

**Sektionsleitung: PD Dr. Olaf Bärenfänger (Leipzig) / Prof. Dr. Claudia Harsch (Bremen)**

**Prof. Dr. Norbert Schlüter (Leipzig):**

**THE LANGUAGE MAGICIAN – Leistungsbewertung mit dem digitalen Zauberstab**

THE LANGUAGE MAGICIAN ist ein Sprachprogramm, das Lehrkräften in den Klassen 4 und 5 hilft, den Lernfortschritt ihrer Schülerinnen und Schüler im Fremdsprachenunterricht in digitaler Form einzuschätzen. Das Spiel steht seit Mai 2018 kostenlos online und als App zur Verfügung stehen. In zwei aufeinander aufbauenden Sprachniveaus lösen die Lerner Sprachaufgaben in den einzelnen Kompetenzbereichen Hörverstehen, Leseverstehen und Schreiben sowie in Kombinationen dieser Fertigkeitsbereiche. Der besondere Vorteil dieses Programms besteht darin, dass alle Kombinationen der fünf beteiligten Sprachen getestet werden können, d.h. sowohl Englisch, als auch Deutsch, Spanisch, Italienisch und Französisch können jeweils als Test- und/oder Umgebungssprache gewählt werden. Die entwickelten Sprachniveaus wurden ausführlich getestet. Durch die Rückmeldungen der beteiligten Wissenschaftler konnten beide Sprachniveaus optimiert werden. Auch in den kommenden Jahren sollen die vom Programm erzeugten Daten in anonymisierter Form ausgewertet werden und so zu einem besseren Verständnis von Sprachlernprozessen und der zu erreichenden Sprachkompetenz der im Spiel integrierten Sprachen beizutragen. Seit der Freischaltung wurde das Programm bereits von mehr als 10.000 Lernern weltweit verwendet. Das Erasmus Plus Projekt wurde von 2015 bis 2018 von der EU und vom Goethe Institut gefördert. Unter der Leitung des Londoner Goethe Instituts schlossen sich zehn Partner zusammen, um das Lehren und Lernen von Fremdsprachen im Grundschulbereich und frühen Sekundarstufenbereich zu unterstützen. Partner im Projekt sind: Goethe-Institut London (Projektleitung), der Sprachenlehrerverband ALL in Großbritannien (Association for Language Learning), die Bildungsabteilung der Spanischen Botschaft in Großbritannien, das Bildungsministerium der autonomen Region Rioja (Spanien), das Lehrerfortbildungszentrum des spanischen Bildungsministeriums auf Teneriffa (Spanien), die Universität Leipzig (Deutschland), die Universität Reading (England), die Universitäten für Ausländer in Perugia und Siena (Italien) sowie die Universität Westminster (England). Der Vortrag wird sich mit der digitalen Leistungsmessung anhand des in diesem Forschungsprojekt entwickelten Programms beschäftigen und die Möglichkeiten und Grenzen aufzeigen, die sich aus der Digitalisierung von Bewertungsprozessen ergeben.

**Vanessa Appoh (telc):**

**Digitales Prüfen in der schulischen Praxis**

Im Rahmen eines Pilotprojekts hat die telc gGmbH mit *telc English A2-B1 From School to Job* ihre erste computergestützte Sprachprüfung für den Schulmarkt entwickelt. Kernzielgruppe sind Schüler/-innen der 9./10. Klassen, die ihre berufsbezogene Englischkompetenz mit einem offiziellen Zertifikat nachweisen möchten.

Die Entwicklung einer rein digitalen Prüfung stellt Testentwickler vor neue Fragestellungen, weshalb im Pilotprojekt das testmethodische Qualitätsmanagement konsequent mit der aus dem Innovationsmanagement stammenden Methode *Design Thinking* verzahnt wurde. Die Vorgehensweise zur empirischen Erhebung qualitativer und quantitativer Daten bzgl. Anforderungen der Zielgruppe, adäquaten Testinhalten, neuen Prüfungsprozessen und insbesondere zur Sicherung des Validitätsarguments der Prüfung soll im Vortrag erläutert werden. Ein Schwerpunkt ist dabei der mehrstufige Prozess der statistischen Datenerhebung für die testmethodische Bewertung (Standard Setting), in dem zugleich die Akzeptanz der Zielgruppe hinsichtlich des Prüfungsformats, -inhalts sowie der Usability überprüft wurde.

Die digitale, alltagsnähere Testumgebung erfasst Kompetenz- und Lebensbereiche der Lernenden, die in konventionellen Analogformaten bisher nicht oder unzureichend berücksichtigt werden konnten. So ermöglicht der testmethodisch validierte Einsatz multimedialer Mittel (Vlogs, Lerntutorials) sowie multimodaler Aufgaben (mixed skills-item) das Messen der Fertigkeiten Hör-/Sehverstehen und Mediation. Die berufsbezogenen Schreibaufgaben sind domainübergreifend und an der Lebenswirklichkeit der Lernenden orientiert. Sie ermöglichen zudem positive Effekte auf Arbeitsformen im Schulunterricht, wie das Verfassen von Messenger-Nachrichten im Rahmen von BYOD-Konzepten.

Im Vortrag soll zudem das Spannungsfeld zwischen Prüfen und Digitalisierung ausgelotet werden: Wichtige Aspekte sind dabei der Widerspruch zwischen Bedarf an digitalen Testangeboten und der Realität an Schulen sowie didaktische Aspekte der Lernerautonomie in einem digitalen Prüfungssetting. Auch die Frage, wie wichtig GER-Deskriptoren zur Online-Interaktion für eine *high stake*-Prüfung sind, stellt sich vor dem Hintergrund zukünftiger Entwicklungen.

**Dr. Jennifer Schluer (Kassel):**

### **Digitales Feedback mittels Screencasts in der Lehramtsausbildung**

Für das Lehren und Lernen fremder Sprachen sind insbesondere Feedback-Kompetenzen von grundlegender Bedeutung (Hattie, 2009). Jedoch stellt gerade die Bewertung komplexer fremdsprachiger Texte eine Herausforderung für Lehrende dar: Es müssen nicht nur verschiedene Ebenen von Form, Sprache und Inhalt adressiert, sondern auch gegenüber den Lernenden adäquat kommuniziert werden, um den weiteren Lernprozess zu unterstützen (gemäß Vygotsky's ZPD, 1978 u.a.). Dieser potenziellen Mehrschichtigkeit des Feedbacks sind in (hand-)schriftlicher Form jedoch Grenzen gesetzt. Digitale Formen des Feedbacks, wie bspw. über Screencasts, hingegen bieten besondere Vorteile für die Lehrenden und Lernenden: Sie erhöhen nicht nur die Eindeutigkeit, Tiefe und Transparenz des Feedbacks, sondern sprechen durch ihre Multimodalität auch verschiedene Lerntypen an und wirken motivierend (Ali, 2016; Séror, 2012; Silva, 2012; West & Turner, 2016).

Es wurde daher ein Projekt initiiert, in dem Lehramtsstudierende des Fachs Englisch Textprodukte kriteriengeleitet bewerten und Feedback-Kompetenzen auf innovative Weise anwenden: Die Studierenden tauschten untereinander Textprodukte aus, kommentierten und redigierten diese elektronisch in einem Textverarbeitungsprogramm und erstellten anschließend einen Screencast. Darin kommentieren die Studierenden das Textprodukt mündlich, während sie durch das elektronische Dokument navigierten und die positiven und verbesserungswürdigen Aspekte des Textproduktes

erläuterten. Anschließend bereiteten sie das Bildschirmvideo in einem Videobearbeitungsprogramm didaktisch auf, um die Kernaspekte hervorzuheben und eine abschließende Evaluation vorzunehmen.

Die Auswertung der Screencasts sowie der Fragebögen im Pre-Post-Design zeigt, dass Screencast-Feedback eine außerordentlich sinnvolle Ergänzung zu analogen Feedback-Formen ist, da es die fachlichen, didaktischen, medialen, interpersonalen, akademischen und sprachlichen Kompetenzen der zukünftigen Lehrenden und ihrer Lernenden steigern kann. In dem Vortrag werden die wesentlichen Ergebnisse zusammengetragen und Perspektiven für die weitere Forschung und die praktische Anwendung aufgezeigt.

### **Dr. Stéfanie Witzigmann (Heidelberg):**

#### **Videovignetten als diagnostisches Instrument zur Beurteilung mündlicher Sprachproduktion im Französischen**

Die diagnostische Kompetenz von Fremdsprachenlehrkräften umfasst unter anderem die Fähigkeit, mündliche Produktionen zu beurteilen. Diese sind aufgrund ihrer Komplexität allerdings schwierig einzuschätzen, da das Sprechen ein sehr schnell ablaufender Prozess ist (Müller-Hartmann & Schocker-von Ditfurth, 2004), die Reziprozität in einer Interaktion nur bedingt planbar ist und gesprochene Wörter kurzlebig sind (Bygate, 1987). Zudem stellt sich bei der Beurteilung mündlicher Produktionen verstärkt das Problem der Objektivierbarkeit (Luoma, 2004; Nieweler, 2010). Um dennoch einen detaillierteren Einblick in den Lernprozess von Lehrpersonen beim Beurteilen von mündlichen Produktionen zu erlangen, liegt diesem Beitrag eine Studie zu Grunde, in der authentische Videovignetten zur Datenerhebung genutzt wurden. Diese Videovignetten sind bewusst an der Schnittstelle zwischen Grundschule und Gymnasium verortet, damit die Lernenden sowohl von Primar- als auch von Sekundarschullehrkräften beurteilt werden können. In einer ersten Teilstudie wurden die Videovignetten zunächst von Fachdidaktikern (Hochschulprofessor\*innen und Hochschuldozent\*innen der Zielsprache Französisch) bewertet. In der Hauptstudie werden die Videovignetten dann zu gleichen Teilen von erfahrenen Französischlehrkräften und Lehramtsstudierenden im Grundschul- sowie Sekundarbereich beurteilt.

Mündliche Produktionen lassen sich auf inhaltlichen und sprachlichen Ebenen (z.B. Aussprache, Umfang und Differenziertheit des Wortschatzes, grammatikalische Korrektheit aber auch pragmatische oder inhaltsbezogene Elemente) beschreiben. Diese inhaltlichen und sprachlichen Merkmale (Hinweisreize) fließen in unterschiedlichem Maße in die Beurteilung durch die Lehrpersonen mit ein. Der Einfluss variiert dabei bedingt durch die individuellen Merkmale der beurteilenden Lehrperson (u. a. Fachwissen, Lehramtstyp).

Welche Hinweisreize von den verschiedenen Personengruppen bei der Betrachtung von Videovignetten wahrgenommen werden und wie diese Information im Sinne des Brunswik'schen Linsenmodells in einem Gesamturteil gewichtet wird, soll in diesem Vortrag vorgestellt werden.

### **Dr. phil Gabriele Kecker / Sonja Zimmermann (TestDaF):**

**“Zusammenfassen habe ich noch nie gemacht” – Überlegungen zur Prüfungsvorbereitung auf den digitalen TestDaF**

Seit seiner Einführung 2001 hat sich der TestDaF als standardisierter Sprachtest für die sprachliche Hochschulzulassung etabliert (Norris & Drackert 2018). Zur Validierung eines solchen *high-stakes* Tests gehören neben standardmäßig durchgeführten Maßnahmen zur Qualitätssicherung auch die fortlaufende Überprüfung der Konstruktvalidität, also die Frage danach, inwieweit das zu messende Merkmal im Test weiterhin mit den Anforderungen der Realsituation übereinstimmt. Eine zu diesem Zweck durchgeführte Bedarfsanalyse (vgl. Arras 2012; Marks 2015) hat gezeigt, dass der TestDaF in seinem bisherigen Format wichtige sprachliche Anforderungen abbildet, die für die Aufnahme eines Hochschulstudiums erforderlich sind. Gleichzeitig wurde deutlich, dass ein Sprachtest für den Hochschulzugang stärker Verarbeitungsprozesse berücksichtigen sollte, die mehrere Kompetenzen miteinander verbinden. Ausgehend von diesen Überlegungen wurden im TestDaF-Institut daher neben Aufgaben, die weiterhin die Teilkompetenzen weitgehend isoliert erfassen, neue kompetenzübergreifende Formate (Knoch & Sitajalabhorn 2013) entwickelt. Die Umsetzung der Aufgaben erfolgte dabei computerbasiert, um die Einbindung von Medien wie Videos und Audios zu ermöglichen. Bei ersten Erprobungen mit internationalen Teilnehmenden zeigte sich, dass sie ihre bisherigen Strategien zur Bearbeitung von Prüfungsaufgaben nicht auf die neuen Aufgabenformate übertragen konnten. Auch das Medium Computer stellte für viele eine Schwierigkeit bei der Bearbeitung der Aufgaben dar.

Der Beitrag stellt ein Konzept für die Vorbereitung von Teilnehmenden auf den digitalen TestDaF zur Diskussion, das deren Kompetenzentwicklung in den Mittelpunkt rückt. Dazu werden nach einem kurzen Überblick über das neue Prüfungsformat zentrale Kompetenzen für sprachliches Handeln an der Hochschule beschrieben und erläutert, wie sich diese in einzelnen Prüfungsaufgaben widerspiegeln. Zudem werden qualitative Rückmeldungen von Teilnehmenden aus Fragebogen und Interviews sowie von Beurteilern zu den mündlichen und schriftlichen Leistungen aus der Erprobungsphase präsentiert.

**Malgorzata Barras (Freiburg), Katharina Karges (Freiburg), Peter Lenz (Freiburg):**

**„also es ist etwas das man auch im Leben sich identifizieren kann“ – Zur Wirkung digitaler Textsorten in einem computerbasierten Fremdsprachentest**

Spätestens mit der Verbreitung von Tablets und Smartphones haben Lese- und Hörverstehen neue Dimensionen gewonnen. Gedruckte Lesetexte und flüchtige Hörtexte weichen im Alltag digitalen Textsorten. Smartphones machen das (Nach-)Hören immer einfacher und auch die Natur des Leseverstehens verändert sich durch Hypertext, Suchmaschinen, 280 Zeichen lange Tweets, Chats u. Ä. zunehmend. Im Alltag vieler Schüler/innen längst angekommen, wirken sich die genannten Veränderungen auch auf den schulischen Fremdsprachenunterricht und damit auf die Beurteilung sprachlicher Kompetenzen aus. Dieser Kontext war Ausgangspunkt für das Forschungsprojekt "Innovative Formen der Beurteilung", in dem computerbasierte Lese- und Hörverstehensaufgaben für Französisch und Englisch als Fremdsprachen entwickelt wurden. Die Aufgaben basieren auf digitalen Textsorten und sind in quasi-authentische Szenarios (Sabatini et al., 2014) eingebunden. Beispiele für Szenarios sind das Planen und Organisieren einer Reise oder das Sammeln von Informationen für eine Projektarbeit. Durch ihre Nähe zu realweltlichen Aufgabenstellungen sollen diese szenariobasierten Aufgaben u. a. auch die Motivation der Lernenden erhöhen.

Um das Funktionieren der Szenarios und Aufgaben in einem Large-Scale-Kontext zu untersuchen, wurde ein Mixed-Methods-Design verwendet. Die qualitativen Daten wurden durch

Lautdenkprotokolle und retrospektive Interviews mit Schüler/innen erhoben. Die quantitativen Daten umfassen hauptsächlich die Testresultate aus den szenariobasierten Aufgaben, aus „klassischen“ Lese- und Hörverstehensaufgaben sowie aus Wortschatz- und Grammatiktests. Zusätzlich wurden Fragebögen zur Motivation, zu Strategie- und Computereinsatz sowie zum sozialen und sprachlichen Hintergrund eingesetzt. In unserem Beitrag möchten wir aufzeigen, wie die Lernenden auf die digitalen Textsorten reagierten, welche Rolle die Szenarios dabei spielten und inwiefern sich die computerbasierten Testaufgaben auf die Motivation der Schülerinnen und Schüler auswirkten.

**Almut Ketzer-Nöltge, M.A. (Leipzig):**

**Führt ein fremder Akzent zur Fehleinschätzung sprachlicher Kompetenzen? – Eine Onlinestudie mit Lehrenden und Laien**

Zur Strukturierung und Vereinfachung unserer Umwelt kategorisieren wir Menschen und Gegenstände. Eine mögliche soziale Kategorie ist „Herkunft“, deren Zuordnung durch verschiedene Reize ausgelöst werden kann, z. B. „Aussehen“ oder „Spracheigenschaften“. Die vorliegende Studie bedient sich einer digitalen Erhebungsmethode, um zu untersuchen, inwiefern sich ein fremder Akzent auf die Bewertung der Grammatikbeherrschung in der Fremdsprache Deutsch auswirkt und somit als Hinweisreiz im Kategorisierungsprozess fungiert. Im Rahmen einer Onlinestudie beurteilten 45 Laien mit wenig Kontakt und 65 (DaZ)-Lehrende mit regelmäßigem Kontakt zu nichtmuttersprachlichen SprecherInnen die Grammatikbeherrschung von 20 DAZ-SprecherInnen mit der Muttersprache Russisch. Dafür wurden den Befragten jeweils zehn Audioaufnahmen und zehn Transkripte von Audioaufnahmen vorgelegt, bei welchen das Sprachniveau der SprecherInnen kontrolliert wurde und die jeweils zur Hälfte von SprecherInnen mit starkem oder schwachem Akzent stammten.

In einer Varianzanalyse mit den Faktoren Akzentstärke (stark vs. schwach) und Modalität (Audio vs. Transkript) wurde geprüft, inwiefern der Akzent einen Einfluss auf die Grammatikbewertung hat. Die Ergebnisse zeigen einen signifikanten Einfluss der Akzentstärke auf die Bewertung: Audioaufnahmen von SprecherInnen mit starkem Akzent wurden als schlechter bewertet als jene mit schwachem Akzent. Zudem wurde die Grammatikbeherrschung bei Transkripten als schlechter bewertet als bei Audioaufnahmen. Wie erwartet zeigte eine Interaktion der Faktoren Akzentstärke und Modalität in beiden Gruppen (Laien und Lehrende), dass die Einschätzung der Grammatikbeherrschung durch einen Akzent negativ verzerrt wurde. Gleichzeitig zeigte eine Gesamtanalyse, dass die Einschätzung der Lehrenden weniger stark durch den Faktor Akzent beeinflusst wurde. Konsequenzen für die Beurteilung sprachlicher und schulischer Leistungen sowie das Potential digitaler Befragungen für die Prüfung bestimmter Einflussfaktoren auf die Bewertung von Testleistungen, sollen im Beitrag diskutiert werden.

**Simone Stuers (Hagen):**

**Experiences of English language entry requirements at German higher education institutions:  
Today virtual learning, virtual environments and digitalization are burning issues among scholars  
and/or economy**

But no matter how digital we get, or how virtual we become, the very core of learning languages is a process which can be facilitated by computers, but not taken away from a human. Acquiring

knowledge in a foreign language will be essential for future generations due to globalization, technology and the internet. In Germany more and more students decide to pursue their study programme in English – these are not necessarily English teachers but also other disciplines. Nevertheless there must be a reliable measurement of linguistic proficiency to allow prospective students enter academia. Focusing on study programmes in English medium instruction, my presentation will discuss English language entry requirements at German higher education institutions. Since the first international programme was launched in Germany in 1997/98, the number of English-medium and bilingual programmes offered at German universities and universities of applied sciences has risen significantly. While a number of studies on English-medium instruction have been carried out, little has been written about the English language entry requirements that are used in this context. In Germany the parameters for entry requirements are set by the federal government and, institutions have to put them into practice; however there is room for improvement with regard to setting appropriate English language entry requirements. The study used, among other methods, a website search of 426 institutions of higher education in Germany. Institutional homepages were scrutinized for English-medium programmes and their English language entry requirement. The website search suggested that at both bachelor's and master's level, the majority of English-medium programmes do specify English language entry requirements; however, it is noteworthy that in many cases none are explicitly stated. In terms of cut scores, significant variations were found, but the majority equates to B2 (CEFR) for students to start their study programme with English as a medium of instruction.