

Inga Wagner

Fremdsprachliches Leseverstehen anhand von Mehrfachwahlaufgaben testen?

Erfahrungen an der Universität Modena und Reggio Emilia (Italien)

1. Einführung

Am Sprachenzentrum der Universität Modena und Reggio Emilia wurde 2009 mit der Entwicklung eines neuen webbasierten Testprogramms für die Sprachen Englisch, Deutsch, Französisch, Spanisch und Italienisch begonnen, das sich am Gemeinsamen Europäischen Referenzrahmen (GER) orientiert und nach Testfunktion flexibel einsetzbar ist.

Obwohl das Sprachenzentrum seit 2003 computerbasierte Sprachtests für alle Fakultäten durchführt, vornehmlich in Form von englischen Einstufungstests, wurde ein neuer Test notwendig, um dem steigenden und differenzierter werdenden Testbedarf von Seiten der Fakultäten gerecht zu werden.

Es handelt sich um eine Webanwendung zur Itemerstellung, zum Itembanking und Testmanagement, deren Funktionen alle eigenständig verwaltet werden können. Dies bietet die Möglichkeit, unterschiedliche Tests modular und flexibel zusammenzustellen und auch per Computer und von Hand korrigierte Testteile (schriftlicher Ausdruck) zu kombinieren. Je nach Testfunktion wurden Testformate (*test specifications*) für einen Einstufungs-, Aufnahme- und Feststellungstest entwickelt, weitere geplante Testziele sind ein fachspezifischer Englisch-Abschlusstest für einzelne Fakultäten und die Verwendung in der Selbstevaluation. Auch können in Zusammenarbeit mit den Sprachlehrbeauftragten nach Bedarf Lernfortschritts- und Kursabschlusstests erstellt werden.

Weitere Gründe für die Entwicklung eines neuen Testsystems war die Notwendigkeit, die Fertigkeiten Hörverstehen und schriftlichen Ausdruck zu tes-

ten und verstärkt Aspekte der kommunikativen Sprachkompetenz (z.B. Pragmatik) einzubeziehen sowie unterschiedliche Aufgabentypen zu verwenden, um einen Testmethodeneffekt, also das Ausmaß, in dem die Aufgabentechnik selbst das Testergebnis beeinflusst, zu verringern (Alderson et al. 1995: 44). Außerdem wurde aus Gründen der Zeitökonomie insbesondere der Einstufungstest teilweise adaptiv gestaltet, d.h. je nach in der ersten Testbatterie (Screeningtest) erzieltm Ergebnis werden die folgenden Testbatterien aufgerufen.

2. Leseverstehen (testen)

„The number of theories of reading is simply overwhelming: what it is, how it is acquired and taught, how reading in a second language differs from reading in a first language, how reading relates to other cognitive and perceptual abilities, how it interfaces with memory [...].“ (Alderson 2000: 1)

Es gibt eine beeindruckende Anzahl von Theorien über das Lesen und darüber, welche unterschiedlichen Faktoren und Fertigkeiten am Leseverstehen (LV) als Prozess und Produkt beteiligt sind. Die meisten stimmen darin überein, dass es sich um eine komplexe und in hohem Maße individuelle Tätigkeit handelt, die ihrerseits von einer Vielzahl variierender, rezipientenspezifischer Faktoren abhängig ist, wobei man zwischen leserspezifischen Faktoren wie Motivation, Leseziel, Sprach-, Hintergrund- und metakognitives Wissen, etc. und textspezifischen Faktoren, wie beispielsweise Thema und Textsorte, unterscheiden kann (Grotjahn 2001: 13).

Weit verbreitet im Zusammenhang mit Leseverstehen ist außerdem die Unterscheidung unterschiedlicher Lesestile, z.B. in globales, sortierendes, suchendes und detailliertes Lesen (Westhoff 1997: 101), welche in engem Zusammenhang mit der Textsorte und dem Leseziel stehen.

Auch wenn eine strikte Isolierung verschiedener Lesefertigkeiten wahrscheinlich nicht möglich ist (Alderson 2000: 10), da Lesen als globaler Prozess mehrere miteinander vernetzte Fertigkeiten in unterschiedlichem Ausmaß gleichzeitig impliziert, spielt die Unterscheidung verschiedener Teilkomponenten beim Leseprozess eine zentrale Rolle für die Lesedidaktik und das Testen von LV und bildet auch in unserem Fall die Basis für das Testkonstrukt.

Dass die Leistung nicht direkt beobachtbar ist und nicht, wie z.B. beim Schreiben, anhand eines Produktes evaluiert werden kann, ist ein weiterer zentraler Aspekt des Testens von Leseverstehen, welcher dazu führt, dass man

7. Fazit

Der Bedarf nach Sprachtests und -prüfungen unterschiedlichster Art steigt, nicht nur an Universitäten, und der Aufgabentyp Mehrfachwahlaufgabe wird aus diesem immer dichter werdenden Dickicht des Testalltags nicht verschwinden. Dieser Beitrag möchte einen Anstoß dazu geben, den Einsatz von Mehrfachwahlaufgaben zum Testen von Leseverstehen differenzierter zu betrachten. MC erscheint eine geeignete Aufgabentechnik in selten durchgeführten Tests mit einer hohen Anzahl von Kandidaten, wenn eine Wiederverwendung der Items, z.B. durch Speicherung in einer Itembank, gewährleistet ist (Hughes 2003: 59).

Die Verwendung von MC-Aufgaben zum Testen von Leseverstehen beschränkt sich jedoch nicht auf standardisierte *high-stakes* Tests oder solche mit großen Teilnehmerzahlen, sondern zieht sich durch die unterschiedlichsten Testarten und Prüfungskontexte hindurch. Oft erscheint Multiple Choice die offensichtlichste, technisch ‚einfachste‘ Art des Testens, besonders wenn dies am Computer stattfindet. Dennoch sollte bedacht werden, dass gute und benutzbare Items sogar für professionelle Itemautoren alles andere als einfach zu erstellen sind und erst nach der Vorerprobung und statistischen Analyse erkennbar sind (Hughes 2003: 58), wie auch unsere Erfahrung gezeigt hat.

Umso kritischer sollte deshalb hinterfragt werden, ob MC für das jeweilige Testziel und -konstrukt, das Medium, die Anzahl der Kandidaten, den zu erwartenden *washback effect* und das zu testende Niveau eine angemessene Aufgabentechnik zum Testen von Leseverstehen darstellen kann. Entscheidet man sich für den Einsatz von MC-Items, muss ein hoher Zeitaufwand bei der Erstellung, Überarbeitung und Erprobung eingeplant werden, der auch unter testökonomischen Gesichtspunkten gerechtfertigt sein sollte.

Literaturverzeichnis

- Alderson, J.C. (2000): *Assessing Reading*, Cambridge: Cambridge University Press.
- Alderson, J.C./Clapham, C./Wall, D. (1995): *Language Test Construction and Evaluation*, Cambridge: Cambridge University Press.
- ALTE (2007): Materials for the guidance of test item writers. ALTE – The Association of language Testers in Europe. http://www.alte.org/attachments/files/item_writer_guidelines.pdf (12.09.2012).

- Arras, U. (2006): Testen und Beurteilen des Leseverstehens in der Fremdsprache, in *Babylonia*, 14(3–4), 81–86.
- Bolton, S. (1996): *Probleme der Leistungsmessung*, München: Langenscheidt.
- Grotjahn, R. (2001): *Leistungsmessung und Leistungsbeurteilung. Band C: Kapitel 4: Testen der Fertigkeit Leseverstehen*, Hellenic Open University Patras, Postgraduiertenstudium Deutsch als Fremdsprache.
- Haldayna, T.M/Downing S.M/Rodriguez M.C. (2002): A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment, in: *Applied Measurement in Education*, 15 (3), 309–334.
- Hughes, A. (2003): *Testing for language teachers (2nd ed.)*, Cambridge: Cambridge University Press.
- Rupp, A.A./Ferne, T./Choi, H. (2006): How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective, in: *Language Testing*, 23, 441–474.
- Westhoff, G.(1997): *Fertigkeit Lesen*, München: Langenscheidt.

Anhang

SPECIFICATIONS FOR WRITING MC-READING COMPREHENSION ITEMS

Formal aspects:

- phrase options and stems positively; avoid negatives such as NOT or EXCEPT; if used in the stem use cautiously and always ensure that word appears capitalized and bold
- reduce the reading load
- language should be clear and not more difficult than the text
- put the central idea in the stem (don't repeat it in the options)

Distractors:

- make all distractors plausible (while still being incorrect)
- if possible, use typical errors of students to write your distractors
- avoid giving clues to the right answer, such as
 - specific determiners including *always*, *never completely*, and *absolutely*
 - overlapping choices
 - 'word spotting'
 - grammatical clues
 - long correct responses
 - convergence strategy