfüllt. Die DSH am Semesterende bestand er mit einer guten Note (Schriftliche Prüfung: 2,3). Er lernte Deutsch erst in Deutschland, wo er das Zertifikat Deutsch und ein Zeugnis der Mittelstufe I erwarb.
- Der Marokkaner aus dem ersten Semester des Technikkurses erfüllt das Kriterium ebenfalls mit dem Ergebnis TDN 3-4-3-3. Erst nach dem Besuch des technischen Gymnasiums lernte er Deutsch im Heimatland.

Dieses Kriterium wird nicht nur von Kandidaten erfüllt, die sich schon seit längerer Zeit in Deutschland aufhalten oder die im Heimatland viele Jahre Deutsch lernten. Auch ausländische Studienbewerber, die Deutsch erst als Erwachsene nach ihrem Schulabschluss als Vorbereitung auf das Studium in Deutschland erworben haben, gehören zu der Gruppe. In der DSH drei Monate nach dem TestDaF erzielten sie schlechtestenfalls befriedigende Noten.

#### *Zulassung bei durchschnittlich TDN 4?*

Bevor eine abschließende Wertung vorgenommen wird, sollen noch zwei weitere Varianten zur Diskussion gestellt werden. Bei diesen Varianten ist der Durchschnitt der Niveaustufen das Ausschlag gebende Kriterium. Dann besteht die Möglichkeit, ein niedriges Ergebnis in einem Testteil durch ein hohes Ergebnis in einem anderen auszugleichen. Bei der Berechnung des Durchschnitts sollen allerdings nur die TDN 3, 4 und 5 berücksichtigt werden. Wenn in einem Subtest „unter 3“ erzielt wurde, wird das Kriterium nicht erfüllt.

Zur ersten Variante: Wenn Kandidaten von der DSH befreit würden, welche im Durchschnitt eine TDN 4 erzielten, würde sich die Zahl im Vergleich mit dem Kriterium

„mindestens TDN 4 in allen Subtests“ um drei Kandidaten auf insgesamt dreizehn (von 66) erweitern.

- Der Iraker war zum Testzeitpunkt 20 Jahre alt. Nach seinem Schulabschluss im Heimatland ist er im Alter von 18 Jahren nach Deutschland geflohen, wo ihm Asyl gewährt wurde. Zu Zeitpunkt der Ausreise beschränkten sich seine Fremdsprachenkenntnisse auf geringe Englischkenntnisse. Achtzehn Monate nahm er an berufsvorbereitenden Maßnahmen teil, die einen Deutschunterricht mit EDV-Elementen beinhalteten. Anschließend besuchte er eine Sprachschule in Mannheim. Dort erhielt er nach drei Monaten ein Zeugnis über die Grundstufe DaF und weitere drei Monate später ein Mittelstufenzeugnis. Im Studienkolleg bestand der Student das erste Semester mit einer sehr guten Deutschnote. Schwierigkeiten hat er mit dem DSH-Grammatiktest. Nach zwei Jahren in Deutschland und nach zwei Jahren der Teilnahme an Deutschunterricht verschiedener Ausprägung fällt ihm die mündliche Kommunikation offensichtlich leichter als der Umgang mit anspruchsvollen Strukturen. Er bestand die DSH drei Monate im Anschluss an den TestDaF mit der Gesamtnote 1,9.
- Die Studentin aus der Slowakei hält sich bereits seit mehreren Jahren in Deutschland auf. Ihrer eigenen Einschätzung nach ist vor allem die inhaltliche Vorbereitung auf das Studium im Rahmen des Studienkollegs wichtig für sie. Ihre Deutschkenntnisse muss sie lediglich in Bezug auf fachbezogene Sprachverwendungssituationen erweitern.
- Der Student aus Indonesien kann bereits auf eine lange Deutschlernbiografie in seinem Heimatland zurückblicken. Mit dem Studienziel Wirtschaftsinformatik interessiert er sich nach eigener Auskunft allerdings mehr für Computer als für Sprachen. Der Fremdspracherwerb war nie sein Hauptanliegen, sondern lediglich Mittel zum Zweck. Da er nach eigener Auskunft für andere Prüfungsteile in der Feststellungsprüfung nicht viel üben musste, konnte er sich auf die DSH konzentrieren und erzielte ein gutes Ergebnis (2,0).

#### *Zulassung bei durchschnittlich TDN 3,5?*

Wenn man den Durchschnitt auf 3,5 verringert, würden insgesamt 19 der 66 Kandidaten das Kriterium erfüllen, acht mehr als beim Kriterium „Durchschnitt 4“. Auch hier gilt wieder, dass in allen Subtests mindestens die TDN 3 erzielt werden muss. Wer erreicht das Kriterium „durchschnittlich TDN 3,5“?

- Insgesamt zehn von 35 Studierenden aus dem ersten Semester erfüllen dieses Kriterium bereits.
- Auffällig ist auch eine große Gruppe aus dem zweiten Semester des Technikkurses. Während die Spitzengruppe des Wirtschaftskurses die TDN 4 in allen Subtests oder besser erzielt, erfüllen die guten Teilnehmer des Technikkurses lediglich diese Hürde.

- Außerdem beachtenswert ist die Verteilung der Herkunftsländer bzw. -sprachen: Es gibt einzelne Studierende aus China, Korea und Vietnam, welche das Kriterium erfüllen. Übrigens erzielte keiner der siebzehn Studierenden aus Ostasien im TestDaF-Subtest Mündlicher Ausdruck die TDN 4 oder TDN 5.

Diese Zulassungshürde würde den Personenkreis stark erweitern. Studierende mit sehr unterschiedlichen Lernbiografien und Ausgangssprachen sind in der Lage, dieses Kriterium zu erfüllen.

#### **4. Empfehlungen**

*Externen Kandidaten ist der TestDaF zu empfehlen.*

Der TestDaF hat sich in dieser Studie als fairer Test der Sprachkenntnisse erwiesen, der sich einer hohen Akzeptanz von Seiten der Prüfungsteilnehmerinnen und -teilnehmer erfreut. Vor allem Kandidaten, die nicht an institutseigenen Vorbereitungskursen teilnehmen oder die sich im Heimatland auf das Studium in Deutschland vorbereiten, ist die Teilnahme am TestDaF zu empfehlen. Der Vorteil der DSH kommt vor allem zum Tragen, wenn die Kandidaten an einem bestimmten Institut auf ein bestimmtes Studium vorbereitet werden. Die DSH wird jedoch auch von Kandidaten absolviert, die nicht an der hochschuleigenen Vorbereitung teilgenommen haben und die auch kein Studium an der betreffenden Hochschule anstreben. Die internen Kandidaten erzielten in unserer Studie bessere Ergebnisse in der DSH als die externen Kandidaten. Die internen Kandidaten waren mit dem Format der Prüfung vertraut und kannten die Leseerwartungen der Prüfer. Die externen Kandidaten verfügten (laut Prüfungsergebnis) über Deutschkenntnisse auf niedrigerem Niveau, möglicherweise waren sie auf die Prüfung nicht so gut vorbereitet wie die internen Kandidaten. Vor allem im Prüfungsteil Textproduktion waren die Leistungen deutlich niedriger, was an dem Fachbezug dieses Subtests liegen kann. Die Vermutung liegt nahe, dass sie bei einer dezentralen Prüfung, welche von Prüfern in Anlehnung an den Unterricht gestaltet wird, benachteiligt sind.

*Keine Zulassung bei TDN „unter 3“ in einem Subtest.*

Es scheint ein Konsens vorzuliegen, dass Studienbewerber nicht von der DSH zu befreien sind, wenn sie in einem Subtest nicht die TDN 3 erzielt haben („unter 3“). Ich halte diesen Konsens für unstrittig, denn die Einschätzung „unter 3“ beschreibt ein breites Spektrum an Leistungen. Wegen der fehlenden Differenzierung auf dieser Stufe kann die Einstufung „unter 3“ nicht als Basis für eine Entscheidung gelten. Im Rahmen der Untersuchung haben freilich neun der zehn Kandidaten, welche im TestDaF in einem Prüfungsteil nicht die TDN 3 erzielten, die DSH mit Ergebnissen im Bereich befriedigend bis ausreichend bestanden. Allerdings hatte der zeitliche Abstand zwischen den Prüfungen wohl zur Folge, dass die Leistungen der Kandidaten sich verbesserten.

*Allgemeine Zulassung zum Fachstudium bei durchschnittlich TDN 4 im TestDaF.*

Bei welchem TestDaF-Ergebnis sollen Studienbewerber von der DSH befreit werden bzw. zum Fachstudium zugelassen werden? Das Kriterium „TDN 5 in allen Subtests“ erscheint vor dem Hintergrund dieser Studie als zu hoch. In der vorliegenden Untersuchung hätten lediglich drei äußerst sprachbegabte und auch in den anderen Fächern überdurchschnittlich leistungsstarke Studentinnen das Kriterium erfüllt. Das entspricht fünf Prozent einer Stichprobe, zu der zweifelsohne mehr Kandidaten gehören, welche zum Zeitpunkt der Durchführung als studierfähig zu bezeichnen waren. Eine Erweiterung des Kriteriums auf „TDN 5 im Leseverstehen und TDN 4 in den anderen Subtests“ hatte bei der vorliegenden Stichprobe keine Auswirkungen, was an der niedrigen Anzahl derjenigen liegt, welche im LV die TDN 5 erzielten (drei von 66 Kandidaten). Es steht fest: Diese Kriterien liegen deutlich über „DSH bestanden“.

Das Kriterium „TDN 4 in allen Subtests“ erzielten insgesamt zehn Kandidaten. Es wurde (mit Ausnahmen) von sprachgewandten Osteuropäerinnen erreicht, die sich bereits lange in Deutschland aufhalten und/oder über eine gründliche Ausbildung aus ihrem Heimatland verfügen. Selbst unter ausdrücklicher Berücksichtigung der in Kapitel 2 genannten Einschränkungen kann mit großer Sicherheit behauptet werden, dass dieses Kriterium über dem Ergebnis „DSH bestanden“ liegt.

Eine Anerkennung von TestDaF-Zeugnissen mit „durchschnittlich TDN 4“ würde die Zielgruppe nur leicht verändern, aber Kandidaten eine Zulassung ermöglichen, die eine TDN 3 in einem Subtest erzielten und diese durch eine TDN 5 in einem anderen Subtest ausgleichen. Dies dürfte in Ausnahmefällen zutreffen. Der Personenkreis erweitert sich um Kandidaten, bei denen es ebenfalls kaum Zweifel über die ausreichenden Deutschkenntnisse gibt. Auch Studienbewerber, bei denen der Fremdspracherwerb nicht Mittelpunkt ihrer bisherigen Ausbildung darstellte, gehören zu dieser Gruppe. Sie dürften in der Lage sein, ihr Studium zu absolvieren, benötigen aber die Chance, TDN 3 in einem Prüfungsteil auszugleichen. Dies trifft beispielsweise auf Studierende zu, welche aufgrund ihrer Herkunftssprache Schwierigkeiten mit der Aussprache des Deutschen haben.

*Zulassung bei durchschnittlich TDN 3,5 kann sinnvoll sein.*

Ist ein Drittel bis ein Viertel der Studierenden aus dem ersten Semester des Studienkollegs bereits studierfähig? Eine Gruppe dieser Größenordnung könnte zum Studium zugelassen werden, wenn das Kriterium „durchschnittlich TDN 3,5“ lauten würde. Ob ausländische Studienbewerber ein Studienkolleg besuchen müssen, hängt nicht von den Sprachkenntnissen, sondern von der Vorbildung im Heimatland ab. Ich halte es daher für möglich, dass viele Studierende des Studienkollegs bereits im ersten Semester die sprachlichen Zulassungshürden erfüllen können. Die DSH-Ergebnisse der Kandidaten aus dem zweiten Semester waren schlechtestenfalls befriedigend (3,2), das Bestehen der DSH stand bei keinem Kandidaten in Frage. Schließlich ist zu bedenken, dass selbst die sprachlich fortgeschrittensten Kandidaten aus China, Vietnam oder Korea ein höheres



Kriterium nicht erzielen. Vor dem Hintergrund dieser Untersuchung spricht einiges dafür, Studierende bereits mit einem TestDaF-Zeugnis von durchschnittlich TDN 3,5 zum Fachstudium zuzulassen. Unter Beachtung der oben erwähnten Einschränkungen (vgl. Kapitel 2), gibt diese Studie wenig gesicherte Hinweise für eine weitere Öffnung der Zulassungskriterien. Die Ergebnissen der vorliegenden Studie schließen ein derartiges Ergebnis nicht aus. Weitere Studien sind nötig, um festzustellen, ob und unter welchen Umständen Kandidaten mit dem Ergebnis „TDN 3 in allen Subtests“ von der DSH befreit werden sollten. Dabei sollte es sich nicht nur um Vergleichsstudien, sondern auch um qualitative Langzeitstudien handeln.

### *Zulassung am TestDaF orientieren, nicht an der DSH.*

Die Frage nach der Äquivalenz der Prüfungsergebnisse ist für Kandidaten und Zulassungsstellen relevant. Das eindeutige Ergebnis der DSH kann problemlos in das Zulassungsverfahren Eingang finden. Der TestDaF ist testmethodisch ausgereift und mit der DSH in diesem Aspekt nicht zu vergleichen. Der Umgang mit dem Ergebnis ist aber weniger eindeutig, daher auch anspruchsvoller. Möglicherweise ist der TestDaF mit der starken Differenzierung im oberen Leistungsspektrum über das Ziel hinausgeschossen: Diese Differenzierung mag als Sprachstandsdiagnose in einigen Fällen interessant sein, als Hochschulzugangsprüfung ist sie überflüssig. Dennoch sollte sich eine Beschreibung der sprachlichen Zulassungsvoraussetzungen in Zukunft am standardisierten TestDaF ausrichten und nicht an dem „Prüfungsgerüst“ DSH. Das hätte zur Folge, dass der Nachweis der Sprachkenntnisse verlässlicher und damit auch gerechter wird. In Form von qualitativen Studien, die in dieser Untersuchung bereits in begrenztem Umfang durchgeführt wurden, könnten Erkenntnisse über einer Zulassungspraxis gewonnen werden, die auch die Differenzierungsmöglichkeiten des TestDaF nutzt.

### *Die Stärken beider Tests nutzen.*

Welche Rollen können die Tests in Zukunft übernehmen? Die DSH ist in Bezug auf die Gütekriterien dem TestDaF unterlegen. Obwohl eine Standardisierung der DSH wünschenswert wäre, kann eine Erhöhung der Qualität allein die Zukunft der DSH nicht sichern. Zu den Stärken der DSH gehört die Durchführung durch Prüfer, welche die Kandidaten kennen und mit den „Abnehmern“ in engem Kontakt stehen. Der DSH-Grammatiktest, den man ohne gezielte Vorbereitung kaum bewältigen kann, ist ein Hinweis darauf, dass die DSH wie eine Kursabschlussprüfung eingesetzt wird – allerdings ohne dass alle Kandidaten an einem spezifischen Vorbereitungskurs teilnehmen. Dies ist für eine Sprachstandsprüfung zwar ungewöhnlich, bietet aber auch Chancen in Bezug auf eine besonders zielgruppengenaue Gestaltung. Die DSH ist nicht isoliert von der hochschulnahen Studienvorbereitung zu bewerten. Als flexibles Instrument zur Förderung der Eigenständigkeit, mit dem institutionelle Besonderheiten und besondere Zielgruppen berücksichtigt werden können, hat die DSH weiterhin eine Berechtigung. Der TestDaF

macht den Studienzugang für ausländische Studienbewerber offener, transparenter und fairer. Eine Einbindung in studienvorbereitende Maßnahmen steht noch aus.

## 5. Literatur

- Alderson, J. Charles 1993. The Relationship Between Grammar and Reading in an English for Academic Purposes Test Battery. In Douglas, Dan; Chapelle, Carol (Hrsg.). *A New Decade of Language Testing Research: Selected Papers From the 1990 Language Testing Research Colloquium*. Alexandria, Virginia: Teachers of English to Speakers of Other Languages (TESOL), 203-219.
- Alderson, J. Charles 2002. *Criteria for test comparability studies*. Vortrag auf der Fachtagung „DSH und TestDaF: Politische Implikationen und wissenschaftliche Erforschung“ vom 22.-23.6.2002 in Braunschweig.
- Arras, Ulrike; Eckes, Thomas; Grotjahn, Rüdiger 2002. C-Tests im Rahmen des ‘Test Deutsch als Fremdsprache’ (TestDaF): Erste Forschungsergebnisse. In Grotjahn, Rüdiger (Hrsg.). *Der C-Test. Theoretische Grundlagen und praktische Anwendungen* (Bd. 4). Bochum: AKS-Verlag, S. 175-209.
- Bernhardt, Elizabeth 1999. If reading is reader-based, can there be a computer adaptive test of reading? In Chalhoub-Deville, Micheline (Hrsg.). *Issues in Computer-adaptive Testing of Reading Proficiency*. Cambridge: Cambridge University Press, 1-10.
- Bortz, Jürgen; Döring, Nicola 2002<sup>3</sup>. *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin et al.: Springer.
- Clapham, Caroline 1996. *The development of IELTS: a study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.
- Fachverband Deutsch als Fremdsprache (FaDaF) 2001. *DSH-Handbuch für Prüferinnen und Prüfer*. Münster: Fachverband Deutsch als Fremdsprache.
- Grotjahn, Rüdiger 1995. Der C-Test: State of the Art. *Zeitschrift für Fremdsprachenforschung (ZFF)*, 6 (2), 37-60.
- Grotjahn, Rüdiger 1997. Der C-Test: Neuere Entwicklungen. In Gardenghi, Monica, O'Connell, Mary (Hrsg.). *Prüfen, Testen, Bewerten im modernen Fremdsprachenunterricht*. (Bayreuther Beiträge zur Glottodidaktik 8.) Frankfurt am Main et al.: Lang, 117-128.
- Grotjahn, Rüdiger 2000. Determinanten der Schwierigkeit von Leseverstehensaufgaben: Theoretische Grundlagen und Konsequenzen für die Entwicklung des TestDaF. In Bolton, Sybille (Hrsg.). *TestDaF: Grundlagen für die Entwicklung eines neuen Sprachtests. Beiträge aus einem Expertenseminar*. (Standpunkte zur Sprach- und Kulturvermittlung 8, Werkstattberichte des Goethe-Instituts.) München: Goethe-Institut; Köln: Gilde-Verlag, 7-56.
- Grotjahn, Rüdiger; Kleppin, Karin 2001. TestDaF: Stand der Entwicklung und einige Perspektiven für Forschung und Praxis. In: Aguado, Karin; Riemer, Claudia (Hrsg.). *Wege und Ziele: Zur Theorie, Empirie und Praxis der Deutschen als Fremdsprache (und anderer Fremdsprachen)*. *Festschrift für Gert Henrici zum 60. Geburtstag*.

- (Perspektiven Deutsch als Fremdsprache 15.) Baltmannsweiler: Schneider Verlag Hohengehren, 419-434.
- Kenyon, Dorry M. 2000. Tape-mediated Oral Proficiency Testing: Considerations in Developing Simulated Oral Proficiency Interviews (SOPIs). In Bolton, Sybille (Hrsg.). *TestDaF: Grundlagen für die Entwicklung eines neuen Sprachtests. Beiträge aus einem Expertenseminar.* (Standpunkte zur Sprach- und Kulturvermittlung 8, Werkstattberichte des Goethe-Instituts.) München: Goethe-Institut; Köln: Gilde-Verlag, 87-106.
- Kniffka, Gabriele; Üstünsöz-Beurer, Dörthe 2001. TestDaF: Mündlicher Ausdruck. Zur Entwicklung eines kassettengesteuerten Tesformats. *Fremdsprachen Lehren und Lernen* 30, 127-149.
- Koreik, Uwe; Schimmel, Dagmar 2002. Hörverstehenstests bei der DSH, der Feststellungsprüfung und TestDaF – eine Vergleichsstudie mit weiterführenden Überlegungen zu TestDaF und DSH. *InfoDaF.* (In Druck)
- TestDaF-Institut 2001. *Informationen für Hochschulen in Deutschland.* Hagen: TestDaF-Institut/Gesamthochschule Hagen.
- TestDaF-Institut 2002. *Handreichungen zur Zulassung ausländischer Studierender mit TestDaF-Zeugnis.* (Faltblatt.) Hagen: TestDaF-Institut/Gesamthochschule Hagen.

Anschrift des Autors

Christian Krekeler

Fachhochschule Konstanz, Brauneggerstr. 55, D-78462 Konstanz,

Tel.: 07531 206-395, Fax: 07531 206-87395, Email: [krekeler@fh-konstanz.de](mailto:krekeler@fh-konstanz.de)

Thomas Niehr

## **International vergleichende Diskurs- und Argumentationsanalyse. Vorstellung eines Forschungsprogramms**

### **1. Einleitung**

Ich möchte an dieser Stelle ein Projekt vorstellen, in dem es um einen Vergleich der Migrationsdiskurse in Deutschland, Österreich und der Schweiz geht – genauer gesagt: um einen Vergleich der Argumentationen innerhalb dieser Migrationsdiskurse.<sup>1</sup> Ich möchte also vergleichen, wie in den öffentlichen Diskursen zur Migration in Deutschland, Österreich und der Schweiz argumentiert wurde. Dies mag auf den ersten Blick vielleicht als ein wahnwitziges Unterfangen erscheinen, aber lassen Sie mich das Thema ein wenig genauer eingrenzen, damit das Ganze etwas plausibler wird.

### **2. Das zugrunde liegende Diskursmodell**

Zugrundelegen möchte ich ein Diskursmodell, das dem von Busse und Teubert (vgl. Busse/Teubert 1994) nahe kommt. Ohne auf Einzelheiten näher einzugehen, möchte ich hier nur kurz skizzieren, was ich unter Diskurs verstehe. Diskurs ist demnach ein Geflecht von thematisch zusammengehörigen Aussagen, die über Textkorpora zu erschließen sind. Diese Aussagen sind zu verstehen als „Gesprächsbeiträge“, als „Komponenten eines Zeitgesprächs“ (Hermanns 1995, S. 91), das es zu rekonstruieren gilt.

Quantitativ sind dabei die Grenzen prinzipiell nach oben wie nach unten verschiebbar. „Filmtechnisch-metaphorisch ausgedrückt ist der Begriff Diskurs ein Zoom: geeig-

---

<sup>1</sup> Dieser Text ist die schriftliche Fassung eines Vortrages, den ich am 14.11.2000 im Essener Linguistischen Kolloquium gehalten habe. Die am mündlichen Vortrag orientierte Form habe ich weitgehend beibehalten. Durch die Tatsache, dass der hier vorgestellte Diskursvergleich Gegenstand meiner an der Heinrich-Heine-Universität Düsseldorf entstandenen Habilitationsschrift ist, muss ich es an dieser Stelle bei eher pauschalen Bemerkungen zum Forschungsprogramm belassen und für Einzelheiten und detailliertere Ergebnisse auf meine demnächst erscheinende Habilitationsschrift verweisen. Ulrich Schmitz und Hermann Cölfen danke ich für die Gelegenheit, meinen Ansatz in ihrem Kreis vorzustellen und zu diskutieren.

net größte wie auch kleinste Mengen, Obermengen oder Untermengen, von thematisch-dialogisch-intertextuell verknüpften Texten zu bezeichnen.“ (Hermanns 1995, S. 89) Ob ich also den gesamten in der Bundesrepublik Deutschland geführten Vergangenheitsbewältigungs-Diskurs oder aber thematisch bzw. zeitlich eingegrenzte Bereiche (z.B. Historikerstreit, Goldhagen-Debatte, Walser-Bubis-Debatte, Wende-Zeit, 90er Jahre) untersuchen möchte, bleibt meiner Entscheidung überlassen.

Schon unter pragmatischen Gesichtspunkten empfiehlt es sich, vor Erstellung der Textkorpora, auf denen die Analyse basieren soll, Auswahlkriterien zu entwickeln, mit denen die potentielle Textmasse auf ein zu bewältigendes Maß reduziert werden kann, ohne dass grob gegen das Gebot der Repräsentativität verstoßen wird. Das gilt vor allem für internationale Diskursvergleiche, die mit Problemen wie der Unzugänglichkeit von Textmaterial bei Erstellung der Korpora und mit mangelnden Auswertungskapazitäten konfrontiert sind.

Um dementsprechend bearbeitungsfähige Korpora zu erhalten, ist eine Konzentration in vielerlei Hinsicht möglich und nötig.<sup>2</sup> Beschränkungen können bzw. müssen auf folgenden Ebenen erfolgen:

- Thema: z. B. Eingrenzung auf einen oder einige Diskurse
- Zeit: z. B. Auswahl bestimmter, als besonders signifikant erscheinender chronologischer Abschnitte
- Textsorte: z. B. Beschränkung auf bestimmte zugängliche Texte möglichst einheitlicher Art
- Analyseebene: z. B. Konzentration auf bestimmte sprachliche Teilaspekte, etwa Lexik, Metaphorik oder Argumentation.

Die Bestimmung der auszuwählenden Zeitabschnitte sollte ohne große Schwierigkeiten erfolgen, da sie sich an Eckdaten eines (Teil-)Diskurses oder aber an diskursexternen Faktoren (z. B. die Adenauer-Ära) orientieren kann. Die textsortenspezifische Beschränkung auf einen bestimmten Diskursbereich, z. B. den öffentlich-politischen Sprachgebrauch, kann bzw. muss aus forschungspraktischen Gründen durch eine Konzentration auf ausgewählte Printmedien noch verstärkt werden. Andererseits dürfen arbeitsökonomische Faktoren das Kriterium der Repräsentativität auch nicht ad absurdum führen. So hat Fritz Hermanns einmal deutlich gemacht, dass jedes Textkorpus, das einen Diskurs(-ausschnitt) repräsentieren soll, eine Gratwanderung zwischen ökonomisch Machbarem und wissenschaftlich Verantwortbarem ist, und dies gilt ganz besonders für den internationalen Diskursvergleich.

Das heißt nichts anderes, als dass im Zeitalter der Massenmedien fast immer eine Auswahl aus der Fülle der vorliegenden Texte zu treffen ist, es sei denn, es wird ein deutlich abgegrenzter, abgeschlossener und überschaubarer Teildiskurs zur Analyse herangezogen. Diese Auswahl wird jeweils zu begründen sein und stellt bereits im Vorfeld der Analyse ein hermeneutisches Verfahren dar. Die Auswahl muss vielfältigen Bedin-

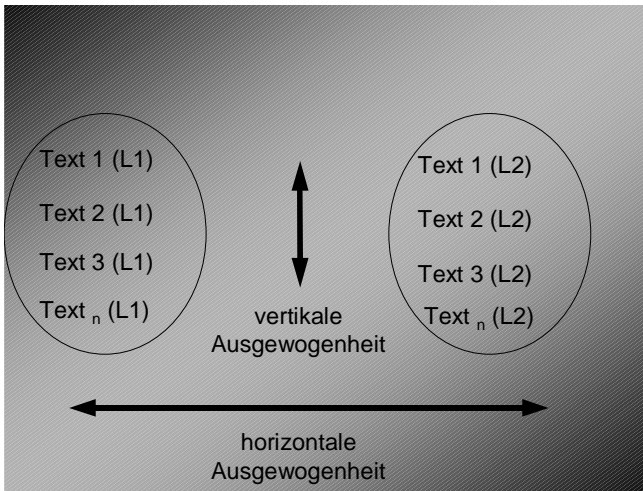
<sup>2</sup> Vgl. dazu ausführlich Böke/Jung/Wengeler/Niehr (2000:15ff).

gungen genügen, die alle in dem Begriff „Repräsentativität“ zusammengefasst sind. Ein repräsentatives Textkorpus bietet Gewähr dafür, dass weder wesentliche Diskurskomponenten fehlen, noch dass bestimmte Komponenten überbetont werden. Betrachtet man den zu analysierenden Teildiskurs als argumentative Auseinandersetzung über ein Thema, so bedeutet Repräsentativität nicht zuletzt die angemessene, sprich: richtig gewichtete Berücksichtigung aller argumentativen Positionen, die im Diskurs eingenommen worden sind.

Wie ist eine solche Repräsentativität zu erreichen? Besteht das Textkorpus aus massenmedialen Texten, z.B. aus Zeitungstexten, so ist darauf zu achten, dass möglichst das gesamte politische Meinungsspektrum abgedeckt wird. Natürlich wird man eine Auswahl treffen müssen, aber es sollte gewährleistet sein, dass die wichtigsten Meinungsbildenden Organe berücksichtigt werden, um sicherzugehen, dass nicht bestimmte politische Positionen von vornherein aus dem Blick geraten. Gerade auch bei quantitativen Analysen ist darauf zu achten, dass nicht durch das Übergewicht bestimmter Zeitungen deren politische Ausrichtung überproportional vertreten ist und auf diese Weise die Ergebnisse verfälscht werden. Freilich gilt hier wie bei allen hermeneutischen Verfahren, dass das Vorgehen nicht operationalisierbar ist, sondern notwendig ein interpretativer Akt bleiben muss.

Überträgt man diese Überlegungen auf die Zusammenstellung eines internationalen Textkorpus, so ergibt sich daraus die Folgerung, dass zunächst beide Textkorpora (ausgehend von einem bilateralen Vergleich) diesen Ausgewogenheits-Kriterien genügen müssen. Darüber hinaus müssen die Textkorpora miteinander sinnvoll vergleichbar sein. Dies bedeutet, dass sie nach ähnlichen Kriterien aufgebaut sein sollten. Wenig sinnvoll wäre beispielsweise ein Vergleich von Boulevardzeitungen des einen Landes mit ‚seriösen‘ Zeitungen eines anderen Landes. Ebenso wenig macht es Sinn, die eher linken Zeitungen eines Landes mit den politisch entgegengesetzten Zeitungen eines weiteren Landes zu vergleichen. Solche Vergleiche, in denen sich gleichzeitig zu viele Parameter ändern, die in der Analyse gesondert berücksichtigt werden müssten, führen zu Ergebnissen, die vielleicht etwas über die jeweiligen Medien aussagen, nicht aber zum Diskursvergleich dieser Länder beitragen können. Es müssen also für beide Länder in sich ausgewogene und zudem gleichartige Textkorpora zusammengestellt werden, um einen sinnvollen und aussagekräftigen Diskursvergleich durchführen zu können.

Abb. 1: Ausgewogene Textkorpora



Wie diese Grafik verdeutlichen soll, bedarf es sowohl vertikaler Ausgewogenheit innerhalb der Textkorpora als auch horizontaler Ausgewogenheit zwischen den zu vergleichenden Textkorpora. Die vertikale Ausgewogenheit sorgt dabei jeweils für eine repräsentative Widerspiegelung des Diskurses, die horizontale Ausgewogenheit für die Vergleichbarkeit der Textkorpora. Das in dieser Abbildung idealisierte Modell darf freilich nicht zu der Vorstellung verleiten, horizontale Ausgewogenheit meine eine Eins-zu-eins-Entsprechung der jeweiligen Medien. Abgesehen davon, dass eine solche nie zu erreichen sein wird, ist sie auch gar nicht erforderlich. Es reicht vielmehr – wie bereits oben erwähnt – eine annähernde *Gleichartigkeit* der in den Textkorpora vertretenen Medien, um die Voraussetzung für horizontale Ausgewogenheit zu erfüllen.

Natürlich sind Textkorpora nicht auf Zeitungstexte beschränkt. Grundsätzlich können alle Textsorten öffentlicher Diskurse verglichen werden. Neben Zeitungskorpora ist z.B. an Parlamentaria<sup>3</sup>, Parteiprogramme, amtliche Akten, literarische, philosophische und historische Schriften, wissenschaftliche Kontroversen wie auch an die in den elektronischen Medien gesendeten Beiträge (Talkshows, Interviews, Reportagen etc.) zu denken. Für jeden internationalen Vergleich gilt dabei jedoch das Gebot der vertikalen wie auch der horizontalen Ausgewogenheit.

Um diese Vielzahl von theoretischen Erwägungen etwas plastischer zu machen, möchte ich Ihnen jetzt schildern, welche Teildiskurse ich miteinander verglichen habe, wie ich bei der Zusammenstellung der Textkorpora konkret vorgegangen bin und natürlich, welche sprachlichen Phänomene ich näher untersucht habe.

<sup>3</sup> Vgl. dazu meine vergleichende Analyse von Parlamentsdebatten; Niehr 2000.

### **3. Das konkrete Korpus**

Es wurde von der Vorannahme ausgegangen, dass es sinnvoll sei, Diskursstücke zu vergleichen, die in Deutschland, Österreich und der Schweiz eine annähernd gleiche Bedeutsamkeit haben und hatten. Andernfalls hätte die Gefahr bestanden, eine inadäquate Vergleichsbasis zur Grundlage zu nehmen und somit zu Ergebnissen zu gelangen, die nicht auf realen Gegebenheiten basieren. Aufgrund dieser Vorüberlegungen schien es sinnvoll, innerhalb der Migrationsdiskurse den Gastarbeiter- und Asyldiskurs zum Vergleich heranzuziehen. Beide Themen werden bzw. wurden in allen drei Ländern kontrovers diskutiert. Die diesen Diskussionen zugrundeliegenden sachgeschichtlichen Hintergründe sind in Deutschland, der Schweiz und Österreich zwar nicht identisch, aber immerhin doch prinzipiell gleichartig, so dass ein Vergleich der drei Länder möglich scheint. Innerhalb dieser Diskurse, die ja in allen drei Ländern jeweils über viele Jahre geführt wurden, wurde folgende zeitliche Auswahl getroffen:

- Gastarbeiter-Diskurse: 1965 bis 1967 und 1972 bis 1974
- Asyldiskurs 1979 bis 1983.

Für den Gastarbeiter-Diskurs wurden diese Zeiträume ausgewählt, da einerseits eine relativ frühe Phase des Diskurses von Interesse schien, andererseits gerade der sachgeschichtliche Hintergrund des Jahres 1973 (Ölkrise) auf seinen sprachlichen Niederschlag im Gastarbeiter-Diskurs untersucht werden sollte. Dass die Ölkrise Deutschland, die Schweiz und Österreich gleichermaßen traf, ist offensichtlich; wie sich die daraus resultierende Konjunkturfalte und die durch sie bedingten Folgen für den Arbeitsmarkt im Einwanderungs-Diskurs niederschlugen, wird erst der diskursanalytische Vergleich zeigen können. Für den Asyldiskurs habe ich eine relativ frühe Phase ausgewählt, da die späteren Phasen (z.B. der Anfang der 90-er Jahre im Zusammenhang mit der Änderung des Art. 16 GG) von der Linguistik bereits relativ ausführlich erforscht worden sind.

Als nächstes ist die Zusammensetzung des Textkorpus von Interesse. Von den verschiedenen Korpora, die im Rahmen des DFG-Forschungsprojektes „Die Einwanderungsdiskussion im öffentlichen Sprachgebrauch nach 1945. Eine Diskursgeschichte im internationalen Kontext“ an der Heinrich-Heine-Universität Düsseldorf zusammengetragen wurden,<sup>4</sup> berücksichtigt meine Analyse eines: nämlich das rund 6.500 Artikel zum Thema „Migration“ umfassende Pressekorpus.

Systematisch ausgewertet wurden darin die bundesdeutschen Tageszeitungen „Frankfurter Rundschau“ (FR), „Frankfurter Allgemeine Zeitung“ (FAZ), „Süddeutsche Zeitung“ (SZ), „Die Welt“ (inkl. „Die Welt am Sonntag“), „Rheinische Post“ (RP) sowie die wöchentlich erscheinenden Publikationen „Die Zeit“ und „Der Spiegel“. Für den hier angestrebten Vergleich wurden aus diesem Korpus nur die Tageszeitungen und „Die Zeit“ berücksichtigt. Wählt man dann die entsprechenden Zeitabschnitte aus, so

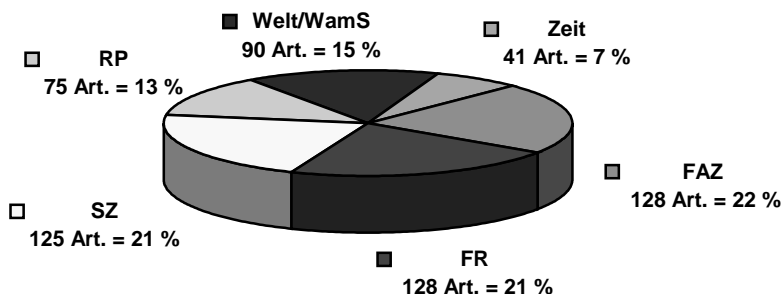
---

<sup>4</sup> Neben dem hier erwähnten Pressekorpus liegt beispielsweise noch ein Debattenkorpus mit relevanten Bundestagsdebatten vor.



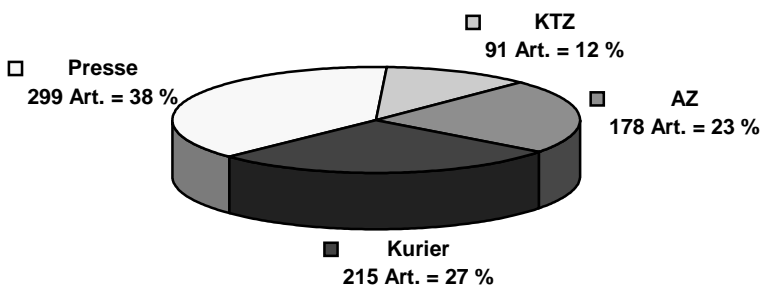
entsteht für die Bundesrepublik Deutschland ein Textkorpus aus 587 Artikeln, die sich wie folgt auf die einzelnen Presseorgane verteilen:

Abb. 2: Das bundesrepublikanische Textkorpus



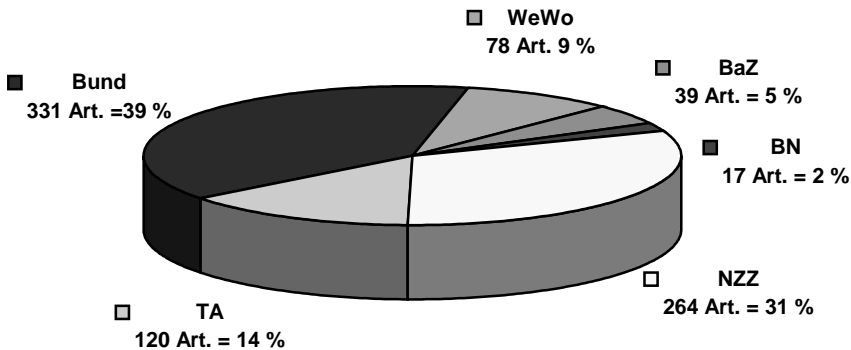
Diesem bundesdeutschen Korpus stehen ein österreichisches und ein Schweizer Textkorpus gegenüber, bei denen ich mich bemüht habe, ein ähnliches Verhältnis von regionalen und überregionalen, von bürgerlichen (wirtschaftsfreundlichen) und eher linken Organen herzustellen:

Abb. 3: Das österreichische Textkorpus



Diese insgesamt 783 Artikel wurden in verschiedenen thematisch sortierten Archiven in Wien und Klagenfurt kopiert. Entsprechend finden wir für die Schweiz ein Textkorpus, das aus insgesamt 849 Artikeln besteht, die sich wie folgt verteilen:

Abb. 4: Das Schweizer Textkorpus



Viele Schweizer Artikel wurden im Schweizer Sozialarchiv in Zürich, das über eine hervorragend sortierte Sammlung verfügt, kopiert. Ergänzt wurde dies durch die Bestände diverser Bibliotheken in Deutschland.

Natürlich stellen Sie sich jetzt zu Recht die Frage, wie man ein solches Korpus, das immerhin mehr als 2.200 Zeitungsartikel umfasst, linguistisch bearbeiten kann. Bei der Auswertung solcher Korpora denkt man ja üblicherweise eher an inhaltsanalytische, also quantitative Verfahren als an eine qualitativ linguistische Analyse. Und in der Tat tut man zunächst auch einmal gut daran, einige Anleihen bei der Inhaltsanalyse zu machen, zumindest was die technische Bearbeitung großer Textkorpora angeht. Doch dazu später.

#### 4. Die linguistische Auswertung der Textkorpora

Zunächst bin ich Ihnen noch Auskunft darüber schuldig, auf welche sprachlichen Phänomene ich denn mein Augenmerk gerichtet habe. Meine Vorannahme war, dass eine annähernd vollständige linguistische Analyse sowohl die Lexik, Argumentation als auch Metaphorik eines Diskurses zugrundelegen müsste. Nun kann man zeigen – das kann ich hier in diesem Rahmen nicht – dass tief greifende Veränderungen auf einer dieser Ebenen nicht ohne Veränderungen auf den jeweils anderen Ebenen stattfinden. Mit anderen Worten: Wenn sich in einem Diskurs entweder Metaphorik, Argumentation oder Lexik tief greifend verändern, dann kann man davon ausgehen, dass auch Veränderungen auf den jeweils anderen Ebenen stattfinden. Denn diese Veränderungen finden ja aufgrund einer veränderten Sichtweise statt, die sich dann sowohl in der Lexik, Metaphorik als auch Argumentation widerspiegelt. Betrachten Sie beispielsweise Ausdrücke wie *Ströme*, *Flut*, *eindämmen*: Eine solche Metaphorik gibt es nur in einem tendenziell aus-

länderkritischen Diskurs; ändert sich Haltung den Migranten gegenüber, so muss sich auch die Metaphorik (und damit die Argumentation) ändern.<sup>5</sup> Entsprechende Rückschlüsse können Sie aus der verwendeten Lexik ziehen, wenn ein Diskursteilnehmer von *Fremdarbeitern*, *Gastarbeitern* oder etwa *ausländischen Mitbürgern* spricht.

Man kann also den Schluss ziehen, dass die Analyse nur einer dieser Ebenen ausreicht, um weitergehende sprachliche Veränderungen in einem Diskurs zu erkennen. Ich habe mich nun bei meiner Analyse auf die Argumentation innerhalb der Migrationsdiskurse konzentriert. Deshalb möchte ich zunächst der Frage nachgehen, wie man überhaupt die Argumentation eines Diskurses beschreiben kann. Ist es nicht so, dass es prinzipiell eine unbeschränkte Zahl von Argumenten gibt, so dass ein Vergleich ohnehin zum Scheitern verurteilt ist? Dies ist zumindest ein Argument, das der Erforschung von Einzelargumenten entgegen gehalten wird. Doch m. E. lässt sich dieses Argument auch entkräften. Denn die These, nach der sich konkrete Argumente wegen der prinzipiellen Unendlichkeit ihrer Zahl nicht vergleichen lassen, überzeugt in der Tat nur für den Fall, dass man versuchen wollte, Inkompatibles zu vergleichen. Die Argumente jeweils eines Diskurses (wie z. B. Einwanderung, Atomenergie, Gleichberechtigung etc.) sind zwar auch prinzipiell von unendlich großer Zahl; dies spricht jedoch keineswegs dagegen, diese Argumente unter eine endliche Anzahl sinnvoller Kategorien zu subsumieren, um so die Grundlage für Vergleichbarkeit zu schaffen.

Der erste Analyseschritt besteht also darin, alle Texte auf das Vorkommen von Argumenten zu untersuchen. Es entsteht so eine vorläufige Liste vorkommender Argumente, die natürlich mit fortschreitender Lektüre immer wieder revidiert werden muss. Ebenfalls im Laufe der Auswertung und mit fortschreitender Leseerfahrung kann man dann eine Anzahl prototypischer Argumente formulieren, unter die sich die real vorkommenden Argument-Modellierungen subsumieren lassen. Dies ist auch deshalb möglich, weil bei der Analyse von Diskursen nicht die einzelnen Texte sondern vielmehr die strukturellen Eigenheiten der jeweiligen Diskurse von Interesse sind. Sowohl die Stammtisch-Parole „Ausländer raus!“ wie die von einem Politiker in wohlgeformten Sätzen vorgetragene Forderung nach „Rückführung der Ausländerzahl“ würden also in diesem Analysemodell als Vorkommen des gleichen Arguments gewertet.

Bei der prinzipiell unbegrenzten Zahl konkreter Argument-Formulierungen kommt es darauf an, eine praktikable Anzahl prototypischer Argumente so zu formulieren, dass eine Vielzahl tatsächlicher Argumente-Formulierungen unter sie subsumiert werden kann. Ein zu hoher Abstraktionsgrad, d. h. eine zu geringe Zahl prototypischer Argumente, ließe allerdings die eigentlichen Inhalte der tatsächlichen Argumente sowie ihre Bandbreite aus dem Blickfeld geraten. Andererseits droht die Gefahr, eine zu große Zahl prototypischer Argumente zu formulieren, so dass umgekehrt die gemeinsame inhaltliche Basis der Einzelargumente nicht deutlich würde.

<sup>5</sup> Zur Analyse der in Diskursen verwendeten Metaphern vgl. die Arbeiten von Böke (1996), (1997).

Insofern stellt auch diese Art der Kategorisierung von real vorkommenden Argumenten natürlich bereits eine Interpretation dar.

Wenn man nun nach erfolgter Textlektüre eine solche Liste von prototypischen Argumenten zusammengestellt hat, dann ist es eine reine Fleißarbeit, das Vorkommen dieser Argumente in Deutschland, Österreich und der Schweiz auszuzählen und zu vergleichen. Doch leider ist es mit dieser Fleißarbeit nicht getan, denn reale Argumentation – das wissen Sie – funktioniert zwar durchaus, aber nicht so. In realen Argumentationen gibt es häufig keine klare Gegenüberstellung von Pro- und Contra-Argumenten, und erst recht stellt sie sich nicht als eine Abfolge von aufeinander bezogenen Argumenten dar. Leider ist die Angelegenheit komplizierter. Denn es gibt höchst unterschiedliche Arten, ein Argument zu bestreiten, weil nämlich Argumente für unterschiedliche Konklusionen verwendet werden können. Diese Multifunktionalität von Argumenten führt letztlich dazu, dass eine starre Gegenüberstellung von Pro- und Contra-Argumenten nicht durchzuhalten ist. An einem formalisierten Beispiel kann dies verdeutlicht werden.

Gegeben seien das Pro-Argument A und das Contra-Argument B. Wenn man sich nun vergegenwärtigt, auf welche Weise A und B in realen Argumentationen bestritten bzw. bekräftigt werden, stößt man zumindest auf folgende Möglichkeiten.

Funktion	Realisierung <sup>6</sup>	lfd. Nr.
Bestreiten von A	$\neg A$	①
	B!	②
Bekräftigen von A	A!	③
	$\neg B$	④
Bestreiten von B	$\neg B$	⑤
	A !	⑥
Bekräftigen von B	B!	⑦
	$\neg A$	⑧

Es fällt sofort auf, dass das Bestreiten von A ① und ② formal dem Bekräftigen von B ⑦ und ⑧ entspricht, wie auch das Bestreiten von B ⑤ und ⑥ dem Bekräftigen von A ③ und ④ entspricht. Zur Verdeutlichung setze ich im Folgenden für

A: *Asylsuchende werden politisch verfolgt und müssen bei uns aufgenommen werden.*

und für

B: *Asylsuchende kommen aus wirtschaftlichen Gründen und müssen deshalb nicht bei uns aufgenommen werden.*

Einsetzungen in die Beispiele ① bis ④ würden dann lauten:

<sup>6</sup> Diese Methode, Propositionen durch „labels“ zu versehen, geht auf Josef Klein (1993, 1995) zurück. „A!“ steht dann für die Behauptung der Proposition A, „ $\neg A$ “ für die Negierung derselben etc.

- ① *Asylsuchende werden nicht politisch verfolgt und müssen nicht bei uns aufgenommen werden.*
- ② *Asylsuchende kommen aus wirtschaftlichen Gründen und müssen deshalb nicht bei uns aufgenommen werden.*
- ③ *Asylsuchende werden politisch verfolgt und müssen bei uns aufgenommen werden.*
- ④ *Asylsuchende kommen nicht aus wirtschaftlichen Gründen und müssen deshalb bei uns aufgenommen werden.*

Ich habe diese Beispiele so ausführlich zitiert, weil sie deutlich machen, dass die analytische Beschreibung derartiger Argumente zwangsläufig in Schwierigkeiten gerät, wenn sie meint, mit „A“ und „B“, also mit einfachen entgegengesetzten Pro- und Contra-Argumenten reale Argumentationen angemessen und vollständig beschreiben zu können. Wie die Beispiele – wenn auch nur sehr unzureichend – zeigen, gibt es weitere Möglichkeiten, Position für oder gegen ein Argument zu beziehen.

Wie aber lässt sich eine Formulierung wie ① analytisch erfassen? Einerseits stellt sich das Problem, dass „ $\neg$  A“ nicht identisch mit „B!“ ist. Andererseits aber lässt sich „ $\neg$  A“ nicht als Pro-Argument behandeln, wiewohl die zugrundeliegende Proposition der des Pro-Argumentes „A“ entspricht.

Ein weiteres Problem ergibt sich, wenn man die reale Verwendung von Argumenten unter dem Aspekt der „Urheberschaft“ betrachtet. Hier zeigt sich nämlich ein analoges Problem. In einem Artikel des Schweizer Tages-Anzeigers vom 02.02.65 unter dem Titel „Wir müssen Schlagwörter überprüfen und Vorurteile abbauen“ heißt es beispielsweise (Hervorhebungen im Original):

Klärendes Licht in die finstere Psychose einer speziellen „Italienerkriminalität“ wirft die Untersuchung von Dr. Jürg Neumann (Juris-Verlag, Zürich). Neumann hat ein fremdenpolizeiliches Aktenmaterial von 200 Fällen sowie über 650 Gerichtsfälle bearbeitet. Er kommt zu dem Schluß, *dass für eine überdurchschnittliche Kriminalität italienischer Arbeitnehmer im Vergleich zu den Schweizern gesamthaft keine Anhaltspunkte vorliegen.* Die Vorstellung vom diebischen Italiener zum Beispiel, meint der Verfasser, werde weitgehend durch die schlechten Erfahrungen schweizerischer Touristen in Italien genährt. Bei den *Sittlichkeitsdelikten* ist die Kriminalitätsziffer der italienischen Arbeitskräfte in der Regel niedriger als die schweizerische, wenn sie auch seit 1960, hauptsächlich wegen des immer stärkeren Ueberwiegens der Männer, im Anstieg begriffen ist.

Im übrigen ist eine enge Wechselbeziehung zwischen *den Kriminalitätsziffern der Italiener und dem Vorurteil der einheimischen Bevölkerung* wahrscheinlich. Wer im Ausländer den potentiellen Verbrecher sieht, schaut ihm auch mehr auf die Finger und erstattet Anzeige bei Kleinstvergehen, die ein Schweizer ungeschoren hinter sich gebracht hätte.

Die spezielle Argumentations-Konstellation dieses Artikels besteht also darin, ein Contra-Argument zu zitieren und es durch andere Argumente zu widerlegen.

Da dies eine in Argumentationen häufig anzutreffende Praxis ist, muss bei der Analyse diesem Phänomen Rechnung getragen werden. Einerseits geht es nicht an, zitierte

Argumente in der quantitativen Analyse zu vernachlässigen, andererseits können sie in einer qualitativen Analyse nicht den tatsächlich als Argument verwendeten Argumenten gleichgestellt werden. Deshalb bedarf es auch hier einer differenzierten Betrachtung, in der der Unterschied von originärer und zitierender Verwendung berücksichtigt wird.

Dies möchte ich an vier Beispielen erläutern.

- 1) Man mag über die Gastarbeiter denken was man will, wir brauchen sie dringend, um unsere Produktion aufrecht zu erhalten, um unseren Lebensstandard nicht absinken zu lassen, um alle jene Aufgaben zu erfüllen, für die man keine heimischen Kräfte mehr bekommt. (AZ 17.12.72)
- 2) Ein völlig einheitlicher und freizügiger Arbeitsmarkt wird abgelehnt, denn die daraus folgenden Vergünstigungen (Stellen-, Berufs- und Kantonswechsel, Familiennachzug) würden nach Dr. Schwarbs Ansicht neue Animositäten in unserem Volk gegen die Ausländer provozieren. (TA 01.11.72; neutral zitierend)
- 3) Die Vorschläge von Elisabeth Blunschy, Präsidentin der Caritas Schweiz, doch erst das Personal aufzustocken und erst später, wenn nötig, eine Gesetzesrevision anzustreben, verhallen ohne Wirkung. Hier aber wäre zuerst anzusetzen gewesen. (Basler Zeitung 30.11.83; positiv zitierend)
- 4) Immer wieder hört man die Klage, die Ausländer würden Spitäler und Schulen füllen und den Schweizern Wohnungen wegnehmen. Wäre die schweizerische Bevölkerung in dem Mass gewachsen, dass die Wirtschaft auf Fremdarbeiter hätte verzichten können, so wären Schulen und Spitäler ebenso überfüllt [...]. (TA 11.10.73; negativ zitierend)

Zu unterscheiden sind demnach mindestens die folgenden Argument-Verwendungen:

- originäre („argumentative“) Verwendung: Beispiel 1
- neutral zitierende Verwendung: Beispiel 2
- affirmativ bzw. positiv zitierende Verwendung: Beispiel 3
- negativ zitierende Verwendung: Beispiel 4

Natürlich ist besonders das negative Zitat für eine präzise Analyse eines Diskurses von Bedeutung. Denn einerseits taucht in ihm ja ein bestimmtes Argument auf (in unserem Beispiel das Argument der mangelnden Kapazitäten), andererseits kann dieses negative Zitieren nicht mit der neutral zitierenden und erst recht nicht mit der argumentativen oder positiv zitierenden Verwendung des Argumentes gleichgesetzt werden. Bei der Analyse ist daher die Trennung dieser vier Verwendungsweisen unerlässlich.

## **5. Ergebnisse**

Hat man nun seine Daten beisammen, so kann man anfangen, die Argumentationsstruktur der einzelnen Diskurse zu untersuchen. Man kann in einem ersten Schritt rein quantitativ arbeiten und zunächst einmal auszählen, welche „Argumentationsrichtung“ die einzelnen Zeitungsartikel haben, d. h. ob sie pro oder contra Migration argumentieren, ob sie abwägend argumentieren oder aber vielleicht gar keine explizite Argumentation erkennen lassen. Man kann weiterhin untersuchen, welche Argumente in einem Diskurs überhaupt vorkommen. Hat man die Daten entsprechend der eben beschriebenen Unterscheidung in der Argumentverwendung aufbereitet, kann man danach untersuchen, wel-

ches Gewicht bestimmte Argumente haben. So lässt sich beispielsweise im Gastarbeiter-Diskurs der BRD feststellen, dass das wichtigste Argument, das in den Jahren 1965-1967 gegen die Anwesenheit bzw. Anwerbung von Gastarbeitern vorgebracht wurde, das der hohen Kosten war. Gleichzeitig taucht das Argument der Gastarbeiter-Kriminalität vergleichsweise häufig auf. Die qualitative Analyse zeigt bei diesem Argument jedoch, dass es vorwiegend negativ zitiert wird. Anders ausgedrückt: Das Argument der Gastarbeiter-Kriminalität kommt zwar häufig vor, wird jedoch meist explizit bestritten. Hat man die einzelnen Diskurs-Abschnitte für sich quantitativ und qualitativ analysiert, so steht einem internationalen Vergleich der Diskursstücke nichts mehr im Wege. Ohne hier auf argumentative Einzelheiten eingehen zu können, kann ich einige wenige Ergebnisse exemplarisch referieren.

Die Gastarbeiter-Diskurse Deutschlands, Österreichs und der Schweiz stellen sich als relativ uneinheitliche Gebilde dar. Zwar gibt es in Deutschland und in der Schweiz strukturelle Parallelen bei chronologischen Trends wie auch in der Argumentverwendung, dennoch findet man auch deutliche Unterschiede, hervorgerufen durch Argumente, die nur in jeweils einem der drei Länder von Bedeutung sind. Diese spezifische Ausgestaltung der Diskurse dürfte mit den wirtschaftlichen Rahmenbedingungen der drei Länder korrespondieren. Auch in den jeweiligen Wirtschaftspolitiken gab es Gemeinsamkeiten (wie beispielsweise die Anwerbepolitik) und Unterschiede (wie beispielsweise das spezielle Vorgehen der Schweiz, um einen bestimmten Gastarbeiter-Anteil nicht zu überschreiten). Einen Reflex auf diese Unterschiede finden wir in der unterschiedlichen Ausgestaltung der Diskurse widerspiegelt.

Die Asyldiskurse (insbesondere der BRD und der Schweiz) sind sich dagegen ähnlicher. Einerseits sorgt sicherlich die gleichartige abendländisch-christliche Tradition dafür, dass die Diskussion um das Asylrecht so ausführlich stattfindet. Andererseits sind wirtschaftliche Erwägungen für Industrienationen natürlich immer dann von Bedeutung, wenn die Kosten humanitärer Maßnahmen ein bestimmtes Maß überschreiten, ohne dass gleichzeitig im Gegenzug wirtschaftliche Anreize bestünden (Erschließung neuer Märkte etc.). Die relative Sonderrolle Österreichs lässt sich wiederum durch andere Faktoren der Sachgeschichte erklären. Österreich unterscheidet sich dadurch von Deutschland und der Schweiz, dass es zunächst von Asylsuchenden hauptsächlich als „Durchgangsstation“ genutzt wird. Damit mag es zusammenhängen, dass in Österreich die wirtschaftlichen Aspekte der Asylgewährung kaum anders gewichtet werden als die ethischen. Sieht Österreich Asylsuchende kommen und danach auch regelmäßig wieder gehen, so könnte dies dazu führen, dass in der öffentlichen Diskussion die wirtschaftlichen Aspekte für das eigene Land relativ stark betont werden und zu Mutmaßungen über Fluchtgründe und den eventuellen Missbrauch des Asylrechts kein besonderer diskursiver Anreiz besteht.

## **6. Fazit**

Abgesehen von solchen inhaltlichen Einzelergebnissen zur Migrationsdiskussion in der Bundesrepublik Deutschland, der Schweiz und Österreich scheint mir ein wichtiges Ergebnis darin zu bestehen, dass auch die germanistische Linguistik eine Methodik zur Analyse große Textkorpora bereitstellen kann. Derartige Analysen waren bislang eine Domäne der verschiedenen Spielarten der Inhaltsanalyse<sup>7</sup>, die m. E. allerdings in keiner ihrer Varianten auf speziell linguistische Fragestellungen befriedigende Antworten liefern kann. Insbesondere die qualitative Unterscheidung von Argument-Verwendungen ist hier zu nennen, die einerseits Argumente nicht aus ihrem Kontext reißt und andererseits die vielfachen in der Realität vorkommenden Spielarten von Argumentation zumindest annäherungsweise darstellen kann.

## **7. Verwendete Siglen**

AZ	=	Arbeiter-Zeitung
BaZ	=	Basler Zeitung
BN	=	Basler Nachrichten
FAZ	=	Frankfurter Allgemeine Zeitung
FR	=	Frankfurter Rundschau
KTZ	=	Kärntner Tageszeitung
NZZ	=	Neue Zürcher Zeitung
RP	=	Rheinische Post
SZ	=	Süddeutsche Zeitung
TA	=	Tages-Anzeiger
Welt/WamS	=	Welt/Welt am Sonntag
WeWo	=	Weltwoche

## **8. Literatur**

- Böke, Karin / Jung, Matthias / Niehr, Thomas / Wengeler, Martin (2000): Vergleichende Diskurslinguistik. Überlegungen zur Analyse national heterogener Textkorpora. In: Niehr/Böke, S. 11-36.
- Böke, Karin (1996): Überlegungen zu einer Metaphernanalyse im Dienste einer „parzellierten“ Sprachgeschichtsschreibung. In: Böke, Karin / Jung, Matthias / Wengeler, Martin (Hrsg.) (1996): Öffentlicher Sprachgebrauch. Praktische, historische und interdisziplinäre Perspektiven. Georg Stötzel zum 60. Geburtstag gewidmet. Opladen, S. 431- 452.
- Böke, Karin (1997): Die „Invasion“ aus den „Armenhäusern Europas“. Metaphern im Einwanderungsdiskurs. In: Jung, Matthias / Wengeler, Martin / Böke, Karin (Hrsg.) (1997): Die Sprache des Migrationsdiskurses. Das Reden über „Ausländer“ in Medien, Politik und Alltag. Opladen, S. 164-193.

---

7 Vgl. dazu die Überblicksdarstellungen von Fröh (1991), Lisch/Kriz (1978) und Merten (1995).



- Busse, Dietrich/Teubert, Wolfgang (1994): Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der Historischen Semantik. In: Busse, Dietrich/Hermanns, Fritz/Teubert, Wolfgang (Hrsg.) (1994): Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik. Opladen, S. 10-28.
- Früh, Werner (1991): Inhaltsanalyse. Theorie und Praxis. 3., überarb. Aufl., München.
- Hermanns, Fritz (1995): Sprachgeschichte als Mentalitätsgeschichte. Überlegungen zu Sinn und Form und Gegenstand historischer Semantik. In: Gardt, Andreas / Mattheier, Klaus / Reichmann, Oskar: Sprachgeschichte des Neuhochdeutschen. Tübingen, S. 69-101.
- Kienpointner, Manfred (1992): Alltagslogik. Struktur und Funktion von Argumentationsmustern. Stuttgart.
- Klein, Josef (1993): Ein 3-Ebenen-Modell zur vergleichenden Analyse argumentativer Texte, dargestellt im Netz-Diagramm-Format und exemplifiziert an Zeitungskommentaren. In: Sandig, Barbara/Püschel, Ulrich (Hrsg.): Stilistik. Bd. III: Argumentationsstile. Hildesheim, S. 77-111. [Germanistische Linguistik 112-113].
- Klein, Josef (1995): Asyl-Diskurs. Konflikte und Blockaden in Politik, Medien und Alltagswelt. In: Reiher, Ruth (Hrsg.): Sprache im Konflikt. Zur Rolle der Sprache in sozialen, politischen und militärischen Auseinandersetzungen. Berlin, New York, S. 15-71.
- Lisch, Ralf / Kriz, Jürgen (1978): Grundlagen und Modelle der Inhaltsanalyse. Bestandsaufnahme und Kritik. Reinbek bei Hamburg.
- Merten, Klaus (1995): Inhaltsanalyse. Einführung in Theorie, Methode und Praxis. Opladen.
- Niehr, Thomas (2000): Parlamentarische Diskurse im internationalen Vergleich. Eine Fallstudie zu den jüngsten Asylrechtsänderungen in Deutschland und der Schweiz. In: Niehr/Böke, S. 109-134.
- Niehr, Thomas / Böke, Karin (Hrsg.) (2000): Einwanderungsdiskurse. Vergleichende diskurslinguistische Studien. Wiesbaden.