

Im zweiten Block des Freitagnachmittags präsentierten Jiang Mingbao (Beijing Yuyan Xueyuan), Noël Dutrait (Université de Provence), Yu Youlan (University of London), Jean-François Vergnaud und M. Pompidor (Université de Montpellier 3) neue Ansätze zum Unterrichtseinsatz von neuen technischen Medien und Video-Lehrmaterialien. Jin Si-yan (Université Paris 7) berührte in ihrem Referat die nach wie vor offenen Fragen in der audio-visuellen Vermittlung der chinesischen Sprache und der oft inakzeptablen Vermischung umgangssprachlicher (*kouyu*) und schriftsprachlicher (*shumianyu*) Ausdrucksebenen.

Der zweite Tag begann mit zwei parallelen Sitzungen, auf denen einerseits die Praxis des Chinesischunterrichts in Europa durch Wu Shuhsung (Universität Tübingen), Li Suxian (Universität Heidelberg) und Yang Dan (Université Paris 8) anhand konkreter Kursmodelle und -beispiele, andererseits einzelne Methoden zum Erwerb mündlicher und schriftlicher Fähigkeiten in den Beiträgen von Françoise Philippe (Ecole La Jonchère), Gao Wanli (University of London), Chi Chihuan (Taipei Language Institute) und Cheng Ting Au (City University of Hong Kong) besprochen wurden.

Die Referenten der anschließenden Plenarsitzung befaßten sich mit den stil-, dialekt-, regional- oder geschlechtsbedingten Variationen der Retroflexanlaute *sh*, *zhi* und *chi* und ihrer differenzierten Vermittlung im Unterricht (He Dehua, Jing University, Taiwan), mit den Prinzipien eines systematisierten Unterrichts der gesprochenen Sprache (Lü Bisong, Beijing Yuyan Xueyuan) und mit Fehlerquellen und Interferenzen bei französischen Chinesischlernenden (E.N.S. de Fontenay-Saint-Cloud).

Der letzte Block des Samstagnachmittags war lexikalischen und syntaktischen Fragen gewidmet, wobei insbesondere die Rolle der Morpheme im Lernprozeß (Monique Hou, Université du Havre), Beispiele lexikalischer Unterschiede zwischen Taiwan, Hongkong und Festlandchina (Yao Rongsong, Guoli Taiwan Shifan Daxue), statistische und prinzipielle Feststellungen zum Schrifterwerb (Ye Deming, Guoli Taiwan Shifan Daxue) und chinesisch-französische syntaktische Vergleiche (Huang Liyi, Guoli Zhengzhi Daxue, Taiwan) zur Sprache kamen.

Angesichts der problemlosen Organisation und der allseitigen Zufriedenheit der so zahlreich erschienenen Teilnehmer blieb letztlich nur die Frage offen, weshalb nicht schon früher in Paris eine solche internationale Tagung abgehalten worden war. Jedenfalls verabschiedete man sich in der Erwartung einer nächsten derartigen Veranstaltung und verstärkter Impulse für die Vermittlung chinesischer Sprache und Kultur in Europa.

Peter Kupfer

Evaluation einer Prüfung: Ergebnisse einer Untersuchung zur HSK

Birgit Ziermann

1. Einleitung

Dieser Beitrag soll die wichtigsten Ergebnisse einer Untersuchung zur HSK (*Hanyu Shuiping Kaoshi*) zusammenfassen, die ich in umfangreicherer Form im Rahmen meiner Magisterarbeit vorlegte.¹ Die HSK ist die erste Standardprüfung des Chinesischen für Nichtmuttersprachler, die in der VR China erstellt wurde. Sie ist ferner in ein umfangreiches Curriculum für den Unterricht in Chinesisch als Fremdsprache (ChF) eingebunden, das entwickelt wurde, um diesen landesweit stärker zu vereinheitlichen. Die erste Durchführung der Prüfung in Europa - am 4. Juni 1994 in Hamburg - aktualisierte die Frage, ob die HSK für deutsche Chinesischler ein geeignetes Meßinstrument ihres Sprachniveaus ist und inwieweit das inhaltliche Curriculum der Prüfung auch als Vorgabe für eine stärkere Vereinheitlichung des Chinesischunterrichts in Deutschland brauchbar wäre. Diesen Fragen gehe ich zum einen auf theoretischer Ebene nach, zum anderen konnte durch eine Erhebung während der HSK in Hamburg die praktische Erfahrung der Prüflinge mit der HSK und dem ChF-Unterricht in Deutschland in die Bewertung der Prüfung miteinbezogen werden. Zunächst sollen also einige Ergebnisse der theoretischen Überlegungen dargestellt werden, um dann die aus den Erhebungsergebnissen abzuleitenden Schwächen der HSK in diesem theoretischen Rahmen zu erläutern.

2. Entwicklung der HSK

Für die Güte einer Standardsprachprüfung ist entscheidend, wie sie entwickelt wurde. Wichtig ist dabei, ob die klassischen Gütekriterien Objektivität, Reliabilität und Validität hinreichend erfüllt sind.²

In der Testtheorie spricht man von drei Aspekten der Objektivität: Durchführungs-, Auswertungs- und Interpretationsobjektivität. Bei der HSK ist die Durchführungsobjektivität garantiert, weil die äußeren Bedingungen für alle Prüflinge gleich sind und die Testanweisungen jeden Prüfling detailliert über den Verlauf

¹ Die Magisterarbeit wurde an der Universität Hamburg eingereicht und von Prof. Dr. Stumpfheldt und Dr. Bos betreut. Für deren Unterstützung möchte ich mich auf diesem Wege herzlich bedanken. Die nachfolgenden Ausführungen beziehen sich nur auf die HSK für die Grund- und Mittelstufen und berücksichtigen nicht die inzwischen ausgearbeitete HSK für die Oberstufe.

² Zu diesen Begriffen siehe Schelten (1980:98ff.)

der Prüfung instruieren. Da die Prüfung zum größten Teil aus Mehrfachwahl-Fragen besteht und maschinell ausgewertet wird, ist auch Auswertungsobjektivität gegeben. Das komplexe Einstufungssystem der HSK in verschiedene Ränge und Stufen gewährleistet außerdem Interpretationsobjektivität, denn das System legt fest, welchem sprachlichen Niveau ein bestimmtes Prüfungsergebnis entspricht.

Das derzeitige Benotungssystem der HSK wirft allerdings trotzdem einige Fragen auf. Das System umfaßt acht Ränge, verteilt auf drei Stufen: Unterstufe (2 Ränge), Grundstufe (3 Ränge) und Mittelstufe (3 Ränge). Diesen Prüfungsstufen sind drei Stufen der curricularen Vorgaben in Form eines Wortschatz- und Grammatikkataloges zugeordnet sowie die Anzahl der zu absolvierenden Unterrichtsstunden auf jeder Stufe und in welche Sprachkurse Prüflinge mit einem bestimmten Rang einzuordnen sind. So sollen die Vokabeln der Stufe A, die einfachsten, im 1. Semester des 1. Jahres gelernt werden, die der Stufe B im 2. Semester des 1. Jahres, und die Stufe C soll im 2. Jahr absolviert werden.³ Hat ein Prüfling bei der HSK aber das niedrigste Zertifikat der Grundstufe erworben, so sind die chinesischen Testkonstruktoren der Ansicht, daß er, obwohl in der Prüfung nur ein geringer Teil des Wortschatzes der Stufe B unter Beweis gestellt wurde, bereits am Unterricht des 2. Jahrgangs teilnehmen und die Vokabeln der Stufe C kompensieren kann. Ein ähnlicher Widerspruch besteht bei der Einteilung der Prüflinge der Mittelstufe in Jahrgangsstufen. Eine solche Einteilung kann m. E. nur zustande kommen, wenn von den Testkonstruktoren angenommen wird, daß die Ergebnisse eines Prüflings bei der HSK niedriger sind als sein tatsächliches Niveau.

Das zweite wichtige Gütekriterium für einen Test ist die Reliabilität. Sie bezieht sich auf die Zuverlässigkeit und Treffsicherheit des Meßinstruments. Um festzustellen, ob dieses Kriterium erfüllt ist, gibt es verschiedene Verfahren. Der Test oder ein Paralleltest kann hierzu einer Gruppe von Prüflingen zweimal vorgelegt werden, um zu prüfen, ob die Ergebnisse reliabel, d. h. zuverlässig, sind. Die Konsistenzanalyse, ein kompliziertes mathematisches Verfahren, ist eine zweite Möglichkeit zur Reliabilitätskontrolle. In Forschungsmaterialien zur HSK wurden Ergebnisse einer Reliabilitätskontrolle veröffentlicht. Die angegebenen Werte bewegen sich in dem Rahmen, der von der westlichen Testtheorie für einen guten Standardtest vorgegeben ist. In den chinesischen Forschungsmaterialien bleibt allerdings unerwähnt, nach welchem Verfahren der entsprechende Wert ermittelt wurde.

Validität (Gültigkeit) besitzt ein Test dann, wenn er tatsächlich mißt, was er zu messen vorgibt. Festgestellt wird dies meist, indem die gleichen Prüflinge an dem zu untersuchenden Test und einer anderen zuverlässigen und gültigen Sprachprüfung teilnehmen und beide Ergebnisse in Beziehung gesetzt werden. Durch dieses

³ Die Einteilung ist auf den ChaF-Unterricht in der VR China bezogen, kann also nicht mit dem Wortschatzerwerb an verschiedenen Institutionen der Bundesrepublik verglichen werden.

Verfahren wird die sog. kriterienbezogene Validität bewiesen, d. h. der Test wird einem Außenkriterium gegenübergestellt. Für die HSK wurde ein solcher Vergleich 1986 ansatzweise durchgeführt. Die HSK-Ergebnisse einiger Sprachkurse von Auslandsstudenten wurden den Semesterabschlußprüfungen dieser Klassen gegenübergestellt. Allerdings machten die Untersucher die Testergebnisse der beiden Prüfungen nicht statistisch vergleichbar und stellten keinen Korrelationskoeffizienten auf, was eigentlich für eine korrekte Validitätskontrolle nötig wäre. Die Validität der HSK kann also anhand dieser Daten zwar angenommen werden, ist aber nicht eindeutig bewiesen.

Ein weiterer Faktor, der entscheidend die Qualität einer Prüfung beeinflusst, sind die Kriterien, die bei der Erstellung der Prüfungsfragen angelegt werden. In den Forschungsmaterialien zur HSK legen die Autoren ihre Auswahlkriterien zu den Fragen der Prüfungsteile Leseverständnis und Hörverständnis dar. Diese Kriterien entsprechen im wesentlichen denen, die in westlichen Standardsprachprüfungen üblich sind.⁴ Außerdem ist unverzichtbar, daß die nach diesen Kriterien entworfenen Testfragen erprobt werden. Hierbei müssen vor allem Schwierigkeitsgrad und Trennschärfe, d. h. der Grad, zu dem eine Frage zwischen guten und schlechten Prüflingen differenziert, überprüft werden. Für diese Faktoren gibt es Richtlinien der Testtheorie, die besagen, wann eine Testfrage als zu schwierig oder noch brauchbar gelten kann. Gemessen an diesen Richtlinien enthält die HSK einige wenige zu schwierige Aufgaben sowie Aufgaben, bei denen weitere Verbesserung möglich wäre.⁵

Tatsächlich sind vereinzelt Beispielfragen zu finden, deren Beantwortung Schwierigkeiten aufwerfen dürfte:

Textausschnitt:

Berichten zufolge vergrößert sich die Wüstenfläche der gesamten Erde jährlich um 50.000 bis 70.000 Quadratkilometer, während sich die bewaldete Fläche des Erdballs seit dem 19. Jh. von 5,5 Milliarden Hektar auf die augenblicklichen 2,8 Milliarden Hektar reduzierte. Die täglich zunehmende Wasser- und Luftverschmutzung tut noch das ihre zur Verschlechterung des Lebensraumes der Menschheit hinzu.

⁴ Für die Kriterien, die bei der Erstellung der HSK-Prüfungsfragen angelegt werden, siehe Liu Yinglin (1989:142ff., 170ff.). Zum Vergleich können die Kriterien, die beim Test of English as a Foreign Language (TOEFL) gelten, herangezogen werden. Siehe hierzu z.B. Peirce (1992:665-691)

⁵ Die Testkonstruktoren wählen laut Liu Yinglin (1989:195ff.) für die HSK Testaufgaben aus, deren Schwierigkeitsgrad zwischen 0,15 und 0,85 liegt. Schelten (1980:135f.) gibt Richtlinien zwischen 0,2 und 0,8 an. Ein Schwierigkeitsindex von 0,15 besagt, daß 15% der Prüflinge die Aufgabe lösen konnten. Die Trennschärfe muß den Materialien zur HSK zufolge höher als 0,3 sein. Schelten bezeichnet Fragen mit einem Trennschärfekoeffizienten zwischen 0,3 und 0,39 als brauchbare Testaufgaben, bei denen weitere Verbesserung möglich sei.

Frage: Nach dem Lesen dieses Abschnittes fühlt man sich:

A: empört B: unruhig C: freudig erregt D: gleichgültig⁶

Als richtige Antwort auf diese Frage wird B angegeben. Dies ist jedoch aus keiner Information des Textes zu begründen. Denkbar wäre ebenso, daß ein Leser auf die Tatsache der immensen Umweltzerstörung empört reagiert. Bei dieser Fragestellung fließen m. E. kulturelle Konflikte mit ein. Sie kann nicht als angemessen gelten, da das Lesen eines solchen Testabschnitts zumindest für westliche Prüflinge nicht an ein festgelegtes Gefühl gekoppelt sein kann. Diese werden vermutlich davon ausgehen, daß sie als Individuen frei entscheiden können, wie sie auf eine solche Information reagieren.

Eine Aufgabenstellung wie die obige ist jedoch wahrscheinlich innerhalb der HSK ein Einzelfall. Die Anzahl der Fragen, die weitere Verbesserung nötig hätten, dürfte so gering sein, daß eine eventuelle falsche Beantwortung keinen entscheidenden Einfluß auf das Gesamt-Prüfungsergebnis nehmen wird. Um hierüber sichere Aussagen treffen zu können, wäre es jedoch nötig, Original-Prüfungsaufgaben in eine Untersuchung mit einzubeziehen.⁷

3. Prüfungsform im Vergleich zu anderen Sprachprüfungen

Die Qualität einer Sprachprüfung kann nicht ohne Vergleich mit anderen Sprachprüfungen beurteilt werden. Dieser Abschnitt faßt kurz die Ergebnisse einer Gegenüberstellung mit anderen gängigen Sprachprüfungen aus den USA und Deutschland zusammen. Für das Englische wird dabei der TOEFL - Test of English as a Foreign Language - herangezogen, außerdem Chinesischprüfungen aus den USA sowie für das Deutsche das Zertifikat DaF (Deutsch als Fremdsprache) und die Goethe-Diplome.⁸

Ein Vergleich der Aufgabenformen dieser Prüfungen zeigt, daß Mehrfachwahl-Aufgaben hauptsächlich bei den standardisierten Tests aus den USA üblich sind. Dagegen prüfen das Zertifikat DaF und die Goethe-Diplome auch oder ausschließlich durch schriftliche Beantwortung von Fragen, Formulierung eines Briefes, durch Diktat oder Umformungsaufgaben.

Der zeitliche Rahmen der HSK erzeugt für den Prüfling in zweierlei Hinsicht Streß: Zum einen könnte die gesamte Prüfungsdauer von ca. 2 Stunden 25 Minu-

⁶ Siehe Liu Yinglin (1989: 164)

⁷ Einige Beispiele aus den Forschungsmaterialien zur HSK sowie aus der Probeprüfung weisen Mängel auf. Ob diese Materialien jedoch gerade Fragen enthalten, die eventuell bei einem Protest aufgrund verschiedener Mängel ausgeschlossen oder revidiert wurden, ist nicht bekannt.

⁸ Bei den Chinesischprüfungen aus den USA handelt es sich um den Pre-Chinese Proficiency Test (Pre-CPT) sowie den Chinese Proficiency Test (CPT). Erster kann nach 1-2 Jahren, der CPT nach 2-3 Jahren Sprachunterricht auf dem College absolviert werden. Siehe hierzu Center of Applied Linguistics (1992) sowie Jiang Xixiang (1994: 14ff.)

ten ohne Pause bewirken, daß die Prüflinge in den letzten Prüfungsteilen nicht mehr voll konzentriert sind. Zum zweiten ist die Zeit pro Frage in den einzelnen Prüfungsteilen knapp bemessen. So müssen im Hörverständnis die Fragen z. B. in durchschnittlich 15-20 Sekunden durchgelesen und beantwortet werden. Dieser große Zeitdruck ist eigentlich für eine Niveauprüfung unangebracht, da hier allein durch die Fragestellung erreicht werden sollte, daß kaum ein Prüfling alle Aufgaben richtig lösen kann. Der Vergleich mit anderen Sprachprüfungen zeigt jedoch, daß tendenziell Prüfungen mit hohem Schwierigkeitsgrad, wie z. B. die Goethe-Diplome, auch zeitlich hohe Anforderungen an den Prüfling stellen. Beim Zertifikat DaF, das in der Schwierigkeit etwa der Grundstufe der HSK gleichkommt, entspricht die Gesamtprüfungszeit der der HSK, allerdings ist beim Zertifikat DaF eine kurze Pause eingeschoben. Auch die beiden Chinesischtests aus den USA zeigen, daß dem Prüfling je nach Niveau eine längere Testzeit zur Verfügung steht: Der Pre-CPT, der etwa den niedrigsten Rängen der Grundstufe HSK entspricht, dauert 1 Stunde 30 Minuten; der CPT, dessen Schwierigkeitsgrad knapp der Mittelstufe HSK entspricht, dauert 2 Stunden. Das Problem bei der HSK ist aber, daß Prüflinge verschiedener Niveaus, nämlich der Grund- und Mittelstufe, in einer gemeinsamen Prüfung getestet werden. Der zeitliche Rahmen der Prüfung scheint aber eher an den Studenten mit höheren Fertigkeiten bemessen und könnte anderen Lernern zu schwierig erscheinen und daher demotivierend wirken.

Der Vergleich mit dem TOEFL zeigt, daß die HSK in der Aufgabenanzahl und dem zeitlichen Rahmen der einzelnen Prüfungsteile sowie der gesamten Prüfung nahezu identisch ist mit dem ehemals gebrauchten fünfteiligen TOEFL. Diese TOEFL-Form wurde Ende der 80er Jahre revidiert und in eine dreiteilige Form umgewandelt, da sich erwies, daß der Prüfungsteil "Writing Ability", der den Mehrfachwahl-Fragen des komplexen Lückentests der HSK entspricht, keine Informationen über das Sprachniveau des Prüflings liefert, die nicht schon durch andere Prüfungsteile gewonnen würden.⁹ Durch diese Revision wurde die Gesamtprüfungszeit des TOEFL auf 1 Stunde 50 Minuten verkürzt. Auch bei der HSK ist durch die Zusammenhänge zwischen den einzelnen Prüfungsteilen anzuzweifeln, ob dieser Teil des komplexen Lückentests für die Gesamteinschätzung der Prüflinge nötig ist. Ein Argument für die Beibehaltung dieses Prüfungsteiles könnte sein, daß eine höhere Aufgabenanzahl grundsätzlich eventuelle Mängel in der Aufgabenkonstruktion ausgleichen kann.

Der Vergleich mit dem zeitlichen Rahmen anderer Prüfungen zeigte auf jeden Fall, daß die Gesamtlänge der HSK die der herangezogenen Prüfungen aus den USA übersteigt. In der Beantwortungszeit für die einzelnen Prüfungsfragen entsprechen die Vorbilder aus den USA und die HSK sich aber weitgehend, der TOEFL setzt sogar beim Hörverständnis noch niedrigere Beantwortungszeiten für die einzelnen Fragen an.

⁹ Siehe hierzu Pike (1979) sowie Hale u. a. (1988).

4. Prüfungsinhalte und ihre Festlegung

Ursprünglich war bei der Entwicklung der HSK vorgesehen, daß sich die Prüfungsinhalte aus vier verschiedenen Lernzielkatalogen zusammensetzen: ein Katalog für Wortschatz, einer für Grammatik, für Sprachfunktionen und -intentionen und für kulturelles Hintergrundwissen. In den theoretischen Materialien zur HSK wird zwar erwähnt, daß die letzten beiden Kataloge nötig seien, bis zum gegenwärtigen Zeitpunkt liegen sie aber nicht vor.¹⁰ Der Lernzielkatalog für Wortschatz wurde auf der Grundlage von dreierlei Materialsorten erstellt: Als Basis benutzten die Autoren ein 1985 erschienenes Frequenzwörterbuch der chinesischen Sprache. Sie betonen in diesem Zusammenhang die Wissenschaftlichkeit, Zuverlässigkeit und Objektivität, die ihrer Meinung nach ein solches sprachstatistisches Werk habe.¹¹ Zur Ergänzung wurden außerdem verschiedene Frequenz-Wortlisten herangezogen, zwei davon sind direkt auf den ChaF-Unterricht bezogen, eine weitere enthält häufige Wörter der Informationsverarbeitung im Ingenieurwesen. Zum anderen wurden 16 verschiedene Lehrwerke für den ChaF-Unterricht verarbeitet.

In der hiesigen Sprachlehrforschung wird seit den 70er Jahren grundsätzlich die Entwicklung von Sprachminima auf der Basis von sprachstatistischen Untersuchungen kritisiert.¹² Am Anfang der Entwicklung eines Grundwortschatzes sollte statt dessen eine explizite Definition der Lernziele und des Adressatenkreises stehen. Das Europaratsprojekt und das Zertifikatssystem der Volkshochschulen waren zwei Konzepte, die ungefähr zeitgleich in den 70er Jahren diesen neuen Denkansatz umsetzen wollten. Beide zielten darauf ab, die Forderungen nach genauer Lernzieldefinition in Form von Situationskatalogen zu verwirklichen. Bei der Entwicklung des Wortschatzes für die HSK wurde diese Stufe übersprungen. Damit wurden die ausgewählten Wörter nicht systematisch auf einen Kommunikationsrahmen bezogen. In den theoretischen Materialien zur HSK ist zwar erwähnt, welche Bereiche des Alltagslebens die Prüfungsfragen abdecken sollten. Diese Ausführungen sind allerdings, verglichen mit dem Kommunikationsrahmen für das Zertifikat DaF, vage und dürftig formuliert. Der Adressatenkreis wurde bei der HSK zwar spezifiziert, über die Kommunikationsbedürfnisse der Zielgruppe wurden jedoch keine detaillierten Aussagen gemacht. Der relativ weit gefaßten Zielgruppe, die sowohl Studenten der Naturwissenschaften als auch der Geisteswissenschaften, Angestellte von Firmen und Angehörige der sogenannten Minderheiten umfaßt, müßten z. T. unterschiedliche Kommunikationsbedürfnisse unterstellt werden. Einige der im folgenden dargestellten Mängel der Wortliste hätten evtl. durch Abstimmen der Wörter auf einen Kommunikationsrahmen verhindert werden können. Außerdem hätte der fehlende Katalog für kulturelles

Hintergrundwissen auf die Problematik einiger Prüfungsfragen aufmerksam machen können.

Der Lernzielkatalog für den Wortschatz der HSK ist alphabetisch angeordnet. Bei den Einträgen wird angegeben, welcher Stufe und Wortart die Vokabel zugehört. Nicht einheitlich ist die Handhabung, wenn ein Wort unterschiedlichen Wortarten zugeordnet werden kann. Teilweise führen die Herausgeber solche Wörter mehrmals in unterschiedlichen Funktionen auf, zum Teil auch unterschiedlichen Stufen zugeordnet. Einer großen Anzahl Wörter wiesen die Autoren hingegen mehrere Wortarten und damit oft auch Bedeutungen zu.

Eine stichprobenartige Untersuchung anhand weniger Beispiele zeigte, daß die Herausgeber der Wortliste tatsächlich bemüht waren, die Vokabeln einer Stufe zu einem funktionalen Wortschatz zusammenzustellen. In 22 Fällen der Stufe A überprüfte ich, ob für die angeführten Wörter die entsprechenden Verben, Objekte oder Zählwörter ebenfalls auf dieser Stufe auftauchen. Nur in drei Fällen waren die dazugehörigen Zählwörter nicht auf derselben Stufe erwähnt. Bedauerlich ist generell, daß nicht angegeben wird, welche Substantive jeweils mit den Zählwörtern verbunden werden können.

Außerdem untersuchte ich stichprobenartig den Wortschatz nach inhaltlichen Gesichtspunkten, indem ich 550 der 2202 Einträge auf der Stufe C auf ihr Vorkommen in einem Wörterbuch für Naturwissenschaft und Technologie überprüfte. Hierbei stellte sich heraus, daß 125 dieser 550 Einträge, also 22.73%, in diesem Wörterbuch aufgeführt waren. Darunter sind auch Wörter wie *baohu* - "beschützen", *bili* - "Proportion, Maßstab" oder *bianji* - "Redakteur", also relativ gemeinsprachliche Vokabeln. Jedoch zeigen Wörter wie *bandaoti* - "Halbleiter", *danbaizhi* - "Eiweiß", *dianzi* - "Elektron", *fenzi* - "Molekül", *jiepou* - "sezieren" bei den Vokabeln der Stufe C eine Tendenz zu naturwissenschaftlichen Inhalten. Dagegen sind für das Alltagsleben so wichtige Wörter wie *dasao* - "aufräumen, reinigen", *fangan* - "Restaurant", *koushi* - "mündliche Prüfung" oder *liuxue* - "im Ausland studieren" ebenfalls erst auf der Stufe C erwähnt. Die Ausdrücke naturwissenschaftlichen Inhalts auf der Stufe C aufzunehmen ist vor allem widersprüchlich, weil die Wörter dieser Stufe den Zertifikaten der Mittelstufe zugeordnet sind, also den Zertifikaten, die für eine Zulassung zum geisteswissenschaftlichen Studium nötig sind. Konsequenter wäre, sie auf Stufe B zu erwähnen. Die Wörter dieser Stufe werden den Zertifikaten der Grundstufe zugeordnet, welche für die Aufnahme eines naturwissenschaftlichen Studiums benötigt werden.

Der Lernzielkatalog für Grammatik ist eine nach drei Stufen getrennte Auflistung grammatischer Muster. Die Herausgeber stellen auf der Stufe A systematisch die Grundzüge der chinesischen Grammatik dar. Die grammatischen Phänomene werden zunächst kurz erklärt, und häufig werden Beispiele aufgeführt. Die Autoren bemühten sich außerdem, die Beispielsätze tatsächlich mit dem entsprechenden Wortschatz der Stufe A zu bilden. Auf Stufe B und C werden keine Beispiele mehr für die einzelnen grammatischen Phänomene angeführt.

¹⁰ Siehe Liu Yinglin (1989:50, 79, 188).

¹¹ Siehe Liu Yinglin (1989:66f.).

¹² Siehe z. B. Raasch (1972) und Kühn (1979).

Bedauerlich ist, daß die Erklärungen und Unterteilungen für die einzelnen grammatischen Phänomene der Stufe A von Chinesischlernern dieser Stufe nicht verstanden werden können. Die verwendeten Begriffe sind zum größten Teil erst bei den Vokabeln der Stufe B und C, viele grammatische Begriffe auch gar nicht in der Wortliste zu finden. Die Unterteilungen und Erklärungen sind also mehr für Lehrbuchautoren, Testverfasser und Lehrkräfte gedacht als für den Lerner oder Selbstlerner. Die Herausgeber der Materialien zur HSK weisen aber ausdrücklich darauf hin, daß die Lernzielkataloge für Wortschatz und Grammatik ebenso als Nachschlagewerke für den Lerner gedacht seien. M. E. wird bisher diesem Anspruch weder der Lernzielkatalog für den Wortschatz, noch der für die Grammatik gerecht.

Der Lernzielkatalog für den Wortschatz kann nur mit Hilfe eines Wörterbuches als Nachschlagewerk benutzt werden. Selbst dann bleibt noch fraglich, welche Bedeutung eines Wortes gemeint ist, da dies nicht durch Beispiele verdeutlicht wird.

Der Lernzielkatalog für die Grammatik wäre für den Lerner nur dann eine sinnvolle Ergänzung zum Katalog für Wortschatz, wenn alle grammatischen Phänomene einer Stufe angeführt wären. Dann könnte der Lerner diese Ausdrücke nicht alphabetisch, wie in der Wortliste, sondern nach grammatischer Funktion unterteilt lernen oder wiederholen. Tatsächlich werden aber auf jeder Stufe nur einige Beispiele für die angeführten Punkte gegeben. So sind z. B. in der Wortliste 40 Konjunktionen für Stufe B aufgeführt, im Lernzielkatalog für Grammatik dagegen unter der Wortart "Konjunktion" für diese Stufe nur acht Beispiele enthalten. Der Lernzielkatalog für Grammatik wurde allerdings nur einmal veröffentlicht, 1988. In dieser Veröffentlichung wurde nicht erwähnt, daß es sich hier um die prüfungsrelevanten Inhalte für die HSK handelt. Dieses Buch sollte lediglich die Stufenstandards für den ChaF-Unterricht festlegen. Das läßt den Schluß zu, daß die Herausgeber des Katalogs sich eventuell selbst bereits der Mängel bewußt sind und diesen noch nicht als ausgereift ansehen.

Der Lernzielkatalog für den Wortschatz stellt aber in sich eine beachtliche Leistung dar. Bislang existieren keine ähnlichen curricularen Vorgaben, die einen Anhaltspunkt für die inhaltliche Gestaltung gerade des ChaF-Unterrichts für Fortgeschrittene geben könnten. Natürlich wäre notwendig, diese Sprachminima einer genauen Untersuchung zu unterziehen und gegebenenfalls zu revidieren.

Die Prüfungsfragen der HSK basieren auf den beschriebenen Sprachminima. Dabei können Wortschatz wie Grammatik in geringem Maße über die Lernzielkataloge hinausgehen. Wortschatz- und Grammatikteile, die nicht in den Lernzielkatalogen enthalten sind, dürfen jedoch 5% der Aufgabenstämme nicht überschreiten. Im Prüfungsteil Grammatik dürfen keine Wörter verwendet werden, die nicht im Lernzielkatalog auftauchen. Solche Wörter dürfen nur in relativ schwierigen Fragen der anderen Prüfungsteile vorkommen.¹³ Diese Regelung ist m. E. ein un-

¹³ Siehe Liu Yinglin (1989:137).

nötiger Unsicherheitsfaktor für den Prüfling. Für das Zertifikat DaF gilt dagegen, daß in den Listen nicht angeführte Wörter nur vorkommen dürfen, wenn sie entweder durch den Kontext zu erschließen, durch eine Fußnote erklärt oder für das Verständnis und die Beantwortung der Prüfungsfrage nicht relevant sind.

5. Erhebung im Rahmen der HSK vom Juni 1994

5.1. Datenerhebungsverfahren

Die erste Durchführung der HSK in Deutschland am 4. Juni 1994 in Hamburg bot die Gelegenheit, sowohl die Meinung der Prüflinge als auch ihr tatsächliches Abschneiden in die Bewertung der Prüfung miteinzubeziehen. Die HSK hat den Anspruch, das sprachliche Niveau der Prüflinge objektiv einzuschätzen. Durch den ausgegebenen Fragebogen konnten Daten zum Umfang der sprachlichen Ausbildung der Prüfungsteilnehmer erhoben werden, um so zu überprüfen, ob die Prüfungsergebnisse dem tatsächlichen Niveau der Teilnehmer entsprechen. Das tatsächliche sprachliche Können wäre allerdings nur durch eine andere Prüfung oder Lehrerurteile über die Prüflinge verlässlich zu messen gewesen. Die erhobenen Daten zum Umfang der sprachlichen Ausbildung boten jedoch einen gewissen Anhaltspunkt zum Wissensstand der Teilnehmer. Ein schlechtes Abschneiden bei der HSK trotz intensiver Sprachausbildung könnte auf Mängel der Prüfung, aber ebenso auf Mißstände bei der Sprachausbildung hinweisen. Daher enthielt die Erhebung Fragen zur Meinung der Teilnehmer über die HSK und über den ChaF-Unterricht. So sollte klarer werden, ob die Schwierigkeiten mit der HSK auf die Sprachausbildung oder auf die Prüfung selbst zurückzuführen sind.

Der ausgegebene Fragebogen bestand aus 40 Fragen, die nach inhaltlichen Gesichtspunkten zu drei Fragekomplexen zusammengestellt wurden. Teil 1 erhob Daten zum modernen Sprachstudium in Deutschland, Teil 2 zu Sprachstudien in China, der 3. Teil 15 Fragen zur HSK sowie 5 Fragen zum ChaF-Unterricht in Deutschland. Bezüglich der HSK wurde nach Teilnahmemotiven, Problemen mit den einzelnen Prüfungsteilen, Meinungen zur HSK als objektives Meßinstrument verschiedener Fertigkeiten, Streß und Gedächtnisleistung gefragt. Außerdem sollten die Probanden beurteilen, inwieweit ihrer Meinung nach die bei der HSK geforderte Reaktionsgeschwindigkeit, die Grammatikkenntnisse, der Wortschatz, das Hörverständnis, kursorisches Lesen und Abstraktionsvermögen im Chinesischunterricht in Deutschland vermittelt würden. Weiterhin wurde nach ihrer Einstellung zu Prüfungen innerhalb des Sinologie-Hauptstudiums und zu stärkerer Strukturierung des Studienganges gefragt.

Grundsätzlich wurden sowohl offene Fragenformen als auch Fragen mit vorgegebenen Antwortmöglichkeiten, sogenannte geschlossene Fragen, verwendet. Bei den geschlossenen Fragen sollten die Teilnehmer auf einer fünfstufigen Skala den Intensitätsgrad der vorangegangenen Behauptung, Frage oder Problemstel-

lung beurteilen. Die meisten Fragen sollten mögliche Probleme der Prüfungsteilnehmer herausfinden, waren also hypothesengewinnender Art. Einige Fragestellungen griffen jedoch auf die in den vorangegangenen Ausführungen bereits herausgearbeiteten Mängel der Prüfung zurück und hatten insofern hypothesenprüfenden Charakter.

5.2. Untersuchungsgruppe und Rücklauf

An der Prüfung in Hamburg nahmen 93 Personen teil. Hiervon kamen 54 Teilnehmer aus Hamburg (58%) und 39 von auswärts (42%). Die große Zahl der Hamburger Studenten ist zum einen durch den Prüfungsort zu erklären. Zum anderen war von den Dozenten des Sinologischen Seminars für die Prüfung geworben sowie ein Kurs zur intensiven Vorbereitung auf die HSK abgehalten worden.

An der Erhebung nahmen insgesamt 62 Personen teil. Von diesen Probanden waren 37 aus Hamburg, 25 von auswärts. Wenn man den Rücklauf dieser zwei Gruppen getrennt betrachtet, so haben 68,5% der Hamburger Teilnehmer den Fragebogen beantwortet und 64,1% der von auswärts angereisten. Auf die Gesamtgruppe der Prüflinge bezogen beträgt der Rücklauf 66,6%.

Über 75% der an der Erhebung Beteiligten studieren Sinologie im Hauptfach. Die Antworten zum letzten Teil des Fragebogens, Fragen zu Prüfungen innerhalb der Sinologie, können demnach als repräsentativ für Sinologie-Studenten gelten. Die 15 Teilnehmer, die aus anderen sinologischen Seminaren als dem Hamburger kamen, sind fünf verschiedenen Seminaren zugehörig, nämlich jenen an den Universitäten Bonn, Tübingen, Würzburg und Köln. Das heißt, daß weder ihr Abschneiden in der Prüfung noch ihre Meinungen repräsentativ für die unterschiedlichen sinologischen Seminare ist. Die restlichen zehn Personen, die Übersetzer- und Dolmetscherstudiengänge absolvieren, studierten an vier verschiedenen Institutionen in den Städten Bonn, Berlin, Gernersheim und München.

6 der 62 Probanden befanden sich zum Zeitpunkt der Erhebung noch im Grundstudium, 18 Teilnehmer im 5. bis 8. Semester, 31 Personen im 9. bis 12. Fachsemester und 5 Probanden im 13. bis 17. Semester. Fast alle Teilnehmer waren bereits in China gewesen, lediglich sieben Personen gaben an, noch keinen Chinaaufenthalt hinter sich zu haben. 93% derjenigen, die bereits in China waren, hatten Sprachkurse besucht. Die Länge des Aufenthaltes sowie die Dauer der Sprachkurse lagen meistens bei zwei Semestern.

5.3. Erhebungsergebnisse

Sämtliche Ergebnisse der Erhebung darzustellen ginge über den Rahmen dieses Artikels hinaus. Neben einfachen Häufigkeitsverteilungen wurden detaillierte Analysen durchgeführt, die Zusammenhänge zwischen der Beantwortung verschiedener Fragen und den Prüfungsergebnissen aufzeigen. Erhebungs- oder

Prüfungsergebnisse wurden nach verschiedenen Lernergruppen, wie z. B. Sinologiestudenten, Studierende von Dolmetscher- oder Übersetzerstudiengängen, Studenten des Grund- und Hauptstudiums etc. differenziert und auf Mittelwertunterschiede überprüft. Außerdem wurden Korrelationskoeffizienten zwischen einzelnen Untersuchungsergebnissen aufgestellt und auf ihre Signifikanz kontrolliert. Hier sollen lediglich die wichtigsten Befunde unter Verzicht auf genaue Angabe der entsprechenden statistischen Ergebnisse zusammengefaßt werden.

Aus den Erhebungsergebnissen können Kritikpunkte und positive Werturteile zur HSK abgeleitet werden. So ist aus der hohen Korrelation zwischen den Prüfungsergebnissen und der gesamten Semesterwochenstunden-Anzahl Sprachunterricht der Probanden zu schließen, daß die HSK offensichtlich tatsächlich die Niveauunterschiede zwischen den einzelnen Prüflingen mißt. Dies ist auch daraus abzulesen, daß die Wortschatzprobleme bei den einzelnen Prüfungsteilen für die Probanden weniger wurden, je mehr Sprachunterricht sie bereits absolviert hatten.

Die von den Testkonstrukteuren angegebenen Mindestunterrichtsstunden für die Grund- und Mittelstufe der Prüfung sind allerdings für die Studenten chinabezogener Studiengänge in Deutschland nicht zutreffend. Hier müßte vielmehr von einer Mindestunterrichtsstundenzahl von 500 Semesterwochenstunden (SWS) Sprachunterricht für ein Zertifikat der Grundstufe ausgegangen werden. Für deutsche Sinologiestudenten ist ferner zu sagen, daß sie innerhalb der Prüfung begrenzt auf Wortschatz treffen, der nicht Teil ihrer Sprachausbildung ist. Dieser Umstand wurde jedoch vom Großteil der Prüfungsteilnehmer nicht als großes Problem eingeordnet.

Der Prüfungsteil Hörverständnis wurde von den Teilnehmern häufig kritisiert. Vor allem Probanden mit einer höheren SWS-Anzahl beurteilten die HSK als weniger gutes Meßinstrument im Bereich Hörverständnis. Außerdem wurde in diesem Prüfungsteil die Zeit häufig als nicht ausreichend empfunden. Dieses Problem verringerte sich nicht signifikant für Probanden mit höherer SWS-Anzahl. Je eher die Probanden ein zeitliches Problem bei diesem Prüfungsabschnitt hatten, desto eher gaben sie an, bei der gesamten HSK unter Streß gestanden zu haben. Je höher dieser Streß war, desto weniger wurde die HSK als objektives Maß für ihre tatsächlichen Fertigkeiten im Bereich Hörverständnis und für die sprachlichen Fertigkeiten in Alltagssituationen angesehen. Die erreichte Punktezahl in diesem Prüfungsteil verbesserte sich im Vergleich zu anderen Prüfungsteilen weniger stark bei höherer SWS-Anzahl, längerem Chinaaufenthalt oder mehr Sprachunterricht in China. Außerdem waren immerhin neun Personen der Meinung, daß unangemessene Fragestellungen beim Hörverständnis ein sehr großes Problem darstellen. Tendenziell waren Personen mit höherer SWS-Anzahl eher der Meinung, daß die Fragestellungen dieses Prüfungsteils unangemessen sind. All diese geschilderten Ergebnisse deuten darauf hin, daß die Prüfungssituation im Hörverständnis nicht der tatsächlichen Alltagssituation in China entspricht und daher von den Studenten als unangemessen empfunden wird.

Der Vergleich mit den amerikanischen Sprachprüfungen ließ, wie in Abschnitt 3 dieser Ausführungen dargestellt, allerdings erkennen, daß hier ungefähr gleiche oder sogar straffere Zeitbegrenzungen für die Beantwortung der einzelnen Fragen bestehen als bei der HSK. Das bedeutet natürlich noch nicht, daß diese Prüfungen die wirklichen Gegebenheiten im Alltagsleben widerspiegeln. Tatsächlich ist auch innerhalb der Sprachlehrforschung umstritten, inwieweit gerade längere Hörpassagen geeignet sind, die Fähigkeiten von Studenten zu prüfen, Vorlesungen in der Fremdsprache zu folgen. Trotzdem zeigt der zeitliche Rahmen der amerikanischen Prüfungen, daß in der Praxis diese Art der Zeitbeschränkung bei standardisierten Sprachprüfungen durchaus üblich ist.

Der Prüfungsteil Grammatik wurde von den Probanden generell nicht als problematisch angesehen. Die Probleme mit diesem Prüfungsteil wurden auf der Skala nicht sehr hoch eingestuft, ferner war ein Großteil der Probanden der Meinung, daß die HSK ein "mittelmäßiges" oder "ziemlich gutes" Meßinstrument für ihre tatsächlichen Grammatikkenntnisse ist. Daß die meisten Teilnehmer hier nur "mittelmäßig" als Werturteil abgaben, ist wahrscheinlich im Zusammenhang mit der freien Kritik an diesem Prüfungsteil zu sehen. Hier kritisierten einzelne Probanden, daß selten gebrauchte grammatische Ausdrücke abgefragt würden.

Beim Prüfungsteil Leseverständnis empfanden die Erhebungsteilnehmer vor allem die knapp bemessene Zeit und die nachlassende Konzentration als problematisch. Je eher sie die Zeit innerhalb des Prüfungsteils als zu kurz empfanden, desto weniger ordneten die Teilnehmer die HSK als objektives Meßinstrument für ihre sprachlichen Fertigkeiten in Alltagssituationen ein. Hier wird also, ähnlich wie beim Prüfungsteil Hörverständnis, der starke Zeitdruck als unangemessen eingeschätzt. Die Kritik am Leseverständnis ist aber nicht so stark wie am ersten Prüfungsteil. Die Teilnehmer ordneten nämlich die HSK im Bereich Leseverständnis als "ziemlich gutes" oder "sehr gutes" Meßinstrument ein. Außerdem ist aus den Prüfungsergebnissen zu erkennen, daß die Gruppe der Erhebungsteilnehmer hier am besten im Vergleich mit den anderen Prüfungsteilen abschnitt. Die Punktezahl im Leseverständnis ist ferner das einzige Prüfungsergebnis, das sich mit höherer Fachsemesterzahl verbesserte. Diese Ergebnisse deuten darauf hin, daß der Prüfungsteil für die Probanden generell nicht sehr problematisch war. Ausnahme ist hierbei die nachlassende Konzentration aufgrund der langen Gesamtprüfungszeit. Außerdem tauchten laut den Urteilen der Probanden in diesem Prüfungsteil einige unangemessene Fragestellungen auf. Dies konnte bereits aus den theoretischen Ausführungen zur Entwicklung der Prüfung vermutet werden.

Im komplexen Lückentest schnitten die Teilnehmer insgesamt schlechter ab als in allen anderen Prüfungsteilen. Als vorwiegende Probleme wurden der Wortschatz, die knapp bemessene Zeit und die nachlassende Konzentration angegeben. Daß der Wortschatz hier als problematisch eingestuft wurde, zeigt, daß die Fragestellungen, ähnlich wie beim Prüfungsteil Leseverständnis, auf Wortschatzkenntnisse ausgerichtet sind.

Bereits die Ausführungen in Abschnitt 3 stellten fest, daß der ehemals gebrauchte fünfteilige TOEFL um den Prüfungsteil, der dem Mehrfachwahl-Fragen-Teil des komplexen Lückentests der HSK entspricht, gekürzt wurde, da dieser Teil offensichtlich keine Informationen zur Einschätzung der Prüflinge hinzufügt, die nicht auch schon durch den Prüfungsteil Leseverständnis gewonnen werden könnten. Auch bei den Erhebungsergebnissen aus Hamburg zeigte sich, daß bei den Korrelationen zwischen den einzelnen Prüfungsteilen die höchste die zwischen den Prüfungsteilen Leseverständnis und komplexer Lückentest war. Das läßt erkennen, daß diese beiden Prüfungsteile tatsächlich ähnliche Fertigkeiten messen.

Gerade da die Teilnehmer der Erhebung im Leseverständnis, verglichen mit anderen Prüfungsteilen, die besten Ergebnisse vorwiesen, ist es um so erstaunlicher, daß sie im komplexen Lückentest am schlechtesten abschnitten. Das ist durch die nachlassende Konzentration innerhalb dieses Prüfungsteiles, durch eventuell mehrere Möglichkeiten beim Einsetzen der Schriftzeichen (von denen jedoch nur eine als richtig anerkannt wird) oder durch mangelnde Übung der Teilnehmer auf dem Gebiet der Schreibfertigkeit zu erklären. Die Teilnehmer der HSK in den USA schnitten ebenfalls in diesen Prüfungsteil am schlechtesten ab.¹⁴ Das läßt vermuten, daß die Prüfungsergebnisse der Erhebungsteilnehmer in Hamburg beim komplexen Lückentest weniger auf ihre eigenen Fertigkeiten sondern vielmehr auf die Gegebenheiten des Prüfungsteils zurückzuführen sind.

Aus den praktischen Erfahrungen der Prüflinge in Hamburg kann geschlossen werden, daß der zeitliche Rahmen der HSK eine unnötige Belastung für die Prüflinge darstellt. Eine Möglichkeit, diese Überlastung zu reduzieren, wäre, nach dem Vorbild der TOEFL-Revision den Mehrfachwahlfragen-Teil des komplexen Lückentests zu streichen. Die entsprechenden Korrelationen in den Forschungsmaterialien zur HSK dagegen machen klar, daß der Prüfungsteil des selbständigen Einsetzens von chinesischen Schriftzeichen Informationen über die Prüflinge bereitstellt, die kein anderer Prüfungsteil bietet. Dieser Prüfungsteil sollte also auf jeden Fall beibehalten werden. Allerdings wäre hier denkbar, mehrere Möglichkeiten, die grammatisch korrekt sind, als richtig gelten zu lassen.

Eine Begründung für die Beibehaltung des Mehrfachwahlfragen-Teils des komplexen Lückentests könnte sein, daß durch Testverlängerung die HSK reliabler gemacht werden soll. Da die Stellen zur Entwicklung der HSK sicherlich über einen kleineren personellen und finanziellen Rahmen verfügen, als jene zur Entwicklung des TOEFL, könnten die Voraussetzungen für die Entwicklung weniger Fragen mit höherer Trennschärfe noch nicht gegeben sein. Eine weitere Möglichkeit, um den Konzentrationsmangel der Prüflinge im letzten Prüfungsteil abzuschwächen, wäre, nach dem Vorbild des Zertifikats DaF eine kurze Pause einzufügen.

¹⁴ Zur Auswertung der HSK in den USA siehe Pan Zhaoming/Shi Xudong (1993)

Aus den vorliegenden Erhebungsergebnissen können außerdem einige Umstände des ChaF-Unterrichts in Deutschland abgelesen werden.

Das Sprachstudium an den unterschiedlichen Institutionen ist uneinheitlich sowohl im Lehrmaterial als auch in der Sprachunterrichtsstundenzahl im Grundstudium. Einen Großteil ihrer gesamten SWS Sprachunterricht absolvierten die Erhebungsteilnehmer im chinesischsprachigen Ausland. Innerhalb des Hauptstudiums in Deutschland werden nicht mehr alle vier Sprachfertigkeiten gleichmäßig weiterentwickelt. Der Schwerpunkt liegt eindeutig bei der Sprachfertigkeit Lesen. 84% der Erhebungsteilnehmer besuchten 1-35 Textlektüreveranstaltungen à 2 Semesterwochenstunden innerhalb ihres bisherigen Hauptstudiums. Dagegen besuchten nur 50% einen Konversationskurs, 29% eine Übung zum Hörverständnis und 26% einen Schreibkurs innerhalb des Hauptstudiums. Das heißt, daß die Fertigkeiten des Sprechens, Schreibens und Hörens außeruniversitär weiterentwickelt werden müssen. Auch kursorisches Lesen wird selten systematisch vermittelt.

Dementsprechend sahen die Urteile der Studenten zum ChaF-Unterricht in Deutschland aus. Die Vermittlung von Grammatik- und Wortschatzkenntnissen wurde positiv beurteilt. Sinologen beurteilten die Vermittlung von Grammatik besser als angehende Übersetzer und Dolmetscher. Die Fertigkeiten Hörverständnis, kursorisches Lesen, Abstraktionsvermögen und Reaktionsschnelligkeit werden nach Ansicht der Probanden jedoch eher wenig vermittelt.

Sinologen des Grundstudiums beurteilten die Vermittlung der Reaktions-schnelligkeit positiver als Sinologen im Hauptstudiums. Dies deutet bereits an, daß im Hauptstudium häufig keine Semesterprüfungen mehr stattfinden. Die angegebenen Teilnahmemotive für die HSK sowie die Meinung der Probanden zu Prüfungen innerhalb der Sinologie zeigen, daß sie Prüfungen generell wünschen. Die wichtigsten Teilnahmemotive waren ein Bedürfnis nach Einschätzung der eigenen Kenntnisse, die Materialien zur HSK zu nutzen, um die eigenen Fertigkeiten zu verbessern oder zu strukturieren und das Zertifikat. Diese Motive lassen bereits darauf schließen, daß die Studenten genau wissen wollen, was zu lernen ist, und daß sie das Gelernte anhand von Prüfungen überprüfen wollen, um ihren Kenntnisstand einschätzen zu können.

95% der Teilnehmer gaben an, daß sie nochmals an der HSK teilnehmen oder Kommilitonen eine Teilnahme empfehlen würden. Auch hier wurde als Begründung hauptsächlich das Bedürfnis nach Einschätzung angegeben. Direkt nach einer sprachlichen Abschlußprüfung innerhalb der Sinologie sowie nach häufigeren studienbegleitenden Prüfungen und einer vermehrten Strukturierung des Hauptstudiums befragt, äußerte sich die Mehrzahl der Probanden positiv. Stärkere Strukturierung und damit genaue Lernvorgaben wurden besonders begrüßt. Die HSK als Abschlußprüfung einzusetzen lehnten dagegen die meisten Teilnehmer ab.

Abgesehen von den erwähnten Mängeln bietet der Lernzielkatalog für den Wortschatz dem fortgeschrittenen Lerner eine gute Möglichkeit, seine vorhande-

nen Sprachkenntnisse zu erweitern, systematisch zu wiederholen und zu strukturieren. Da die Erhebungsergebnisse ein starkes Bedürfnis der Teilnehmer nach solchen konkreten Lernvorgaben zeigten, könnte dieser Katalog also Studenten höherer Semester empfohlen werden, wenn sie Bedürfnis nach strukturierten Vorgaben für ihre sprachlichen Kenntnisse äußern.

Literaturverzeichnis

- Center for Applied Linguistics (CAL). 1992. *Chinese Proficiency Test and Preliminary Chinese Proficiency Test - Combined Examinee Handbook*. Washington
- Deutscher Volkshochschul-Verband e.V./Goethe-Institut (Hrsg.). 1977. *Das Zertifikat Deutsch als Fremdsprache*. Bonn
- Educational Testing Service (Ed.). 1992. *TOEFL Test and Score Manual 1992-93 Edition*. Princeton, New Jersey
- Goethe-Institut - Geschäftsstelle Deutsches Sprachdiplom (Hrsg.). 1994. *Kleines Deutsches Sprachdiplom, Großes Deutsches Sprachdiplom - Informationen für Lehrer und Prüfer*. München
- Guojia Hanyu Shuiping Kaoshi Weiyuanhui Bangongshi (Hrsg.). 1993. *Hanyu Shuiping Kaoshi (HSK) kaoshi dagang* (Leitfaden für die Prüfung zum Nachweis chinesischer Sprachkenntnisse). Beijing
- Hale, Gordon/Stansfield, Charles/Rock, Donald/Hicks, Marilyn/Butler, Frances/Oller, John Jr. 1988. *Multiple-Choice Cloze Items and the Test of English as a Foreign Language*. Educational Testing Service. Princeton. Research Report No. 26
- Jiang Xixiang. 1994. "‘Yuyan jineng’ ji qi zai jiaoxue yu ceyan fangmian de yingyong - jian lun ‘Zhongwen yuyan ceyan’" (Die sprachlichen Fertigkeiten und ihr Gebrauch in Unterricht und in Prüfungen - unter gleichzeitiger Berücksichtigung von Prüfungen der chinesischen Sprache). In: *Huawen Shijie*, Taipei, No. 73, Sept. 1994, 10-19
- Kaden, Klaus. 1995. *Prüfung zum Nachweis chinesischer Sprachkenntnisse - Hanyu Shuiping Kaoshi HSK - Dokumente - Stand 1993*. Herausgegeben vom Fachverband Chinesisch e.V.
- Kühn, Peter. 1979. "Kritik der bisherigen Grundwortschatzlexikographie". In: *Zielsprache Deutsch*. Nr. 4, 34-42
- Liu Yinglin (Hrsg.). 1989. *Hanyu Shuiping Kaoshi (HSK) yanjiu: fu HSK kaoshi dagang ji yangti, HSK changyong cihui yilanbiao* (Forschungen zur HSK: mit Prüfungsleitfaden sowie Probeprüfungsbogen und dem Verzeichnis häufiger Wörter für die HSK). Beijing
- Pan Zhaoming/Shi Xudong. 1993. *Di-si jie guoji Hanyu jiaoxue taolunhui: Guanyu zai Meiguo juxing HSK shikao de fenxi baogao*. (IV. Internationales

- Symposium für ChaF: Analyse der Probeprüfung der HSK in Amerika). Unveröffentlichtes Referat des IV. internationalen Symposiums für ChaF. Beijing Peirce, Bonny Norton. 1992. "Demystifying the TOEFL Reading Test". In: *TESOL Quarterly* 26, 4, winter, 665-691
- Pike, Lewis W. 1979. *An Evaluation of alternative item formats for testing English as a foreign language*. Educational Testing Service, Princeton, TOEFL Research Report No. 2
- Raasch, Albert. 1972. "Neue Wege zu einem Grundwortschatz". In: *Praxis des neusprachlichen Unterrichts*, Nr. 19, 235-244
- Schelten, Andreas. 1980. *Grundlagen der Testbeurteilung und Testerstellung: Teststatistik und Testtheorie für Pädagogen und Ausbilder in der Praxis*. Heidelberg
- Zhongguo Duiwai Hanyu Jiaoxue Xuehui/Hanyu Shuiping Dengji Biaozhun Yanjiu Xiaozu. 1988. *Hanyu shuiping dengji biaozhun he dengji dagang*. (Stufenstandards und Stufenprogramme für chinesische Sprachkenntnisse). Beijing

Summary

The present article summarizes the most important results of a research about the HSK. Development, content and form of this examination are discussed on a theoretical level. Furthermore the results of a questionnaire delivered to the examinees of the first HSK in Europe are summed up.

Aspekte der Wortstellung in der chinesisch-deutschen Übersetzung

Dorothea Wippermann

1. "Stellungssinismen" in deutschen Übersetzungen aus dem Chinesischen

Im Zusammenhang mit der chinesisch-deutschen Übersetzung findet die Wortstellung¹ vor allem dann Beachtung, wenn es darum geht, sogenannte "Stellungssinismen" in der deutschen Zielsprache zu kritisieren oder zu vermeiden. So wird immer wieder auf deutsche Stilfehler hingewiesen, die durch eine zu enge Anlehnung an die chinesische Wortstellung zu erklären sind, und empfohlen, sich bei der Übersetzung bewußt von den Vorgaben der chinesischen Wortstellung zu entfernen.² Typische Stellungssinismen ergeben sich z. B., wenn deutsche Nebensätze an unpassender Stelle ins deutsche Satzgefüge eingeschoben werden, so in Satz (1a), der in einer veröffentlichten Besprechung³ kritisiert wurde:

- (1a) [...] nur weil er heute morgen, nachdem er um drei Uhr aufgewacht war, plötzlich nicht mehr hatte schlafen können; [...]
(Mao Dun, "Eine ideale Frau", S. 137)

Der unterstrichene Nebensatz wurde mitten in den übergeordneten Satz hineingezwängt, was der Stellung im chinesischen Ausgangssatz (1b) entspricht:

- (1b) [...] 无非因为早晨三点钟醒过来后,忽然不能再睡, [...]
[...] wufei yinwei zaochen san dianzhong xingguolai hou, huran bu neng zai shui,
[...]
(Mao Dun, "Chuangzao", S. 3)

Übersetzungskritiker und -didaktiker haben bezüglich solcher Fälle eine Korrektur durch Umstellung nahegelegt. Danach wäre der eingeschobene Nebensatz im vorliegenden Beispiel nach hinten zu rücken:

¹ Der Terminus "Wortstellung" wird hier in einem sehr allgemeinen Sinne verwendet. Tatsächlich geht es bei den besprochenen Phänomenen nicht generell um die Abfolge von Einzelwörtern, sondern einerseits um die Abfolge von Satzgliedern bzw. Wortgruppen oder Konstituenten innerhalb von Sätzen/übergeordneten Konstituenten und andererseits um die Abfolge von Einzelwörtern innerhalb von Konstituenten (vgl. Gück 1993:698).

² Siehe z. B. Reissinger (1985:55f.); Kautz (1990:107). Entsprechende Empfehlungen wurden auch für die umgekehrte Übersetzungsrichtung "Deutsch-Chinesisch" ausgesprochen, z. B. von Yang Shouguo (1989:146-150). Siehe außerdem die Kritik an "Stellungssinismen" in einigen sinologischen Übersetzungskritiken: Kettelhut (1992:93f.); Höger u. a. (1992:107f.); Hilfrich-Xiang/Wang-Jin (1992:146). Zur Befreiung von der Wortfolge des Originals als Mittel zur Erlangung eines guten deutschen Stils siehe auch Chang Peng (1991:141).

³ Kettelhut (1992:93).