

ERWIN TSCHIRNER
UNIVERSITÄT LEIPZIG

**ZUR BEWERTUNG MÜNDLICHER KOMPETENZ:
DAS ACTFL ORAL PROFICIENCY INTERVIEW
UND DER GEMEINSAME EUROPÄISCHE
REFERENZRAHMEN FÜR SPRACHEN**

ZWEI WICHTIGE ZIELE DES EUROPARATS SIND DIE BEWAHRUNG DER SPRACHLICHEN und kulturellen Vielfalt Europas und die Förderung von Kommunikation und Interaktion zwischen Europäern unterschiedlicher Muttersprachen. Dadurch soll gegenseitiges Verständnis und Kooperation zwischen den Bürgern Europas gefördert und eine größere Mobilität zwischen den Mitgliedstaaten ermöglicht werden (Council of Europe 2000). Da es in den 40 Mitgliedsstaaten des Europarats 35 offizielle Sprachen und insgesamt 185 Einzelsprachen gibt, fällt der europäischen Sprachenpolitik hierdurch eine zentrale Rolle zu. Um die sprachlichen und kulturellen Barrieren zu überwinden, die auch aus unterschiedlichen Bildungssystemen resultieren, wurde der Gemeinsame europäische Referenzrahmen für Sprachen (GER) geschaffen, der sprachenübergreifend als Basis für die Entwicklung von Lehrplänen, Prüfungen, Lehr- und Lernmaterialien dienen soll. Die drei übergreifenden Ziele des GER sind die folgenden:

- to promote and facilitate cooperation and mutual information among educational institutions in different countries
- to provide a sound basis – a common yardstick – for the mutual recognition of language qualifications
- to assist learners, teachers, course designers, examining bodies and educational administrators to reflect on their current practice and to situate and coordinate their efforts (Council of Europe 2000: 1)

Diese drei Ziele waren auch die Ziele der Tagung „Sprachpolitik und Zertifizierung“, die vom 6. bis 8. Dezember 2005 in Leipzig stattfand und die Kolleginnen und Kollegen der Fächer Deutsch als Fremdsprache und Polnisch als Fremdsprache aus Deutschland und aus Polen zusammenbrachte. Im Mittelpunkt der Tagung stand die integrative Bewertung sprachlicher Kompetenzen. Eine der zentralen Fertigkeiten, die Kommunikation, Interaktion und Mobilität in einem zusam-

menwachsenden Europa ermöglichen, ist sicherlich die mündliche Kompetenz. Ich möchte mich deshalb in meinen Beitrag mit der Fertigkeit Sprechen beschäftigen, die besonders schwer zu beurteilen ist, nicht nur weil hier wie beim schriftlichen Ausdruck eine integrative Bewertung notwendig ist, die das ziel- und handlungsorientierte Zusammenspiel unterschiedlicher sprachlicher Systeme (Aussprache, Grammatik, Lexik, Pragmatik, soziolinguistische und -kulturelle Kompetenz) umfasst, sondern auch weil die Flüchtigkeit der Rede und der Einfluss von Gefühlen, die in der mündlichen Interaktion eine besondere Rolle spielen, eine zuverlässige und valide Bewertung deutlich schwieriger machen.

Um das Ziel des „common yardstick“ – eine sprachenübergreifende Garantie, dass Zertifikate belastbare Handlungskompetenzen verbürgen – auch für den mündlichen Ausdruck zu gewährleisten, müssen neue mündliche Prüfungen entwickelt werden, die internationalen Standards an Validität und Reliabilität genügen. Deshalb lohnt sich ein Blick über die Grenzen Europas hinaus in die USA, wo bereits seit 50 Jahren mit großem Erfolg eine sprachenübergreifende Prüfung mündlicher Handlungskompetenz, das „Oral Proficiency Interview“ (OPI), auf der Grundlage von Profilbeschreibungen eingesetzt wird, dessen Validität und Reliabilität in mehreren Untersuchungen bestätigt wurde.

Während die Skala des Europarats vor allem durch Lehrende validiert wurde (Schneider 2001), kann die OPI-Skala auf eine jahrzehntelange, intensive Prüfungs- und Bewertungstätigkeit zurückblicken (Liskin-Gasparro 2003). Als die „Mutter“ vieler profilbasierter mündlicher Kompetenzskalen hat die OPI-Skala auch den Gemeinsamen Europäischen Referenzrahmen beeinflusst und lässt sich deshalb besonders gut mit ihm vergleichen. Darüber hinaus sind die Erfahrungen, die in den USA und in anderen Ländern mit der OPI-Skala gemacht wurden, aus mindestens zwei weiteren Gründen für die europäischen Länder, die die Skalen des Europarats adoptieren, interessant. Zum einen sind dies didaktisch-methodische und curriculare Erfahrungen (Rifkin 2003), zum anderen sind es Erfahrungen mit Testergebnissen, z.B. welche Niveaus nach wie vielen Lernstunden erreichbar sind (Tschirner/Heilenman 1998), was Lerner auf unterschiedlichen Niveaus können (über die Niveaudeskriptoren hinaus) und welche Niveaus für welche Tätigkeiten notwendig sind (Swender 2003).

Im Folgenden gebe ich einen kurzen Überblick über die Entwicklung des OPI und der ACTFL *Oral Proficiency Guidelines*, beschreibe die Kompetenzprofile der 10 OPI Niveaus und stelle Korrelationen mit der mündlichen Skala des Europarats her. Anschließend stelle ich das Prüfungs- und Bewertungssystem des OPI vor, um ein Beispiel dafür zu geben, wie mündliche Handlungskompetenz auf der Grundlage von Deskriptoren zuverlässig evaluiert werden kann. Schließlich gehe ich auf Ergebnisse von Studien ein, die sich mit der zeitlichen Einordnung von Kompetenzniveaus beschäftigt haben und mit der Art von Sprache, die auf unter-

schiedlichen Niveaus produziert wird, um diese Erfahrungen auch für die mündliche Skala des Europarats fruchtbar zu machen, insbesondere im Hinblick auf curriculare Entscheidungen und für die Entwicklung von Lehrmaterialien.

DIE ACTFL-SKALA UND DAS *ORAL PROFICIENCY INTERVIEW*: EIN RÜCKBLICK

Die ACTFL *Leitlinien mündlicher Handlungsfähigkeit (Oral Proficiency Guidelines)* basieren ursprünglich auf dem Kriterienkatalog des *Foreign Service Instituts* (FSI), das 1952 vom US-amerikanischen Außenminister beauftragt wurde, eine Prüfung zu entwickeln, die die mündliche Handlungsfähigkeit amerikanischer Diplomaten und Beamter, die im Ausland eingesetzt werden sollten, zuverlässig bewerten konnte. Gefordert wurde eine Prüfung, die mündliche Kompetenz auf allen Niveaustufen evaluieren konnte, vom Nullanfänger bis zur fehlerfreien Beherrschung der Fremdsprache. Die ersten Prüfungen wurden holistisch und rein numerisch bewertet, von 0 = keine Kompetenz bis 5 = muttersprachliche Kompetenz, bis 1958 die ersten Niveaubeschreibungen entwickelt wurden, die 5 Kategorien zugeordnet wurden: Aussprache, Verständnis, Flüssigkeit, Grammatik und Wortschatz. Damit nahm die mündliche Prüfung des FSI eine entscheidende Vorreiterrolle ein. Andere mündliche Prüfungen begannen erst in den späten 70-er und frühen 80-er Jahren Skalen zu entwickeln, die beobachtbares Verhalten auf unterschiedlichen sprachlichen Kompetenzniveaus beschrieben und diese Beschreibungen zu Evaluationszwecken verwendeten (Fulcher 1997).

Die Skala und das Interviewformat des FSI wurde in den sechziger Jahren von immer mehr US-amerikanischen Behörden übernommen, u.a. dem *Defense Language Institute* (DLI), dem Peace Corps und dem CIA, die 1968 unter dem Namen *Interagency Language Roundtable* (ILR) eine standardisierte Version für alle staatlichen Stellen und Behörden entwickelten. Die ILR-Prüfung breitete sich in den 70-er Jahren vor allem aus drei Gründen auch an Schulen und Universitäten aus. Sie hatte als direkter Test eine hohe Augenscheinvalidität, die *Interrater-Reliabilität* war sehr hoch, mit einem Reliabilitätskoeffizienten von 0,84 und höher, und sie passte gut zur kommunikativen Methode, die in den 70-er Jahren die didaktische Diskussion zu dominieren begann (Barnwell 1987).

Anfang der achtziger Jahre überarbeitete ACTFL zusammen mit dem ILR und dem *Educational Testing Service* (ETS) die Skala erneut, um sie besser auf die sprachlichen Verhältnisse an Schulen und Universitäten anzupassen. ACTFL erweiterte die unteren Niveaus, legte die oberen Niveaus zusammen und präzierte sowohl das Interviewformat als auch die Beurteilungskriterien. Darüber hinaus wurden Leitlinien zur Erfassung des Hörverständnisses, des Leseverständ-

nisses, der Schreibfertigkeit und des kulturellen Verständnisses verfasst. Diese ACTFL Leitlinien kamen 1982 in einer vorläufigen und 1986 in einer endgültigen Fassung heraus (ACTFL 1986).

In den 90-er Jahren etablierten sich die ACTFL Leitlinien und das darauf basierende OPI als die *de facto* Standardprüfung mündlicher Kompetenz in Nordamerika, sowohl im schulischen und universitären Bereich wie auch in der Geschäftswelt. Die Anstrengungen ACTFLs, das OPI und die Leitlinien an Schulen und Universitäten bekannt zu machen und zu verbreiten, war so erfolgreich, dass die 90-er Jahre als die Jahre des *proficiency movement* in die Geschichte der Fremdsprachendidaktik in Nordamerika eingegangen sind. Der Erfolg der Leitlinien und ihr weitergehender methodisch-didaktischer Einfluss wird auch daran deutlich, dass die Leitlinien mündlicher Handlungskompetenz 1999 in einer überarbeiteten Fassung herauskamen (Breiner-Sanders/Lowe/Miles/Swender 2000), die 2001 von den überarbeiteten Leitlinien für die Fertigkeit Schreiben (Breiner-Sanders/Swender/Terry 2002) ergänzt wurden und in den kommenden Jahren mit Neufassungen für die Fertigkeiten Lesen und Hören komplettiert werden sollen.

DIE MÜNDLICHEN SKALEN VON ACTFL UND EUROPARAT IM VERGLEICH

Die Skala von ACTFL unterscheidet 10 Niveaustufen: *Novice Low, Mid und High, Intermediate Low, Mid und High, Advanced Low, Mid und High und Superior*¹. *Novice* Sprecher können mittels einzelner Wörter, Wortlisten, auswendig gelernter Wendungen und einiger selbst formulierter Neuzusammensetzungen gelernter sprachlicher Elemente einfache Fragen zu sehr einfachen alltäglichen Dingen beantworten. Während Sprecher der Niveaus *Novice Low* und *Novice Mid* nur einzelne Wörter auflisten oder aneinander reihen und nur sehr wenige Fragen beantworten können, wird auf dem Niveau *Novice High* eine, wenn auch sehr begrenzte, aktive Kompetenz sichtbar, wobei schon erste Sätze gebildet und sehr einfache Transaktionen aus dem Alltagsleben halbwegs bewältigt werden können. Die Stufe *Novice High* entspricht ungefähr dem Niveau A1.

Intermediate Sprecher können durch Fragen und Antworten Auskünfte einholen oder geben, einfache sprachliche Interaktionen und Transaktionen initiieren, durchführen und beenden, eigenständig formulieren und ihre Anliegen geduldigen Gesprächspartnern mitteilen, indem sie satzwertige Äußerungen produzieren

¹ Die folgenden Skalenbeschreibungen entstammen der deutschen Übersetzung der ACTFL proficiency guidelines – speaking, revised 1999 (Tschirner 2001a).

und aneinander reihen. Sie können sich an einfachen Gesprächen zu Alltagsthemen und Themen mit Bezug auf die persönliche Umgebung beteiligen und sie können Grundbedürfnisse und grundlegende soziale Anforderungen der Zielkultur befriedigen. Dadurch kommen sie auf einfache Weise in der Zielkultur zu recht. Während *Intermediate Low* Sprecher dies auf sehr einfache Art und Weise vermögen, in ihren Antworten dabei selten über ein oder zwei zusammenhängende Sätze hinausgehen, wobei auch die Themenvielfalt noch stark eingeschränkt ist, reihen *Intermediate Mid* Sprecher deutlich mehr Sätze aneinander, können über mehr einfache und persönliche Themen sprechen und wirken in einfachen Alltagssituationen oft recht fließend und kompetent. *Intermediate Mid* entspricht ungefähr dem Niveau A2.

Intermediate High Sprecher beginnen bereits, ihre Äußerungen mit Nebensätzen anzureichern und entwickeln rudimentäre Textstrukturen. Ihr Themenspektrum ist deutlich erweitert und sie können bereits, wenn auch teilweise nur mit Mühe und mit sprachlichen Ungenauigkeiten, die groben Umrisse einer Geschichte erzählen und über aktuelle Themen sprechen. Das Niveau *Intermediate High* entspricht einem guten B1-Niveau.

Advanced Sprecher können sich aktiv an formellen und informellen Gesprächen über Themen von persönlichem und öffentlichem Interesse beteiligen, können unvorhergesehene Probleme im Gesprächsverlauf mit unterschiedlichen sprachlichen Mitteln lösen, können zusammenhängende Texte (Beschreibungen, Erzählungen, Berichte u.ä.) über Gegenwärtiges, Vergangenes und Zukünftiges, unter angemessener Markierung aspektueller Verhältnisse und mit angemessener Sicherheit und sprachlicher Korrektheit produzieren und können allgemein die sprachlichen Anforderungen von Schule, Universität oder Beruf erfüllen. *Advanced Low* Sprecher tun dies auf eher einfache Weise mit Schwierigkeiten im Ausdruck und einer gewissen Holprigkeit (fehlenden Flüssigkeit), während *Advanced Mid* Sprecher die erwähnten Sprachhandlungen auf generell überzeugende und flüssige Art durchführen können. *Advanced Low* entspricht einem niedrigen B2-Niveau und *Advanced Mid* einem hohen B2-Niveau.

Advanced High Sprecher erfüllen bereits viele Anforderungen des *Superior* Niveaus, allerdings nicht durchgängig und teilweise mit deutlichen Schwierigkeiten. *Advanced High* entspricht ungefähr dem Niveau C1. *Superior* Sprecher können professionelle Anforderungen des Berufslebens bzw. der akademischen Welt erfüllen. Sie können mit Hilfe ausführlicher und gut strukturierter Diskussionsbeiträge Standpunkte erläutern und begründen und überzeugende Hypothesen entwickeln. Sie können Themen auf konkrete und abstrakte Weise erörtern und sprachlich ungewohnte und unvorbereitete Situationen bewältigen, wobei sie ein hohes Maß an sprachlicher Korrektheit einhalten können. *Superior* entspricht ungefähr dem Niveau C2.

Tabelle 1. Die Skalen von Europarat und ACTFL im Vergleich

Europarat	A1	A2	B1	B2	C1	C2
ACTFL	Nov High	Int Mid	Int High	Adv Mid	Adv High	Superior

DAS PRÜFUNGS- UND BEWERTUNGSVERFAHREN DES ORAL PROFICIENCY INTERVIEWS

Ziel einer mündlichen Prüfung sollte es sein, eine repräsentative Auswahl der Sprachhandlungen, die eine Person durchführen kann, auf ökonomische Weise zu liefern. Gleichzeitig sollte darauf geachtet werden, wie einzelne Aufgaben mit den Merkmalen der getesteten Person interagieren: mit ihrer sprachlichen Kompetenz, mit ihrem Sach- und Fachwissen und mit ihren Gefühlen. Schließlich soll das Prüfungsgespräch eine positive Rückwirkung auf die getestete Person haben, auf die Art und Weise, wie sie ihre eigene Kompetenz wahrnimmt, wie zufrieden sie mit ihr ist, und darauf, welche Schlüsse sie im Hinblick auf ihre weitere fremdsprachliche Entwicklung zieht (Bachman/Palmer 1996). Im Folgenden soll an der Struktur des Prüfungsgesprächs dargestellt werden, wie das OPI diese Ziele zu erfüllen versucht.

Das OPI beginnt mit einer unbewerteten Aufwärmphase². Sie ist dafür gedacht, das Gespräch auf dem Niveau zu beginnen, mit dem der Kandidat oder die Kandidatin keine Probleme hat, damit er oder sie sich zu Beginn des Prüfungsgesprächs kompetent fühlen kann. Gleichzeitig bewirkt diese Phase, dass sich die Testperson „aufwärmt“, d.h. beginnt, wieder in der Fremdsprache zu denken. Ein weiteres Ziel dieser Phase ist es, Prüfer und Prüfling miteinander bekannt zu machen. Dies soll dazu führen, dass das Gespräch authentischer wird. Man unterhält sich, um sich weiter kennen zu lernen. Dazu soll dadurch eine freundliche und angenehme Atmosphäre geschaffen werden. Die Aufwärmphase wird weiterhin dazu benutzt, um zu erfahren, wofür sich die getestete Person interessiert, um Themen für die Prüfungsphase des Interviews zu gewinnen. Schließlich bekommt der Prüfer im Rahmen dieser Phase bereits erste Hinweise darauf, auf welchem Niveau sich die getestete Person befinden könnte.

Im Hauptteil der Prüfung werden vier bis fünf Themen behandelt, die sich auf den unteren Niveaus oft auf elementare alltägliche Erfahrungen bzw. Autobiographisches beschränken, auf den höheren Niveaus jedoch mit dem Testkandidaten ausgehandelt werden müssen. Dieses Aushandeln der Themen hat vor allem das Ziel, die sprachliche Kompetenz der Kandidaten in den Mittelpunkt zu stellen und zu vermeiden, dass Testergebnisse durch unterschiedliches fachliches Wissen bzw. durch unterschiedliche Gefühle beeinflusst werden. Gleichzeitig verstärkt das Aushandeln der Themen mit den Kandidaten die Authentizität des Gesprächs. Schließlich führt es auch zu

² Die folgende Darstellung des OPI beruht auf dem ACTFL OPI Tester Training Manual (Swender 1999).

einer größeren Testökonomie, weil vermieden wird, Themen anzuschneiden, über die die Kandidaten nicht Bescheid wissen bzw. nichts zu sagen haben.

Zu jedem Thema werden zuerst Fragen und Aufgaben gestellt, die der Kandidat vermutlich problemlos beantworten kann (Boden), und dann Fragen und Aufgaben, die er wahrscheinlich nur mit Mühe oder gar nicht beantworten kann (Decke). Man erhält so eine Profilbeschreibung, die die Person eindeutig zwischen zwei Niveaus platziert. Der Wechsel zwischen für die Kandidaten einfachen und schwierigen Fragen führt sie immer wieder auf ein Niveau von Mühelosigkeit zurück. Dies verschafft ihnen zum einen immer wieder Erfolgserlebnisse, ein Gefühl von Kompetenz und damit eine positive emotionale Grundstimmung, und zum anderen notwendige Ruhepausen für die schwierigeren Fragen und Aufgaben. Die Proben auf dem nächsthöheren Sprachniveau, welches die Testperson noch nicht oder nicht gut beherrscht, haben das Ziel, die Prüfung auch nach oben hin bewertbar zu machen, denn nur in der Zusammenschau zwischen gelösten und nicht gelösten Aufgaben lassen sich Prüfungsteilnehmer präzise einordnen. Ein weiterer Vorteil des Wechsels zwischen lösbaren und komplizierten Aufgaben ist das Feedback, das die getestete Person dadurch erhält. Sie bekommt ein relativ deutliches Bild ihrer Fähigkeiten, darüber, was sie kann und was sie nicht kann. Dies erhöht zum einen die Validität der Prüfung für die getestete Person (Augenscheinvalidität), zum anderen vermittelt es ein persönliches Profil, das als Ausgangspunkt weiterer Lernwege dienen kann.

Weil in einem Prüfungsgespräch nicht alle Sprechhandlungen und sozialen Kontexte auf authentische Weise elizitiert werden können, werden die Kandidaten, meist gegen Ende des Interviews, gebeten, mit dem Prüfer ein Rollenspiel durchzuführen, um z.B. einfache oder komplexe Transaktionen aus dem Alltag zu simulieren, um ein weiteres Register in das Prüfungsgespräch einzuführen oder um das Autoritätsgefälle in der Prüfung umzukehren, damit die Testperson mehr Eigeninitiative zeigen kann. Das OPI endet mit einer Abkühlphase, in der die Kandidaten auf das Niveau zurückgebracht werden, das sie am besten beherrschen. Diese Phase dient dazu, das Prüfungsgespräch für die Testperson mit einem Erfolgserlebnis zu beenden, um eine positive Rückwirkung auf das weitere Lernen zu haben.

Das ACTFL OPI wird holistisch (integrativ) bewertet. Die Leistung eines Kandidaten wird nicht nach Punkten bewertet, sondern es wird verglichen, zu welcher Profilbeschreibung sie am besten passt. Die Prüfung selbst wird von einem einzelnen Prüfer abgenommen, der sich während der Prüfung keine Notizen macht, um sich auf das Prüfungsgespräch konzentrieren zu können. Die Prüfung wird auf Band aufgenommen und im Anschluss vom Prüfer mindestens einmal komplett angehört, bevor er seine Bewertung abgibt. Das Band wird dann von einem zweiten Bewerter angehört, der die Einschätzung des ersten Prüfers nicht kennt (blinde Doppelbewertung). Wenn die Bewertungen nicht übereinstimmen, wird das Band an einen dritten Bewerter geschickt, der seine Beurteilung ebenfalls blind abgibt. Prüferinnen und Prüfer wer-

den in einem aufwändigen Verfahren, das bis zu einem Jahr dauern kann, geschult und müssen eine Prüfung ablegen, bevor sie von ACTFL zertifiziert werden.

Die Zuverlässigkeit des ACTFL OPI wurde in vielen Studien untersucht. Der Reliabilitätskoeffizient nach Pearson erreichte in allen Studien einen Wert von mindestens 0,8 und oft einen deutlich höheren (Kenyon/Tschirner 2000). Dieser Wert wird international bei wichtigen Prüfungen als die untere Grenze der Zuverlässigkeit angesehen (vgl. für schriftliche Prüfungen Hamp-Lyons 1990). Surface und Dierdorff (2003) untersuchten die Zuverlässigkeit des OPI in einer großangelegten Studie mit 5881 OPIs in 19 unterschiedlichen Sprachen und ermittelten einen durchschnittlichen Reliabilitätskoeffizienten von 0,978, wobei keine Sprache ein schlechteres Ergebnis als 0,94 erreichte.

DER FAKTOR ZEIT

Eine wichtige curriculare Entscheidung ist es, wie viel Zeit man für das Erreichen eines bestimmten Niveaus vorsieht. Zwar ist die Anzahl der Unterrichtsstunden keine Garantie dafür, dass ein bestimmtes Niveau erreicht wird, trotzdem können hochspezialisierte und -automatisierte Kenntnisse und Fähigkeiten wie professionelle fremdsprachliche Handlungsfähigkeit, ähnlich wie professionelle Kompetenzen in der Musik und im Sport, nur durch tägliches, mehrstündiges Üben und Anwenden über einen mehrjährigen Zeitraum hindurch erworben werden. Weil das OPI von Beginn an mit einem großen methodisch-didaktischen Interesse verbunden war, wurde der Faktor Zeit relativ früh untersucht.

Die ältesten zeitlichen Vorgaben für Kurse zur Entwicklung mündlicher Handlungsfähigkeit stammen vom *Foreign Service Institut* (FSI) (Thompson 1996), die nach wie vor verwendet werden und die von empirischen Studien im Großen und Ganzen bestätigt wurden. Das FSI unterscheidet zwischen vier Sprachgruppen, je nach Schwierigkeit ihres Erlernens für Muttersprachler des Englischen.

Tabelle 2 zeigt, wie viele Unterrichtsstunden in intensiven, handlungsorientierten Fremdsprachenkursen des FSI mit bis zu vier Teilnehmern pro Kurs für die einzelnen Sprachgruppen veranschlagt werden. Zur Sprachgruppe 1 gehören germanische und romanische Sprachen wie Dänisch, Niederländisch, Norwegisch, Schwedisch, Französisch, Italienisch, Portugiesisch, Rumänisch und Spanisch, ebenso wie die Kreolsprachen Suaheli und Haitisch. Zur Sprachgruppe 2 gehören u.a. Bulgarisch, Dari, Deutsch, Farsi, Hindi, Indonesisch, Malayisch, Neugriechisch und Urdu. Zur Sprachgruppe 3 gehören Amharisch, Bengalisch, Burmesisch, Finnisch, Hebräisch, Khmer, Laotisch, Nepalisch, Pilipino, Polnisch, Russisch, Serbokroatisch, Sinhalesisch, Tamilisch, Thai, Tschechisch, Türkisch, Ungarisch und Vietnamesisch. Zur Sprachgruppe 4 gehören Arabisch, Chinesisch, Japanisch und Koreanisch.

Tabelle 2. Unterrichtsstunden und Kompetenzniveaus nach Sprachgruppen

	240 h	480 h	720 h	1320 h	2400–2760 h
Gruppe 1	IM	IH	AL		
Gruppe 2		IM	IH	AM	
Gruppe 3		IL	IM	AL	
Gruppe 4		NH	IL	IH	AM

Empirische Studien bestätigen im Großen und Ganzen diese Stundenvorgaben, deuten aber darauf hin, dass sie oft nicht genügen und dass andere Variablen, wie z.B. ein Auslandsaufenthalt, eine große Rolle spielen können. Eine Studie mit Fremdsprachenlernern am FBI (Walker/Williams/Navarrete 1988) zeigte, dass bei Sprachen der Schwierigkeitsstufe 1 das Niveau AL erst nach durchschnittlich 1080 Stunden erreicht wird und bei Sprachen der Schwierigkeitsstufe 3 erst nach 1410 Stunden. Bei Sprachen der Schwierigkeitsstufe 4 schließlich wird selbst nach 1410 Stunden im Allgemeinen nur das Niveau IH erreicht.

Magnan (1986) stellte fest, dass Studierende an der University of Wisconsin in Madison nach einem Jahr Französischunterricht (ca. 150 Stunden) im Allgemeinen auf dem Niveau IL waren und nach zwei Jahren (ca. 300 Stunden) auf dem Niveau IM. Studierende mit dem Bachelorabschluss Französisch als Nebenfach (ca. 500 Stunden) waren meist auf dem Niveau IH, wobei alle geprüften Studierenden angaben, zwischen einigen Wochen und sechs Monaten in einem französischsprachigen Land verbracht zu haben. Studierende mit Bachelorabschluss Französisch als Hauptfach (ca. 700 Stunden) waren meist auf dem Niveau AL/AM. Zusätzlich zu ihrem Studium verbrachten diese Studierenden im Durchschnitt 6 Monate in einem französischsprachigen Land. Tschirner (1996) und Tschirner/Heilenman (1998) ermittelten für Deutsch als Fremdsprache bei Studierenden ähnliche, wenn auch teilweise niedrigere Ergebnisse, nämlich IL nach einem Jahr Deutsch (ca. 120 Stunden) und IL bzw. IM nach zwei Jahren. Thompson (1996) erhielt für Russisch als Fremdsprache an Universitäten deutlich niedrigere Ergebnisse, nämlich NM nach einem Jahr, NM/IL nach zwei Jahren, IM/IH nach drei Jahren, IH/AL nach vier Jahren und AM/AH nach fünf Jahren. Rifkin (2003) stellte für Russisch noch niedrigere Werte fest, nämlich nach einem Jahr (150 Stunden) im Durchschnitt NH, nach zwei Jahren (300 Stunden) IL, nach drei Jahren (450 Stunden) IM und nach vier Jahren (520 Stunden) ohne Auslandsaufenthalt ebenfalls IM.

Die Beobachtung, dass die Schwelle von *Intermediate* zu *Advanced* oft nur nach einem längeren Auslandsaufenthalt überschritten wird, wird immer wieder gemacht, besonders bei Sprachen, die nicht der Gruppe 1 angehören. Rifkin (2003) selbst fand ein Niveau von AL oder besser nur bei Studierenden vor, die mindestens ein Semester in Russland studiert hatten. Selbst ein Auslandsstudium ist allerdings keine Garantie dafür, dass die Schwelle zu *Advanced* überschritten wird. Nach Brecht, Davidson und Ginsberg (1993) und Davidson (1998, 2002)

erreicht nur ein gutes Drittel von Russischstudierenden nach einem Auslandsstudium von einem oder zwei Semestern das *Advanced* Niveau. Etwas einfacher scheinen es Studierende von romanischen Sprachen zu haben. O'Connor (1988) prüfte die Sprechkompetenz von 30 Studierenden am Middlebury College vor und nach einem Studienjahr in Frankreich. Vor der Abreise nach Frankreich waren diese Studenten IM oder IH und nach ihrer Rückkehr IH oder AL/AM. Milleret (1990) stellte bei Portugiesischstudierenden, die an einem sechswöchigen Intensivprogramm in Brasilien teilnahmen, ebenfalls einen Fortschritt von IM zu IH fest.

Swender (2003) fasst die Ergebnisse von fünf amerikanischen Universitäten zusammen, die über einen Zeitraum von 5 Jahren (1998–2002) ihre Studierenden (insgesamt 501) regelmäßig zum Abschluss ihres Bachelorstudiums einer Fremdsprachenphilologie von ACTFL testen ließen. Viele dieser Studierenden hatten dabei an einem Austauschprogramm, im Normalfall ein akademisches Jahr, teilgenommen. 17% der Studierenden der Fächer Hispanistik, Französisistik und Italianistik (insgesamt 442 Studierende) waren am Ende ihres Bachelorstudiums noch IM (Europarat: A2), 57% waren IH oder AL (B1) und 26% waren AM (B2) oder besser. Ähnliche Ergebnisse erzielten Studierende der Germanistik und Russistik (insgesamt 39 Studierende)³. 26% waren IM (A2), 44% waren IH/AL (B1) und 30% waren AM (B2) oder besser. Weniger gute, aber doch recht passable Ergebnisse erzielten Studierende der Sinologie und Japanologie (insgesamt 20 Studierende). 35% waren IM (A2), 50% waren IH/AL (B1) und 15% waren AM (B2) oder besser.

Diese und andere Untersuchungen zum Erwerb der Sprechkompetenz in universitären Bachelorstudiengängen ergeben im Allgemeinen folgendes Muster. Nach ca. 300 Unterrichtsstunden werden je nach Zielsprache die Kompetenzniveaus NH bis IM erreicht. Um die Niveaus IH/AL und AM zu erreichen bedarf es meist eines Auslandsaufenthalts. Bei Abschluss eines Bachelorstudiums (ca. 700–800 Stunden) erreicht die Mehrzahl der Studierenden der Fremdsprachenphilologien im Allgemeinen, vor allem wenn sie einen Teil ihres Studiums im Land der Zielsprache verbringen, das Niveau IH/AL, ein gutes Viertel (mit Auslandsstudium) auch das Niveau AM (außer bei Sprachen der Gruppe 4).

MERKMALE LERNERSPRACHLICHER TEXTE AUF UNTERSCHIEDLICHEN KOMPETENZNIVEAUS

Neben der Frage nach Unterrichtsstunden bzw. Länge von Auslandsaufenthalten rücken seit den neunziger Jahren verstärkt curriculare Fragen in den Mittelpunkt der Aufmerksamkeit, insbesondere Fragen nach einer sinnvollen grammatischen

³ Zum Vergleich die Prozentzahlen für Deutsch (32 Studierende): IM = 28%, IH/AL = 41%, AM+ = 31% und Russisch (7 Studierende): IM = 14%, IH/AL = 57%, AM+ = 29%.

Progression zur Stärkung der Sprechkompetenz. Magnan (1988) stellte bei Ausgangssprache Englisch und Zielsprache Französisch fest, dass z.B. Kongruenzfehler bei Verben ab dem Niveau IH deutlich abnehmen und dass der Fehlerquotient bei Artikelwörtern, Stellung und Kongruenz von Adjektiven und korrekter Verwendung von Objekt- und Relativpronomen erst ab *Advanced* unter die Zwanzigprozentmarke fiel.

Tschirner (1996) regte an, sich auf einer gegebenen Stufe auf die grammatischen Phänomene zu konzentrieren, die auf dem jeweils nächsthöheren Niveau kontrolliert werden. Auf dem *Novice* Niveau wäre die Grammatik des *Intermediate* Niveaus zu lehren und zu üben, auf dem *Intermediate* Niveau die Grammatik des *Advanced* Niveaus usw. Dazu stellte er auf Grundlage der Niveaubeschreibungen theseartig folgende Grammatikelemente zusammen, deren mündliche Beherrschung auf den Niveaus *Intermediate* und *Advanced* gegeben sein müsste.

- Intermediate: Präsens, Subjekt-Verb-Kongruenz, Verbalklammer, Inversion in Wort- und Satzfragen
- Advanced: Perfekt, Kasus, Inversion in Aussagesätzen, Verbstellung in Nebensätzen.

Um Teile dieser Hypothesen empirisch zu überprüfen, analysierte Tschirner (1996) 40 OPIs auf den Niveaus NM, NH, IL und IM und stellte fest, dass der Prozentsatz korrekter Verwendung von Verbkongruenz und Verbalklammer von Niveau zu Niveau stetig ansteigt, während weder die Inversion in Aussagesätzen noch die Endstellung des finiten Verbs in Nebensätzen auf diesen Niveaus erworben sind.

Tabelle 3: Durchschnittlicher Prozentsatz korrekt bei drei oder mehr obligatorischen Kontexten pro Interview mit Anzahl der Teilnehmer, die drei oder mehr obligatorische Kontexte produzierten, in Klammern. Die Anzahl der Prüfungen pro Niveau steht hinter den Abkürzungen für die Niveaus in Klammern. Durchschnittliche Anzahl der mit einem Subjekt verwendeten Verben (ohne sein, haben, werden und Modalverben) und durchschnittlicher Prozentsatz korrekter Verwendung der Kongruenz bei diesen Verben.

	NM (N = 3)	NH (N = 7)	IL (N = 10)	IM (N = 20)
Verbalklammer		40% (5)	64% (10)	92% (20)
Inversion		17% (4)	25% (9)	34% (18)
Nebensatzstellung				24% (14)
Kongruenz	56%	64%	75%	88%
Sätze	10,33	29,43	32,20	42,26

Nach Tabelle 3 wird spätestens auf dem Niveau IM die Verbalklammer in 92% der obligatorischen Kontexte richtig verwendet. Weder die Inversion in Aussagesätzen (34% korrekt) noch die Verbendstellung in Nebensätzen (24% korrekt) scheinen auf diesem Niveau erworben zu sein. Die Verbkongruenz im Präsens in

Hauptsätzen wiederum scheint spätestens auf dem Niveau IM erworben zu sein. Bei einer durchschnittlichen Anzahl von 42,26 Sätzen pro Interview, d.h. Äußerungen mit mindestens einem Subjekt und einem Verb, wurden 88% dieser Verben mit der korrekten Endung versehen. Auffällig ist, dass selbst auf dem niedrigsten Niveau der Untersuchung, NM, die Kongruenz bereits in mehr als der Hälfte der Sätze, im Durchschnitt 10,33 pro Interview, korrekt verwendet wird und diese korrekte Verwendung auf den folgenden Niveaus stetig ansteigt. Dies hat möglicherweise damit zu tun, dass sich die untersuchte Stichprobe aus Studierenden zusammensetzte, die Deutsch an einer amerikanischen Universität studierten und damit sowohl explizites Grammatikwissen besaßen als auch wahrscheinlich eine höhere Sprachaufmerksamkeit als z.B. Einwanderer. Daneben mögen, insbesondere auf den unteren Niveaus, eine Reihe von Subjekt-Verb-Verbindungen als lexikalische Einheiten verwendet worden sein.

Eine spätere Studie mit 20 amerikanischen Studierenden nach 4 Semestern Deutsch (Tschirner/Meerholz-Härle 2001) ergab ein sehr ähnliches Bild (vgl. Tabelle 4). Auch hier erscheint die Verbalklammer spätestens mit dem Niveau IM (91% korrekt) erworben, während in zwei Drittel der Sätze weder die Inversion in Aussagesätzen (32% korrekt) noch die Nebensatzstellung (37% korrekt) richtig verwendet werden. Subjekt-Verb-Kongruenz wiederum wird bereits auf dem Niveau NH zu 70% richtig verwendet und erreicht auf dem Niveau IM bereits eine korrekte Verwendung von 90%.

Tabelle 4: Durchschnittlicher Prozentsatz korrekt bei drei oder mehr obligatorischen Kontexten pro Interview mit Anzahl der Teilnehmer, die drei oder mehr obligatorische Kontexte produzierten, in Klammern. Die Anzahl der Prüfungen pro Niveau steht hinter den Abkürzungen für die Niveaus in Klammern. Durchschnittliche Anzahl aller mit einem Subjekt verwendeten Verben durchschnittlicher Prozentsatz korrekter Verwendung der Kongruenz bei diesen Verben.

	NH (N = 5)	IL (N = 9)	IM (N = 6)
Verbalklammer	56% (5)	81% (9)	91% (6)
Inversion	5% (5)	30% (8)	32% (6)
Nebensatzstellung	0% (2)	5% (3)	37% (6)
Kongruenz	70% (5)	76% (9)	90% (6)
Sätze	18,6	36,6	43,7

Eine weitere Studie beschäftigte sich mit den Niveaus IM bis Superior (Norris 1996) und bestätigte auch auf diesen Niveaus die von Tschirner (1996) aufgestellten Thesen für den Erwerb der Wortstellung und der Kongruenz. Norris (1996) analysierte 44 Simulated Oral Proficiency Interviews (SOPI), eine kassettengestützte Variante des OPI (vgl. Kenyon/Tschirner 2000). Nach Tabelle 5 sind sowohl die Verbalklammer (84%) als auch die Kongruenz (90%) auf dem Niveau IM weitgehend erworben, während die Nebensatzstellung (21%) noch nicht er-

worben ist. Ab dem Niveau IH steigert sich die Verwendung von Nebensätzen nicht nur erheblich, sondern die Nebensatzstellung (86%) kann als weitgehend erworben gelten.

Tabelle 5. Durchschnittlicher Prozentsatz korrekt bei vier und mehr obligatorischen Kontexten pro Prüfung für drei syntaktische Regeln (Verbalklammer, Kongruenz, Nebensatzstellung) und durchschnittliche Anzahl der Nebensätze pro Interview. Die Anzahl der Prüfungen pro Niveau steht hinter den Abkürzungen für die Niveaus in Klammern (Norris 1996).

	IM (3)	IH (10)	A (11)	AH (4)	S (2)
Verbalklammer	84%	98%	97%	98%	100%
Kongruenz	90%	94%	98%	99%	100%
Nebensatzstellung	21%	86%	89%	90%	98%
Nebensätze	4,33	14,30	19,18	17,75	22,50

Weder Tschirner (1996) und Tschirner/Meerholz-Härle (2001) noch Norris (1996) untersuchten den Erwerb des Kasus. Die Daten des Genfer DiGS-Projekts (Diehl u.a. 2000) scheinen jedoch darauf hinzudeuten, dass der Kasuserwerb erst in Angriff genommen wird, wenn der Erwerb der Wortstellung im Großen und Ganzen abgeschlossen ist. Sollten sich diese aus Aufsatzdaten gewonnenen Ergebnisse auch für die gesprochene Sprache zeigen, würde dies die ursprünglichen Hypothesen von Tschirner (1996) weiter bestätigen.

AUSBLICK

Die vorliegenden Forschungsergebnisse im Zusammenhang mit dem ACTFL OPI werfen eine Reihe von Fragen auch für die Skalen des GER auf. Zum einen deuten diese Ergebnisse darauf hin, dass sich die Entwicklung der mündlichen Kompetenz von A1 bis C2 nicht in überschaubare Blöcke von jeweils ca. 300 Unterrichtsstunden einteilen lässt, sondern dass die Entwicklung auf den höheren Niveaus deutlich langsamer verläuft und möglicherweise ohne die intensive Erfahrung eines längeren Auslandsaufenthalts trotz vieler weiterer Unterrichtsstunden auf dem Niveau B1 stagniert. Da es zum GER noch keine zuverlässigen und validen mündlichen Sprachstandsprüfungen gibt, besteht die Gefahr, dass Lernern nach Bestehen z.B. der Zentralen Mittelstufenprüfung des Goethe-Instituts einfach bescheinigt wird, dass sie auch im Mündlichen das Niveau B2 erreicht haben, auch wenn dies gar nicht der Fall ist. Sicherlich sollte man auch für den GER davon ausgehen, dass am Ende eines jeweiligen Kursblockes deutlich unterschiedliche Niveaus in den vier Fertigkeiten erreicht werden können. Wenn ein Lehrwerk verspricht, dass in 300 Stunden in allen vier Fertigkeiten das gleiche Niveau erreicht wird, kann das nur reines Wunschdenken sein.

Zum anderen deuten die Ergebnisse von Studien zum OPI darauf hin, dass die grammatische Progression der Lehrwerke, die auf die jeweilige Niveaustufe vorbereiten sollen, zumindest was die mündliche Handlungsfähigkeit betrifft oft fern jeglicher Realität ist. Wenn grammatische Strukturen wie Nebensatzstellung, Inversion in Aussagesätzen, Perfekt und Kasus erst ab dem Niveau B2 in spontan gesprochener Sprache in einem nennenswerten Umfang richtig verwendet werden, stellt sich die Frage, ob es nicht sinnvoll wäre, sich auf den Niveaus A1 und A2 auf Funktionen und Strukturen zu beschränken, die auf diesen Niveaus auch gelernt werden können. *Profile Deutsch* (Glaboniat u.a. 2002) z.B. reiht sich in seiner Grammatikprogression nahtlos in deutsche Lehrwerkstraditionen ein, schlägt vor, den Akkusativ auf A1 einzuführen, den Dativ auf A2, das Perfekt auf A2, das Präteritum auf B1, den Großteil der Nebensätze auf A2 usw. Zwar sagen Glaboniat u.a. ausdrücklich, dass es sich nicht um eine Erwerbsgrammatik handelt und dass es nicht darum geht vorherzusagen, was ein Lerner auf einem bestimmten Niveau kann (S. 154). Nur stellt sich dann die Frage, warum die Grammatikprogression völlig losgelöst von sprachlicher Handlungskompetenz entwickelt wurde. Auch wenn die empirische Basis der Kannbeschreibungen schmal ist und die von grammatischen Kompetenzen noch schmaler, führt sicherlich kein Weg daran vorbei, genau diese empirische Basis zu vergrößern und nicht traditionelle Progressionen einfach auf die GER-Niveaus zu verteilen. Dies führt wahrscheinlich dazu, dass für unterschiedliche Fertigkeiten unterschiedliche Grammatikprogressionen entwickelt werden müssen, darunter eine Grammatikprogression speziell für den mündlichen Ausdruck (Tschirner 2001).

Das ACTFL OPI ist für die europäische Diskussion im Zusammenhang mit dem GER sowohl als Prüfung interessant als auch für die curriculare Planung und die Entwicklung von Lehrmaterialien. Es ist als Prüfung interessant, weil sich die ACTFL Skala ohne größere Schwierigkeiten mit der Skala des GER korrelieren lässt und damit ein Instrument zur Verfügung steht, mit dem GER-Prüfungen validiert und erreichte Niveaus kontrolliert werden können. Es ist für die curriculare Planung interessant, weil es eine Fülle empirischer Studien gibt, die versucht haben zu erfassen, wie lange es dauert, bis ein bestimmtes Niveau erreicht wird und welche sprachlichen Strukturen auf unterschiedlichen Niveaus beherrscht werden. Es würde der europäischen Curriculum- und Lehrmaterialienentwicklung sicher nicht schaden, wenn sie diese Studien nicht nur wahrnehmen würde, sondern deren Ergebnisse gegebenenfalls auch einarbeiten würde.

LITERATUR

- American council on the teaching of foreign languages. *Proficiency Guidelines*, 1986, Yonkers, NY, ACTFL.
 Bachman L., Palmer A., 1996: *Language testing in practice*, Oxford, Oxford University Press.

- Barnwell D., 1987: *Oral proficiency testing in the United States*, in „British Journal of Language Teaching“ 25, p. 35–42.
- Brecht R.D., Davidson D., Ginsberg R.B., 1993: *Predictors of foreign language gain during study abroad*, Washington, National Foreign Language Center.
- Breiner-Sanders K., Lowe Jr. P., Miles J., Swender E., 2000: *ACTFL proficiency guidelines – speaking, revised*, in: „Foreign Language Annals“ 33, p. 13–18.
- Breiner-Sanders K., Swender E., Terry R., 2002: *Preliminary proficiency guidelines – writing revised*, in: „Foreign Language Annals“ 35, p. 9–15.
- Council of Europe, 2000: *Common European Framework of Reference for Languages. Learning, Teaching, Assessment*. <http://culture2.coe.int/portfolio/documents/0521803136txt.pdf>.
- Davidson D., 1998: *Summer, semester, and academic year proficiency profiles*, paper presented at the Annual AATSEEL Conference, San Francisco.
- Davidson D., 2002: *When just being there is not enough*, paper presented at the Symposium on Language Gain in Study Abroad. Global Forum on Language, Literature and Culture, University of Wisconsin, Madison.
- Diehl E., Christen H., Leuenberger S., Pelvat I., Studer T., 2000: *Grammatikunterricht: Alles für der Katz? Untersuchungen zum Zweitspracherwerb Deutsch*, Tübingen, Niemeyer.
- Fulcher G., 1997: *The testing of L2 speaking*, in: Clapham C., Corson D., ed., *Encyclopedia of language and education, vol. 7: Language testing and assessment*, Dordrecht, Kluwer, p. 75–85.
- Glaboniat M., Müller M., Schmitz H., Rusch P., Wertenschlag L., 2002: *Profile deutsch*, Berlin m.in., Langenscheidt.
- Hamp-Lyons L., 1990: *Second language writing: Assessment issues*, in: Kroll B., ed.: *Second language writing: Research insights for the classroom*, Cambridge, Cambridge University Press, p. 69–87.
- Kenyon D., Tschirner E., 2000: *The rating of direct and semi-direct oral proficiency interviews: Comparing performance at lower proficiency levels*, in: „Modern Language Journal“ 84, p. 85–101.
- Liskin-Gasparro J., 2003: *The ACTFL Proficiency Guidelines and the Oral Proficiency Interview: A brief history and analysis of their survival*, in: „Foreign Language Annals“ 36, p. 483–490.
- Magnan S., 1986: *Assessing speaking proficiency in the undergraduate curriculum: Data from French*, in: „Foreign Language Annals“ 19, p. 429–438.
- Magnan S., 1988: *Grammar and the ACTFL Oral Proficiency Interview: Discussion and data*, in: „Modern Language Journal“ 72, p. 266–276.
- Milleret M., 1990: *Assessing the gain in oral proficiency from summer foreign study*, paper presented at the Annual AATSP Conference.
- Norris J., 1996: *A validation study of the ACTFL guidelines and the German Speaking Test*, Magisterarbeit, University of Hawaii.
- O'Connor N., 1988: *Oral proficiency testing of junior year abroad: Implications for the Undergraduate curriculum*, paper presented at the Annual MLA Conference.
- Rifkin B., 2003: *Oral proficiency learning outcomes and curricular design*, in: „Foreign Language Annals“ 36, p. 582–588.
- Schneider G., 2001: *Kompetenzbeschreibungen für das Europäische Sprachenportfolio*, in: „Fremdsprachen Lehren und Lernen“ 30, S. 193–214.
- Surface E., Dierdorff E., 2003: *Reliability and the ACTFL Oral Proficiency Interview: Reporting indices of interrater consistency and agreement for 19 languages*, in: „Foreign Language Annals“ 36, p. 507–519.
- Swender E., ed., 1999: *ACTFL Oral Proficiency Interview: Tester training manual*, Yonkers, NY, ACTFL.
- Swender E., 2003: *Oral proficiency testing in the real world: Answers to frequently asked questions*, in: „Foreign Language Annals“ 36, p. 520–535.

- Thompson I., 1996: *Assessing foreign language skills: Data from Russian*, in: "Modern Language Journal" 80, p. 47–65.
- Tschirmer E., 1996: *Scope and sequence: Rethinking beginning foreign language instruction*, in: „Modern Language Journal“ 80, p. 1–14.
- Tschirmer E., 2001a: *Die ACTFL Leitlinien mündlicher Handlungsfähigkeit*, in: „Fremdsprachen Lehren und Lernen“ 30, S. 116–126.
- Tschirmer E., 2001b: *Kompetenz, Wissen, mentale Prozesse: Zur Rolle der Grammatik im Fremdsprachenunterricht*, in: Funk H., Koenig M. (Hrsg.), *Kommunikative Didaktik in Deutsch als Fremdsprache – Bestandsaufnahme und Ausblick. Festschrift für Gerhard Neuner*, München, Iudicium, S. 106–125.
- Tschirmer E., Heilenman, L.K., 1998: *Reasonable expectations: Oral proficiency goals for intermediate students of German*, in: „Modern Language Journal“ 82, p. 147–158.
- Tschirmer E., Meerholz-Härle B., 2001: *Processability theory: Data from oral proficiency interviews*, paper presented at EUROSLA 11, Paderborn.
- Walker M., Williams M., Navarrete O., 1988: *Aptitude and language learning of FBI Special Agents*, paper presented at the Interagency Language Roundtable Invitational Symposium on Language Aptitude Testing, Washington.