



Mündliche kommunikative Kompetenz in der Fremdsprache messen: Herausforderungen und Chancen

Assessing Oral Communicative Competence in the Foreign Language: Challenges and Opportunities

Received 26 March, 2021

Revised 20 June, 2021

Accepted 23 June, 2021

Jan Stevener¹

Abstract

Kommunikative Kompetenz in der Fremdsprachendidaktik greift auf die Theorie der Sprechakte (Searle 1971, Austin 1972) als erfolgreiches sprachliches Handeln zurück. Dies soll in kommunikativen Situationen geschehen, in Lerner während eines Aufenthaltes in Deutschland kommen könnten (Huneke & Steinig, 2013, p. 208). Der kommunikative Ansatz bestimmte entscheidend die Gestaltung des Europäischen Referenzrahmens für Sprachen und dieser wiederum alle wichtigen Sprachprüfungen. Anhand von Beispielen im Prüfungsteil Sprechen des TestDaF und des Zertifikat B1 des Goethe-Instituts soll exemplarisch beleuchtet werden, ob und inwieweit mündliche kommunikative Kompetenz in diesen Prüfungen verlässlich und valide gemessen werden kann.

Schlüsselwörter: Testen und Prüfen, kommunikative Kompetenz, Gemeinsamer Europäischer Referenzrahmen für Sprachen, TestDaF.

¹ Assistant Lecturer, Humanities and Language Division, Mahidol University International College (MUIC). e-mail: jan.ste@mahidol.edu



Abstract

Communicative competence in foreign language didactics draws on the theory of speech acts (Austin, 1972, Searle, 1971) as successful linguistic action. This should take place in communicative situations that the learner could come into during a stay in Germany (Huneke & Steinig, 2013, p. 208). The communicative approach decisively determined the design of the Common European Framework of Reference for Languages and this in turn all major language tests. Using examples in the speaking section of the TestDaF and the B1 certificate from the Goethe Institute, it is intended to illustrate whether and to what extent oral communicative competence can be reliably and validly measured in these exams.

Keywords: Testing and evaluation, Communicative competence, Common European Framework of Reference for Languages, TestDaF.

1. Einleitung

Der DaF Unterricht in Thailand hat den kommunikativen Ansatz erst relativ spät entdeckt. Dies ist wohl auch darauf zurückzuführen, dass beispielsweise den Schulen erst in den 2000er Jahren freigestellt wurde, welches Lehrwerk verwendet wird. Mit der Entscheidung für bestimmte Lehrwerke wurde dann auch der darin berücksichtigte methodische Ansatz übernommen. Während aktuell verwendete Lehrwerksreihen wie „Berliner Platz Neu“ oder „Studio 21“ durchaus dem kommunikativen Ansatz verpflichtet scheinen, ist festzustellen, dass in der Test- und Prüfungspraxis oft doch wieder grammatikzentriert geprüft wird (vgl. Rodi 2013a, 2013b). Das Überprüfen der kommunikativen Kompetenz stellt Lehrende darüberhinaus oft vor praktische Probleme, denn das sprachliche Handeln in möglichst authentischen Verwendungssituationen ist in der Prüfungspraxis oft schwer umsetzbar. Auf der anderen Seite bieten grammatikzentrierte Prüfungen den Lehrenden eine gewisse Sicherheit bei der Durchführung und Bewertung, da relativ einfach und sicher entschieden werden kann, ob eine Aufgabe richtig oder falsch gelöst wurde.



Mit Blick auf die großen formellen Prüfungen wie TestDaF oder die verschiedenen Prüfungen des Goethe Instituts ist jedoch die Erhebung kommunikativer Kompetenz beabsichtigt. Der relative Bedeutungsverfall grammatischer Kompetenz ist auch daran zu erkennen, dass neuere formelle Prüfungen wie TestDaF zwar die Teilfertigkeiten Schreiben, Lesen, Hören und Schreiben erfassen, es jedoch keinen expliziten Grammatikteil mehr gibt. Moderne Lehrwerke stellen kommunikative Kompetenz in den Mittelpunkt, und vor allem diese wird als Gradmesser des sprachlichen Niveaus im Gemeinsamen Europäischen Referenzrahmen für Sprachen (Europarat, 2001), der den meisten Lehrwerken und Prüfungen zugrunde liegt, betrachtet. In diesem Beitrag will ich daher Prinzipien des kommunikativen Ansatzes darstellen und erörtern, ob und inwieweit diese in gültige Prüfungsaufgaben in den Aufgaben zum Sprechen im TestDaF und im Goethe-Zertifikat B1 des Goethe Instituts überführt wurden.

2. Kommunikative Kompetenz

Im Folgenden soll dargestellt werden, welchen theoretischen Prinzipien und Ansätzen die kommunikative Kompetenz verpflichtet ist, denn im Zusammenhang des Testens und Prüfens ist Klarheit über den Untersuchungsgegenstand unabdingbar (Stevener, 2018, p. 29). Die Wende zur kommunikativen Kompetenz als übergeordnetes Lernziel begann mit Piepho (1974), der diese einerseits als kommunikatives Handeln und andererseits als Diskurstüchtigkeit definierte. Kommunikatives Handeln wird als die Fähigkeit beschrieben, sich in einem thematischen und situativen Kontext mit angemessenen Strategien und Mitteln verständlich zu machen und in der Lage zu sein, andere zu verstehen. Insbesondere dieser Aspekt hat starken Einfluss auf die Methodendebatte im Fach Deutsch als Fremdsprache gehabt.

Diskursfähigkeit wird als metakommunikative Fähigkeit bestimmt, sprachliches Handeln zu problematisieren, zu erklären und zu durchschauen. Dieser Anspruch ist verständlich vor dem Hintergrund, dass Piepho seine Überlegungen für das westdeutsche Schulsystem der 1970er Jahre formulierte und den Sprachunterricht als einen emanzipativen Erziehungsprozess zum mündigen Bürger sah. Dieser Anspruch ist jedoch beim Unterrichten einer Fremdsprache mit erwachsenen



Lernern weniger relevant. Ferner griff das Konzept der kommunikativen Kompetenz auf die Sprechakttheorie nach Searles (1971) und auf Austins Text mit dem prägnanten Titel „How To Do Things With Words“ (Austin, 1962) zurück. Canale und Swain (1980) und Canale (1983) beschreiben kommunikative Kompetenz weiterführend als ein Zusammenspiel von Wissen und Können in den folgenden Bereichen: *Grammatische Kompetenz*, d.h. die Beherrschung grammatischer Regeln und Strukturen, *soziolinguistische Kompetenz*, d.h. korrekter Umgang mit kommunikativen Kontexten und Situationen, *Diskurskompetenz*, d.h. Texte dekodieren und produzieren können und *Strategiekompetenz*, d.h. die Fähigkeit, mangelnde sprachliche Fertigkeiten auszugleichen und mit Schwierigkeiten umzugehen.

Auch Ellis (1994) stellt fest: “Communicative competence includes knowledge the speaker-hearer has of what constitutes appropriate as well as correct language behavior and also of what constitutes effective language behavior in relation to particular communicative goals ... it includes both linguistic and pragmatic knowledge” (p. 13). Dazu gehört u.a. auch das Wissen, illokutive Sprechakte auszuführen (Ellis, 1994, p. 165), die Canale (1983, p. 7) zur soziolinguistischen Kompetenz zählt. Legutke (2013, p. 73) weist darauf hin, dass sich in der Praxis starke und schwache Versionen dieses Ansatzes finden. In der starken Version bemühe man sich „um eine grundlegende Neukonzeption des fremdsprachlichen Klassenzimmers als Ort kommunikativen Handelns“, während in der schwachen Version formorientierter Unterricht und Lehrwerke mit kommunikativen Elementen angereichert werden. „Kommunikative Kompetenz wurde zum Schlagwort und Werbeträger für Lehrbuchverlage“ (2013, p. 73) und damit einhergehend verlor sich die pädagogisch-emanzipative Bedeutung des Konzepts. Legutke spricht hier von einer Trivialisierung des Konzepts. Hallet (2008) bestimmt die Diskursfähigkeit in einem auch für den außerschulischen Unterricht besser anwendbarem Sinn als die Fähigkeit, an fremdsprachlichen gesellschaftlichen Prozessen und Diskursen teilnehmen und kulturell partizipieren zu können.

2.1 Kompetenz, Performanz und Kann-Beschreibungen im Gemeinsamen Europäischen Referenzrahmen für Sprachen



Für die Frage, wie kommunikative Kompetenz zu testen sei, liefern die Kategorien von Canale und Swain (1980, p. 34) bereits brauchbare Hinweise. In Hinblick auf die Erhebung kommunikativer Kompetenz durch Prüfungen geben sie zu bedenken: „Communicative testing must be devoted not only to what the learner knows about the second language and about how to use it (competence), but also to what extend the learner is able to actually demonstrate this knowledge in a meaningful communicative situation (performance)“. Des Weiteren argumentieren sie, dass integrative Tests eher Performanz messen können als *discrete point tests*, also Tests, die einzelne Strukturen oder Elemente testen. Die letzteren, die eher Wissen messen, geben jedoch den Lernenden Aufschluss über den Lernstand. Dies ist für die Entwicklung und Anwendung der Strategiekompetenz eine wichtige Voraussetzung. Die Hinwendung zur Performanz zeigt sich deutlich bei der Entwicklung des Gemeinsamen Europäischen Referenzrahmens für Sprachen (im folgenden „GER“ abgekürzt) und dessen Erweiterung im Begleitband (Europarat, 2020), dessen Bestimmungen von Niveaustufen anhand von sogenannten Kann-Beschreibungen vorgenommen werden. Diese Kann-Bestimmungen liefern Test-Entwicklern und Lehrenden ein Instrument, um Sprachtests zu entwickeln, denn der GER soll für unterschiedliche Sprachen eine Basis nicht nur der Entwicklung von Curricula und die Beschreibung von Lehr- und Lernzielen bieten, sondern auch Methoden und Verfahren der Evaluation ermöglichen. Legutke (2013, p. 73) weist darauf hin, dass dem GER die von Canale/Swain (1980) und Canale (1983) beschriebenen Kompetenzen zugrundeliegen.

Die Kann-Beschreibungen des GER werden als sprachliche Handlungen beschrieben. Hauptziel des fremdsprachlichen Unterrichts ist demnach die sprachliche Handlungsfähigkeit (Ende, Grotjahn, Kleppin & Mohr, 2013). Grundvoraussetzung für (sprachliche) Kompetenz sind „die Summe des (deklarativen) Wissens, der (prozeduralen) Fertigkeiten und der persönlichkeitsbezogenen Kompetenzen, die es einem Menschen erlauben, Handlungen auszuführen“ (Europarat, 2001, p. 21). Diese Kompetenzen ermöglichen kommunikative Sprachkompetenzen, die nach GER *linguistische Kompetenzen* (Lexik, Semantik, Grammatik etc.), *soziolinguistische Kompetenzen* (Höflichkeit, Kennzeichnung sozialer Beziehungen,



Redewendungen und Zitate, Registerunterschiede und Varietäten) und *pragmatische Kompetenz* (bezügl. Diskurs und Funktion) umfassen (Roche, 2013, p. 217). Die Struktur der Kann-Beschreibungen folgt dem Muster: *kann* + [*was*] + [*wie gut*] + [*Verb*] + [*Bedingungen und Einschränkungen*] (Quetz, 2013, p. 46).

In Hinblick auf die Tauglichkeit solcher Kann-Beschreibungen muss darauf hingewiesen werden, dass der GER wegen Ungenauigkeiten und terminologischer Unschärfen kritisiert wurde (Harsch, 2007). Hier sei auf des „z.B.“ und das „usw.“ in vielen Beschreibungen hingewiesen. Entlastend sei aber gesagt, dass der GER keinen normativen Anspruch vertritt, sondern sich nur als Hilfestellung zur Kallibrierung von Kompetenzstufen versteht.

2.2 Interkulturelle Kompetenz

In den 1990er Jahren rückten aufgrund von Migration, Globalisierung und höherer grenzüberschreitender Mobilität zunehmend interkulturelle Aspekte fremdsprachlicher Kommunikation in den Mittelpunkt. Die hier dargestellten kommunikativen Kompetenzen wurden ergänzt um interkulturelle Aspekte, die sich auf kognitive, affektive, ethische und willensbezogene Kompetenzen bezogen (Hu, 2013, p. 76). Interkulturelle Kompetenzen finden sich im GER, der interkulturelle kommunikative Kompetenz als wichtiges Lehrziel für die Fremdsprache benennt: „In einem interkulturellen Ansatz ist es ein zentrales Ziel fremdsprachlicher Bildung, eine günstige Entwicklung der gesamten Persönlichkeit des Lernenden und seines Identitätsgefühls als Reaktion auf die bereichernde Erfahrung des Andersseins anderer Sprachen und Kulturen zu fördern“ (2001, p.14). Genannt werden u.a. die Entwicklung von Neugier und Offenheit gegenüber dem Fremden und kultureller Vielfalt, kulturangemessenes Verhalten sowie Bewusstsein über die Sicht- und Wahrnehmungsweisen des eigenen und des fremdkulturellen Landes. Es ist jedoch festzustellen, dass der GER ein Erfolg bezüglich der kallibrierten Niveaustufen A1 – C2 ist, der interkulturelle Ansatz dahinter jedoch zumeist übersehen wurde. Hu (2013, p. 77) weist darauf hin, dass interkulturelle Domänen zu den wenig operationalisierten und schwer überprüfbaren Bereichen gehören. Prüfungen, die auf den GER verweisen, müssen also auch dahingehend betrachtet werden, ob und wie interkulturelle



Kompetenz gemessen wird. Im Mai 2020 erschien als Erweiterung des GER ein Begleitband (Europarat, 2020), der die vorhandenen 54 Skalen auf 80 Skalen erweiterte. Viele dieser Skalen sind für den vorliegenden Untersuchungsgegenstand irrelevant (Gebärdensprache, Niveaus unterhalb A1, online Kommunikation etc). Eine allgemeine Übersicht der Ergänzungen findet sich in Wisniewski, Tschirner & Bärenfänger (2019). Die akademische Auseinandersetzung mit den neuen Skalen hat jedoch erst in Teilbereichen wie Strategien (Wisniewski, 2019) und Literatur und Wortschatz (Jazbec & Kacjan, 2019) begonnen, und bislang gibt es noch keine große Sprachprüfung, die sich auf die neuen Skalen beruft, was auch daran liegen könnte, dass es noch keine Anpassung des linguistischen Materials und Beispiele durch eine Revision von „Profile deutsch“ (Glaboniat, Müller, Rusch, Schmitz & Wertenschlag, 2013) gibt.

3. Testkonstrukt kommunikative Kompetenz

Ziel des kommunikativen Fremdsprachenunterrichts ist die Vermittlung von bestimmten sprachlichen und pragmatischen Kompetenzen, und der Erfolg der Sprachvermittlung kann durch eine Erhebung dieser Kompetenzen überprüft werden. Der GER ist hier trotz aller Unschärfen ein wertvolles Instrument, denn er ermöglicht Aussagen darüber, was die Lernenden schon können sollen. Diese Aussagen bilden die Grundlage für eine lernzielgerechte Überprüfung dieser Kompetenzen. Für die Prüfungspraxis ist jedoch festzustellen, dass man Kompetenzen nicht direkt messen kann, sondern nur beobachtbares Verhalten. Die Kompetenzen, so wie sie von den Theoretikern beschrieben wurden, sind vor allem theoretische Konstrukte. Um diese Kompetenzen messen zu können, muss derjenige, der einen Test erstellt, sich darüber im Klaren sein, welche bestimmten Merkmale als beobachtbares Verhalten der jeweiligen Kompetenz zugeschrieben werden können. Auf diese Weise wird auch der Zusammenhang von Theorie und Praxis transparenter und implizite theoretische Grundannahmen können kritisch reflektiert werden. Grotjahn und Kleppin (2015, p. 87) sehen die genaue Beschreibung des Testkonstruktes auch als Voraussetzung für die Auswahl passender Aufgabenformate. Das Messen sprachlicher



Handlungsfähigkeit muss also Auskunft darüber geben, wie erfolgreich Prüflinge „lebensweltlich relevante Fragestellungen sprachlich lösen können“ (Grotjahn & Kleppin, 2015, p. 17).

Die Kompetenzen, die diesem Verhalten zugrunde liegen, müssen über bestimmte Prüfungsaufgaben sichtbar gemacht werden. Die Aufgaben sollen dabei bestimmte beobachtbare Aktivitäten (*Operationen*) ausführen. Auf Grund der gezeigten Aktivitäten, d.h.

der Performanz, sind dann Rückschlüsse auf die jeweilige Kompetenz, d.h. das Testkonstrukt, möglich. Zu beachten ist, dass das Testkonstrukt in den Prüfungsaufgaben ausreichend repräsentiert ist. Grotjahn und Kleppin (2015, p. 47) geben das folgende Beispiel: „Eine Kollegin möchte in ihrem Unterricht überprüfen, ob ihre Lernenden sich in Deutschland schriftlich für einen Praktikumsplatz bewerben können. Sie berichtet von ihrer Prüfungsvorbereitung: „Ich habe in einem deutschen Bewerbungsbrief, den ich im Internet gefunden habe, eine Reihe von schwierigen Wörtern entfernt. Die Schülerinnen und Schüler sollten die fehlenden Wörter ergänzen“. In dieser Aufgabe werden vor allem Lexikkenntnisse überprüft, die zwar ein Teilbereich des intendierten Testkonstrukts (Fähigkeit, einen Bewerbungsbrief auf einen Praktikumsplatz zu schreiben) darstellen, jedoch weitere erforderliche Fähigkeiten (Textkohärenz, pragmatische Angemessenheit etc.) nicht erfassen; es besteht damit eine Konstruktunterrepräsentation. Stevener (2018) verweist darauf, dass beispielsweise der Schreibkompetenz (*Konstrukt*) die Fähigkeit „Korrektheit“ zuzuordnen ist, die beobachtet werden kann, indem die Fehler gezählt werden, dass jedoch mit dem Testkonstrukt weitere Eigenschaften wie Komplexität, Angemessenheit, Kohärenz etc. verbunden sind, die schwieriger quantitativ zu erfassen sind. Klarheit über die dem Testkonstrukt zugeschriebenen Eigenschaften sind daher die Grundlage für die Erstellung guter Prüfungen.

4. Gütekriterien von Sprachprüfungen

Darüber hinaus sollten zentrale Gütekriterien auf den Test angewendet werden, denn erst mit Hilfe solcher Gütekriterien wird deutlich, welcher Erkenntnisgewinn mit dem jeweiligen Test überhaupt zu erzielen ist. Stevener (2018) nennt als die wichtigsten Gütekriterien Validität,



Reliabilität und Objektivität. *Validität* bedeutet Gültigkeit; misst der Text wirklich das, was er zu messen vorgibt? Können die gezeigten Ergebnisse eindeutig auf das untersuchte Testkonstrukt zurückgeführt werden oder gibt es alternative Erklärungen? Je mehr alternative Erklärungen möglich sind, desto geringer ist die interne Gültigkeit des Test. Die externe Validität hingegen fragt danach, inwieweit das in der spezifischen Prüfungssituation gezeigte Verhalten auch auf reale lebensweltliche Situationen ausserhalb der Prüfungssituation verallgemeinbar ist.

Reliabilität bezieht sich auf die Zuverlässigkeit einer Prüfung. Wenn derselbe Prüfling den (struktur-) gleichen Test zweimal hintereinander ablegt, sollte auch das gleiche Ergebnis erzielt werden, da die Kompetenz des Lernernden sich nicht verändert haben dürfte.

Objektivität wird vor allem durch eine hohe Standardisierung bei der Durchführung und Bewertung von Prüfungen erreicht. Bei der Durchführung kann beispielsweise die Objektivität durch zahlreiche Störfaktoren (beim Hörverstehen Aspekte der Akustik oder Besonderheiten unterschiedlicher Sprecher, beim Sprechen oder Schreiben nicht festgelegte Zeitvorgaben, beim Sprechen als irritierend empfundenen Prüerverhalten etc.) beeinträchtigt werden.

Grotjahn und Kleppin (2015) und Stevener (2018) nennen des Weiteren *Praktikabilität*, *Ökonomie*, *Authentizität*, *Fairness*, *Transparenz*, *Trennschärfe* und *Nützlichkeit* als relevante Gütekriterien für Prüfungen. Die Gütekriterien helfen bei der Einschätzung und Erstellung sowohl von informellen Tests als auch von grossen Testbatterien wie dem TestDaF. Das Kriterium der *Ökonomie* fragt danach, wie ökonomisch ein Test in Erstellung, Durchführung und Bewertung ist. Braucht man beispielsweise teure technische Hilfsmittel, zusätzliche Prüfer, besondere Räumlichkeiten (Internet, Raumakustik etc.)? Wieviel Arbeitszeit muss auf den Test verwendet werden und lohnt sich der finanzielle Aufwand? Bei einer Prüfung wie dem TestDaF ist die Einschätzung sicher eine andere als bei der Semesterzwischenprüfung eines universitären Sprachkurses. Auch das Kriterium *Praktikabilität* betrifft die Unterscheidung von Test und Prüfungen. Die vollumfängliche Erfüllung aller Gütekriterien ist gerade bei informellen Tests kaum möglich, da es kaum praktikabel ist, auf die Vorbereitung eines kleinen Tests beispielsweise 20 Stunden Vorbereitung zu verwenden, absolut authentische Situationen zu schaffen oder unter



vollkommen kontrollierten Bedingungen zu testen. Gemeinhin wird im Deutschen zwischen Tests und Prüfungen unterschieden, jedoch werden in der Praxis beide Begriffe nicht klar getrennt (Stevener, 2018). Mit dem Begriff Prüfungen bezeichnet man eher große formelle Prüfungen, während Test sich eher auf kleinere, informelle Prüfungen bezieht. Aus dem Englischen hat sich jedoch das Wort Test soweit etabliert, dass auch Prüfungen wie der TestDaF den Test im Titel tragen. Für den Fertigungsbereich Sprechen sollen im Folgenden die Prüfungsteile zum Mündlichen Ausdruck aus dem TestDaF und dem Goethe-Zertifikat B1 genauer betrachtet werden.

5. Sprechen im TestDaF und im Goethe-Zertifikat B1

Die Entscheidung, exemplarisch den TestDaF und das Goethe-Zertifikat B1 zu untersuchen, fiel vor allem auf Grund der hohen gesellschaftlichen Relevanz dieser formellen Prüfungen. Denkbar wäre auch ein Vergleich TestDaF mit der Prüfung C1 Hochschule von telc, jedoch sind die zwei ausgewählten Prüfungen deutlich verbreiteter als die telc Prüfung. Im Rahmen der Flüchtlingskrise seit 2015 wird verstärkt auf die Bedeutung der sprachlichen Fertigkeiten für eine erfolgreiche Integration hingewiesen. In der Praxis hat sich dabei die Niveau Stufe B1 als Zielvorgabe etabliert, die bei vielen Berufen als Voraussetzung gilt (Klapsa, 2019), und dieses Niveau wird vor allem mit dem Goethe-Zertifikat B1 nachgewiesen. Der TestDaF hingegen liegt auf den Niveaustufen B2 bis C1 (Kecker, 2011) und entscheidet darüber, ob internationale Studienbewerber genügend Deutsch nachweisen können, um an einer deutschen Hochschule studieren zu können. Damit sind diese Prüfungen sogenannte *high-stakes-exams*, denn das Bestehen oder Nichtbestehen kann erhebliche persönliche Konsequenzen für die Prüflinge haben. Des Weiteren unterscheiden beide Prüfungen sich deutlich in der Operationalisierung, daher lassen sich verschiedene Optionen des Designs kontrastiv darstellen.

5.1 Sprechen im Zertifikat B1 des Goethe Instituts



Das Prüfungsmodul Sprechen im Goethe-Zertifikat B1 besteht aus drei Teilen. Im ersten Teil sollen zwei Prüflinge in einer Paar-Prüfung etwas gemeinsam planen oder aushandeln. Im zweiten Teil soll es eine kurze Präsentation zu einem vorgegeben Thema geben, zu der dann als dritten Teil Fragen gestellt und beantwortet werden sollen. In einer 15-minütigen Vorbereitungszeit können sich die Prüflinge Notizen machen, sollen dann aber in der Prüfung möglichst frei sprechen. Betreffs der *Praktikabilität* und *Ökonomie* ist die Prüfung gut durchführbar, denn es wird nur ein Raum benötigt und mindestens ein Prüfer. Die Prüfung ist dadurch standardisiert, dass die Prüflinge die gleiche Vorbereitungszeit (15 Minuten) erhalten und es Vorgaben für die Redezeit gibt: 3 Minuten für Planung, je 3 Minuten für die Präsentation, 2 Minuten für Fragen bzw. Antworten. Die Einhaltung dieser Zeitvorgaben ist jedoch nicht systemisch, es kann also je nach Prüfer oder Umständen zu abweichenden Zeiten kommen. Insofern ist die *Objektivität* nicht vollständig gegeben. Die Präsentation, der mit 40 Punkten am höchsten bewertete Prüfungsteil, folgt immer dem selben Aufbau. Ein Thema soll 5 Folien folgend präsentiert werden, die dem Muster Themenvorstellung – meine persönlichen Erfahrungen – Situation in meinem Heimatland bzgl. des Themas – Vor- und Nachteile plus meine eigene Meinung – Abschluss und Dank folgen (vgl. Balser, Glotz-Kastanis, Papadopoulou, Paradi-Stai & Vosswinkel, 2014). Dafür vorgesehen sind 3-4 Minuten pro Kandidat. Anschließend sollen 1-2 Fragen zur Präsentation gestellt und beantwortet werden.

5.2. Testtheoretische Bewertung des Goethe-Zertifikat B1

Teil 1 wird als Paarprüfung durchgeführt, und hier hängt die Performanz in der Prüfung auch von der Performanz des Partners ab. Spricht der Partner undeutlich, kann der Prüfling möglicherweise keine oder eine unpassende Antwort geben, spricht der Partner sehr gut, kann der Prüfling Ausdrücke und „Chunks“ kopieren. Norton (2005) hat sich mit Paarprüfungen im Cambridge Speaking Test beschäftigt und konnte zeigen, dass die Performanz von Probanden beeinflusst wird, wenn Sprecher mit hohen und niedrigeren linguistischen Fähigkeiten zusammen getestet werden oder wenn die Probanden sich kennen, was z.B. auf die Prüfungstermine zutrifft, die das Goethe-Institut speziell für Studierende bestimmter Hochschulen anbietet. Des Weiteren



konnte Norton einen Einfluss der Paar-Zusammensetzung auf die Menge des Gesagten als auch einen Einfluss des Geschlechts auf die Teilnahme am Gespräch nachweisen. Die Befunde zum Cambridge Speaking Test sollten teilweise auf auch die B1 Prüfung übertragbar sein. Auch Grotjahn und Kleppin (2015, p. 154) merken zu Paarprüfungen an, dass Störeffekte möglich sind, wenn Prüflinge sich im Ausmaß der gegenseitigen Unterstützung deutlich unterscheiden oder sich gegenseitig benachteiligen, z.B. durch Unterbrechungen. Insofern muss die *Reliabilität* als eher gering eingestuft werden. Andererseits ist jedoch diese Situation unter dem Kriterium der *Authentizität* als besser einzustufen, denn in lebensweltlichen Situationen trifft man durchaus nicht nur Muttersprachler auf C2 Niveau, und auch der Begleitband des GER hat das Wort „Muttersprachler“ aus dem GER gestrichen. In der Prüfung zum Sprechen wird also eine höhere Authentizität mit einer geringeren Reliabilität erkaufte.

Im Sinne der *externen Validität* ist dies einerseits sinnvoll, denn kommunikative Kompetenz soll als sprachliches Handeln in möglichst authentischen Situationen stattfinden. Je authentischer die Situation ist, umso wahrscheinlicher kann das sprachliche Handeln auch auf lebensweltliche Situationen ausserhalb der Prüfungssituation übertragen werden. Absurderweise ist die *interne Validität* oder *Konstruktvalidität* jedoch dadurch etwas geringer, denn faktisch wird nicht nur das Sprechen, sondern als dessen Voraussetzung auch das Hörverstehen miterhoben.

Kommunikative Kompetenz erfordert jedoch auch eine soziolinguistische Kompetenz, die hier über eine informelle Sprachverwendung (etwas mit einem Bekannten planen) und eine formelle Verwendungssituation operationalisiert wird. In den Bewertungskriterien ist jedoch nicht klar, wie bestimmte pragmatische Kompetenzen (Ellis, 1994) oder Strategiekompetenzen (Canale, 1983) bewertet werden, beispielsweise wenn der Partner unverständliche oder unangemessene Äusserungen produziert. In der Operationalisierung sind diese nicht in dem Sinne berücksichtigt, dass diese elizitiert würden. Um eine Aussage über Strategiekompetenz auf Niveau B1 zu erhalten, müssten diese aber erhoben werden.

Der Gemeinsame Europäische Referenzrahmen für Sprachen (Europarat, 2001, pp. 88-89) stellt dafür sogar Skalen bereit. Für B1 liegen Skalen für *Sprecherwechsel*, *Kooperieren* und



um Klärung bitten vor. Auf dem B1 Niveau gibt es für *Kooperieren* die Beschreibung „kann andere auffordern, sich am Gespräch zu beteiligen“ und für *um Klärung bitten* die Beschreibung „kann andere bitten zu erklären oder genauer zu erläutern, was sie gerade gesagt haben“. In einer Paarprüfung und einer Präsentation gibt es jedoch kaum Möglichkeiten, andere außer dem designierten Partner zur Äußerung aufzufordern. Eine Bitte um Klärung wäre jedoch zumindest im Prüfungsteil *Fragen* nach der Präsentation möglich. Des Weiteren werden auch interkulturelle Kompetenzen nicht erhoben. Dies mag daher kommen, dass der GER trotz Bezugnahme auf kommunikative Kompetenz zuerst keine Skalen zur interkulturellen Kompetenz anbot (Kecker, 2011) und diese erst für den Begleitband 2020 entwickelt wurden.

5.3 Mündlicher Ausdruck im TestDaF

Im TestDaF hat man eine andere Operationalisierung zur Erhebung der mündlichen kommunikativen Kompetenz gewählt. Es gibt 7 Aufgaben, von denen jeweils 2 auf dem gleichen Schwierigkeitsniveau liegen und eine Aufgabe als Aufwärmübung nicht bewertet wird. Der TestDaF erhebt damit die drei TestDaF Niveaustufen (TDN) 3, 4 und 5, welche die GER Stufen B2 und C1 abdecken. TDN 4 liegt zwischen B2 und C1. Kecker (2011) konnte diese Zuordnung trotz einiger Grauzonen im Allgemeinen bestätigen, und für den Teil Mündlicher Ausdruck sieht sie „die Verbindung der TestDaF-Aufgaben zu den GER-Stufen B2 – C1 insgesamt bestätigt“ (2011, p. 185). Die Prüflinge hören in den Aufgaben zuerst die Beschreibung einer kommunikativen Situation im Kontext Hochschule. Ergänzt wird dies mit einer Problembeschreibung oder der Themenstellung eines fiktiven Gesprächspartners in dieser Situation. Die Vorbereitungszeit auf die sprachliche Äußerung ist fest vorgegeben und beträgt je nach Schwierigkeitsgrad der Aufgabe zwischen 30 Sekunden und 3 Minuten. Nach einem akustischen Signal hören die Prüflinge eine Aufzeichnung mit der Aufforderung des fiktiven Gesprächspartners. Danach haben die Prüflinge je nach Aufgabe 1 bis 2 Minuten Sprechzeit zur Verfügung. Das Ende der Sprechzeit wird wieder durch ein akustisches Signal angezeigt. Die Prüflinge sprechen nicht mit wirklichen Partnern, sondern hören Anweisungen und Gesprächspartner von einer Aufzeichnung und nehmen ihre eigenen Sprachhandlungen per



Computer oder Band auf. Eine ausführlichere Beschreibung findet sich bei Kecker (2011) und Koreik (2005).

Der jeweilige Schwierigkeitsgrad wird durch die Faktoren *Art der Sprechhandlung* (Informationen einholen, Diagramm beschreiben etc.), *Beziehung zu den Adressaten* (formelles oder informelles Register), *Äußerung zum eigen- oder fremdkulturellen Kontext* und *Art der kommunikativen Situation* (Hochschulveranstaltung, unter Freunden etc.) bestimmt. Kecker (2011, p. 175) konnte belegen, dass für die Erstellung der Aufgaben hauptsächlich zwei Skalen des GER verwendet wurden. Diese sind die Skala „Mündliche Produktion allgemein“, sozusagen eine Meta-Skala, und die Skala für „Zusammenhängendes monologisches Sprechen: Erfahrungen beschreiben“. Die Aufgaben 4 und 6, die TDN5/GER C1 zugeordnet sind, erfordern Sprechen vor einem Publikum. Daher wurde auch ein Teil der GER Skala „Vor Publikum sprechen“ berücksichtigt, allerdings „wurden nur die Deskriptoren berücksichtigt, die sich nicht auf den Umgang des Sprechers mit Nachfragen aus dem Publikum beziehen“.

5.4. Testtheoretische Bewertung Mündlicher Ausdruck im TestDaF

Die Anwendung einiger Gütekriterien zeigt Vor- und Nachteile dieser Operationalisierung. Unter dem Gesichtspunkt der *Objektivität* weist der Test eine deutlich höhere Standardisierung als das Goethe-Zertifikat B1 auf. Zeitangaben sind bindend und können auf Grund des technischen Formats nicht überschritten werden. Weltweit hören alle Probanden die gleichen Sprecher an festgesetzten Terminen. Die technischen Voraussetzungen (Text durch Kopfhörer) garantieren bestmögliche Verständlichkeit. Auch die *Reliabilität* ist als sehr hoch einzuschätzen, denn Arbeitsanweisungen sind klar formuliert und gekennzeichnet. Die Ziele werden vorab von einem Sprecher erläutert und das TestDaF Institut stellt unter www.testdaf.de kostenlose Musterprüfungssätze und Informationen für die Vorbereitung bereit, so dass auch die *Transparenz* des Verfahrens als hoch einzuschätzen ist. Jedoch ist die *Validität* unterscheidlich zu beurteilen. Konträr zum Goethe-Zertifikat B1 ist hier die *interne Validität* hoch, denn tatsächlich wird fast nur das Sprechen erhoben. Die Arbeitsanweisungen und Äußerungen von Partnern werden zum Hören und Lesen in den Prüfungsunterlagen angeboten, was diesen



Störfaktor deutlich minimiert. Die hohe *Durchführungsobjektivität* der Prüfung reduziert jedoch die *externe Validität*, da die Kommunikationssituation „mit dem Computer monologisch sprechen“ schwer auf lebensweltliche Situationen ausserhalb der Prüfungssituation Übertragbar ist. Es ist davon auszugehen, dass die hohe Künstlichkeit der Kommunikationssituation Prüflinge, die nicht gut mit dem Testformat vertraut sind, irritiert und damit Messfehler ermöglicht.

Die hohe Künstlichkeit der Prüfung steht darüber hinaus im klaren Gegensatz zur Forderung nach *Authentizität*. Da von Prüfungen immer auch ein *washback* Effekt für die fremdsprachliche Unterrichtspraxis ausgeht, muss der TestDaF hier negativ bewertet werden. Der Wunsch nach klaren Messverfahren und der Blick zu naturwissenschaftlichen und eher technischen Messverfahren torpediert letztendlich den kommunikativen Ansatz, der eine „grundlegende Neukonzeption des Klassenzimmers als Ort kommunikativen Handelns“ (Legutke, 2013, p. 73) intendiert und eben keine Laboratmosphäre erzeugen möchte.

Noch schwerer wiegt der Einwand, dass die gesamte Prüfung zum mündlichen Ausdruck monologisch stattfindet. Die für die kommunikative Kompetenz erforderlichen Strategien zum Sprecherwechsel beispielsweise werden in der Operationalisierung vollkommen ausgeklammert, denn in den Aufgaben wird dem Prüfling per se das Wort erteilt. Strategische Kompetenzen wie Umschreibungen und Reparaturen werden dadurch erschwert, dass es kein Feedback gibt. Wählt der Proband das falsche Register, indem er z.B. einen Professor oder eine Mitarbeiterin der Univerwaltung (Aufgaben 1, 4, 6) duzt oder Kommilitonen siezt (Aufgaben 2 und 7), kann auf Grund fehlender Reaktion keine Reparaturstrategie gestartet werden. Schmenk (2005, p. 80) beschreibt kommunikative Kompetenz als „eine rezeptive und produktive Fähigkeit zur sozialen Teilnahme am Dialog mit dem Ziel, eigene Standpunkte zu vertreten und gegebenenfalls zu modifizieren sowie diejenigen anderer anzuerkennen und sich damit auseinanderzusetzen, und dies sowohl im Bereich der Mündlichkeit wie auch Schriftlichkeit“. Der TestDaF jedoch reduziert die kommunikative Kompetenz hier auf die monologische Darlegung eigener Standpunkte und Meinungen. Immerhin könnten mit den 2020 im GER Begleitband neu erschienenen Skalen



zur „Kompensation“ immerhin Umschreibungen, Monitoring und Reparaturen kalibriert werden.

5.5. Interkulturelle Aspekte im TestDaF

Interkulturelle Aspekte werden als Faktoren bei der Bestimmung des Schwierigkeitsgrades berücksichtigt. Aufgabe 2 erfordert die Beschreibung bestimmter Sachverhalte im Heimatland, beispielsweise im Modellsatz 2 „wann junge Leute von zu Hause ausziehen“. Auch Aufgabe 7, die sich auf ein persönliches Anliegen bezieht, kann einen Bezug zur Kultur des Heimatlandes haben. Diese beiden Aufgaben sind TDN3/GER B2 zugeordnet und erfordern neben lexikalischen und grammatischen Strukturen auch die Verwendung von Verknüpfungsmitteln und die Reproduktion von Wissen (Aufgabe 2) bzw. Wissensverarbeitung (Aufgabe 7). Die Sprechhandlungen sind als *berichten/beschreiben/Sachverhalt darstellen* (Aufgabe 2) bzw. *begründet Rat geben/Meinung darlegen* (Aufgabe 7) angegeben (Kecker, 2011, p. 66).

Keine dieser Aufgaben erhebt jedoch interkulturelle kommunikative Kompetenz, denn diese geht über reines Wissen um die eigene und fremde Kultur hinaus. Byram (1997) ergänzt das reine kulturelle Wissen um die Fähigkeiten, andere Kulturen zu verstehen und zu interpretieren, selbständig Neues über diese zu lernen, sowie Offenheit und Neugier gegenüber anderen Kulturen und ein kritisches Bewusstsein darüber, wie in einer Gesellschaft mit Werten umgegangen wird. Die genannten TestDaF Aufgaben tragen das Interkulturelle damit nur im Titel, ohne dass interkulturelle Kompetenz wirklich erfasst würde. Dies mag dem Fehlen entsprechender Skalen im ursprünglichen Gemeinsamen Europäischen Referenzrahmen für Sprachen (Europarat, 2001) geschuldet sein. Die Überarbeitung der Skalen im Rahmen des Begleitbandes 2020 betrifft einerseits die Streichung des Begriffs „Muttersprachler“ in den Beschreibungen, andererseits weitere 26 Skalen. Von diesen Skalen betreffen die folgenden mehr oder weniger die interkulturelle Kommunikation: Kooperative Interaktion unterstützen, gemeinsame Konstruktion von Bedeutung, Interaktion organisieren, plurikulturelle Interaktion ermöglichen, plurikulturelle Kommunikation ermöglichen, auf einem plurikulturellem Niveau



aufbauen, plurilinguales Verstehen, soziolinguistische Angemessenheit und kulturelles Vorwissen. Diese Skalen sind wie gewohnt vage und zumeist nur für die höheren Niveaustufen ab B2 definiert.

Vordergründig sinnvolle Skalen wie zur plurikulturellen Kommunikation definieren jedoch Kompetenzen, die sich empirischer Überprüfbarkeit entziehen, z.B. sich mit Empathie in die Gefühlswelt anderer zu versetzen, die Identifizierung emotionaler Bedürfnisse des anderen zu leisten oder zu einer Gruppe zu gehören und trotzdem Ausgewogenheit und Distanz zu bewahren. Grundsätzlich sollte eine Auswahl von 8-10 GER Skalen in Bezug auf die Lerngruppe und -ziele getroffen werden, denn es ist praktisch kaum möglich, alle Skalen gleichzeitig anzuwenden. Im Goethe-Zertifikat B1 könnten im Hinblick auf die neuen Skalen Adaptionen vorgenommen werden (B1 Kooperative Interaktion beschreibt die Fähigkeit, andere Personen in ein Gespräch einzuladen oder in Gespräche zu intervenieren und würde mehr als 2 Prüflinge erfordern), was prinzipiell im Prüfungsteil 1 möglich wäre. Im monologischen TestDaF Design hingegen gibt es in der bestehenden Form auch mit den erweiterten Skalen keine Möglichkeit, plurikulturelle Kompetenz zu erheben.

Es ist auch denkbar, dass die Ausrichtung des TestDaF auf „harte“ Methoden des Messens und Überprüfens der Erhebung interkultureller Kompetenz entgegen steht. Hu (2013, p. 78) stellt fest, dass in der aktuellen Diskussion auch auf die Verwendung „weicher“ Messverfahren wie z.B. Selbstreflexion oder Verfahren der Kulturwissenschaft (Interview, Diskursanalyse) verwiesen wird. Diese sollten als Desiderata für die Erhebung der produktiven Fertigkeiten im TestDaF gesehen werden. Betreffs von *washback* Effekten zeigt ein kursorischer Blick auf einige aktuelle Lehrwerke, dass die Zertifikate des Goethe Instituts, nicht nur auf dem B1 Niveau, erwartbar einen starken Einfluss auf die Gestaltung der Prüfungsbände haben. So werden die jeweiligen Operationalisierungen des mündlichen Ausdrucks für die Prüfungsbände von *Berliner Platz Neu* (Band 1 von Zertifikat A1 Start Deutsch 1, Band 3 von Goethe-Zertifikat B1: Rodi, 2013a und 2013b) verwendet, problematischerweise allerdings ohne dass diese Formate in den entsprechenden Lehrbüchern ausreichend vorbereitet worden wären. Auch das Testheft



von *Studio 21 B1* (Funk & Kuhn, 2016) orientiert sich am Goethe-Zertifikat B1. Das Lehrwerk selbst integriert im Gegensatz zu *Berliner Platz Neu* entsprechende Aufgaben und Übungen, die auf das Goethe-Zertifikat B1 vorbereiten. Es ist offensichtlich, dass mit der Ausrichtung auf die großen formellen Prüfungen auch deren Schwächen, beispielsweise die fehlende Vermittlung und folglich Erhebung strategischer oder interkultureller Kompetenz, übernommen werden. Die Vorbereitung auf den TestDaF erfolgt meist in gesonderten Vorbereitungskursen, was auf Grund der geringen Praktikabilität und den geforderten technischen Voraussetzungen für den Prüfungsteil mündlicher Ausdruck erklärbar ist.

6. Schlusswort

Die Anwendung der einzelnen Gütekriterien hat gezeigt, dass diesen von TestDaF und der B1 Zertifikatsprüfung sehr unterschiedlich entsprochen wird. Es zeigt sich, dass der TestDaF eher den Gütekriterien harter naturwissenschaftlicher Forschung entspricht (Reliabilität, Objektivität) als die B1 Zertifikatsprüfung des Goethe Instituts. Die Validität des Goethe-Zertifikat B1 unter Berücksichtigung der Forderungen kommunikativer Kompetenz nach Authentizität ist jedoch trotz der formulierten Einwände am Prüfungsformat Paarprüfung deutlich höher als die monologische Sprechprüfung des TestDaF. Die Erweiterung des GER im Begleitband (2020) beschreibt Sprache in Bezug auf Beziehungsgestaltung. Bedenkt man den Einfluss, den der GER auf Unterricht und Testen hat, so scheint es dringend geboten, den Mündlichen Ausdruck im TestDaF zu revidieren, denn der TestDaF überprüft sprachliche Fertigkeiten für ein Studium und insbesondere Hochschulen sind Orte plurikultureller Kommunikation.

Der vorliegende Artikel kann allerdings nur ausschnittsweise erfassen, in welchem Umfang die Forderungen nach kommunikativer interkultureller Kompetenz in Sprachprüfungen berücksichtigt werden. Die Analyse der Prüfungsteile zum Sprechen im Goethe-Zertifikat B1 und im TestDaF legt nahe, dass bislang nur Teilbereiche der kommunikativen Kompetenz erfasst werden. Insbesondere strategische und interkulturelle Kompetenzen werden in diesen Prüfungen



vernachlässigt. Insofern muss man Legutke (2013) dahingehend ergänzen, dass „weiche“ Versionen des kommunikativen Ansatzes sich nicht nur in den Lehrwerken, sondern auch in der Prüfungspraxis durchgesetzt haben. Dies mag in der schwierigen Erhebbarkeit strategischer und interkultureller Kompetenzen begründet sein. Im Begleitband 2020 finden sich weitere Skalen zur Strategie, allerdings sind diese ebenfalls teilweise kaum mit vertretbarem Aufwand oder harten wissenschaftlichen Methoden messbar (u.a. Planungs- und Monitorprozesse, verdecken von Wortschatzlücken, das kaum bemerkbare Ersetzen von Wörtern). Immerhin liegt erstmals eine Skala für Kompensationsstrategien vor, deren Erstellung bereits 2016 von Koeppel (2016, p. 318) gefordert wurde.

Ein gangbarer Weg wäre auch die Verwendung weicherer Erhebungsmethoden (Selbstreflexion, Selbsteinschätzungen). Strategische Kompetenzen, die nach GER „das Identifizieren des Kontextes auf der Basis des Weltwissens“ (2001, p. 77) umfassen, könnten stärker elizitiert werden, indem beispielsweise Situationen geschaffen werden, in denen Missverständnisse, Wortschatzlücken, Wissensmangel, schwierige Rezeptionsbedingungen in der Operationalisierung berücksichtigt werden und dadurch strategische Kompetenzen erhoben werden können. Die Erhebung interkultureller bzw. plurikultureller Kompetenzen in Tests wird auch durch die zusätzlichen Skalen nicht einfacher, denn einerseits liegen Skalen oft nur für die höheren Niveaustufen ab B2 vor, und zum anderen geht es um schwer messbare Zuschreibungen. Hier sollte vertiefend diskutiert werden, in welcher Form strategische und plurikulturelle Kompetenzen anhand der neuen Skalen des Begleitbandes in Prüfungen erhoben werden können.



Literatur

- Austin, J. L. (1962). *How to do Things with Words*. Oxford: Oxford University Press.
- Austin, J. L. (1972). *Zur Theorie der Sprechakte*. Stuttgart: Reclam.
- Balser, A., Glotz-Kastanis, J., Papadopoulou, M., Paradi-Stai, D. & Vosswinkel, A. (2014). *Zertifikat B1 neu. 15 Übungsprüfungen*. München: Hueber.
- Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence*. Philadelphia: Clevedon.
- Canale, M. (1983). From Communicative Competence to Communicative Language Pedagogy. In J.C. Richards, & R.W. Smith (Eds.), *Language and Communication* (pp. 2-28). London: Routledge.
- Canale, M. & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1, 1-47.
- Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.
- Ende, K., Grotjahn, R., Kleppin, K. & Mohr, I. (2013). Curriculare Vorgaben und Unterrichtsplanung. *Deutsch Lehren Lernen* 6. Berlin: Klett-Langenscheidt.
- Europarat (2001). *Gemeinsamer Europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin u.a.: Klett-Langenscheidt.
- Europarat (2020). *Gemeinsamer Europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen. Begleitband*. Berlin u.a.: Klett-Langenscheidt.
- Funk, H. & Kuhn, C. (2016). *Studio [21] B1. Testheft*. Berlin: Cornelsen.
- Glaboniat, M., Müller, M., Rusch, P., Schmitz, H. & Wertenschlag, L. (2013). *Profile deutsch*. Berlin: Klett-Langenscheidt.
- Grotjahn, R. & Kleppin, K. (2015). *Deutsch Lehren Lernen 7: Prüfen, Testen, Evaluieren*. Berlin: Klett-Langenscheidt.
- Hallet, W. (2008). Diskursfähigkeit heute. Der Diskursbegriff in Piephos Theorie der kommunikativen Kompetenz und seine zeitgemäße Weiterentwicklung für die



- Fremdsprachendidaktik. In M.K. Legutke (Ed.), *Kommunikative Kompetenz als fremdsprachendidaktische Vision*, (pp. 76-96). Tübingen: Narr.
- Harsch, C. (2007). *Der Gemeinsame europäische Referenzrahmen für Sprachen: Leistung und Grenzen. Die Bedeutung des Referenzrahmens im Kontext der Beurteilung von Sprachvermögen am Beispiel des semikreativen Schreibens im DESI-Projekt. Inaugural Dissertation*. Universität Augsburg. Retrieved from https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/deliver/index/docId/297/file/DISS_Claudia_Harsch.pdf
- Hu, A. (2013). Interkulturelle Kommunikative Kompetenz. In W. Hallet & F. G. Königs (Eds.), *Handbuch Fremdsprachendidaktik*. 2. Auflage, (pp. 75-79). Berlin: Klett Kallmeyer.
- Huneke, H. W. & Steinig, W. (2013). *Deutsch als Fremdsprache. Eine Einführung*. 6. neu bearbeitete Auflage. Berlin: Erich Schmidt.
- Jazbec, S. & Kacjan, B. (2019). Der gemeinsame Europäische Referenzrahmen für Sprachen - Lücken und Herausforderungen. *Aussiger Beiträge*, 13, 181–203.
- Kecker, G. (2011). *Validierung von Sprachprüfungen. Die Zuordnung des TestDaF zum Gemeinsamen europäischen Referenzrahmen für Sprachen*. Frankfurt am Main: Peter Lang.
- Klapsa, K. (2019). Migranten: „Manche beginnen erst nach 2 Jahren den Integrationskurs“. *Die Welt*, 9.10.2019. Retrieved from <https://www.welt.de/politik/deutschland/article201510966/Migranten-Manche-beginnen-erst-nach-zwei-Jahren-den-Integrationskurs.html>.
- Koeppel, R. (2016). *Deutsch als Fremdsprache – Spracherwerblich reflektierte Unterrichtspraxis*. 3. überarb. Auflage. Baltmannsweiler: Schneider.
- Koreik, U. (2005). *TestDaF und DSH – eine Vergleichsstudie*. Baltmannsweiler: Schneider.
- Legutke, M. K. (2013). Kommunikative Kompetenz und Diskursfähigkeit. In W. Hallet & F. G. Königs (Eds.), *Handbuch Fremdsprachendidaktik*. 2. Auflage, (pp. 70-74). Berlin: Klett Kallmeyer.



- Norton, J. (2005). The paired format in the Cambridge Speaking Tests. *ELT Journal*, 59(4), 287–297.
- Piepho, H.-E. (1974). *Kommunikative Kompetenz als übergeordnetes Konzept im Englisch Unterricht*. Dornburg-Frickhofen: Frankonius.
- Quetz, J. (2013). Gemeinsamer Europäischer Referenzrahmen. In W. Hallet & F. G. Königs (Eds.), *Handbuch Fremdsprachendidaktik*. 2. Auflage, (pp. 45-48). Berlin: Klett Kallmeyer.
- Roche, J. (2013). *Fremdsprachenerwerb Fremdsprachendidaktik*. 3. Auflage. Tübingen: Francke.
- Rodi, M. (2013a). *Berliner Platz 1 Neu. Testheft mit Prüfungsvorbereitung*. Berlin: Klett.
- Rodi, M. (2013b). *Berliner Platz 3 Neu. Testheft mit Prüfungsvorbereitung*. Berlin: Klett.
- Schmenk, B. (2005). Mode, Mythos, Möglichkeiten oder ein Versuch, die Patina des Lernziels „kommunikative Kompetenz“ anzukratzen. *Zeitschrift für Fremdsprachenforschung*, 16(1), 57-87.
- Searle, J. R. (1971). *Sprechakte. Ein sprachphilosophischer Essay*. Frankfurt am Main: Suhrkamp.
- Stevener, J. (2018). Testen und Prüfen in der Fremdsprache: Was macht eine gute Sprachprüfung aus?. *Ramkhamhaeng University Journal Humanities Edition*, 37(1), 21-46.
- Wisniewski, K. (2019). Zur Skalierung von Strategien im Gemeinsamen europäischen Referenzrahmen: Eine kritische Auseinandersetzung. *Deutsch als Fremdsprache*, 3,
- Wisniewski, K., Tschirner, E., & Bärenfänger, O. (2019). Der Begleitband zum „Gemeinsamen europäischen Referenzrahmen für Sprachen“. Einleitung in den Themenschwerpunkt. *Deutsch als Fremdsprache*, 56(2), 67-76.