Zeitschrift: Schweizerische Zeitschrift für Bildungswissenschaften = Revue suisse des sciences de l'éducation = Rivista svizzera di scienze dell'educazione

Herausgeber: Schweizerische Gesellschaft für Bildungsforschung

Band: 27 (2005)

Heft: 1

Artikel: Der "gemeinsame europäische Referenzrahmen für Sprachen" als Grundlage von Bildungsstandards für die Fremdsprachen : methodologische Probleme der Entwicklung und Adaptierung von Kompetenzbeschreibungen

Autor: Schneider, Günther

DOI: https://doi.org/10.5169/seals-786563

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 20.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Der «Gemeinsame europäische Referenzrahmen für Sprachen» als Grundlage von Bildungsstandards für die Fremdsprachen – Methodologische Probleme der Entwicklung und Adaptierung von Kompetenzbeschreibungen

Günther Schneider

Der «Gemeinsame europäische Referenzrahmen für Sprachen» gilt in der Diskussion um die Entwicklung von Bildungsstandards als Beispiel für ein ausgearbeitetes Kompetenzmodell und auch als eine gute Grundlage für die Festlegung von Kompetenzniveaus für die Fremdsprachenfächer an der obligatorischen Schule in der Schweiz im Rahmen des Projekts HARMOS. Nach einer kurzen Darstellung der Wirkung des Referenzrahmens beschreibt der Beitrag das Kompetenzmodell des Referenzrahmens und die Methoden, mit denen die Kompetenzniveaus entwickelt wurden. Anhand von Beispielen wird gezeigt, welche Probleme bei der Auswahl, Adaptierung und Ergänzung von Niveaubeschreibungen entstehen können und in welchen Bereichen noch Entwicklungsarbeit zu leisten ist.

Der Referenzrahmen und seine Wirkung

Der «Gemeinsame europäische Referenzrahmen für Sprachen» hat in kurzer Zeit eine erstaunliche Resonanz und Akzeptanz gefunden. Er ist inzwischen in vielen Ländern Europas als Grundlage für Reformen im Fremdsprachenunterricht, für die Entwicklung von Bildungsstandards sowie für die Entwicklung neuer Curricula, Lehrwerke, Prüfungen oder Selbstbeurteilungsinstrumente genutzt worden.

Die Veröffentlichungen der englischen, französischen und deutschen Fassung im Jahr 2001, denen inzwischen viele weitere Sprachfassungen gefolgt sind, machten das Werk breit bekannt. Im Folgenden wird in der Regel die deutsche Fassung mit dem Titel «*Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*» « (Europarat, 2001, abgekürzt als «Referenzrahmen» oder GER) zitiert.

Im Umkreis des Referenzrahmens sind verschiedene ergänzende und stützende Arbeiten des Europarats entstanden. Man kann sieben Gruppen unterscheiden:

ISSN 1424-3946, Academic Press Fribourg

1. An erster Stelle ist das Europäische Sprachenportfolio (kurz ESP) zu nennen, in dem zentrale Aspekte des Referenzrahmens umgesetzt sind. Die zahlreichen Initiativen zur Entwicklung und Implementierung verschiedener Versionen des Sprachenportfolios haben wesentlich zur Bekanntheit und zur Akzeptanz der Niveaubestimmungen des Referenzrahmens z.B. bei Fremdsprachenlehrerinnen und -lehrern beigetragen. Ein Kernelement aller Portfolioversionen sind Selbstbeurteilungsinstrumente mit Kann-Beschreibungen für die Referenzniveaus des Europarats (Schneider, North & Koch, 2001; Schneider & Lenz, 2001; Lenz, 2004a , 2004b).
2. Als Einführung und Hilfe zum Gebrauch des Referenzrahmens gibt es Handreichungen für verschiedene Benutzergruppen, z. B. Lehrpersonen verschiedener Stufen, Lehrerbildner, Curriculumentwickler, Lehrwerkautoren, Prüfungsanbieter oder Verantwortliche für Qualitätsentwicklung («Guide for Users», Council of Europe, 2002a).
3. In einem Sammelband wurde vom Europarat eine Reihe von Fallberichten zur Verwendung des Referenzrahmen in verschiedenen Kontexten veröffentlicht, die teilweise mit wissenschaftlichen Untersuchungen zur Überprüfung und Weiterentwicklung von Elementen des Referenzrahmens verbunden waren («Case studies», Council of Europe, 2002b).
4. Aufbauend auf dem Referenzrahmen wurden und werden Lernzielbestimmungen für verschiedene Sprachen ausgearbeitet, welche die bekannten Lernzielbeschreibungen des Europarats aus den 70ger und 80ger Jahren wie «Waystage», «Threshold Level», «Un Niveau seuil», «Kontaktschwelle» usw. ersetzen, aktualisieren oder ergänzen (z. B. Glaboniat, Müller, Rusch, Schmitz & Wertenschlag, 2002; Beacco, Bouquet & Porquier, 2004).
5. Der Referenzrahmen berücksichtigt zwar unter den «Allgemeinen Kompetenzen» auch soziokulturelles Wissen, interkulturelles Bewusstsein und interkulturelle Fertigkeiten, enthält aber noch keine Beschreibungen von Niveaus interkultureller Kompetenz. Verschiedene Auftragsarbeiten der Abteilung Sprachpolitik in Strassburg und Projekte des Europäischen Fremdsprachenzentrums in Graz dienen dem Ziel, die Ergänzung dieser Dimension in späteren Ausgaben des Referenzrahmens vorzubereiten (Byram & Zarate 1997; Byram, 2003; Zarate, Gohard-Radenkovic, Lussier & Penz, 2003).
6. In einer Pilotfassung liegt ein Handbuch zur Zuordnung von Sprachprüfungen zum Referenzrahmen vor (Council of Europe, 2003). Verbunden mit der Erprobung des Handbuchs entstehen verschiedene Hilfsmittel für die Zuordnung von Sprachprüfungen und Tests zum Referenzrahmen, z. B. Videoaufzeichnungen, welche mit typischen Lernerleistungen die Referenzniveaus illustrieren, sowie kalibrierte Test-Items zum Hör und Leseverstehen auf den verschiedenen Niveaus. Nach der Erprobung soll das Handbuch 2007 zusammen mit Fallstudien publiziert werden.
7. Eine letzte Gruppe bilden die Arbeiten des Europarats zur Sprachenpolitik, ein Aspekt, der im Referenzrahmen nur kurz und zurückhaltend behandelt

wird («Guide for the development of language education policies in Europe», Beacco & Byram, 2003).

Neben diesen Arbeiten des Europarats gibt es eine Vielzahl von Arbeiten und Dokumenten aus verschiedenen Ländern, die sich explizit auf den Referenzrahmen und besonders auf seine Niveaubeschreibungen stützen. Besonders relevant in unserem Zusammenhang sind das Schweizer Sprachenkonzept (EDK, 1998; Lüdi, 2003) und die Arbeiten zur Entwicklung von Bildungsstandards in Deutschland (Klieme et al., 2003; Beck & Klieme, 2003; KMK, 2004; Vollmer, 2003).

Mit der für manche überraschend grossen Wirkung, die der Referenzrahmen auf die Entwicklung von Prüfungen, Lehrwerken, Kerncurricula oder die Formulierung von Bildungsstandards ausübte, wobei der Rückgriff oder die Berufung auf den Referenzrahmen mehr oder weniger begründet, mehr oder weniger differenziert und mehr oder weniger seriös sein kann, setzte zunächst in Deutschland, später besonders im Kreis der Spezialisten für Sprachtests[1] eine teilweise sehr kritische Auseinandersetzung mit dem Referenzrahmen ein. (Bausch, Christ, Königs & Krumm, 2003; Krumm, 2003; Quetz, 2003; Morrow, 2004).

Wie Schneider (2003) zeigt, zieht sich ein seltsamer Widerspruch durch die Kritik. Zum einen fürchtet man, dass der Referenzrahmen zu stark normierend wirken könne, zum andern ruft man nach mehr Empfehlungen und kritisiert zum Beispiel, dass keine Bildungsziele und Unterrichtsmethoden empfohlen[2] und keine eindeutigen Anweisungen für die Ableitung von Testaufgaben gegeben werden. Aber der Referenzrahmen will als ein Instrument für Praktiker Grundlagen für Entscheidungen durch die Benutzer bereitstellen. Der Referenzrahmen nimmt niemandem die Entscheidungen ab, sondern hilft dabei, indem er aufzeigt, welche Fragen man sich stellen und beantworten sollte, und er gibt Hinweise für die Beantwortung der Fragen. Das gilt auch im Hinblick auf die Entwicklung von Bildungsstandards auf der Basis des Referenzrahmens.

Der nächste Abschnitt erläutert, inwiefern der Referenzrahmen Grundlage für die Entwicklung von Bildungsstandards sein kann.

Das Kompetenzmodell des Referenzrahmens

Kompetenzmodelle, so heisst es in der Expertise im Auftrag des deutschen Bundesministeriums für Bildung und Forschung, «erfüllen in Bezug auf Bildungsstandards zwei Zwecke: erstens beschreiben sie das Gefüge der Anforderungen, deren Bewältigung von Schülerinnen und Schülern erwartet wird (Komponentenmodell); zweitens liefern sie wissenschaftlich begründete Vorstellungen darüber, welche Abstufungen eine Kompetenz annehmen kann bzw. welche Grade oder Niveaustufen sich bei den einzelnen Schülerinnen und Schülern feststellen lassen (Stufenmodell)» (Klieme et al., 2003, 74).

Der Referenzrahmen entspricht diesen Anforderungen, denn er unterscheidet 1. basierend auf anerkannten linguistischen Modellen verschiedene Dimensionen/Komponenten der kommunikativen Kompetenz und gibt 2. Bezugspunkte (Niveaus) für die Einstufung von Lernfortschritten. Die Niveaubeschreibungen wurden, wie im nächsten Teil gezeigt werden soll, auf empirischem Weg skaliert. Zudem berücksichtigt der Referenzrahmen, dass die Entwicklung der kommunikativen Kompetenz nicht nur sprachliche Dimensionen umfasst, sondern auch soziokulturelle Bewusstheit, Erfahrung und Phantasie, Motivation, affektive Beziehungen, Lernen lernen etc.. In der Expertise wird denn auch der Referenzrahmen verschiedentlich als «Beispiel für die Entwicklung eines Kompetenzmodells», so die Überschrift des Anhangs A, angeführt (Klieme et al. 2003, 100, 137, 146-156).

Der Referenzrahmen basiert auf einem handlungsorientierten Ansatz, der so zusammengefasst wird:

> *Sprachverwendung – und dies schließt auch das Lernen einer Sprache mit ein – umfasst die Handlungen von Menschen, die als Individuen und als gesellschaftlich Handelnde eine Vielzahl von Kompetenzen entwickeln, und zwar allgemeine, besonders aber kommunikative Sprachkompetenzen. Sie greifen in verschiedenen Kontexten und unter verschiedenen Bedingungen und Beschränkungen auf diese Kompetenzen zurück, wenn sie sprachliche Aktivitäten ausführen, an denen (wiederum) Sprachprozesse beteiligt sind, um Texte über bestimmte Themen aus verschiedenen Lebensbereichen (Domänen) zu produzieren und/oder zu rezipieren. Dabei setzen sie Strategien ein, die für die Ausführung dieser Aufgaben am geeignetsten erscheinen. Die Erfahrungen, die Teilnehmer in solchen kommunikativen Aktivitäten machen, können zur Verstärkung oder zur Veränderung der Kompetenzen führen* (GER, 21).

Die im Original fett hervorgehobenen Begriffe bilden die wesentlichen Beschreibungskategorien des Referenzrahmens.

«Kompetenzen» werden definiert als «Summe des (deklarativen) Wissens, der (prozeduralen) Fertigkeiten und der persönlichkeitsbezogenen Kompetenzen und allgemeinen kognitiven Fähigkeiten, die es einem Menschen erlauben, Handlungen auszuführen» (GER, 21).

Zu den «Allgemeine Kompetenzen», die im Referenzrahmen als Komponenten beschrieben, für die jedoch keine Kompetenzstufen unterschieden werden, gehören 1. Deklaratives Wissen (savoir), 2. Fertigkeiten und prozedurales Wissen (savoir-faire), 3. Persönlichkeitsbezogene Kompetenzen (savoir-être) und 4. Lernfähigkeit (savoir-apprendre). Im Einzelnen werden unterschieden:

- Weltwissen (z.B. Geographie, Geschichte, Politik)
- Soziokulturelles Wissen (tägliches Leben, interpersonale Beziehungen, Werte, Körpersprache, u.a.)
- Interkulturelles Bewusstsein (Ähnlichkeiten/Unterschiede von Kulturen)
- Praktische Fertigkeiten (z. B. für Alltag, Beruf, Freizeit)
- Interkulturelle Fertigkeiten (Kulturen in Beziehung setzen, Strategien für

Kontakt mit Angehörigen anderer Kulturen verwenden, als kultureller Mittler agieren)
- Persönlichkeitsbezogene Kompetenz (Einstellungen, Motivation, Wertvorstellungen, kognitiver Stil, u.a.)
- Lernfähigkeit (Lerntechniken, mit neuen Erfahrungen umgehen).

Bei der kommunikativen Kompetenz im engeren Sinn, d. h. den kommunikativen Sprachkompetenzen, unterscheidet der Referenzrahmen drei Gruppen:
- Linguistische Kompetenzen (phonologische, lexikalische und grammatische Kompetenz)
- Soziolinguistische Kompetenzen (z. B. Kenntnis gesellschaftlicher Sprachkonventionen, Varietäten)
- Pragmatische Kompetenzen (Kenntnis der Sprachfunktionen, Diskurskompetenz)

Zu den kommunikativen Sprachkompetenzen enthält der Referenzrahmen 13 Einzelskalen (GER, 110-129).

Die nicht der Beobachtung direkt zugänglichen kommunikativen Sprachkompetenzen zeigen sich in «kommunikativen Sprachaktivitäten». Abweichend von der klassischen Einteilung in die vier Fertigkeiten Hören, Lesen, Sprechen und Schreiben unterscheidet der Referenzrahmen die vier sprachlichen Handlungsbereiche Rezeption, Interaktion, Produktion und Sprachmittlung[3], wobei jeweils zwischen mündlich und schriftlich unterschieden wird.

Kapitel 4 des Referenzrahmen, welches das komplexe Kategoriensystem für die Beschreibung von Sprachverwendung liefert, enthält insgesamt 41 Skalen zu den kommunikativen Aktivitäten und den mit ihrer Realisierung verbundenen Strategien:
- Rezeption: 12 Skalen (z. B. mündlich: Radiosendungen verstehen, Fernsehsendungen und Filme verstehen; schriftlich: Korrespondenz verstehen, schriftliche Anweisungen verstehen; Strategien: Hinweise identifizieren und erschliessen)
- Interaktion: 15 Skalen (z. B. mündlich: Informelle Diskussion, Formelle Diskussion und Besprechungen, Informationsaustausch, Korrespondenz, Notizen, Sprecherwechsel, Um Klärung bitten)
- Produktion: 11 Skalen (z. B. mündlich: Zusammenhängendes monologisches Sprechen: Erfahrungen beschreiben, Argumentieren, Vor Publikum sprechen; schriftlich: kreatives Schreiben, Berichte und Aufsätze schreiben; Strategien: Planen, Kompensieren)
- Textarbeit: 2 Skalen (Notizen machen in Vorträgen und Seminaren, Texte verarbeiten).

Die Zahl der Kategorien, zu denen es skalierte Kann-Beschreibungen gibt, ist zwar beeindruckend (53 Skalen für kommunikative Aktivitäten/Strategien und Aspekte der Sprachkompetenz zusammen), aber damit liegen bei weitem nicht zu allen Kategorien des Referenzrahmens Skalen vor. Im Referenzrahmen wird mehrfach ausdrücklich darauf hingewiesen, dass erstens nicht versucht wurde, zu

allen Kategorien Kompetenzbeschreibungen zu entwickeln, und dass zweitens für manche Bereiche die Entwicklung und Skalierung von Kompetenzbeschreibungen, die den Qualitätsansprüchen an eine gemeinsame Referenzskala genügen würden, bisher gescheitert ist.

So fehlen im Referenzrahmen Skalen zur Sprachmittlung, zum soziokulturellen Wissen, zu interkulturellen Fertigkeiten, zu Lernstrategien, zur Lektüre literarischer Texte, zum Telefonieren. Man kann die Liste weiter verlängern, indem man in den Kapiteln 4 und 5 alle Kategorien durchgeht. So gibt es weder eine Skala zur non-verbalen Kommunikation noch zu den kommunikativen Sprachprozessen. Der Referenzrahmen führt zwar beispielsweise die am Lesen beteiligten Prozesse auf:

Um zu lesen, muss der Lernende
- den schriftlichen Text wahrnehmen können (visuelle Fertigkeiten);
- die Schrift erkennen können (orthographische Fertigkeiten);
- die Mitteilung als solche identifizieren können (linguistische Fertigkeiten);
- die Mitteilung verstehen können (semantische Fertigkeiten);
- die Mitteilung interpretieren können (kognitive Fertigkeiten). (GER, 93)

Aber anders als beim PISA-Lesetest, bei dem 3 Skalen entwickelt wurden (Heraussuchen spezifischer Informationen, Entwickeln einer Interpretation des Textes, über den Inhalt eines Textes reflektieren), werden diese Verstehensleistungen im Referenzrahmen nicht in separaten Skalen abgebildet, sondern sie sind als Elemente in den Deskriptoren von Leseverstehensskalen enthalten, die am Gebrauchskontext und Texttyp orientiert sind (z. B. Zur Orientierung lesen, Korrespondenz lesen, Schriftliche Anweisungen verstehen).

In den Beiträgen des Sammelbands «Der Gemeinsame europäische Referenzrahmen für Sprachen in der Diskussion» (Bausch, Christ, Königs & Krumm, 2003) werden von verschiedenen Kritikern u. a. Kann-Beschreibungen und Skalen zu folgenden Bereichen angemahnt:

- zum savoir-être und den allgemeinen Kompetenzen überhaupt,
- zur ästhetischen Dimension,
- zu erzieherischen oder auf die Motivation bezogenen Aspekten,
- zum Fremdverstehen,
- zur plurikulturellen Kompetenz,
- zu sprachenübergreifenden Kompetenzen und echter Mehrsprachigkeit,
- zum Rückgriff auf bereits gelernte Sprachen,
- zu Sprachhandlungen im Kontext unterrichtlicher Inszenierung.

Dass solche Deskriptoren in Niveaubeschreibungen gewünscht werden, ist verständlich. Allerdings ist das offensichtlich leichter gesagt als getan. Zumindest, wenn man den Anspruch erhebt, dass die Niveauzuordnungen nicht nur auf Expertenintuition beruhen dürfen, sondern eine empirische Grundlage haben müssen. Die wenigen Vorschläge der Kritiker zu einzelnen Kann-Beschreibungen enthalten keinerlei Indizien für eine Niveauzuordnung (Schneider, 2003, 99-102).

Auf die Methoden der Skalenentwicklung geht der nächste Abschnitt ein. Hier soll abschliessend noch darauf hingewiesen werden, dass die sechs Niveaus, die der Referenzrahmen unterscheidet und deren Benennung durch die Kürzel A1, A2, B1, B2, C1 und C2 sich inzwischen europaweit durchgesetzt hat, nicht nur durch die erwähnten Einzelskalen, sondern auch durch drei zusammenfassende holistische Skalen bzw. Raster beschrieben werden (GER, Kapitel 3). Nicht ganz unproblematisch ist die viel zitierte Globalskala. Sie taugt eher dazu, bestehende Prüfungen und Diplome zu beschreiben als die Kompetenz realer Lernender. Denn diese entwickeln meist ein unausgeglichenes Profil. Eine differenziertere Beschreibung von Kompetenzprofilen wird durch die beiden Raster ermöglicht. Der Raster zur Selbstbeurteilung, der sich als fester Bestandteil aller Sprachenportfolios bewährt hat, erlaubt eine Profilbildung nach Fertigkeitsbereichen, der Beurteilungsraster zur mündlichen Kommunikation berücksichtigt verschiedene Kompetenzaspekte bzw. Beurteilungskriterien, wie Korrektheit, Flüssigkeit, Kohärenz.[4]

Methoden der Skalenentwicklung

Die detaillierten Kompetenzbeschreibungen und Skalen für die sechs «Gemeinsamen Referenzniveaus» des Europarats haben sich als besonders wirksam erwiesen.

Die Kompetenzbeschreibungen wurden in einem Schweizer Projekt im Rahmen des Nationalen Forschungsprogramms NFP 33 «Wirksamkeit unserer Bildungssysteme» entwickelt. Das in Projektberichten (Schneider & North, 2000; North, 2000, 2002; North & Shneider 1998) und im Referenzrahmen (GER, Anhang B) ausführlicher geschilderte methodische Vorgehen wird hier kurz zusammengefasst. Das Projekt orientierte sich am handlungs- und sprachgebrauchsorientierten Konzept der kommunikativen Sprachfähigkeit, das den Modellen von Canale & Swain (1981), von Bachman (1990) und dem *Gemeinsamen europäischen Referenzrahmen* (Europarat, 2001) zu Grunde liegt, und stützte sich auf die Messtheorie des Rasch-Modells[5].

Man kann folgende Methoden der Entwicklung von Sprachkompetenzskalen unterscheiden:

1. intuitive Verfahren, bei denen Einzelne oder Expertengruppen auf der Basis ihrer Erfahrung die Deskriptoren formulieren und Niveaus zuteilen.
2. qualitative Verfahren, bei denen z. B. Gruppen von Beurteilern Lernerproduktionen anschauen und diskutieren, um Schlüsselmerkmale für die Performanz auf verschiedenen Niveaustufen zu identifizieren
3. quantitative Verfahren: Bei der indirekten Skalierung werden Testergebnisse statistisch analysiert und dann Niveaubeschreibungen zugeordnet; bei der direkten Skalierung wird die Verwendung von Deskriptoren bei der Beurteilung einer statistischen Analyse unterzogen.

4. Methodenkombination
Während die meisten bestehenden Skalen am Schreibtisch entstanden sind und auf Expertenkonsens beruhen, wurde zur Entwicklung der Skalen für den Referenzrahmen eine Methodenkombination angewendet und als quantitatives Verfahren die direkte Skalierung der Deskriptoren, nicht der Weg über die Skalierung von Aufgaben gewählt. Die Entwicklung durchlief die folgenden Phasen:

Intuitive Phase:
- Detaillierte Analyse vorhandener Sprachkompetenzskalen, um auf dem Fachwissen und der Erfahrung, die in diese Skalen eingegangen sind, aufzubauen.
- Zerlegung und Vereinfachung von längeren und komplexen Deskriptoren dieser Skalen; Formulierung neuer Deskriptoren, wo dies für bestimmte Kategorien (z. B. Kommunikationsstrategien) nötig war; Zuordnung der Einzelaussagen zu Kategorien des Referenzrahmens; Aufteilung der Deskriptoren in Datenbanken zur Interaktion, Produktion und Rezeption mit insgesamt rund 2000 Einzelaussagen.

Qualitative Phase:
- Analyse von Tonbandaufnahmen, in denen Lehrende über Videoaufzeichnungen von Lernergesprächen diskutieren, mit dem Ziel zu überprüfen, ob die Kategorien und die Metasprache von Praktikern angemessen in den Deskriptoren vertreten waren.
- Über 30 Workshops, in denen Lehrerinnen und Lehrer aus verschiedenen Sprachregionen und Bildungssektoren, die Englisch, Französisch oder Deutsch als Fremdsprache unterrichten, die Deskriptoren erstens nach Kategorien und/oder Niveaus sortierten, und ausserdem anmerkten, welche Deskriptoren sie für unbrauchbar oder unwichtig oder aber für klar und relevant hielten und welche sich ihrer Meinung nach für die Selbstbeurteilung eigneten.
- Erstellen von einander überlappenden Fragebögen mit Deskriptoren, welche sich in den Workshops bewährt hatten; Übersetzung der englischen Deskriptoren ins Deutsche und Französische.

Quantitative Phase:
- Datensammlung: Lehrpersonen beurteilen mit Hilfe der Fragebögen Lernende aus ihren Klassen; ausserdem beurteilen alle Lehrpersonen gemeinsam eine Gruppe ausgewählter Lernender anhand von Videoaufzeichnungen. In die Erhebung waren rund 300 Lehrerinnen und Lehrer für die Fremdsprachen Englisch, Französisch und Deutsch involviert und rund 2700 Lernende aller Schulstufen (ab Sekundarstufe I) und der Erwachsenenbildung in allen Sprachregionen der Schweiz. Ein Teil der beurteilten Lernenden hat zusätzlich einen Selbstbeurteilungsbogen mit ausgewählten Kompetenzbeschrei-

bungen in «Ich-kann»-Form ausgefüllt. Um eine äussere Kontrolle zu ermöglichen, haben Lernende teilweise Tests und Diplomprüfungen abgelegt.
- Rasch-Analyse mit dem Programm FACETS (Linacre, 1992); Berechnung des Schwierigkeitswerts für jeden einzelnen Deskriptor und Untersuchung im Hinblick auf statistisch signifikante Variation bei der Interpretation von Deskriptoren je nach Bildungssektor, Sprachregion und Zielsprache. Identifizierung derjenigen Deskriptoren, die über unterschiedliche Kontexte besonders stabil verwendet wurden und sich deshalb für die Definition der gemeinsamen Referenzniveaus eignen.

Interpretationsphase:
- Identifizieren von Niveaugrenzen auf der Skala; Aufteilung in die sechs Referenzniveaus A1, A2, B1, B2, C1 und C2.
- Zusammenfassung der Referenzniveaus in einer holistischen Skala und zwei Rastern (Tabellen 1-3 des Referenzrahmens) für die Deskriptoren ausgewählt wurden, deren Schwierigkeitswerte nicht an der unteren oder oberen Grenze eines Niveaus lagen und die sich als ganz besonders konsistent und stabil erwiesen hatten; Zusammenstellung von Beispielskalen für diejenigen Kategorien, die sich als skalierbar erwiesen (Kapitel 4 und 5 des Referenzrahmens).

Der Referenzrahmen und entsprechend das Sprachenportfolio enthalten sowohl aufgabenorientierte als auch beurteilungs- bzw. diagnoseorientierte Deskriptoren[6]. In vielen Deskriptoren aber sind die Aufgabenbeschreibungen mit qualitativen Aussagen verbunden. Denn Checklisten und Skalen für die kontinuierliche Beurteilung durch Lehrende oder für die Selbstbeurteilung funktionieren am besten, wenn sie nicht nur aussagen, was Lernende tun können (aufgabenorientiert), sondern auch, wie gut sie es können (Europarat, 2001, 48).

Die Skalierung der Deskriptoren wurde durch verschiedene Folgeuntersuchungen bestätigt. Hohe Korrelationen für die Schwierigkeitswerte der Deskriptoren ergaben sich sowohl im «Can-do»-Projekt von ALTE , der Association of Language Testers in Europe (Jones, 2002), als auch im EU-Projekt DIALANG für internetbasierte diagnostische Tests in 14 Sprachen (Kaftandjieva & Takala, 2002). Aber es zeigten sich auch einige Abweichungen, z. B. andere Schwierigkeitswerte bei jüngeren Lernern. Dies ist dadurch zu erklären, dass die Kompetenzbeschreibungen des GER an der Welt der Erwachsenenwelt orientiert sind und sich teilweise auf Aufgaben beziehen, die nicht im Erfahrungsbereich von jüngeren Lernern liegen (Jones, 2002, 180).[7]. Auf die Anpassung der Deskriptoren für bestimmte Alters- oder Adressatengruppen geht Teil 4 ein.

Wie erwähnt waren an der Skalenentwicklung Lernende und Lehrende aus der Sekundarstufe I, der Sekundarstufe II, den Universitäten und der Erwachsenenbildung beteiligt, nicht aber aus Kindergarten und Primarstufe. Es handelte sich um Lernende, welche die Fremdsprache (zumindest auch) gesteuert in Institutionen lernten. Nicht berücksichtigt wurde der ungesteuerte Zweitspracher-

werb. Normalfall war zudem die Situation des Erwerbs durch eigentlichen Fremdsprachenunterricht, nicht aber z. B. der Erwerb durch Immersion oder zweisprachigen Unterricht.

Es wäre denkbar, dass unter stärkerem Einbezug von Lernkontexten wie dem des zweisprachigen Unterrichts, wie Little (2003) vermutet, die Skalen etwas anders aussähen. Doch gibt es bisher keine empirischen Untersuchungen in diese Richtung.

Dass die Kompetenzbeschreibungen offensichtlich für eine sehr grosse Zahl von Lernenden und Lehrenden in ganz Europa plausibel sind und von ihnen sinnvoll verwendet werden können, liegt wahrscheinlich an einigen charakteristischen Merkmalen der Deskriptoren des GER. Als «gut» und gut skalierbar erwiesen sich nach Schneider & North (2000, 89) Kompetenzbeschreibungen, die folgende Bedingungen erfüllen:

- *das Können ist positiv formuliert;*
- *sie machen für sich allein genommen Sinn und ermöglichen eine Ja-/Nein-Entscheidung;*
- *ihre Interpretation ist nicht abhängig von anderen Beschreibungen des gleichen Niveaus oder von Beschreibungen angrenzender Niveaus;*
- *Niveauunterschiede sind nicht nur ausgedrückt durch verbale Abstufungen wie «kaum», «selten», «mehrfach»;*
- *sie enthalten wenig Jargon/Fachterminologie*
- *sie sind konkret, klar und kurz.*

Diese Eigenschaften machen die Kompetenzbeschreibungen für die Betroffenen, d. h. auch für Laien, verständlich und attraktiv. Einzelne der positiven Merkmale erscheinen jedoch möglicherweise für bestimmte Benutzergruppen, besonders Fachleute wie Testspezialisten, als Nachteil. In kurzen Deskriptoren beispielsweise können nicht viele der Faktoren aufgeführt werden, die einen Handlungsbereich und ein Niveau mitbestimmen. So enthalten z. B. die Deskriptoren zum Leseverstehen nicht jeweils systematisch Angaben zu Leseintention, Lesesituation, Textsorte, Diskurstyp, Textlänge, Themenbereich, Informationsdichte, sprachlichen Merkmalen der Texte usw., sondern sie fokussieren Schlüsselmerkmale und nennen nur das, was für die sprachliche Handlungskompetenz auf dem entsprechenden Niveau ganz besonders typisch ist.

Testautoren, die Prüfungsaufgaben zu den Deskriptoren des Referenzrahmens entwickeln, müssen jedoch alle Faktoren kontrollieren, welche die Schwierigkeit einer Aufgabe beeinflussen (Bygate, Skehan & Swain, 2001; Ellis, 2000). Für sie stellt sich die Aufgabe, die Deskriptoren zu interpretieren, für ihre Zwecke anzupassen und als Anweisungen für die Ausarbeitung von Test-Items systematisierend zu erweitern (Alderson, 2004).

Probleme der Auswahl, Adaptierung und Ergänzung von Kompetenzbeschreibungen

In den seltensten Fällen werden Benutzer des Referenzrahmens die Skalen völlig unverändert übernehmen wollen und können. Es gibt eine Reihe von Gründen, die Veränderungen durch Auswahl, Adaptierung oder Ergänzung wünschenswert oder notwendig machen können. Sowohl für die Entwicklung von Sprachenportfolios wie für die Entwicklung von Bildungsstandards sind folgende Gründe relevant:

- Wunsch nach mehr Niveaustufen (Zwischenniveaus), z. B. wenn es um ein Sprachenportfolio oder Bildungsstandards für eine bestimmte Schulstufe geht;
- Anpassung an eine bestimmte Erfahrungswelt (z.B. von Kindern und Jugendlichen);
- Ergänzung wichtiger Fertigkeitsbereiche (z.B. Sprachmittlung, interkulturelle Fertigkeiten);
- Verbesserung der Verständlichkeit und stärkere Konkretisierung (z.B. für die Selbstbeurteilung durch die Lernenden oder als Anhaltspunkt für die Formulierung von Aufgaben) (Schneider, 2001; Lenz & Schneider 2002).

Die Erfahrung hat gezeigt, dass solche Veränderungen und Erweiterungen problematisch sein können, wenn sie nicht empirisch validiert werden. Verbunden mit empirischeren Untersuchungen ist wie die schon erwähnten Projekte von ALTE und DIALANG auch das von Kantonen der Deutschschweiz getragene Projekt IEF (= Instrumente für die Evaluation von Fremdsprachenkompetenzen), in dem Beurteilungsinstrumente für die Selbst- und Fremdbeurteilung von Sprachkompetenzen in Französisch und Englisch für Schülerinnen und Schüler der 5.-9. Klasse entwickelt werden. Die Beurteilungsaufgaben beziehen sich auf im Rahmen des Projekts adaptierte und validierte Kompetenzbeschreibungen, die auch für die Checklisten zur Selbstbeurteilung im neuen Sprachenportfolio für 11-15-Jährige, dem sogenannten ESP II, verwendet werden (Studer, Mettler, & Lenz, i. Vorb.).

Oft beruhen Übernahmen und Anpassungen von Elementen des Referenzrahmens nur auf Expertenurteilen. Das gilt für das Gesamtsprachenkonzept und verschiedene Lehrpläne in der Schweiz ebenso wie für Sprachenportfolioversionen in verschiedenen europäischen Ländern[8] und für die Entwürfe von Bildungsstandards und Kernlehrplänen in Deutschland.

Ich möchte einige typische Probleme aufzeigen, die sich in Sprachenportfolios und in Entwürfen für Bildungsstandards und Kernlehrpläne finden:

Problem: Vereinfachung

Problematisch sind Vereinfachungen, bei denen alle Modifikatoren weggekürzt werden und nur noch die Aktivität ohne jegliche niveauspezifische Angabe erhalten bleibt, z. B.

Original:
I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect. (B2)
Problematische Vereinfachung:
I can understand films and documentaries.
Durch solche Vereinfachungen und Verallgemeinerungen wird das Niveau unbestimmter und die Interpretation geht eher in Richtung eines höheren Niveaus.

Problem: Achievement- statt Proficiency- Perspektive

Bei der Anpassung von Deskriptoren an den Lernkontext wird oft der Unterschied zwischen Achievement- und Proficiency-Assessment nicht beachtet wie in den nachfolgenden Beispielen. Achievementtests bzw. achievement-orientierte Beschreibungen beziehen sich auf Unterrichtsziele, darauf, was in einer bestimmten Zeitspanne unterrichtet wurde. Im Unterschied zu dieser Binnenperspektive haben Proficiencytests und -beschreibungen eine Aussenperspektive. Sie beziehen sich unabhängig von der Unterrichtssituation und Lerngeschichte darauf, was jemand zu einem bestimmten Zeitpunkt in Realsituationen mit seinen Sprachkenntnissen tun kann:
Beispiele aus Sprachenportfolios:
Ich kann die Texte und die Aufgaben im Lehrbuch verstehen. (A1)
Ich kann Texte meiner Mitschülerinnen und Mitschüler verstehen und deren Inhalt kommentieren. (B1)
Beispiele aus Entwürfen zu Bildungsstandards:
Die Schülerinnen und Schüler können ausgewählten kurzen Radiosendungen oder mittels Tonträgern dargebotenen einfachen Hörtexten <u>mit bekanntem Sprachmaterial</u> über vertraute Themen die Hauptinformationen entnehmen, wenn relativ langsam und deutlich gesprochen wird.
Die Schülerinnen und Schüler können <u>ausgehend von den im Unterricht behandelten Kommunikationsbereichen und Themen</u> Vergleiche zwischen ihrem eigenen kulturellen Hintergrund und dem von Jugendlichen in anglophonen Ländern anstellen und sind bereit, andere Kulturen zu verstehen und zu akzeptieren.
Solche Formulierungen machen für die (Selbst-)Beurteilung in der Schule durchaus Sinn, denn innerhalb der Klasse gehört es zur gemeinsamen Erfahrung, was damit gemeint ist (welches Lehrbuch benutzt wird, welches Niveau die Mitschüler haben, was bekannt ist und was behandelt wurde). Aber es sind typische Deskriptoren, die eindeutig nicht Indikatoren für ein bestimmtes Niveau sind, sondern für ganz unterschiedliche Niveaus stehen können. Sie erlauben daher auch nicht einzuschätzen, ob ein bestimmtes Referenzniveau erreicht ist. Sie taugen nicht für eine Information nach aussen und nicht für Standards.

Problem: Niveauvermischung

Im KMK-Entwurf, der sich direkt auf die Niveaubeschreibungen im GER bezieht, finden sich seltsamerweise häufiger Kompetenzbeschreibungen, für die zwei Referenzniveaus angegeben sind, z. B.:

Die Schülerinnen und Schüler können klar formulierte Anweisungen, unkomplizierte Anleitungen, Hinweise und Vorschriften verstehen. (B1/ B2)

Es handelt sich trotz dieser Niveauangaben natürlich nicht um einen Deskriptor des GER für zwei Niveaus. Die Formulierungen im GER sind deutlich niveaudifferenzierend:

Kann klar formulierte, unkomplizierte Anleitungen zur Bedienung eines Geräts verstehen (B1)

Kann lange, komplexe Anleitungen im eigenen Fachgebiet verstehen, auch detaillierte Vorschriften oder Warnungen, sofern schwierige Passagen mehrmals gelesen werden können. (B2)

Es ist offensichtlich, dass die KMK-Formulierung und die differenzierenden Formulierungen im GER eine stark unterschiedliche Textauswahl und Aufgabenstellungen nahe legen.

Problem: Kompetenzbeschreibungen für Sprachmittlung

Im Projekt «Profile deutsch» (Glaboniat et al., 2002) wurden erstmals auch Kann-Beschreibungen zur Sprachmittlung (Mediation) formuliert. Um die Verknüpfung mit den Referenzniveaus sicherzustellen wurden bei der Entwicklung dieser Kann-Beschreibungen Elemente aus den skalierten Deskriptoren für Rezeption, Produktion oder Interaktion in die neuen Items eingebaut. Allerdings wurden die so durch Kombination konstruierten neuen Beschreibungen nie empirisch validiert.

Im KMK-Entwurf (2004) finden sich ebenfalls Standards zur Sprachmittlung:

Die Schülerinnen und Schüler können mündlich in Routinesituationen und schriftlich zu vertrauten Themen zusammenhängende sprachliche Äußerungen und Texte sinngemäß von der einen in die andere Sprache übertragen.

Die Schülerinnen und Schüler können

- in Alltagssituationen sprachmittelnd agieren

- persönliche und einfache Sach- und Gebrauchstexte sinngemäß übertragen.

Diese Beschreibungen bleiben so allgemein, dass sie stark unterschiedlich interpretiert werden könnten. Es fehlen Modifikatoren oder Hinweise, welche die sprachliche Leistung einschränkend bestimmen. Dies geschieht im KMK-Entwurf durch die Beschreibung der Überprüfungsaufgaben. Es ist nicht so sehr der Deskriptor, der das Niveau (auch für die Aufgaben) definiert. Niveaueingrenzend sind vielmehr die Aufgabenbeschreibungen.

Wenig hilfreich sind auch die Beschreibungen in den Bildungsstandards von Baden-Württemberg, z. B. für die Realschule, in denen echte Niveauunterschiede kaum ersichtlich sind (Baden-Württemberg, 2004):

Die Schülerinnen und Schüler können
- *in einfachen Alltagssituationen vermitteln. (Klasse 6)*
- *in Alltagssituationen vermitteln; persönliche und einfache Sachtexte zu vertrauten Themen sowie Redeabsichten aus dem Englischen sinngemäß ins Deutsche übertragen. (Klasse 8)*
- *in Alltagssituationen vermitteln; persönliche und einfache Sachtexte (Beschreibungen, Gebrauchsanweisungen, Bedienungsanleitungen) aus dem Englischen sinngemäß ins Deutsche übertragen. (Klasse 10)*

In einer Formulierung wie der folgenden kommt noch das Problem der Bindung an den Unterricht (Achievement) hinzu:

Die Schülerinnen und Schüler können in elementaren zweisprachigen Situationen mündlich vermitteln, in denen ihnen vertraute Inhalte in bekanntem Wortlaut zur Sprache kommen.

Die Problematik solcher Niveaubeschreibungen würde schon bei einer einfachen empirischen Überprüfung durch Sortieraufgaben deutlich werden.

Problem: Kompetenzbeschreibungen für interkulturelle Kompetenzen

Im Referenzrahmen sind die interkulturellen Fertigkeiten umschrieben, jedoch ohne Niveauabstufungen:

Interkulturelle Fertigkeiten umfassen:
- *die Fähigkeit, die Ausgangskultur und die fremde Kultur miteinander in Beziehung zu setzen;*
- *kulturelle Sensibilität und die Fähigkeit, eine Reihe verschiedener Strategien für den Kontakt mit Angehörigen anderer Kulturen zu identifizieren und zu verwenden;*
- *die Fähigkeit, als kultureller Mittler zwischen der eigenen und der fremden Kultur zu agieren und wirksam mit interkulturellen Missverständnissen und Konfliktsituationen umzugehen;*
- *die Fähigkeit, stereotype Beziehungen zu überwinden* (GER, 106).

In Sprachenportfolios finden sich z. B. in Checklisten Kompetenzbeschreibungen wie:

I can discuss and question different value systems (belief and attitudes). (B1)

Das ist natürlich ein wichtiges Lernziel. Problematisch allerdings wird es, wenn solche Beschreibungen einem bestimmten Referenzniveau kommunikativer Kompetenz, in diesem Fall B1, zugeordnet werden. Das ist eine reine Setzung ohne empirische Basis. Solche interkulturellen Kompetenzen sind nicht an sprachliche Kompetenzen eines bestimmten Niveaus gebunden.

Wenig hilfreich für Bildungsstandards ist es, wenn zur Niveaudifferenzierung auf den konkreten Unterricht Bezug genommen wird, z. B.:

Die Schülerinnen und Schüler haben eine elementare interkulturelle Handlungsfähigkeit entwickelt, die sie in unterrichtlich vermittelten Erfahrungsbereichen und Kommunikationssituationen einsetzen können.

Der KMK-Entwurf von Standards für den mittleren Schulabschluss enthält Formulierungen, die als Bildungsziele zweifellos zu befürworten sind, bei denen aber fraglich ist, ob und wie sie als Standards überprüft werden könnten:

Die Schülerinnen und Schüler sind neugierig auf Fremdes, aufgeschlossen für andere Kulturen und akzeptieren kulturelle Vielfalt.

Die Schülerinnen und Schüler sind bereit, sich auf fremde Situationen einzustellen und sich in Situationen des Alltagslebens angemessen zu verhalten.

In Bezug auf interkulturelle Fertigkeiten und soziokulturelles Wissen fehlt bisher empirische Evidenz für Niveaustufen. Zudem stellt sich die Frage, ob diese Bereiche separat oder in Verbindung mit Aufgaben zur Überprüfung der kommunikativen sprachlichen Handlungskompetenz getestet werden sollen. Möglicherweise bringen Erfahrungen im DESI-Projekt (Nold, 2003) nützliche Hinweise dazu.

Ganz ähnliche Probleme wie bei den Kompetenzbeschreibungen zu den interkulturellen Kompetenzen in Sprachenportfolios und in Entwürfen zu Bildungsstandards stellen sich in Bezug auf die so genannte Methodenkompetenz bzw. Lernstrategien.

Problem: Konkretisierung durch Aufgaben

Teilweise enthalten die Kompetenzbeschreibungen Hinweise auf beispielhafte Anwendungssituationen und -aufgaben, z. B.:

Ich kann kurze, einfache Postkarten schreiben, z. B. Feriengrüße. (GER, 36)

In «Profile deutsch» werden die Kann-Beschreibungen systematisch durch die Angabe von Beispielen für Aufgaben/Situationen in den drei Domänen des privaten und beruflichen Lebens und des Bildungsbereichs konkretisiert.

Klieme et al. (2003, 124) fordern zu Recht, dass die Bildungsstandards zusammen mit Musteraufgaben vorgelegt werden sollten.

Bei den vorliegenden deutschen Entwürfen für Bildungsstandards und Kernlehrpläne finden sich als illustrierende Testaufgaben sowohl Übernahmen aus bestehenden internationalen Diplomprüfungen als auch Aufgaben aus Lehrwerken und neu entwickelte Aufgaben.

In Bezug auf die Teile aus Diplomprüfungen ist festzuhalten, dass die Zuordnung zu einem bestimmten Referenzniveau zunächst nur eine Hypothese ist, denn bisher sind solche Zuordnungen noch nicht durch Verfahren, wie sie das Handbuch für die Zuordnung von Prüfungen zum Referenzrahmen vorsieht, erwiesen[9]. Bei den aus Lehrwerken zusammengestellten und den «neuen» Aufgaben im KMK-Entwurf gewinnt man den Eindruck, dass die Aufgabenbeispiele allzu schnell zusammengestellt wurden. Da gibt es beispielsweise die für Tests nicht ganz unproblematischen Aufgabentypen, ein Mind-Map zu erstellen und vorgegebene Sätze zu ergänzen. Zudem ist dem Lesetext eine Fülle von Vokabelangaben beigegeben, was der Forderung nach authentischer Aufgaben widerspricht (KMK, 2004, 36-37). Es findet sich sogar eine Aufgabensituation für spontanes Sprechen, in der wie in längst vergessen geglaubten Zeiten audiovi-

suellen Unterrichts in der Arbeitsanweisung nicht nur den Prüflingen eine Geschlechtsrolle zugeschrieben wird, sondern auch noch vorgegeben wird, dass Prüflinge auf dem Niveau B1 in Englisch eine Schwester englischer Muttersprache haben! (KMK, 2004, 42-43)

Es wird deutlich, dass Testautoren erstens über ein klares Konzept für die Umsetzung von Kompetenzbeschreibungen in Aufgaben verfügen müssen, wie dies u. a. Alderson (2004) anmahnt, und dass zweitens ohne empirische Erprobung nicht von Musteraufgaben gesprochen werden kann.

In den beiden erwähnten Projekten, DESI in Deutschland und IEF in der Schweiz, wird eine theoriegestützte und empirisch abgesicherte Aufgabenentwicklung angestrebt. Sie entsprechen daher dem Postulat von Klieme et al.(2003, 124):

> *Die Konstruktion eines größeren «Pools» von Aufgaben, die zu den inhaltlichen und kognitiven Anforderungen der Kompetenzstufen passen und sie in einer Vielzahl von Kontexten variieren, erfordert jedoch ein sehr systematisches Vorgehen.* [...] *Aussagefähige Tests, die tatsächlich das Kompetenzmodell der Bildungsstandards operationalisieren, bedürfen also einer gezielten Entwicklungsarbeit.*

Solche Entwicklungsarbeit braucht Ressourcen und Zeit.

Konsequenzen für die Entwicklung von Bildungsstandards

Das Kompetenzmodell des Referenzrahmens und die wissenschaftlich abgesicherten Niveaubeschreibungen sind eine gute Grundlage für die Entwicklung von Bildungsstandards für die Fremdsprachen im Rahmen des Projekts HARMOS (EDK, 2004, 13). Aber sie nehmen nicht alle Entscheidungen ab und machen weitere Entwicklungsarbeit nicht überflüssig. Zu den möglichst auch empirisch zu stützenden Entscheidungen gehören:

- die Wahl des erwarteten Kompetenzprofils (das gleiche Niveau für alle Fertigkeitsbereiche oder ein differenziertes Profil?),
- die Bestimmung der Kompetenzbereiche (Einschluss von Sprachmittlung, soziokulturellem Wissen, interkulturellen Fertigkeiten, Lernkompetenzen?),
- die Auswahl, ggf. Adaptierung und Ergänzung von Kompetenzbeschreibungen im Hinblick auf die altersgemässe Erfahrungswelt der Schülerinnen und Schüler,
- die Entwicklung und Zuordnung von Testaufgaben zu den Kompetenzbeschreibungen (Erweiterung des Aufgabenpools von IEF und/oder Übernahme von Teilen aus internationalen (Jugend-)Sprachdiplomen?),
- und schliesslich die Zuordnung zu Jahrgangsstufen (im Projekt HARMOS zum Ende der 6. und 9. Klasse).

Im «Gesamtsprachenkonzept» (EDK, 1998) waren für das Ende der obligatorischen Schulzeit die folgenden Niveaus als Richtziele in Bezug auf die im Raster

zur Selbstbeurteilung unterschiedenen 5 Fertigkeitsbereiche (GER, 36 und Sprachenpass des Sprachenportfolios) anvisiert:

	Hörverstehen	Leseverstehen	Gespräch	Zusammenhängendes	Schreiben
2. Landessprache	**B1+**	**B1+**	**B1**	**B1**	**B1**
Englisch	**A2+**	**A2+**	**A2+**	**A2**	**A2**

Richtziele, Ende 9. Klasse im Sprachenkonzept (EDK, 1998)

In den wegen der Diskussion um die Sprachenfolge (Englisch vor oder nach der 2. Landessprache) nie verabschiedeten Empfehlungen der EDK wurde dann für beide Fremdsprachen das gleiche Niveau der kommunikativen Sprachkompetenz angesetzt, ergänzt durch den Hinweis auf kulturelle Ziele, dies allerdings merkwürdigerweise nur für die zweite Landessprache[10]. Die Fertigkeitsbereiche sind nicht mehr aufgeführt, sondern es wird nur zwischen Rezeption und Produktion unterschieden. Dafür wird das Niveau differenziert bezogen auf Zweige mit Grundansprüchen und mit erweiterten Ansprüchen:

Einstiegsfremdsprache und zweite Fremdsprache			
Rezeptiv		Produktiv	
Grundansprüche	Erw. Ansprüche	Grundansprüche	Erw. Ansprüche
A2+	**B1+**	**A2**	**B1**
Zweite Landessprache: zusätzlich kulturelle Ziele und Inhalte			

Empfehlungen Ende 9. Klasse (EDK, 2001)

In der Deklaration der Westschweizer Erziehungsdirektoren erscheinen zwei neue Aspekte 1. werden Niveauvorgaben auch für das Ende der Primarschule gemacht und 2. wird für das Ende der obligatorischen Schulzeit – bei gleichen Zielen für die 2. Landessprache und Englisch – unterschieden zwischen dem angestrebten Niveau, für dessen Erreichen erst die vorgesehenen Reformen (z. B. bilingualer Unterricht) greifen müssten, und einem tieferen Niveau während einer Übergangszeit.

Minima	Rezeptiv	Produktiv
6. Schuljahr D	A1	A1
9. Schuljahr D+E	B1+	B1
Übergangszeit (max. bis 2010)	A2+	A2

Déclaration relative à la politique de l'enseignement des langues en Suisse romande (CIIP, Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin, 2003)

Solche Überlegungen sind auch für das Projekt HARMOS wichtig, das ja richtigerweise vorsieht, empirisch zu prüfen, welche Ziele erreicht werden (können). Auf jeden Fall ist zu entscheiden, ob dabei vom Status quo ausgegangen werden soll (heutiger Fremdsprachenunterricht) oder ob auch mit dem Ertrag von Innovationen (z. B. zusätzlicher bilingualer Sachunterricht) gerechnet wird. Mir scheint der zweite Ansatz unbedingt wünschenswert. Das würde aber bedeuten, dass für die Evaluation der provisorischen Mindestkompetenzen auch eine genügend grosse Stichprobe aus innovativen Pilotklassen identifiziert würde. Das DESI-Projekt sieht beispielsweise ein Sub-Sample von Lernenden aus Programmen mit partieller Immersion vor (Nold, 2003, 372). Das Spektrum der zu entwickelnden Evaluationsaufgaben müsste entsprechend grösser sein.

Ich möchte abschliessend meine Schlussfolgerungen in zwei Punkten zusammenfassen:

1. Wenn Probleme, wie sie am Beispiel von vorhandenen Entwürfen zu Bildungsstandards gezeigt wurden, vermieden werden sollen, dann dürfen die Entwicklungsarbeiten nicht allein auf der Erfahrung von Experten und Expertinnen beruhen, sondern müssen qualitativ und quantitativ validiert werden.
2. Die Festlegung von Erwartungen, d. h. die Zuordnung eines differenzierten Kompetenzprofils zu Jahrgangsstufen bleibt zu einem rechten Teil eine (zu begründende) Setzung unter Berücksichtigung von Aspekten wie Wünschbarkeit, Realisierbarkeit, Reformwillen, Bereitstellung von Ressourcen. Diese Entscheidung kann nicht Wissenschaftlerinnen und Wissenschaftlern allein überlassen werden, sondern muss demokratisch unter Mitbeteiligung von sprachpolitischen Entscheidungsträgern und Lehrpersonen zustande kommen.

Anmerkungen

1 Die Diskussion wird in Diskussionsgruppen der EALTA (European Association for Language Testing and Assessment) geführt (www.ealta.eu.org).

2 Es sind teilweise die gleichen Vorwürfe, die auch gegenüber Bildungsstandards erhoben werden: der Vorwurf des Reduktionismus (auf Testbares), der Nivellierung und der fehlenden Legitimation (Klieme et al. 2003, 55-58).

3 Zu den sprachmittelnden Aktivitäten gehören z. B. Dolmetschen, Übersetzen und die Umformung oder Zusammenfassung eines Textes für Dritte (GER, 26). Der Referenzrahmen enthält allerdings keine Skalen mit Deskriptoren zur Sprachmittlung. Auf die Schwierigkeit, Deskriptoren für die Sprachmittlung zu formulieren wird in Teil 4 unter der Überschrift «Problem: Kompetenzbeschreibungen für Sprachmittlung» eingegangen.

4 In Analogie zum Raster des Referenzrahmens zu den qualitativen Aspekten der mündlichen Interaktion wurden im «Manual» für die Zuordnung von Sprachprüfungen weitere Raster zusammengestellt: ein Raster zu den qualitativen Faktoren der Rezeption sowie ein Raster zu den qualitativen Faktoren der Produktion (Council of Europe, 2003: Tabellen 4.3 und 4.5). Diese und andere neue Beurteilungsskalen im «Manual» wurden alle durch eine Kompilation von skalierten Deskriptoren gewonnen, die in verschiedenen Einzelskalen des Referenzrahmens enthalten sind.

5 Das probabilistische Rasch-Modell hat in der Sprachtestforschung grosse Bedeutung erlangt. Es geht davon aus, dass die Wahrscheinlichkeit eine Aufgabe zu lösen mit zunehmender Fähigkeit der Personen grösser wird. Es erlaubt, die Schwierigkeitswerte von Aufgaben und die Fähigkeitswerte der Personen auf einer gemeinsamen Skala abzubilden. Die Raschanalyse wird oft zur Skalenentwicklung und auch bei der Testkonstruktion eingesetzt, z. B. um sicherzustellen, dass verschiedene Prüfungssätze die gleiche Schwierigkeit haben. Rasch-Analysen ermöglichen die Schaffung von Aufgabenbanken mit bekanntem Schwierigkeitswert für die einzelnen Testaufgaben.

6 Rein aufgabenorientierte Skalen sind in erster Linie für Testautoren bestimmt. Sie beschreiben Aufgaben, welche die Lernenden in Tests lösen können sollen. Der Akzent liegt darauf, was die Lernenden tun können. Bei beurteilungsorientierten Skalen liegt der Akzent darauf, wie gut die Lernenden etwas tun können. Sie dienen Prüfern bei der Beurteilung von Leistungen und enthalten in der Regel Beschreibungen zu einer begrenzten Auswahl von drei bis fünf qualitativen Aspekten wie z. B. Flüssigkeit, Korrektheit und Aussprache. Diagnoseorientierte Skalen oder Raster können mehr Kategorien berücksichtigen. Sie dienen dazu «ein Profil des aktuelles Lernstands zu erfassen, ein Profil der Bedürfnisse und Ziele in relevanten Kategorien zu erstellen sowie eine Diagnose, was man zur Erreichung dieser Ziele noch abdecken muss» (GER, 47).

7 Das wird bestätigt durch Ergebnisse im IEF-Projekt (Persönliche Mitteilung P. Lenz u. T. Studer).

8 Im Sinne einer Qualitätssicherung für das europäische Sprachenportfolio und als Hilfe für Portfolioentwickler wurde inzwischen im Auftrag des Europarats eine Itembank mit «guten», auf die skalierten Deskriptoren im Referenzrahmen beziehbaren Kann-Beschreibungen zusammengestellt (Lenz & Schneider, 2004 a,b).

9 Dazu heisst es einleitend zu den Aufgabenbeispielen der KMK: «Hinsichtlich der Orientierung am Gemeinsamen europäischen Referenzrahmen und ggf. auch im Lichte der Erfahrungen von künftigen Vergleichsuntersuchungen sind sie (die Aufgaben) zu überprüfen und weiter zu entwickeln» (KMK, 2004, 22-23).

10 In der Deklaration der Westschweizer Erziehungsdirektionen wird diese Einseitigkeit korrigiert: «Les objectifs linguistiques sont complétés par d'autres objectifs portant sur la connaissance de la réalité sociale, culturelle et géographique des pays concernés. Pour l'allemand, une partie de ces objectifs concernent la Suisse alémanique et portent sur le développement d'attitudes d'ouverture vis-à-vis de la culture alémanique et de compétences de

communication avec des locuteurs alémaniques (diglossie). En anglais, un accent est mis sur la compréhension du monde anglophone et de sa diversité (sur les plans social et culturel, notamment) (CIIP, 2003, 3).

Schlagworte: Bildungsstandards - Fremsprachenkompetenz - Kompetenzmodell - sprachkompetenzskala

Literatur

Alderson, J. C. (2004). *The Dutch CEF Construct Group: advice on how to develop specifications for tests at each level of the CEF. http://www.ealta.eu.org/conference/ppt/alderson14may2.ppt.*

Bachman, L. F. (1990). *Fundamental Considerations in Language Testing.* Oxford: Oxford University Press.

Baden-Württemberg (2004). Ministerium für Kultus, Jugend und Sport, Baden-Württemberg (Hrsg.), *Bildungspläne. http://www.bildungsstandards-bw.de/.*

Bausch, K.-R., Christ, H., Königs, F.G. & Krumm, H.-J. (Hrsg.) (2003). *Der Gemeinsame Europäische Referenzrahmen in der Diskussion.* Arbeitspapiere der 22. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts. Tübingen: Narr, [Giessener Beiträge zur Fremdsprachendidaktik].

Beacco, J.-C. & Byram, M. (2003). *Guide for the Development of Language Education Policies in Europe. From Linguistic Diversity to Plurilingual Education.* Main Version. Draft 1. Strasbourg: Council of Europe. *http://www.coe.int/T/E/Cultural_Cooperation/education/Languages/Language_Policy.*

Beacco, J.-C., Bouquet, S. & Porquier, R. (2004). *Niveau B2 pour le Français. Un referentiel.* Paris: Didier.

Beck, B. & Klieme, E. (2003). DESI – Eine Längsschnittstudie zur Untersuchung des Sprachunterrichts in deutschen Schulen. In C. Finkbeiner & S. Fehling (Hrsg), *Evaluation im Brennpunkt – Fremdsprachen lernen und lehren* (S. 380-395). (Empirische Pädagogik 17, (3), Themenheft.) Landau: Verlag Empirische Pädagogik.

Bygate, M., Skehan, P. & Swain, M. (Hrsg.) (2001). *Researching Pedagogic Tasks. Second Language Learning, Teaching and Testing.* Harlow: Pearson Education, [Applied Linguistics and Language Study].

Byram, M. & Zarate, G. (Hrsg.) (1997). *The Sociocultural and Intercultural Dimension of Language Learning and Teaching.* Strasbourg: Council of Europe Publishing, [Education].

Byram, M. (Hrsg.) (2003). *Intercultural Competence.* Strasbourg: Council of Europe Publishing.

Canale, M & Swain, M. (1981). A Theoretical Framework for Communicative Competence. In A.S. Palmer, J.M. Groot & G.A. Trosper (Hrsg.), *The Construct Validation of Tests of Communicative Competence.* Washington D.C.: TESOL.

CIIP (2003). *Déclaration de la CIIP relative à la politique de l'enseignement des langues en Suisse romande. www.ciip.ch/ciip/pdf/cp030403_3%20.pdf.*

Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer.* Strasbourg: Conseil de l'Europe/Didier, [Apprentissage des langues et citoyenneté européenne].

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: CUP. *www.coe.int/portfolio >Documentation.*

Council of Europe (2002a). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Guide for Users.* Strasbourg: Council of Europe, [Language Policies].

Council of Europe (2002b). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case Studies.* (Hrsg. Ch. Alderson). Strasbourg: Council of Europe, [Language Policies].

Council of Europe (2003). *Relating Language Examinations to the Common European Frame-*

work of Reference for Languages: Learning, Teaching, Assessment (CEF). Manual, Preliminary Pilot Version. Strasbourg: Council of Europe, [Language Policies]. *http://www.coe.int/T/E/Cultural_Cooperation/education/Languages/Language_Policy/Manual/].*

EDK, Schweizerische Konferenz der kantonalen Erziehungsdirektoren, Expertengruppe (1998). *Sprachenkonzept Schweiz: Welche Sprachen sollen die Schülerinnen und Schüler der Schweiz während der obligatorischen Schulzeit lernen? http://www.romsem.unibas.ch/sprachenkonzept.*

EDK, (2001). Empfehlungen zur Koordination des Sprachenunterrichts in der obligatorischen Schule. Entwurf 1.6.2001

EDK, (2004). HARMOS. *Zielsetzungen und Konzeption.* Bern: EDK Generalsekretariat, Koordinationsbereich «Obligatorische Schule». *http://www.edk.ch/PDF_Downloads/Harmos/Harmos_Weissbuch_d.pdf.*

Ellis, R. (2000). Task-based Research and Language Pedagogy. *Language Teaching Research 4,* (3), 193-220.

Europarat (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen.* Berlin u.a.: Langenscheidt. *http://www.goethe.de/referenzrahmen.*

Glaboniat, M., Müller, M., Rusch, P., Schmitz, H. & Wertenschlag, L. (2002). *Profile deutsch. Gemeinsamer europäischer Referenzrahmen: Lernzielbestimmungen, Kannbeschreibungen, kommunikative Mittel, Niveau A1, A2, B1, B2.* Berlin: Langenscheidt.

Jones, N. (2002). Relating the ALTE Framework to the Common European Framework of Reference. In Council of Europe. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case Studies* (S. 167-183). (Hrsg. Ch. Alderson). Strasbourg: Council of Europe.

Kaftandjieva, F. & Takala, S. (2002). Council of Europe Scales of Language Proficiency: A Validation Study. In Council of Europe. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case Studies* (S. 106-129). (Hrsg. Ch. Alderson). Strasbourg: Council of Europe.

Klieme, E. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise.* (Hrsg. Bundesministerium für Bildung und Forschung). Bonn: BMBF.

KMK, Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.) (2004). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss.* Neuwied: Luchterhand. *http://www.kmk.org/doc/beschl/aschulw.htm.*

Krumm, H.-J. (2003). Der Gemeinsame europäische Referenzrahmen – ein Kuckucksei für den Fremdsprachenunterricht? In K.-R. Bausch, H. Christ, F.G Königs & H.-J. Krumm (Hrsg.), *Der Gemeinsame europäische Referenzrahmen in der Diskussion* (S. 120 – 126). Tübingen: Narr.

Lenz, P. (2004a). The European Language Portfolio – A Tool for the Promotion of Plurilingualism. In G. Bräuer & K. Sanders (Hrsg.), *New Visions in Foreign and Second Language Education* (S. 123-151). San Diego: LARC.

Lenz, P. (2004b): The European Language Portfolio. In K. Morrow (Hrsg.), *Insights from the Common European Framework* (S. 22-31). Oxford: Oxford University Press.

Lenz, P. & Schneider, G. (2002). Developing the Swiss model of the European Language Portfolio. In Council of Europe Ch. Alderson (Hrsg.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case Studies* (S. 68-86). Strasbourg: Council of Europe.

Lenz, P. & Schneider, G. (2004a). *Introduction to the Bank of Descriptors for Self-assessment in European Language Portfolios.* Strasbourg: Council of Europe, [Language Policy Division]. *www.coe.int/portfolio.*

Lenz, P. & Schneider, G. (2004b). *A Bank of Descriptors for Self-assessment in European Language Portfolios.* Strasbourg: Council of Europe, [Language Policy Division]. *[www.coe.int/portfolio]* .

Linacre, J. M. (1992). *FACETS: A Computer Program for the Analysis of Multi-Faceted Data.* Chicago: MESA Press.

Little, D. (2003). The Common European Framework: Principles, Challenges, Issues. *Neusprachliche Mitteilungen 56,* (3), 130-140.

Lüdi, G. (2003). Konzepte für mehrsprachige Repertoires als Zielvorstellungen für die Schule. Überlegungen zu einem «Gesamtsprachenkonzept» für die Schulen in der Schweiz. In G. Schneider, & M. Clalüna (Hrsg.), *Mehr Sprache – mehrsprachig – mit Deutsch. Didaktische und politische Perspektiven* (S. 276-291). München: Iudicium.

Morrow, K. (Hrsg.) (2004). *Insights from the Common European Framework.* Oxford: Oxford University Press.

Nold, G.M.J. (2003). DESI – A Language Assessement Project in Germany and the Pros and Cons of Large Scale Testing. In C. Finkbeiner & S. Fehling (Hrsg.), (S. 368-379). (Empirische Pädagogik 17, (3), Themenheft). Landau: Verlag Empirische Pädagogik.

North, B. (2000). *The Development of a Common Framework Scale of Language Proficiency.* (Theoretical Studies in Second Language Acquisition 8). New York & Bern: Peter Lang.

North, B. (2002). Developing descriptor scales of language proficiency for the CEF Common Reference Levels. In Council of Europe In Council of Europe Ch. Alderson (Hrsg.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case Studies* (S. 87-105). Strasbourg: Council of Europe.

North, B. & Schneider, G. (1998). Scaling Descriptors for Language Proficiency Scales. *Language Testing 15,* 2, 217-262.

Quetz, J. (2003). A1-A2-B1-B2-C1-C2. Der Gemeinsame europäische Referenzrahmen. *Deutsch als Fremdsprache 40,* 1, 145-155.

Schneider, G. (2001). Kompetenzbeschreibungen für das «Europäische Sprachenportfolio. *Fremdsprachen lehren und lernen (FLuL), 30,* 193-214.

Schneider, G. (2003). Der europäische Referenzrahmen und die Mehrsprachigkeit. In G. Neuner & U. Koithan (Hrsg.), *Tagungsbeiträge und Arbeitsberichte zum Thema «Mehrsprachigkeit im Bereich Deutsch als Fremdsprache»,* Tagungsdokumentation «Internationales Qualitätsnetz Deutsch als Fremdsprache» (S. 87-108). Kassel: Kassel University Press. *http://www.uni-kassel.de/fb9/dafliqn/iqn-tagung_dokumentation_2003.htm.*

Schneider, G. & Lenz, P. (2001). *Guide for Developers of an European Language Portfolio.* Strasbourg: Council of Europe. *www.coe.int/portfolio >Documentation.*

Schneider, G. & North, B. (2000). *Fremdsprachen können – was heisst das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit.* Chur & Zürich: Rüegger.

Schneider, G., North, B. & Koch, L. (2001). *Portfolio européen des langues. Version pour jeunes et adultes – Europäisches Sprachenportfolio. Version für Jugendliche und Erwachsene – Portfolio europeo delle lingue. Versione per giovani e adulti – European Language Portfolio. Version for Young People and Adults.* (Hrsg. EDK). Bern: Berner Lehrmittel- und Medienverlag. *www.sprachenportfolio.ch.*

Studer, T., Mettler, M., & Lenz, P. (i. Vorb.). Das Projekt IEF: Instrumente für die Evaluation von Fremdsprachenkompetenzen (Englisch/Französisch). *Schweizerische Zeitschrift für Bildungswissenschaften.*

Vollmer, H. J. (2003). Ein gemeinsamer europäischer Referenzrahmen für Sprachen: Nicht mehr, nicht weniger. In K.-R. Bausch, H. Christ, F.G. Königs & H.-J. Krumm (Hrsg.), *Der Gemeinsame europäische Referenzrahmen in der Diskussion* (S. 192-206). Tübingen: Narr.

Zarate, G., Gohard-Radenkovic, A., Lussier, D., & Penz, H. (2003). *Médiation culturelle et didactique des langues.* Strasbourg: Centre européen pour les langues vivantes. Graz/Strasbourg: Editions du Conseil de l'Europe.

Le «Cadre européen commun de référence pour les langues» en tant que base des standards de formation en langues étrangères – Problèmes méthodologiques du développement et de l'adaptation de descriptions des niveaux de compétence

Résumé

Le «Cadre européen commun de référence pour les langues» (CEC) est largement reconnu dans la discussion autour du développement de standards de formation comme l'exemple d'un modèle de compétences élaboré. Il est également considéré dans le cadre du projet HARMOS comme une bonne base pour déterminer des niveaux de compétences en langues étrangères à l'école obligatoire en Suisse. Après une brève présentation de l'impact du cadre de référence, la contribution décrit le modèle de compétences du CEC et présente les méthodes qui ont permis l'élaboration des niveaux de compétences. A l'aide d'exemples, l'article rend compte ensuite des problèmes qui se posent quand il s'agit de sélectionner, d'adapter et de compléter les descriptions de niveaux. Il se conclut en indiquant les domaines pour lesquels un travail de développement doit être encore mené.

Mots clés: Standards éducatifs Compétences en langues étrangères; modèles de compétences; Echelle de compétences linguistiques

Il «Quadro comune di riferimento europeo per le lingue» quale base per degli standard della formazione nelle lingue straniere – problemi metodologici dello sviluppo e dell'adattamento di descrizioni di competenza

Riassunto

Nella discussione sullo sviluppo degli standard per la formazione, il «Quadro comune di riferimento europeo per le lingue» (QEL) ha assunto valore di esempio di modello per l'elaborazione delle competenze. Di riflesso nel contesto del progetto HARMOS esso fornisce una buona base per determinare i livelli di competenza richiesti nelle lingue straniere nella scuola dell'obbligo svizzera. Dopo una breve presentazione dell'impatto del QEL, l'articolo ne descrive il modello di competenza e illustra i metodi che hanno permesso l'elaborazione dei livelli di competenza. Alcuni esempi permettono di mettere a fuoco i problemi che si pongono quando si tratta di selezionare, adattare e completare le descrizioni dei livelli. Infine si indicano gli ambiti che restano da sviluppare.

Parole chiave: Standard nella formazione, competenze in L2, modello di competenza, scala delle competenze linguistiche

The «Common European Framework of References for Languages» as the Basis for Standards in Foreign Languages - Methodological Issues in Developing and Adapting Descriptors of Language Proficiency

Summary

The «Common European Framework of References for Languages» (CEF) plays a significant role in the current discussion on the development of standards. It is considered an example of a comprehensive descriptive scheme, and in the context of the HARMOS Project, it should serve as a basis to determine levels of proficiency for foreign languages in the school system. After a brief introduction on the impact of the CEF, this article introduces the descriptive model of language use, and the methods used for the development of proficiency levels. Specific examples are given in order to show what problems can arise when descriptors are selected, adapted or added. Finally, areas are highlighted in which developmental work is still needed.

Key words: standards, language proficieny, competence model, language proficiency scale