

Bildungsstandards im Französischunterricht: Testen und Fördern durch Aufgaben

→ französisch heute 3 (2006): 210–228

L'Allemagne, traumatisée par des études internationales comparant l'efficacité de la formation scolaire, se propose de faire un grand effort pour renverser la tendance. Des tests centraux, du type sommatif, sont considérés comme l'outil central de contrôle-qualité. Cet « assessment » passe par l'apprentissage fondé sur les tâches, d'où l'importance de savoir construire des tâches adéquates qu'elles soient de caractère sommatif ou formatif. Il va sans dire que les seconds préparent aux premiers.

Franz-Joseph Meißner

Kern der tiefgreifenden, gegenwärtig platzgreifenden Reformen im Schulwesen, also nicht nur des Fremdsprachenunterrichts, sind die sogenannten ‚Bildungsstandards‘ und ihre Entwicklung, Implementierung und Überprüfung. Man kann sie als das Steuerungsinstrument einer output-orientierten Bildungspolitik begreifen. ‚Standards‘ alleine reichen indes zur Verbesserung der Leistungsbilanz in Sachen Fremdsprachenunterricht nicht aus, solange die rein quantitativen Daten – z.B. die Diversifikation des real genutzten Sprachenangebots – nicht angemessen realisiert werden. Mit durchschnittlich nur 1,3 belegten Fremdsprachen pro Schüler steht Deutschland am unteren Ende der EU-Vergleichsskala. Im Jahr 2002 unterstrichen die Regierungschefs in Barcelona, „dass die Erziehung zur Mehrsprachigkeit ein wichtiges Mittel der europäischen Integration darstelle, sie könne Verständnis der kulturellen Vielfalt (...) bewirken. Die Regierungschef bedauerten das Fehlen verlässlicher Daten über die Sprachenkenntnisse der jungen Generation in Europa und forderten die Einrichtung eines Indikatorensystems zur sprachlichen Kompetenz in Gestalt der sogenannten ‚Barcelona-Indikatoren‘. Inzwischen dient der ‚Gemeinsame europäische Referenzrahmen für Sprachen‘ als Grundlage nicht nur für die gegenseitige Anerkennung von Sprachprüfungen, sondern auch für die Entwicklung und Überprüfung der ‚Sprachindikatoren‘“ – soweit Schröder et al. (2006: 11). Man muss sich indes fragen, ob schon die Tatsache, dass ‚ausschließlich‘ die Englischkenntnisse in Studien wie DESI (‚Deutsch-Englisch-Studien International‘¹) getestet wurden, bereits eine engführende Fokussierung auf die *lingua franca* signalisiert. Dass es mit der Mehrsprachigkeitsförderung in der Breite, insbesondere in der Sekundarstufe II, in Deutschland nicht allzu weit her ist, zeigten die Erhebungen vom Meißner & Lang (2005) zu den realen Belegzahlen

von Fremdsprachenunterricht in der Jahrgangsstufe 12. In der deutschen Presse wurde es als skandalös empfunden, dass der Unterschied des real vermittelten Fremdsprachenunterrichts in der Klasse 12 sowohl in der Breite der belegten Sprachen als auch in der Qualität des Fremdsprachenlernens (GK vs. LK) zwischen den Bundesländern im Europäischen Jahr der Sprachen über 30 und 40 Prozent betrug². Es war indes nicht davon auszugehen, dass die betroffenen Länder und die politisch verantwortlichen Amtsträger das Ergebnis dieser Studie begrüßten. Eines darf man nie aus den Augen verlieren: Fremdsprachenunterricht ist immer mehr als ausschließlich Unterricht in nur einer und stets derselben Fremdsprache und... wo es keinen Fremdsprachenunterricht mehr gibt, muss/kann man über dessen Qualität nicht reden. – Bei den DESI-Autoren umreißen die angesprochenen Innovationen:

- die Ausarbeitung von Kompetenzmodellen, die entsprechend den Stand fremdsprachendidaktischer und psychologischer Forschung die wichtigsten Teilkompetenzen und deren Abstufungen systematisch beschreiben sowie, damit verbunden,
- die Entwicklung und Skalierung von Tests, mit denen die Einlösung der Standards empirisch abgesichert und überprüft werden kann, und schließlich,
- die auf den empirischen Testergebnissen beruhende Feststellung, wie sich die Schülerinnen und Schüler der Zielpopulation auf verschiedene Kompetenzniveaus verteilen.

Doch was sagt die Kultusministerkonferenz in ihrer ‚Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss (Jahrgangsstufe 10)‘ vom 4.12.2003?

Die Kultusministerkonferenz sieht es als zentrale Aufgabe an, die Qualität schulischer Bildung, die Vergleichbarkeit schulischer Abschlüsse sowie die Durchlässigkeit des Bildungssystems zu sichern. Bildungsstandards sind hierbei von besonderer Bedeutung. Sie sind Bestandteile eines umfassenden Systems der Qualitätssicherung, das auch Schulentwicklung, interne und externe Evaluation umfasst. Bildungsstandards beschreiben erwartete Lernergebnisse. Ihre Anwendung bietet Hinweise für notwendige Förderungs- und Unterstützungsmaßnahmen.

Bildungsstandards greifen allgemeine Bildungsziele auf und benennen Kompetenzen, die Schülerinnen und Schüler bis zu einer bestimmten Jahrgangsstufe an zentralen Inhalten erworben haben sollen. Sie konzentrieren sich auf Kernbereiche eines Faches.

Bildungsstandards formulieren fachliche und fachübergreifende Basisqualifikationen, die für die weitere schulische und berufliche Ausbildung von Bedeutung sind und die anschlussfähiges Lernen ermöglichen. Die Standards stehen im Einklang mit dem Auftrag der schulischen Bildung. Sie zielen auf Persönlichkeitsentwicklung und Weltorientierung, die sich aus der Begegnung mit zentralen Gegenständen unserer Kultur ergeben.

Der Schlüsselbegriff der Standards heißt Kompetenz, dessen Umriss Bernd Tesch (in diesem Heft) nachzeichnet. Der Kompetenzbegriff der KMK gilt für alle Fächer und Lernbereiche. Er ist daher von dem in den Fremdsprachendidaktiken benutzten Begriff der fremdsprachlichen *Kompetenz (etwa kommunikative Kompetenz, Teilkompetenz oder Teilfertigkeit, in-*

Beschlüsse der Kultusministerkonferenz

Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss

Beschluss vom 4.12.2003



*terkulturelle Kompetenz, Methodenkompetenz usw.) zu trennen. Schon diese unvollständige Anreihung von Kompetenzdomänen zeigt, dass Kompetenz eine übergeordnete Qualifikation meint, deren Komposition mehrere untergeordnete Kompetenzen erfordert. Für die pädagogische Praxis heißt dies, dass an vielen Kompetenzkomplexen gearbeitet werden muss, damit das Ziel der KMK-Standards erreicht werden kann. Die Situation wird dadurch nicht einfacher, dass die Fähigkeit zur ‚Kommunikation in der Zielsprache‘ mit nativen Partnern sowie in der *lingua franca*-Situation (Typ Deutscher unterhält sich mit Polin in französischer Sprache) seit den Tagen der ‚kommunikativen Didaktik‘ zu ‚der‘ Legitimation des Fremdsprachenunterrichts schlechthin geworden ist (I. Christ 2006b). Über die Frage, wieviel Literatur und welche lässt sich heute streiten, über die der Befähigung unserer Schülerinnen und Schüler zur Kommunikationsfähigkeit in der Zielsprache hingegen nicht. – Das hat natürlich viel mit der europäischen Ausrichtung unserer Bildungsziele zu tun.*

Ausgelöst wurde die Kompetenzorientierung durch die internationalen Vergleichsstudien zur Effizienz der Bildungssysteme, die Deutschland bekanntlich herb aus dem Dornröschenschlaf rissen. – Parallel hierzu hat der Europarat den ‚Gemeinsamen Europäischen Referenzrahmen für Sprachen‘ vorgelegt. Er liefert inzwischen ‚die‘ Richtschnur für die Messung von qua Unterricht erreichbarer fremdsprachlicher Kommunikationsfähigkeit, aber auch für das Sprachenlernen und dessen Methodik. Die Niveaustufen des GeR sind primär keine Lernziele, sondern Beschreibungen von fremdsprachlichen Kompetenzstufen. Es handelt sich um Maßstäbe, welche die sog. ‚Wissengesellschaft‘ bzw. das lebensbegleitende Lernen benötigen. Lernziele werden diese Stufen dann, wenn die Bildungspolitik jene als Lernziele definiert. Genau das hat die KMK im Hinblick auf die Klasse 10 und B1 getan. – Schon das Inhaltsverzeichnis der Bildungsstandards zeigt, wie sehr sich diese an dem oben umrissenen Kompetenzbegriff orientieren.

- 1 Der Beitrag der ersten Fremdsprache zur Bildung**
- 2 Kompetenzbereiche der ersten Fremdsprache**
- 3 Standards für Kompetenzbereiche der ersten Fremdsprache**
- 3.1 Kommunikative Fertigkeiten**
- 3.2 Verfügung über die sprachlichen Mittel**
- 3.3 Interkulturelle Kompetenzen**
- 3.4 Methodenkompetenzen**
- 4 Aufgabenbeispiele**
- 4.1 Kommentierte Aufgabenbeispiele: Englisch**
- 4.2 Kommentierte Aufgabenbeispiele: Französisch**
- Anhang: Kompetenzstufen des Gemeinsamen Europäischen Referenzrahmens für Sprachen**

Methoden- oder Lernkompetenz

Die Auflistung liefert den Rahmen für die Erstellung kompetenzorientierter Aufgaben. Sie offenbart aber auch in der benutzten Begrifflichkeit liegende mögliche Steuerungsschwächen, denen es nun durch die geplanten Metaaufgaben des IQB vorzubeugen gilt. Das Risiko der Fehlsteuerung sehen wir vor allem beim Begriff der ‚Methodenkompetenz‘. Er scheint missverständlich und wird leicht ‚flach‘, etwa im Sinne von Lerntipps, gedeutet. Dies ist möglich, weil

Methodenkompetenz nicht auf den ersten Blick durchblicken lässt, wer eigentlich als Adressat und was gemeint sein soll und weil in der Fachlehrerschaft aufgrund einer im fachdidaktischen Bereich defizienten Ausbildung der hinter dem Bereich stehende Komplex auf relatives Unwissen stößt. Dies signalisieren auch die Lehrwerke, welche die Strategienkomplexe von Ressourcenmanagement, Kognitivität und Metakognitivität nur bislang unzureichend umsetzten.

Methode meint ‚Lernweg‘. Schließlich bedeutet das griechische Etymon *methódos* im Kern ‚Weg‘ (*hodos*). Zu *methódos* notiert Littré : «Ensemble de procédés raisonnés pour faire quelque chose». Zahl und Art der Wege hängen also von den Zielen ab. Immerhin sagen die Standards klar, was gemeint ist. Demnach sollen die Lernenden der Sekundarstufe I:

- Lerntechniken und –strategien für den Ausbau ihrer Kenntnisse in der jeweiligen Fremdsprache sowie für den Erwerb weiterer Fremdsprachen einsetzen,
- kooperative Formen des Arbeitens und Lernens entwickeln,
- durch Analyse ihres persönlichen Lernstils und die Auswahl besonders geeigneter Lernverfahren zur Selbstständigkeit im Sprachenlernen gelangen,
- verschiedene Verfahren zur Auswertung gesprochener und geschriebener Texte nutzen,
- Verfahren zum anwendungs- und produktorientierten Gestalten von mündlichen und schriftlichen Texten anwenden. (KMK 2003: 10)

Der Erziehungswissenschaftler Jürgen Baumert (1993: 327) notiert zum Konnex von Lernstrategien, Lernmotivation und Selbstwirksamkeitsüberzeugung:

Selbstgesteuertes Lernen (...) ist voraussetzungsvoll und der Erwerb von Lernexpertise entsprechend langwierig. (...) Fasst man diese Gesichtspunkte zusammen, dann verhält sich der erfahrene Lerner proaktiv, planvoll, reflexiv und dem eigenen Lernen gegenüber verantwortungsbewusst. Selbstgesteuertes Lernen ist (...) ein dynamisches Zusammenwirken von *skill and will* und damit eine komplexe Leistung der Selbstregulation. (328).

Damit ist – grob – das Ziel benannt, dass fächerübergreifend in etwa mit der Jahrgangstufe 10 erreicht werden soll/kann. Der Weg dahin ist weit, aber er muss schon frühzeitig beschriftet werden. Je früher definierbare Zwischenziele erreicht werden, desto größer ist der Zugewinn, desto größer der weitere Lernzuwachs. Das bedeutet, dass die Fachdidaktiken (fachbezogene) Zwischenziele unterschiedlicher Reichweite beschreiben und diese empirisch überprüfen müssen. Dies haben sie aber bislang in unserem Bereich nicht wirklich getan! Gefragt sind vor allem Strategien kurzer und mittlerer Reichweite und Lerntechniken, weil sich diese auf konkrete Lernfragen beziehen. Diese werden in konkreten Lernkontexten ausgebildet. Routinisierung und Modifikationen sind erforderlich. Im Unterschied zu natürlichen Lernsituationen sichert schulisches Lernen thematische und inhaltliche, im Fall der Sprachen, sprachliche Kontinuität. Zu erinnern ist allein an den Terminus des Spiralcurriculums, der bislang vor allem auf das rein sprachliche Lernen bezogen wurde. Die Strategieforschung nennt im Bereich der kognitiven Strategien vor allem Memorier-, Elaborations- und Transformationsstrategien. Elaborationsstrategien erlauben die Sinnkonstruktion für und an neuen Lernstoffen mit den weiteren Kennzeichen der Wissensvernetzung und des Transfers. Im fachdidaktischen Zusammenhang ist daran zu erinnern, dass die Interkomprehensions- oder Mehrsprachigkeitsdidaktik hier eine sehr wichtige *raison d'être* findet: Sie führt nicht nur zu rezeptiven Mehrsprachigkeit, sondern gerade zu Sprachenlernkompetenz. Dies liegt daran, dass interkomprehensiv gesteuerte zentrale Erkenntnisse der Lerntheorie umsetzt, wie Doyé (2006) zeigt. Die Fokussierungen für die Aufgabenkonstruktion lauten hier: Reflexivität gegenüber

den neuen Informationen, Aktivierung/Revision des relevanten Vorwissens und Reflexivität, Abgleich des neuen Wissens mit vorhandenen Wissensbeständen und reflexive Sensibilität gegenüber dem eigenen Lernverhalten (Selbstmonitoring). Eine wichtige Aufgabe der Lernaufgaben ist es, Schüler zu diesen Schritten anzuleiten und zu befähigen. Hier besteht zur Zeit, was die Aufgabenentwicklung angeht, deutlicher Nachholbedarf. – Es war von den Defiziten der fremdsprachendidaktischen Forschung die Rede. Qualitativ hochwertige, nicht nur didaktische Forschung ist langwierig. Die Praxis aber muss handeln, sie kann nicht zuwarten, bis die Wissenschaft abgesicherte Erkenntnisse liefert (wenn sie denn kommen). Deshalb benötigt die Aufgabenentwicklung Vorgaben, wie das Lernen des Lernens von Sprachen zu optimieren ist. Immerhin ist die Fremdsprachendidaktik hier durchaus nicht unwissend.

Kompetenzorientierung und GeR

Der GeR ist nun nicht die erste europäische wegweisende Publikation auf dem Feld der Förderung der Mehrsprachigkeit der europäischen Bürgerinnen und Bürger. Zu erinnern ist zumindest an die Schwellenniveaulisten (*Niveau Seuil, Threshold level ...*). *Neu ist allerdings das Maß der Verbindlichkeit, dass der GeR heute für die Bildungspolitik der europäischen Staaten und der nachgeordneten Lehrpraxen hat.*

Der Entwicklung des GeR gingen die Untersuchungen der sogenannten ALTE-Gruppe voraus. Die *Association of Language Testers in Europe* beleuchtete auf empirischer Grundlage Möglichkeiten zur Entwicklung von Kriterien des Qualitätsmanagements im Bereich der Kenntnis und des Erlernens fremder Sprachen (Bolton 2004). ALTE fand schon früh das Interesse der führenden Anbieter von Fremdsprachenunterricht in der Erwachsenenbildung. Heute gehören die großen nationalen Sprachinstitutionen der europäischen Staaten der ALTE-Gruppe an; so das Goethe-Institut, das Instituto Cervantes usw. Eine Reihe privater Anbieter kommen hinzu, die sich verpflichtet haben, dem ALTE-Standard zu entsprechen und dies durch geeignete Maßnahmen nachweisen. Die Zahl der so oder so mit ALTE assoziierten Organisationen wächst. Das folgende Bild gibt einen Ausschnitt hiervon wieder. – Das Gewicht Europas wird außerhalb der Europäischen Union durchaus gesehen. Im Bereich der Sprachen bemerkt man dies daran, dass der GeR inzwischen schon mehr ist als ein rein innereuropäischer Vergleichsmaßstab. Dies zeigen die Entwicklungen in Russland, Asien, im Vorderen Orient und in Südamerika. – All dies führt jedoch auch wieder zum Schulwesen zurück. Denn es ist heutzutage offensichtlich nur schwer, wenn überhaupt, möglich, den ALTE/GeR-Qualitätsstandards aufzugeben oder gar wieder zu verlassen. Aber wer wollte dies?

Dies heißt nicht, dass die Entwicklungen nicht ihre Kritiker hätte. Am deutlichsten wurde dies bislang wohl in den Arbeitspapieren der 22. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts formuliert (Bausch et al. 2003). Die Arbeitspapiere sind seitdem ein Referenzpunkt für die weitere Auseinandersetzung mit dem GeR. Auch die ‚Bildungsstandards‘, aber ebenso die ‚Aufgabenorientierung‘ waren Gegenstand der Frühjahrskonferenz (Bausch et al. 2005 sowie 2006). Der Wichtigkeit, die gemeinhin dem Thema zumessen wird, spiegelt sich auch in den Fachzeitschriften: Hiervon gibt im Anschluss an *Der fremdsprachliche Unterricht Englisch* auch das vorliegende *französisch heute* ein Beispiel. Auch andere Zeitschriften des Fremdsprachenunterrichts haben das Thema längst angesteuert.

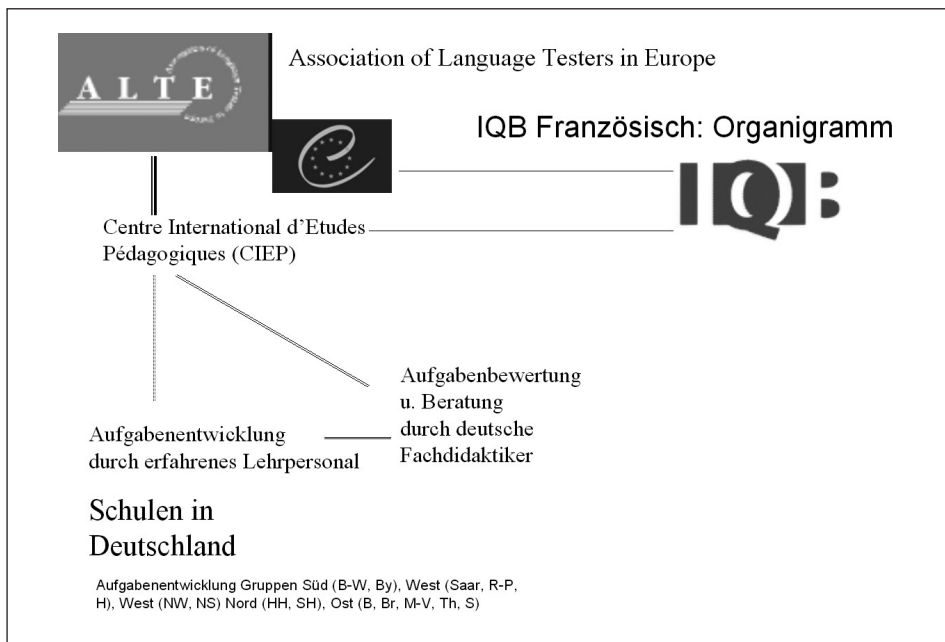
Die Bildungspolitik hat sich durch die Bildung des ‚Instituts für Qualitätsentwicklung im Bildungswesen‘ auf diese Entwicklung eingestellt, wie der Beitrag von Tesch spiegelt. Zum ande-

ren hat sie aber auch mit der Entwicklung von Kerncurricula und Bildungsstandards reagiert. Die *Vereinigung der Französischlehrerinnen und -lehrer* war von Anfang an in die das Französische betreffenden Entwicklungen einbezogen. Und die Bildungsstandards für die erste Fremdsprache für den Mittleren Bildungsabschluss im Fach Französisch sind das Produkt dieses, von der KMK initiierten Projekts. In ihren Leistungsbeschreibungen nehmen die Standards selbstverständlich auf den GeR Bezug. Es entspricht der Offenheit der KMK, dass sie auch international auf ganz unterschiedlichen Ebenen die Zusammenarbeit sucht. Das spiegelt auch das IQB.

Die Bildungsstandards erklären die Aufwertung von zwei Phänomenen, die den Unterricht seit jeher begleiten: die Qualitätssicherung von Prüfungen – jetzt: deren Zentralität nach Standards – und die Steuerung der Praxis durch zielgerichtete Aufgaben. – Dies wiederum führt zu der Kernfrage: Wie kann man kommunikative Kompetenzen beschreiben?

Kompetenzstufenbeschreibung nach dem GeR

Der GeR unterscheidet die sprachliche Kompetenz in drei, jeweils sequentiell gedoppelten Stufen: elementare Sprachverwendung (A1, A2), selbstständige Sprachverwendung (B1, B2) und kompetente Sprachverwendung (C1, C2), mit den in deutschen Fachkreisen inzwischen gängigen englischen Bezeichnungen *Breakthrough* (A1), *Waystage* (A2), **Threshold** (B1), *Vantage* (B2), *Effective Proficiency* (C1) und *Mastery*. (C2). Auch C2 deckt noch keine muttersprachliche Kompetenz ab. Es handelt sich eben um ein Raster für die Verwendung von Fremdsprache, was eine entsprechende Unterscheidung notwendig macht (vgl. Meißner in diesem Heft). Als Beschreibungsinstrumentarium für kommunikative Kompetenzstufen erfasst er unterschiedliche Zielgruppen; dies betreffen: alle Altersstufen, alle Schulformen, alle



Lernstufen, Lerner aller möglichen Nationen und Zielsprachen, allgemeine Lernkontexte und fachsprachlich ausgerichtete. Dem entspricht, dass der GeR einiges nicht bietet: keine Themenlisten (wohl Sprechsituationstypen), keine Vokabel- und Grammatikminima. Ihre Auswahl überlässt der GeR den Aufgabenentwicklern.

Der GeR blickt etwa mit dem pädagogischen Zeigefinger weniger auf Defizienzen – also auf das, was noch zur Erreichung einer muttersprachennahen Kompetenz fehlt –, statt dessen auf das, was ein Lerner schon kann. Es wäre nun naiv anzunehmen, er täte dies allein aus pädagogischen und motivationalen Erwägungen heraus. Vielmehr folgen die *can do*-Deskriptoren dem Gedanken, dass Europäer einander in ihren unterschiedlichen Sprachen verstehen sollen, und dem Wissen, dass es dorthin ein weiter Weg ist. Anders gesagt: Der GeR ist Teil der Europäischen Sprachenpolitik zu Gunsten der Mehrsprachigkeit (I. Christ 2006a). Wie dieses Ziel aussehen kann, beschreibt etwa Umberto Eco in *Auf der Suche nach der vollkommenen Sprache* (dt. 1994):

Ein Europa von Polyglotten ist kein Europa von Menschen, die viele Sprachen perfekt beherrschen, sondern im besten Fall eines von Menschen, die sich verständigen können, indem jeder die eigene Sprache spricht und die des anderen versteht, ohne sie fließend sprechen zu können., wobei er, während er sie versteht, wen auch nur mit Mühe, zugleich ihren ‚Geist‘ versteht, das kulturelle Universum, das ein jeder ausdrückt, wenn er die Sprache seiner Vorfahren und seiner Tradition spricht. (Eco 1994: 355)

Den GeR interessiert auch die pädagogische Frage, wie der Weg zu dieser Zielvorstellung aussehen mag. In jedem Fall wird es notwendig unterschiedliche Etappen zu definieren. Genau dies leistet der GeR. Dem entspricht auch, dass es der GeR nicht bei der Nennung der vier Grundfertigkeiten belässt.

Dem entsprechend, aber auch in Kenntnis der realen Erfahrungen im Aufbau von Kompetenz in europäischen Sprachen, demzufolge sich die Kompetenzen in unterschiedlichem Tempo entwickeln lassen, misst der GeR die Fertigkeiten getrennt voneinander. Er misst umfassend dann nur in der Summe dieser Fertigkeiten. Anders gesagt: Global C1 heißt: C1 Hörverstehen, C1 Lesen, C1 Sprechen und C1 Schreiben sowie interkulturell-kommunikative Kompetenz. Ein solches Muster ist aber für einen Unterricht, der lernerorientiert ist und jeden nach seinem eigenen Tempo lernen lässt, untypisch. Diese Bemerkung meint nicht lernerseitige Trägheit oder Motivationsschwäche, sondern sie berücksichtigt die individuelle Arbeitsweise des mentalen Lexikons bzw. die Art der Entwicklung von Lernaltersprache. Das Muster wird so deshalb nicht erreicht, weil sich die Teilfertigkeiten von Lerner zu Lerner in unterschiedlichem Tempo entwickeln und der Intake auch bei gleichem Input individuell verschieden wächst. Ein Muster ‚A2 alle Lerner‘ nach der Lernzeit von einem Jahr ist entweder falsch, weil es verschweigt, was die einzelnen wirklich können, oder aber es zeigt, dass die schnellen Lerner (scheinbar) ‚ausgebremst‘ wurden. Realistischer erscheint etwa das folgende Muster. Es spiegelt den fiktiven Leistungsstand einer Klasse nach zwei Lernjahren:

Stufe	HV	Lesen	Sprechen	Schreiben
A1	5 Schüler	1 Schüler	8 Schüler	3 Schüler
A2	10 Schüler	13 Schüler	6 Schüler	11 Schüler
B1	3 Schüler	4 Schüler	4 Schüler	4 Schüler

Fiktive Verteilungskurve innerhalb einer Klasse

Schon hier sei darauf hingewiesen, dass die Arbeit nach dem GeR eine intensive Beobachtung der einzelnen Schülerleistungen verlangt sowie eine nach Fertigkeiten getrennte Binnendifferenzierung. Die Aufgaben tragen dem durch den Charakter der Mantelaufgabe Rechnung. Wir verstehen hierunter eine Gesamtaufgabe mit auf unterschiedliche Fertigkeiten bezogene Teilaufgaben. – Die aufgelistete Viererskala der Teilfertigkeiten deckt nur die primär-sprachlichen und noch nicht die Kompetenzbreite des GeR ab. Sie abstrahiert um das Lernen des Lernens (*savoir apprendre*).

Grundwissen und Kernbegriffe zu Testen, Prüfen und Fördern

Bildungsstandards und Zentrale Abschlussprüfungen beabsichtigen die Vergleichbarkeit von Abschlüssen und die Qualitätssicherung, d.h. die Überprüfung von Lernzielen bzw. dem Bildungsziel. Das Maß der Entsprechung wird mit der ‚Note‘ korreliert. Bildungsstandards und zentrale Prüfungsabschlüsse (für große Populationen) erfordern daher summative ‚Tests‘.

Tests sind Messinstrumente, die bestimmten Gütekriterien genügen müssen. Die Hauptgütekriterien von Tests (und auch von anderen Messverfahren) sind Objektivität, Reliabilität und Validität. Daneben werden als weitere Kriterien z.B. Authentizität, Ökonomie und Nützlichkeit sowie die Aufgabentrennschärfe genannt. Die Gütekriterien gelten stets in Abhängigkeit von der spezifischen Verwendung von Tests. (Grotjahn 2006: 222).

Die Testtheorie unterscheidet grundsätzlich zwischen summativen und formativen Tests bzw. Aufgaben. Erstere testen breite Populationen – bekannte Beispiele TESTDAF oder DELF –, formative Aufgaben sind Lernaufgaben, die den Lernzuwachs auf/nach einer Lernstrecke überprüfen. Lernaufgaben können sich auch primär auf die intendierte Verursachung eines konkreten Lernzuwachses beziehen. In engem qualitativen Zusammenhang hierzu stehen die großen Lernparadigmen: Behaviorismus, Imitationslernen, Instruktivismus, Konstruktivismus, *Learning bei Doing* usw. Die Testentwicklung umfasst natürlich die Entwicklung von Aufgaben und Aufgabenbatterien bzw. von Test-Items.

Wie werden summative Tests entwickelt? – Nachdem der Testentwickler den Einsatzzweck und den Einsatzort des Tests definiert hat, beantwortet er die Frage der **Testvalidität**. Dies korreliert mit mehreren Eingangsfragen: Was soll geprüft werden? Besteht eine eindeutige Kohärenz zwischen dem Was und dem Wie der Prüfung? Werden alle Fertigkeiten der Zielvorgaben berücksichtigt? Auf der Aufgaben- bzw. Item-Ebene korreliert dies mit der **Trennschärfe** der Items und der involvierten Verfahren. Grundsätzlich darf nur das geprüft werden, was geprüft werden soll, und zwar unter möglichst gleichen Bedingungen.

Man sollte gerade im Hinblick auf erzieherische Kontexte allerdings auch nicht übersehen: „Man bekommt, was man testet; was man nicht testet, bekommt man nicht.“ (bei Gage & Berliner 1996: 620) Es sind also stets die erzieherischen Ziele im Auge zu halten, was eine Relativierung der messbaren Kompetenzen gegenüber den nicht gemessenen/messbaren verlangt. Dies erfordert eine weitere Bemerkung: Fremdsprachenbezogene Kompetenztests sind Performanztests. Da die Grundgröße ‚kommunikative Kompetenz‘ eine nur schwer klar zu umreißende komplexe Größe darstellt, kommt es darauf an, diese Grundgröße zu ‚operationalisieren‘. Genau dies versuchen die Kompetenz-Deskriptoren des GeR.

GeR: Globalskala

Kompetente Sprachverwendung	C2	Kann praktisch alles, was er/sie liest oder hört, mühelos verstehen. Kann Informationen aus verschiedenen schriftlichen und mündlichen Quellen zusammenfassen und dabei Begründungen und Erklärungen in einer zusammenhängenden Darstellung wiedergeben. Kann sich spontan, sehr flüssig und genau ausdrücken und auch bei komplexeren Sachverhalten feinere Bedeutungsnuancen deutlich machen.
	C1	Kann ein breites Spektrum anspruchsvoller, längerer Texte verstehen und auch implizite Bedeutungen erfassen. Kann sich spontan und fließend ausdrücken, ohne öfter deutlich erkennbar nach Worten suchen zu müssen. Kann die Sprache im gesellschaftlichen und beruflichen Leben oder in Ausbildung und Studium wirksam und flexibel gebrauchen. Kann sich klar strukturiert und ausführlich zu komplexen Sachverhalten äußern und dabei verschiedene Mittel zur Textverknüpfung angemessen verwenden.
Selbstständige Sprachverwendung	B2	Kann die Hauptinhalte komplexer Texte zu konkreten und abstrakten Themen verstehen; versteht im eigenen Spezialgebiet auch Fachdiskussionen. Kann sich so spontan und fließend verständigen, dass ein normales Gespräch mit Muttersprachlern ohne größere Anstrengung auf beiden Seiten gut möglich ist. Kann sich zu einem breiten Themenspektrum klar und detailliert ausdrücken, einen Standpunkt zu einer aktuellen Frage erläutern und die Vor- und Nachteile verschiedener Möglichkeiten angeben.
	B1	Kann die Hauptpunkte verstehen, wenn klare Standardsprache verwendet wird und wenn es um vertraute Dinge aus Arbeit, Schule, Freizeit usw. geht. Kann die meisten Situationen bewältigen, denen man auf Reisen im Sprachgebiet begegnet. Kann sich einfach und zusammenhängend über vertraute Themen und persönliche Interessengebiete äußern. Kann über Erfahrungen und Ereignisse berichten, Träume, Hoffnungen und Ziele beschreiben und zu Plänen und Ansichten kurze Begründungen oder Erklärungen geben.
Sprachverwendung	A2	Kann Sätze und häufig gebrauchte Ausdrücke verstehen, die mit Bereichen von ganz unmittelbarer Bedeutung zusammenhängen (z. B. Informationen zur Person und zur Familie, Einkaufen, Arbeit, nähere Umgebung). Kann sich in einfachen, routinemäßigen Situationen verständigen, in denen es um einen einfachen und direkten Austausch von Informationen über vertraute und geläufige Dinge geht. Kann mit einfachen Mitteln die eigene Herkunft und Ausbildung, die direkte Umgebung und Dinge im Zusammenhang mit unmittelbaren Bedürfnissen beschreiben.
	A1	Kann vertraute, alltägliche Ausdrücke und ganz einfache Sätze verstehen und verwenden, die auf die Befriedigung konkreter Bedürfnisse zielen. Kann sich und andere vorstellen und anderen Leuten Fragen zu ihrer Person stellen – z. B. wo sie wohnen, was für Leute sie kennen oder was für Dinge sie haben – und kann auf Fragen dieser Art Antwort geben. Kann sich auf einfache Art verständigen, wenn die Gesprächspartnerinnen oder Gesprächspartner langsam und deutlich sprechen und bereit sind zu helfen.

In großen, summativen Tests werden all diese Fragen nicht allein von den Prüfungsautoren entschieden, sondern auch von externen Prüfern bzw. einem Test-Prüfungskollegium.

Unterhalb dieser hohen Schwelle findet sich die Augenschein-Validität (face validity), die einem Test-Item ad-hoc von Prüfern und Prüfling zugesprochen wird. Dies betrifft auch die **Transparenz**, welche vor allem im Hinblick auf Inhalt, Testformat und Bewertungskriterien bestehen muss.

Grotjahn erinnert an **kriterienbezogene Validität**, an **inhaltliche** und **Konstruktvalidität**. Ersterer überprüft die Kohärenz der Ergebnisse innerhalb eines Tests, die Kontextvalidität erfasst die Korrelation zwischen den Inhalten eines Tests und einem definierten ‚Lernstoff‘. Die kriterienbezogene und die Kontextvalidität haben somit eine besondere Funktion für ‚prognostische Tests‘. Sie sagen voraus, inwieweit ein Lerner eine vorgegebene Lernstrecke erfolgreich durchlaufen kann. Derlei ist bei ‚Einstufungstests‘ wichtig. Die Konstruktvalidität überprüft, inwieweit die „beobachteten Testergebnisse gültige Indikatoren von zugrunde liegenden theoretischen Konstrukten sind“ – so etwa der Fähigkeit zur lautlichen Dekodierung (Grotjahn 2002: 223).

Dem Entwurf der Aufgabe/des Items folgt die **Reliabilitätsprobe**. Diese beginnt mit der Eingangsfrage: Kommen die unterschiedlichen Probanden immer zu einem gleichen Ergebnis, unabhängig von Prüfern und Prüfungsorten? Die zuletzt erfassten Kriterien betreffen bereits die **Objektivität**: Inwieweit ist ein Messverfahren von den Prüfern abhängig? Idealerweise kommen die Prüflinge unabhängig voneinander (in Unkenntnis voneinander und an verschiedenen Orten) immer zu einem einheitlichen Ergebnis. Hier gibt es gewisse, definierte Toleranzgrenzen. In der Praxis der Durchführung summativer Tests bedarf es einer **Prüferschulung** und einer **Prüferzertifizierung**. Prüfungssysteme nach dem ALTE-Standard leisten dies.

Weitere wichtige Kriterien betreffen die Test-Ökonomie und die Authentizität. Letztere erhöht die Augenscheinvalidität. So wird es als ‚realitätsnah‘ (positiv) empfunden, wenn ein Item gemäß seines Materials und seiner Aufgabenführung mit einer kommunikativen interkulturellen Funktion korreliert. – Angesichts der fundamentalen Rolle des GeR füge ich des Weiteren das Kriterium der GeR-Konformität eines Tests hinzu. Die zuletzt genannten Eigenschaften stehen ihrerseits in engem Bezug zur **Adressatenrelevanz**. Hier lautet die Eingangsfrage: Sind die Lernziele und die curricularen Inhalte sowie das Format der Aufgaben den Teilnehmern vertraut? Selbstverständlich kann ein wenig reliabler und nicht-objektiver Test nicht valide sein (Grotjahn 2006: 223; Tagliante 2005: 31).

Die Psychometrie überprüft mit mathematischen bzw. statistischen Mitteln, also mit hoher Zuverlässigkeit, die Reliabilität und Validität von Tests. Dies betrifft Items, Testdurchführung und Testbewertung.

Die Eingangsfragen zur **Validität** wurden bereits genannt. Anschlussfragen lauten: Wie wird das Sprachcurriculum definiert? Es folgt die allgemeine Beschreibung der linguualen Lernziele (Operationalisierbarkeit), etwa die Auflistung des Wortschatzes, der Morphologie, der Syntax usw.; im pragmatisch-kommunikativen Bezirk die Erwähnung von Sprechsituationen, Themen, Sprechintentionen (materialisierbar durch Redemittel); aber dann auch der kommunikativen Strategien bzw. des interkulturellen Kommunikationsverhalten. Hier greifen durchaus die ‚ich kann-Beschreibungen‘ im Rahmen der Skalierung des GeR. Im Bereich des Mittleren Bildungsabschlusses dürfen die Lehrwerke der nächsten Generation hier eine neue Gewichtigkeit bekommen.

Folgefragen zur **Reliabilität** lauten dann: Sind alle Testversionen auf dem gleichen Niveau? Wird Reliabilität ‚systematisch‘ gesichert? Wird z.B. die Herstellung von Reliabilität dokumentiert? Welche Rolle spielt das Sprachcurriculum? Finden Vortests statt? Erfolgt eine Item-Analyse? Erfolgt eine Schulung der Prüfer? Erfolgt eine Lizenzierung der Prüfer? Wer kreditiert? Wie ist dessen Fachkompetenz? Sichert die Kreditierung die Aussagefähigkeit des Tests? Auch derlei Fragen entscheiden über die Qualität eines auf Qualitätssicherung bedachten Bildungswesens!

Aus dem folgenden Dokument des APIEL-Tests geht hervor, was die Attraktivität eines Sprachtests ausmacht:

Zunächst wird die Note erklärt: „4 = well qualified (...) a level of language proficiency sufficient to to conduct university studies in English.“ Ein *College Handbook* informiert über die 3.200 Universitäten und Colleges, welche das AP als Eingangsvoraussetzung für ein Studium anerkennen. Auch dies ist wichtig, denn der Prüfling will natürlich wissen: Was kann ich mit dem Zertifikat anfangen?

Soweit zu den Merkmalen summativer Tests. Zur *évaluation sommative* heißt es bei Tagliante (2005: 190): «S’opère en fin de formation et a pour fonction essentielle le classement des élèves les uns par rapport aux autres. Elle évalue un ensemble de paramètres désignés dans le programme d’enseignement.» Damit ist sie von der *évaluation formative* zu unterscheiden: «...s’opère pendant la formation et sert à donner de l’information sur l’état des performances et sur la réussite de l’enseignement.» Formative Tests verlangen Lernaufgaben, für welche die oben genannten Kriterien nur sehr eingeschränkt gelten, was sich schon allein mit dem Ökonomie-Prinzip erklärt.

Task Based Language Learning and Teaching

La notion de tâche provient de la recherche anglo-saxonne qui s’intéresse depuis longtemps à l’apprentissage fondé sur les tâches (Nunan 1989; Willis 1976). Ces dernières années, elle a été largement utilisée par le Cadre européen commun de références pour les langues.

Pour Nunan, la tâche est un ensemble structuré d’activités devant faire sens pour l’apprenant ; il s’agit de se confronter à un support authentique, par des activités de compréhension, de production, d’interaction avec les pairs, la visée pragmatique étant privilégiée par rapport à la forme linguistique. Concevoir une tâche implique la prise en compte de six paramètres: les *objectifs*, les *supports*, les *activités*, les *rôles* respectifs de l’enseignant et des apprenants, le *dispositif*. Willis ajoute un paramètre au cadre tracé par Nunan, celui du résultat de la tâche. Selon lui, une tâche communicative consiste à échanger des significations pour parvenir à un *résultat*. (Cuq 2003: 234)

Natürlich sind Aufgabenorientierung und *Task Based Language Learning* nicht dasselbe, dennoch haben sie vieles gemeinsam. Dies folgt vor allem aus zwei Gründen: aus der Handlungsorientierung des *Task Based Learning*, die schon ihrer inneren Natur nach in einem sehr engen Verhältnis zur Kompetenzorientierung steht (Tesch 2006). Ihrerseits ist also die Aufgabenorientierung sehr gut begründet, wie der von Bausch et al. (2006) herausgegebene Sammelband dokumentiert. Mit Blick auf Skehan (1998) weist Edmondson (2006: 55) „sinnvollen“ Aufgaben folgende Merkmale zu:

- Bedeutung hat Priorität gegenüber Form.
- Ein kommunikatives Problem wird angesprochen und gelöst.

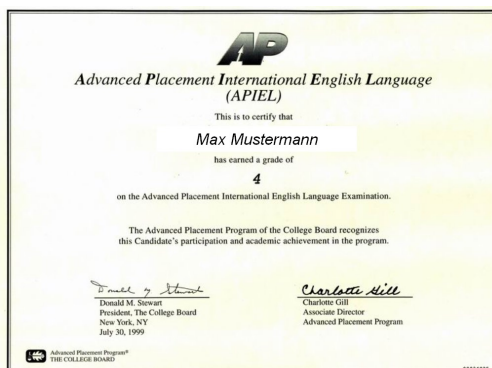
- Eine Vergleichbarkeit mit Handlungen außerhalb des Klassenzimmers ist gegeben.
- Es ist wichtig, dass die Aufgabe vollendet wird.
- Die Aufgabe wird bezüglich ihrer Ergebnisse evaluiert.

Sprachlernen und Spracherwerb sind bekanntlich hochkomplexe Vorgänge. Dementsprechend nennt die Sprachlehrforschung nicht nur Gründe, die Aufgabenorientierung als Grundlage einer besseren Kompetenzentwicklung zu begrüßen. Dies

beginnt schon bei der Frage nach den geeignetsten Aufgaben: “It may be impossible to claim that one task is better than another on psycholinguistic grounds.” (Ellis 2003: 127 bei Edmondson 2006b: 149). Der Grund für eine vorsichtige Beurteilung von ‘tasks’ liegt in der Individualität der Lerner: “Tasks cannot be externally defined and classified because the ‘activity’ that results from a task will vary from learner to learner and also within a single learner on different occasions.” (Ellis 2003: 199) Offensichtlich wohnt Aufgaben ein Potential inne, „unterschiedliche Beiträge zum Fremdspracherwerb zu leisten“ (Edmondson 2006b: 150). – Ein weiterer kritischer Punkt betrifft die Opposition ‚sequentiell‘ von Übungsvorgängen nach PPP-;Muster (pre-, present-, post) oder ‚integriert‘. Hierbei ist zu sehen, dass Sprechen z.B., erst recht dialogisches Sprechen, eine ganze Reihe parallel verarbeiteter mentaler Prozesse in Gang setzt. Der natürliche Spracherwerb zeigt vor allem integrierte, d.h. verschaltete Aktivierung verschiedener Teilfertigkeiten und ihrer Subroutinen. Das PPP-Muster erlaubt zwar eine gezielte Fokussierung auf bestimmte Fertigkeiten und Teilprozesse dieser Fertigkeiten, wodurch man sich eine zielgerichtete Kognitivierung erhofft, sie verhindert jedoch leicht die Verschaltung unterschiedlicher Fertigkeiten. Der Komplex ist bislang nicht erforscht. Empfehlung an die Praxis der Aufgabenentwicklung: Beide Muster, PPP und Integrierte Aufgaben alternieren lassen. Und eines ist wohl immer richtig: Auf die natürliche Sprachbenutzung schauen und die entsprechenden Schlüsse ziehen (wenn möglich). Nach Prüfung der umfangreichen Forschungsliteratur notiert Edmondson:

Folgende kognitive Prozesse und Ereignisse stehen in Verbindung mit den Zielen des Fremdsprachenunterrichts und können daher bei didaktischen Entscheidungen mitberücksichtigt werden:

- Grammatische Einsichten in Bezug auf eine Sprache beziehungsweise in Bezug auf Sprachen insgesamt werden gefördert.
- Ein Gewinn an Einsichten bezüglich menschlicher Kommunikation insgesamt und interkulturellen Unterschieden findet statt.
- Die Internalisierung/Speicherung einer sprachlichen Routine kommt zustande.
- Die Wahrnehmung eines neuen Aspekts der Zielsprache (Struktur, Verwendung, Lexik usw.) wird ermöglicht.
- Ein Vergleich zwischen einem Aspekt der Zielsprache und anderen Aspekten der Zielsprache bzw. anderer Sprachen wird gezogen.
- Ein sprachlicher Ausdruck, der neues sprachliches Material enthält, wird gespeichert.
- Verbindungen zwischen verschiedenen Aspekten der Zielsprache werden hergestellt (Vernetzung, Einsichten, Aha-Effekt).



- Neue Details über die Systematik der Zielsprache werden entdeckt und/oder verstanden.
- Die Übung und Automatisierung expliziten Wissens über die Systematik der Zielsprache findet statt.
- Der kontrollierte Einsatz von explizitem Wissen bei der fokussierten Sprachproduktion/Rezeption wird geübt (das Wissen steht z.B. gedruckt auf einem Arbeitsblatt).
- Explizites Wissen wird in Verbindung mit verschiedenen sprachlichen Umgebungen aktiviert und eingesetzt.
- Der kontrollierte Abruf und Einsatz von explizitem Wissen in der Sprachproduktion/Rezeption (das dem Lerner schon im Gedächtnis zur Verfügung steht) wird gefördert/ermöglicht/verlangt.
- Implizites Wissen wird aufgebaut.
- Implizites Wissen wird explizit zur Kenntnis genommen.
- Implizites/explicites Wissen wird in verschiedenen semantischen/pragmatischen/situativen Umgebungen aktiviert (produktiv/interpretativ) und somit vernetzt.
- Eine Diskrepanz zwischen dem eigenen Können und dem kommunikativen Bedarf oder dem vorliegenden Input wird zur Kenntnis genommen.
- Der Zugang zu implizitem und explizitem sprachlichen Wissen wird automatisiert.
- Relativ komplexe sprachliche Blöcke werden als ‚Chunks‘ gespeichert oder mindestens so gespeichert, dass sie relativ automatisch wiederhergestellt werden können.
- Eine Diskrepanz zwischen der eigenen Sprachproduktion und dem eigenen sprachlichen Wissen wird bemerkt (Selbst-Monitoring).
- Durch Monitoring bzw. eine andere Form der Selbstwahrnehmung wird die Selbsteinschätzung als Fremdsprachenlerner bzw. als Benutzer einer Fremdsprache erhöht: Dies führt vermutlich zu einer positiveren Einstellung und Motivation und daher indirekt zu weiteren Fortschritten beim Fremdspracherwerb. (Edmondson 2006b: 151/152)

Die hier aufgezählten Punkte stellen insgesamt eine Kriterienliste da, welche sich aus psycholinguistischer und spracherwerbstheoretischer Sicht auf ‚Lernaufgaben‘ beziehen lassen. Je mehr dieser Kriterien durch eine Aufgabe erfasst werden desto besser.

Kompetenzaufgaben zur Mündlichkeit nach ALTE-Standard

Doch kommen wir zu der Frage: Wie wird etwa Hörverstehen, wie Sprechen überprüft? Während die erste Frage in einem eigenen Aufsatz (in diesem Heft) breit angegangen wurde, blicken wir zu Beantwortung der zweiten kurz auf den Referenzrahmen und auf die Materialien des WBT (Weiterbildungstestsysteme GmbH)³.

Die Fertigkeit des Sprechens überprüft das WBT im Rahmen einer Einzel- oder Paarprüfung. In der Einzelprüfung übernimmt ein Prüfer die Rolle des Gesprächspartners. Eine gewisse Problematik ergibt sich aufgrund der Doppelrolle des Prüfers (ohne Peer), denn die Mehrfachbelastung seines Arbeitsspeichers durch die Simultanität von Sprechen und Hören, Gesprächssteuerung und Fokussierung (Protokollierung) auf die (Lerner)sprachliche Performanzleistung des Prüflings (*fluency*, formale Korrektheit, pragmatische Angemessenheit) führt (leicht) zu einer Überforderung des Prüfers bzw. zu dessen selektiver Wahrnehmung. Gilt prüfungsseitig das ‚vier Augen-Prinzip‘, so ist das Risiko geringer.

Ein Gespräch umfasst bei WBT drei Teile: 1) Kontaktnahme, 2) Gespräch über ein Thema, 3) gemeinsam eine Aufgabe lösen. In Teil 1 ist beabsichtigt, dass die Partner einander näher kom-

men können und eine Partnerhypothese aufgebaut werden kann (vgl. Meißner in diesem Heft). Demgemäß stehen hier Fragen im Vordergrund bezüglich von: Name, Wohnort, Erfahrung am Wohnort, Freizeit, Auslandserfahrung... Das dann folgende Prüfungsgespräch dauert 5 bis 6 Minuten. Ziel ist es: über etwas sprechen, Ideen und Meinungen darüber austauschen. (WBT Deutsch 1999: 390f)⁴. Ein Beispiel aus dem Arbeitsleben:

Eine Arbeitskollegin wird in einer anderen Firma arbeiten. Sie haben die Aufgabe, zusammen mit Ihrem Gesprächspartner/Ihrer Gesprächspartnerin eine Abschiedsparty zu organisieren. Überlegen Sie sich, was alles zu tun ist und wer welche Aufgaben übernimmt.

Als Kriterien für die Bewertung werden genannt:

- Ausdrucksfähigkeit (nach Inhalt, Rolle, Wortschatz, Verwirklichung der Sprechabsicht)
- Aufgabenbewältigung (Gesprächsbeteiligung, kommunikative Strategien inkl. Kompensationsstrategien, Diskursstrategien)
- Formale Richtigkeit (Syntax, Morphologie, Tempi...)
- Aussprache und Intonation.

Dem WBT erschien es offenbar günstig, innerhalb eines Positiv-Punkte-Systems zu bewerten. Die für die Zielsprache Italienisch vorgesehene, uns vorliegende Fassung nennt für den Verlauf des Gesprächs:

Criterion		A	B	C	D
1	Capacità espressiva	4 punti	6 punti	2 punti	0 punti
2	Padronanza del compito	4 punti	6 punti	2 punti	0 punti
3	Competenza linguistica	4 punti	6 punti	2 punti	0 punti
4	Pronuncia e intonazione	4 punti	6 punti	2 punti	0 punti

Bezüglich der vier Fertigkeiten erläutert die der Zielsprache Italienisch gewidmete Version:

Capacità espressiva und padronanza del compito

A	adeguata sotto tutti gli aspetti
B	adeguata sotto la maggioranza degli aspetti
C	per lo più inadeguata
D	completamente inadeguata

Competenza linguistica

A	Non fa nessun errore o fa solo occasionali
B	Fa alcuni errori che non impediscono la comprensione delle sue asserzioni
C	Fa errori che impediscono la comprensione di importanti parti delle sue asserzioni
D	Fa una quantità di errori che rendono praticamente incomprensibile le sue asserzioni

Pronuncia

A	Non impediscono la comprensione
B	Possono occasionalmente impedire la comprensione e richiedono concentrazione particolare per parte dell'ascoltatore
C	Impediscono considerevolmente la comprensione
D	Rendono la comprensione (quasi) impossibile

In der Summe sind für alle Fertigkeiten 300 Punkte erreichbar. Es ist offensichtlich, dass das System, mit dem man gute Erfahrungen gemacht zu haben vorgibt, die Schulung der Prüfer verlangt. – Auch dies ist eine neue Aufgabe für die Lehrerfortbildung!

Ausblick

Die Erstellung von Kompetenzaufgaben und die Weiterentwicklung des herkömmlichen Unterrichts zur Kompetenzorientierung ist ein schwieriges Geschäft. – Lehrende haben es nun vor allem mit weniger aufwendigen Lernaufgaben zu tun. Diejenigen unter ihnen, welche mit der Entwicklung von summativen Kompetenzaufgaben betraut werden (Zentralprüfung für den Mittleren Bildungsabschluss oder Zentralabitur), werden allerdings Aufgaben vorlegen müssen, die dem Kriterium der Standardisierung bzw. von formativen Tests entsprechen. Dass indes Aufgaben aus zentralen, landesweit durchgeführten Prüfungen eine Rückwirkung auf die Art der Lernaufgaben haben, ergibt sich schon aus der indirekten Rückbindung der letzteren an die ersteren. Denn Lernaufgaben überprüfen nicht nur das auf einer Lernstrecke erworbene Wissen und Können (Lernerfolgsüberprüfung), sondern bereiten auch mehr oder weniger auf die finalen Prüfungen und deren Formate vor. Lernaufgaben haben daher eine retro- ‚und‘ prospektive Funktion. Deshalb sei abschließend erwähnt, dass die Vorbereitung von Lernenden auf Prüfungen auch immer Vorbereitung auf Prüfungsformate sind. Ein dosiertes Maß von *teaching to the test* ist daher unerlässlich und notwendig. – Das erste Gefahrenmoment besteht, wie vielfach moniert, darin, dass *teaching to the test* zum latent leitenden und dominanten Prinzip des Unterrichts werden kann; das zweite, dass *teaching to the test* in sachlich unangemessener Weise erfolgt, weil es lehrseitig an ‚Testkompetenz‘ fehlt; die dritte darin, dass die Kernfunktion von Tests und Standards von den Bildungsbehörden, den Lehrenden und der Öffentlichkeit fehlgedeutet wird. Dies tritt bei einer unkritischen ‚Gläubigkeit‘ an Test- und Messergebnissen und bei mangelnder Testkompetenz ein:

Wir leben in einer Welt, die den Beweis anhimmt (...) und wir wollen den Beweis dafür, dass die Schulen unseres Landes etwas taugen. (...) Diese Beweissucht hat sich in höchst dramatischer Weise in den Bildungssektor hineingedrängt – und das sollte Pädagogen dazu zwingen, einer fundamentalen Wahrheit Bedeutung zu schenken. In einem beweisorientierten Unternehmen haben jene, die den Mechanismus zur Beweissammlung in der Hand haben, auch das ganze Unternehmen in der Hand. Da im Bereich der Pädagogik unser Hauptmechanismus zur Beweissammlung der Test ist, ist es einleuchtend, dass all jene, die etwas mit Erziehung zu tun haben, sich Kenntnisse über die Grundlagen des Messens im Unterricht verschaffen sollten. (Popham 1981: 5-6, nach Gage & Berliner 1996: 623)

Und wir fügen den Finalsatz hinzu: ... nicht allein um zu testen, sondern auch um Tests einzuschätzen und gegebenenfalls relativieren zu können. Nochmals: Hierzu gehört auch die Einsicht, dass der aus dem Bezirk der Ökonomie stammende Begriff der *Output*-Orientierung im Bildungsbereich stark verkürzt – und dadurch die angestrebte hohe Qualität gefährdet. Denn die Qualität des ‚Bildungsprodukts‘ ist immer abhängig vom Bildungsprozess. Gerade das Lernziel ‚Optimierung der Lernkompetenz‘ verlangt von den Lernenden die Auseinandersetzung mit den eigenen Lernprozessen; von den Lehrenden, den Lernenden hierbei kompetent zu helfen und ihren Lernen in seinen Voraussetzungen und Schritten einsichtiger zu machen. Diejenigen Länder, welche bei den internationalen Vergleichsstudien gut abschlossen, verfügen über Bildungssysteme, die dies leisten. Sie bieten abgestuft und zielgenau *opportunity to learn*-Standards. Diese Standards umfassen sowohl den Stoff als auch das Lernen am Stoff. Der GeR weiß dies.

Deutschland im Bildungsumbruch? Erinnern wir daran, wie zur Zeit in Deutschland (noch) etwa die Abitursprüfung gehandhabt wird: Während landauf landab die Überprüfung der Lese- und Schreibkompetenz keine sonderlichen Schwierigkeiten bereitet, sieht es für die Mündlichkeit überwiegend anders aus. Angesichts des größeren prüfungstechnischen Aufwandes und des Versäumnisses, die Schüler durch Lernaufgaben entsprechend vorzubilden, mussten sich mehrere Bundesländer zu einem Aufschub der Einführung mündlicher Fertigkeitstests entschließen. Dies kann durchaus als ein Kompromiss an eine ‚unzulängliche Lehrpraxis‘ gedeutet werden – sofern man den GeR ernst nimmt. Damit ist kein Vorwurf an die Lehrenden gemeint, sondern eine nachhaltige Empfehlung an die Lehrerfortbildung, die Materialentwicklung und die Notwendigkeit einer zielgerichteten Qualitätssicherung im Bereich der Fortbildung. Sollten allerdings die Länder den jetzigen Prüfungszustand dennoch perpetuieren, so hätten sie zugleich auch die eigenen Versuche zur Qualitätssicherung konterkariert. Es wäre absurd, Qualitätssicherung durch die Formulierung von Standards zu implementieren, aber die qualitätsgerechte Umsetzung der Standards zu versäumen.

Nachbemerkung zur Spezifik des Französischen im Vergleich zu Englisch: Reformverdrossene könnten die Bedeutung der Standards für den Französischunterricht als erster Fremdsprache mit dem Argument abtun wollen, dass das Lernerkontingent von Französisch insgesamt nur im Umfang von ca. 5 Prozent der Gymnasiasten (die Französisch als erste Fremdsprache wählen) betroffen sei. Abgesehen davon, dass diese Zahl insoweit der Ergänzung bedarf, als es in Baden-Württemberg, Berlin, Hessen, Rheinland-Pfalz und dem Saarland im Bereich der Grundschule durchaus starke Anteile von Französischlernern gibt, dass man über 80 bilinguale deutsch-französische Züge zählt, in denen Französisch durchaus in der 5. Jahrgangsstufe unterrichtet wird, dass sich ebenfalls Realschulen mit Französisch als erster Fremdsprache finden lassen, ist das Argument dennoch nicht haltbar. Denn es ist durchaus nicht zu leugnen:

- dass die Standards für die 1. auf die 2., 3. und 4. Fremdsprachen rückwirken.
- dass sich Qualitätssicherung nicht nur auf einen Teil der Schulfächer begrenzen lässt.
- dass, da valide Tests Ausdruck von Qualitätssicherung sind, kein Fach den Eindruck erwecken wollen wird, auf ‚Qualität‘ zu verzichten. Jedes Fach wird daher, wie auch immer, auf den Trend zur Qualitätssicherung positiv reagieren müssen.
- dass, da summative Tests mit einer breiten Publikumserwartung korrelieren, die direkten (Lerner) und indirekten Abnehmer von Fremdsprachenunterricht (Firmen...) Standards fordern werden. Daher werden absehbar Bildungsstandards und GeR auch für die 2., 3. und für die nachfolgenden Fremdsprachen eine zentrale Rolle spielen.

- dass, *last but not least*, die an kompetenzorientiertem Unterricht gewöhnten Schüler ein entsprechendes implizites methodisches Vorverständnis auf den Unterricht der weiteren Fremdsprachen projizieren.

Das Schulfach Französisch ist, natürlich von Lernkontext zu Lernkontext unterschiedlich, von dem hier behandelte ‚Trend‘ als Grundschulfremdsprache, als 1., 2., 3. und manchmal auch 4. Fremdsprache betroffen – im Tertiärbereich allemal.

Anmerkungen

¹ Im Netz abrufbar unter: www.desi.de.

² Zum Pressecho der Studie u.a.: H. Schmoll: Schwach in den Sprachen. Deutsche Schüler erfüllen den EU-Standard von zwei modernen Fremdsprachen nicht. *Frankfurter Allgemeine Zeitung*. 31.12.2004, 12. A. Pribošček: Schüler lernen zu wenig Sprachen. *Rheinische Post*. 18.01.2005, 1.

³ Das Zertifikat Deutsch wurde in gemeinsamer Verantwortung des Deutschen Instituts für Erwachsenenbildung (DIE), des Goethe-Instituts, der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren, dem Österreichischen Sprachdiplom Deutsch und dem WBT entwickelt.

⁴ Vgl. zu FLE-Übungen: Tagliante 2001; Veltcheff & Hilton 2003; Noël-Jothy 2004; Pons & Karcher 2006:

Literatur

- Baumert, J. (1993): Lernstrategien, motivationale Orientierung und Selbstwirksamkeits-überzeugen im Kontext schulischen Lernens. *Unterrichtswissenschaft* 21(4), 327–354.
- Bausch, K.-R., Burwitz-Melzer, E., Königs, F. G. & Krumm, H.-J. (Hrsg.) (2005): *Bildungsstandards für den Fremdsprachenunterricht auf dem Prüfstand. Arbeitspapiere der 25. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts*. Tübingen.
- Bausch, K.-R., Burwitz-Melzer, E., Königs, F. G. & Krumm, H.-J. (Hrsg.) (2006): *Aufgabenorientierung als Aufgabe. Arbeitspapiere der 25. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts*. Tübingen.
- Bausch, K.-R., Christ, H., Königs, F. G. & Krumm, H.-J. (Hrsg.) (2003): *Der Gemeinsame europäische Referenzrahmen für Sprachen in der Diskussion. Arbeitspapiere der 22. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts*. Tübingen.
- Beacco, J.-C., Bouquet, S. & Porquier, R. (2004): *Niveau B2 pour le français. Un référentiel*. Paris.
- Bolton, S. (2004): Qualitätsmanagement bei den Prüfungen der Association of Language Testers in Europe (ALTE). *Fremdsprachen und Hochschule* 69, 9-22.
- Bolton, S. (1997): Tests im Wandel theoretischer Prämissen. In: Gardenghi & O’Connell, 17-24.
- Burwitz-Melzer, E. & Quetz, J. (2006): Mehrsprachigkeit und Interkulturalität in Sprachenportfolios. In: Martinez et al., 203-214.
- Byram, M. (ed.) (2004): *Routledge Encyclopedia of Language Teaching and Learning*. London/New York.
- Christ, I. (2006a): Lassen Bildungsstandards Raum für Mehrsprachigkeit? In: Martinez et al., 255-268.
- Christ, I. (2006b): Wozu lernt man heute fremde Sprachen? In: P. Scherfer & D. Wolff (Hrsg.): *Vom Lehren und Lernen fremder Sprachen: Eine vorläufige Bestandsaufnahme*. Frankfurt a.M., 39-68.
- Doyé, P. (2004): Gütekriterien für den Sprachunterricht. *Neusprachliche Mitteilungen* 57/3, 130-134.
- Doyé, P. (2006): Allgemein-pädagogische Aspekte einer Didaktik der Interkomprehension. In: Martinez et al., 13-27.
- Eco, U. (1994): *Die Suche nach der vollkommenen Sprache*. München.
- Edmondson, W. (2006a): Konversationsanalyse und Lehrerverhalten im Fremdsprachenunterricht. In: Jung, 51-57.
- Edmondson, W. (2006b): *Taking Tasks to Task*. Aufgabenbasierte Forschung und Lernaufgaben im Fremdsprachenunterricht. In: H.-P. Timm (Hrsg.): *Fremdsprachenlernen und Fremdsprachenforschung: Kompetenzen, Standards, Lernformen, Evaluation. Festschrift für Helmut Johannes Vollmer*. Tübingen, 139-156.
- Gage N. L. & Berliner, D. C. (1996): *Pädagogische Psychologie*. 5. vollständig überarbeitete Auflage. (Herausgegeben und aus dem Amerikanischen übersetzt von Gerhard Bach). Weinheim/München.
- Gardenghi, M. & O’Connell.M. (Hrsg.) (1997): *Prüfen, Testen, Bewerten im modernen Fremdsprachenunterricht*. Frankfurt a.M.

- Grotjahn, R. (Koord.) (2001): Leistungsmessung und Leistungsevaluation. *Fremdsprachen Lehren und Lernen* 30.
- Klieme, E. (2005): Bildungsqualität und Standards. Anmerkungen zu einem umstrittenen Begriffspaar. *Friedrich Jahresheft XXIII*, 6–7.
- KMK (2003): *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss. Beschluss vom 4.12.2003*. KMK.
- Küster, L. (2006): Auf dem Ordnungswege: Zu Risiken und Nebenwirkungen der Bildungsstandards für die erste Fremdsprache. *Der fremdsprachliche Unterricht Englisch* H. 81, 18–21.
- Long, M. H. & Norris, J. M. (2004): Task-based teaching and assessment. In: Byram, 597–603.
- Martinez, H., Reinfried, M. & Bär, M. (Hrsg.) (2006): *Mehrsprachigkeitsdidaktik gestern, heute und morgen. Festschrift für Franz-Joseph Meißner zum 60. Geburtstag*. Tübingen: Narr 2006.
- Meißner, F.-J. (2006): Linguistische und didaktische Überlegungen zur Entwicklung von Kompetenzaufgaben im Lernbereich Mündlichkeit (Schwerpunkt Hörverstehen). *französisch heute* 37, (in diesem Heft)
- Meißner, F.-J. & Lang, A. (2005): Fremdsprachenunterricht in Deutschland in der Jahrgangsstufe 12 des Gymnasiums im Europäischen Jahr der Sprachen (2000/1). *Zeitschrift für Fremdsprachenforschung* 16(2), 187–216.
- Müller-Hartmann, A. & Schocker-von Dithfurth (Hrsg.) (2005): *Aufgabenorientierung im Fremdsprachenunterricht. Task-Based Language Learning and Teaching. Festschrift für Michael K. Legutke*. Tübingen.
- Müßigmann, C. & Philipp, E. (2006): Aus der Praxis: Aufgabenentwicklung Französisch im Rahmen der Bildungsstandards. *französisch heute* 37 (in diesem Heft).
- Noël-Jothy, F. & Sampsonis, B. (2006): *Certifications et outils d'évaluation en FLE*. Paris.
- North, B. (2005): Assessing Spoken Performance in relation the Common European Framework of Reference. *Babylonia* 2/05, 46–49.
- North, B. (dir.) (2006): *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEF). (Manual Pilot Version)*. Strasbourg.
- Nunan, D. (1989): *Designing Tasks for the Communicative Classroom*. Cambridge.
- Pons, S. & Karcher, G. (2006): *TEF. Test d'évaluation de français. 250 activités*. Paris.
- Popham, W. J. (1990): *Modern Education Measurement*. Engelwoods Cliffs.
- Schelten, A. (1980): *Grundlagen der Testbeurteilung und Testerstellung*. Heidelberg.
- Schröder, K., Harsch, C. & Nold, G. (2006): DESI – Die sprachpraktischen Kompetenzen unserer Schülerinnen und Schüler im Bereich Englisch. Zentrale Befunde. *Neusprachliche Mitteilungen* 59(3), 11–32.
- Skehan, P. (1998): *A cognitive approach to language learning*. Oxford.
- Tagliante, C. (2001): *La classe de langue*. Paris.
- Tagliante, C. (2002): Techniques d'évaluation en langues. Le contexte européen. *Le français dans le monde* 324, 30–33.
- Tagliante, C. (2005): *L'évaluation et le Cadre européen commun*. Paris.
- Tesch, B. (2006): Bildungsstandards für die erste Fremdsprache Englisch/Französisch: Kompetenzen, Kompetenzmodelle und kompetenzorientierter Unterricht. *französisch heute* 37 (in diesem Heft).
- Veltcheff, C. & Hilton, S. (2003): *L'évaluation en FLE*. Paris.
- Vollmer, H. J. (2006): Bildungsstandards von oben – Bildungsstandards von unten. *Der fremdsprachliche Unterricht Englisch* H. 81, 12–17.
- WBT-Deutsch (1999): *Zertifikat Deutsch. Lernziele und Testformat*. Frankfurt a.M.
- WBT-Italienisch (2002): *Obiettivi didattici e struttura del Certificato*. Frankfurt a.M.
- Willis, J. (1996): *A Framework for Task-Based Learning*. London.

<http://coe>

<http://www.alte.org>

<http://coep.fr>

<http://www.primlangues.fr>

<http://recherche.aix-mrs.iufm.fr/publ/voc/n1/amigues4/>