

Beurteilen versus Zählen [Korn]

Korn, Matthias (2020). Beurteilen versus Zählen. Mutmachende Perspektiven für die Bewertung schriftlicher Übersetzungsleistungen in den alten Sprachen*. Ars docendi, 5, dicembre 2020.

Matthias Korn (università di Lipsia) ci parla di un problema assai attuale nella didattica delle lingue antiche nella scuola: del problema della valutazione adeguata dei compiti in classe, soprattutto della valutazione adeguata della traduzione: Come si riesce a valutare meglio e in un modo abbastanza corretto gli errori (talvolta anche molto differenti) di grammatica nella traduzione?

1) Der Rahmen

Schulrechtliche Vorgaben:

Feste Vorgaben zur Leistungsbewertung in der Form einer verbindlichen Fehler-Noten-Relation gibt es in den Ländern der BR Deutschland nur im Abiturprüfungsbereich. Falls darüber hinaus Vorgaben erfolgen, sind sie abhängig vom jeweiligen Landesrecht und dann im jeweiligen Kerncurriculum bzw. Kernlehrplan festgelegt.¹

Es gibt aber durchaus auch Länder ohne derartige Vorgaben.² Falls keine curricularen Vorgaben bestehen, dann sind Verabredungen der Fachkonferenz / des Fachbereichs z.B. zu Fehlergewichtung, Korrekturzeichen und Bewertung maßgeblich. Beim Beispiel Sachsen erfolgt dann häufig der Rückgriff auf eine abiturbezogene Landesverwaltungsvorschrift, die sich an den Fehlergrenzen der Einheitlichen Prüfungsanforderungen in der Abiturprüfung (EPA) Latein orientiert.

Über allem aber stehen in Deutschland die allgemeinen Notendefinitionen laut Beschluss der Kultusministerkonferenz (KMK) von 1968: Zum Beispiel

- gut (2): Die Note „gut“ soll erteilt werden, wenn die Leistung den Anforderungen voll entspricht.
- ausreichend (4): Die Note „ausreichend“ soll erteilt werden, wenn die Leistung zwar Mängel aufweist, aber im Ganzen den Anforderungen noch entspricht.

Doch der Begriff der Anforderung ist gerade für die Übersetzungsleistung bislang nicht hinlänglich objektiviert bzw. nicht intersubjektiv nachvollziehbar definiert. Damit liegt er im Ermessensspielraum jeder einzelnen Lehrkraft, was die Sache nicht erleichtert.

* Überarbeitete Fassung eines Impulsreferats, das am 12. September 2020 in einem Webinar des Landesverbandes Nordrhein-Westfalen im Deutschen Altphilologenverband gehalten wurde.

¹ Beispiel: Nordrhein-Westfalen

² Beispiel: Sachsen

Schulfachliche Vorgaben:

Bei der Leistungsbewertung erfolgt in Deutschland noch häufig die Orientierung an der 10-Prozent-Marke, ganz oft auch in modifizierter Form, wie es von Scholz und Weber propagiert wird. Diese Marke selbst kommt aus den EPA, hat aber über das Abitur hinaus allenfalls Relevanz für die gymnasiale Oberstufe. Und gerade die Oberstufe spielt angesichts der Entwicklung der Zahl lateinlernender Schüler*innen im bundesdeutschen Durchschnitt eine immer geringer werdende Rolle, weil hier in der Regel nur Einbringungspflicht für *eine* fortgeführte Fremdsprache besteht, deren am häufigsten gewählte natürlich Englisch ist. Außerdem ist diese Marke fachdidaktisch gesehen völlig willkürlich gesetzt. Dazu kommt, dass die Lehrkräfte Fehler ganz unterschiedlich definieren und gewichten, d.h. Fehler ist hier lange nicht gleich Fehler. Und damit sind wir mittendrin im Schlamassel.

2) Grundsätze und Dilemmata der Leistungsbewertung

Grundsätze:

- Leitkriterium: Textkompetenz, genauer Textverständnis,
- Bewertungskriterium: Grad der Sinnentstellung,
- Messeinheit: Satz oder Teilsatz,
- Maßstab für die Bewertung (im Folgenden auch ‚Messverfahren‘ genannt):
 - Zählverfahren: Gezählt werden die Fehler,
 - Beurteilungsverfahren: Beurteilt wird der Grad der Sinnentsprechung;
 - Hybridverfahren: Verbindung von Zähl- und Beurteilungsverfahren.

Dilemmata:

Diese sind – je nach Verfahren – ganz unterschiedlich gelagert:

Beim Fehlerzählverfahren sind es die Uneinheitlichkeit in der Frage, was als Fehler verstanden wird, die Unterschiedlichkeit in der Fehlergewichtung (halb, ganz, doppelt) und die Arbitrarität in der Frage des Verhältnisses von Fehlerzahl und Anforderung (z.B. Frage der 10%-Marke).

Beim Verfahren der Beurteilung von Sinnentsprechung sind es die Fragen, welcher Äquivalenzbegriff der Sinnentsprechung zugrunde liegt, wie Sinnentsprechung zuverlässig und intersubjektiv nachvollziehbar gemessen werden kann und wie das Messergebnis zuverlässig und intersubjektiv nachvollziehbar bewertet werden kann.

3) Ein fachdidaktisches Zwischenresümee

In der aktuellen Schulpraxis scheint in Deutschland ganz überwiegend die konventionelle, Fehler zählende Bewertung auf der Grundlage der Negativkorrektur angewandt zu werden. Daneben gibt es – freilich seltener – die Grundlage der Positivkorrektur z.B. nach Clasen, die von einem – wie auch immer gebildeten – Punkte,konto‘ ausgeht, von dem bei den unterschiedlichen Fehlerarten unterschiedliche Punktzahlen ‚abgebucht‘ werden. Beide

Korrekturweisen fokussieren somit letztlich lediglich Fehler, was infolge des o.g. Dilemmas zu erheblichen Verzerrungen führen kann. Diese Verzerrungen schaden der Reputation des Faches sehr, weil sie u.a. zu ganz unterschiedlichen Bewertungen der Schülerleistungen führen – aus meiner Sicht ein Nagel für den Sarg des Lateinunterrichts in Deutschland.

4) Drei alternative Bewertungsweisen

Im Folgenden sollen drei alternative Bewertungsweisen neueren Datums vorgestellt und eingeschätzt werden, deren Bestreben dahin geht, vom (rein quantifizierenden) Fehlerzählen weg zu einem (qualifizierenden) Beurteilungsverfahren hin zu kommen.

Das sogenannte Duisburger Modell ist ein Korrektur- und Bewertungskonzept des Landesverbandes (LV) Nordrhein-Westfalen im Deutschen Altphilologenverband (DAV), das bisher nur im Regierungsbezirk Düsseldorf amtlich zugelassen ist.

Das Verfahren der österreichischen ‚Standardisierten kompetenzorientierten schriftlichen Reifeprüfung‘ (SKRP) wird seit 2015 in der Matura für die Bewertung der Übersetzungsleistung aus den alten Sprachen angewandt und liegt inzwischen auch in allen vorangehenden Klassenstufen der AHS der Leistungsbewertung zugrunde.

Und schließlich die Weise der Bewertung der Übersetzungsleistung in der Abschlussprüfung zum Erwerb des ‚International Baccalaureate Diploma‘ (IBD) in den Fächern Latein und Griechisch, die an ‚International Schools‘ und ‚IB-World Schools‘ zur Anwendung kommt.

5) Das ‚Duisburger Modell‘

In einem ersten Arbeitsdurchgang werden die Schülerübersetzungen im Hinblick auf das Textverständnis qualifizierend den Stufen A (Notenbereich 1-2), B (3-4) und C (5-6) zugeordnet. Dann erfolgt eine konventionelle Negativkorrektur. Dabei werden Folge- und Wiederholungsfehler herausgerechnet; bei Fehlernestern zählt nur der ursächlich erste Fehler; außerdem gibt es pro Wort maximal einen halben Fehlerwert. Sinnkonforme Abweichungen zählen nicht als Fehler. Besondere Leistungen im Bereich der Modulation und Transposition können zur Verbesserung um eine Notenstufe führen.

Im Ganzen erscheint das ‚Duisburger Modell‘ aber noch nicht mehr als ein ‚verkleidetes‘ Fehlerzählverfahren zu sein, weil der qualifizierende, erste Durchgang systemisch und damit messtechnisch zu wenig elaboriert / untersetzt ist und dessen Ergebnis eigentlich ganz ohne Konsequenzen für die Leistungsbewertung bleibt. Positiv einzuschätzen sind aber die Bestrebungen, die Eskalation von Fehlerzahlen zu vermeiden und gelungene Lösungen bei der Bewertung zu berücksichtigen. Außerdem ist das ‚Duisburger Modell‘ textlinguistisch und übersetzungswissenschaftlich solide und modern fundiert. So wird z.B. als Bewertungskriterium die denotative Äquivalenz nach Koller / Henjum gewählt. Damit sind Streitereien bei Umstellungen auf der grammatischen Ebene (Transpositionen, z.B. die Übersetzung ‚Die Verschwörer haben Caesar getötet‘ für ‚Caesar a coniuratis interfectus est‘) ein für alle Mal vom Tisch.

Alles in allem bleibt aber das ‚Duisburger Modell‘ aus meiner fachdidaktischen Sicht derzeit noch hinter den Möglichkeiten zurück, die der eingeschlagene, richtige Weg der Ansetzung von denotativer Äquivalenz als Bewertungskriterium eröffnet. Mit Blick u.a. darauf hat sich im LV Nordrhein-Westfalen des DAV eine Arbeitsgruppe gebildet, die das ‚Duisburger Modell‘ weiterentwickeln will.

6) Das österreichische Bewertungsmodell

..... legt der Bewertung der Übersetzungsleistung eine Mischkalkulation zugrunde:

- a) 33,3 % (12/36) zählt die standardisierte Beurteilung der Sinnerfassung des fremdsprachigen Textes auf der Grundlage einer beigegebenen Inhaltszusammenfassung. Der Text wird dafür in zwölf etwa gleichlange Sätze / Teilsätze / Kola zerlegt. Die Inhaltszusammenfassung gibt nur die wesentlichen Inhalte in knapper Form wieder und verzichtet auf periphere Details.
- b) 50 % (18/36) zählen an sogenannten Checkpoints die Verstöße gegen morphologische, syntaktische und lexikalische Richtigkeit auf der Grundlage einer Negativkorrektur der Schülerübersetzung.
- c) 16,6% (6/36) sind der ebenfalls standardisierten, dreistufigen Beurteilung der unterrichtssprachlichen Kompetenz der Schüler*innen vorbehalten und betreffen Morphologie, Syntax, Wortwahl und Ausdruck in der Unterrichtssprache.

Im Ganzen geht der österreichische Ansatz aus meiner Sicht in die richtige Richtung, aber die Anteile sind nicht der Bedeutung des Leitkriteriums ‚Textverständnis‘ entsprechend ponderiert: b) und c) sind zu hoch, a) zu niedrig angesetzt.

7) Das Bewertungsmodell für das ‚International Baccalaureate‘

..... legt der Bewertung der Übersetzungsleistung ebenfalls eine Mischkalkulation zugrunde:

50% (90/180) zählen die Verstöße gegen die grammatische und lexikalische Richtigkeit auf der Grundlage einer standardisierten, vierstufigen Beurteilung, weitere 50% (90/180) zählt die ebenfalls standardisierte Beurteilung der Sinnerfassung des fremdsprachigen Textes auf der Grundlage einer allgemeinen, ebenfalls vierstufigen Qualifizierung, die dem ‚Textverständnis-ABC‘ des ‚Duisburger Modells‘ ähnlich ist. Der Text wird dafür in 30 unterschiedlich lange Teilsätze / Kola zerlegt, deren Länge – abhängig vom sprachlichen Anforderungsniveau – zwischen 2 und 6 lateinischen Wörtern liegt.

Dieser Ansatz scheint mir im Ganzen am gründlichsten durchdacht zu sein. Es fragt sich allerdings, warum die beiden Bereiche überhaupt noch getrennt sind, zumal Verstöße vor allem gegen die lexikalische Richtigkeit ja in aller Regel einer vollständigen Sinnerfassung entgegenstehen.

8) Meine fachdidaktische Vision: Die zukünftige Bewertung der Übersetzungsleistung in Klassenarbeiten und Klausuren

Gestatten Sie, dass ich Ihnen zum Schluss dieser Ausführungen meine fachdidaktische Vision zur Bewertung der Übersetzungsleistung in Klassenarbeiten und Klausuren vorlege.

Um es gleich vorwegzunehmen: Die Zukunft hier hat längst begonnen. Das zeigt ein Blick auf den Anteil, den in den drei von mir oben vorgestellten Modellen die Bewertung der Übersetzungsleistung bei Klassenarbeiten und Klausuren an der Gesamtbewertung dieser Klassenarbeiten und Klausuren hat – hier werden nämlich bereits jetzt deutliche Unterschiede erkennbar:

- In Nordrhein-Westfalen auf der Grundlage des Kernlehrplans von 2019: 50% bis maximal 66,6%,
- in Österreich: 60%,
- beim ‚International Baccalaureate‘: 35%.

Aus meiner Sicht spiegelt die Gewichtung beim ‚International Baccalaureate‘ das wider, was mittel- und langfristig die Trias der fachbezogenen Strategien (Dekodierung, Interpretation und Übersetzung / Rekodierung) auch in Deutschland für sich beanspruchen muss, wenn die alten Sprachen in der Schule überleben wollen, nämlich eine Drittelparität, d.h. $D : I : \ddot{U} / R = 1 : 1 : 1$. Dabei könnten sich Dekodierung, Interpretation und das sprachbildende Potenzial von Übersetzung bzw. Rekodierung in paritätischer und einer die Schülermotivation fördernden Weise ergänzen, ohne dass Schüler*innen und Schüler, die nicht so gut übersetzen können, sich aber für die Antike und die alte Sprache interessieren, leistungsmäßig außen vor bleiben würden. Eine lediglich aus Tradition mitgeführte Dominanz der Übersetzung als Fachleistung wäre zwar gewiss Alleinstellungsmerkmal des altsprachlichen Unterrichts, aus dem genannten Grund aber eben auch ein – weiterer – Sargnagel.

In diesem Zusammenhang hat das Fehlerzählen als Messverfahren keinen Platz. Bewertungskriterium muss sein die Sinnentsprechung auf der Grundlage der denotativen Äquivalenz, Messverfahren muss sein die standardisierte Beurteilung des Grades der Sinnentsprechung.

Den Kritikern, die dabei die sprachbildende Kraft der Übersetzung bzw. Rekodierung verloren gehen sehen, sei entgegnet, dass diese mit ihrer mikroskopischen, jeden Fehler zählenden Arbeitsweise ja im Unterricht neben Dekodierung und Interpretation ihren Raum einnehmen kann, nicht aber auch in der Klassenarbeit bzw. Klausur finden muss, die ja ein unterrichtlicher Sonderfall ist – und schon gar nicht in dominierender Weise.

9) Abschließende Zusammenfassung der weiterführenden Perspektiven

Wollen wir, dass Lateinunterricht auch in der zweiten Hälfte dieses Jahrhunderts an deutschen Schulen eine nennenswerte Rolle spielt und vielleicht sogar in Deutschland und in Österreich aus Gründen der Bildungsgerechtigkeit an nicht-gymnasialen Schulformen bzw. auf nicht-gymnasiale Schulformen ausgeweitet wird, ist dafür eine Orientierung an den folgenden Perspektiven erforderlich:

- 1) Bei der Bewertung der Übersetzungsleistung in Klassenarbeiten und Klausuren wird künftig die Beurteilung der Sinnerfassung die maßgebliche Rolle spielen;
- 2) dieser Beurteilung wird das Kriterium der denotativen Äquivalenz zugrunde liegen;
- 3) die Beurteilung der Sinnerfassung kann durchaus in standardisierter Form erfolgen; trennschärfere intersubjektiv nachvollziehbare Messinstrumente wären aber natürlich wünschenswert;
- 4) gewichtungsmäßig wird sie sich an der Drittelparität der fachbezogenen Strategien orientieren;
- 5) wenn die Beurteilung der Sinnerfassung in standardisierter Form erfolgen kann, so kann dies auch für die Beurteilung der grammatischen und lexikalischen Richtigkeit gelten. Dann wären wir das Fehlerzählen ganz los – denn schließlich hat das ja auch mit Kompetenzorientierung eigentlich gar nichts zu tun.

10) Ausblick

Das Thema ‚Die Bewertung der Übersetzungsleistung in Klassenarbeiten und Klausuren‘ (mit anschließender Kleingruppenarbeit) wird eines von vier Schwerpunktthemen der 5., internationalen Fachtagung ‚Perspektiven für den Lateinunterricht‘ sein. Diese findet am 25. und 26. November 2021 an der Christlichen Schule Johanneum in Hoyerswerda/Sachsen statt. Die Leser*innen der ‚Ars docendi‘, die an näheren Informationen und eventuell an einer Teilnahme interessiert sind, werden uns im Interesse einer weiteren Entwicklung der länderverbindenden Zusammenarbeit herzlich willkommen sein und wenden sich bitte per Email an Matthias.Korn@uni-leipzig.de

11) Ausgewählte Literatur

Clasen, Adolf (1976): Probleme der Übersetzungsklausur, dargestellt an der Positivkorrektur. In: Bayer, Karl (Hrsg.) (1976): Leistungsmessung im altsprachlichen Unterricht. Auer: Donauwörth, S. 47-75.

Doepner, Thomas / Hesse, Godehard / Keip, Marina / Kurczyk, Stephanie (2017): Würdigung und Sinnverständnis. Kompetenzorientierte Übersetzungsbewertung nach dem Duisburger Modell. In: Der altsprachliche Unterricht Latein Griechisch, 60. Jahrgang, Heft 4+5, S. 60-69.

International Baccalaureate Organization (UK) Ltd (Hrsg.) (2014): Classical Languages Guide. First Examinations 2016. Selbstverlag: Cardiff. <im Internet erhältlich>

Koller, Werner / Henjum, Kjetil Berg (2020): Einführung in die Übersetzungswissenschaft. UTB, 9. grundl. überarb. u. aktual. Auflage.

Korn, Matthias (2020): Die Sicht eines Fachdidaktikers auf die aktuelle Stellung des Unterrichtsfachs Latein in Deutschland. In: cursor Latein4EU. Zeitschrift für Freunde der lateinischen Sprache und Kultur, Nr. 16 – August 2020, S. 16-18.

Lošek, Fritz / Niedermayr, Herrmann (2017): Die standardisierte kompetenzorientierte schriftliche Reifeprüfung in Österreich. Ein neues Modell für das Überprüfen von

Textverständnis. In: Der altsprachliche Unterricht Latein Griechisch, 60. Jahrgang, Heft 4+5, S. 98-104.

Scholz, Ingvelde / Weber, Christian (2015): Denn sie wissen, was sie können.

Kompetenzorientierte und differenzierte Leistungsbeurteilung im Lateinunterricht. V&R: Göttingen, 2., verbesserte Auflage.

Sekretariat der Kultusministerkonferenz (Hrsg.) (2005): Einheitliche

Prüfungsanforderungen in der Abiturprüfung Latein. Beschluss der Kultusministerkonferenz vom 01.02.1980 i.d.F. vom 10.02.2005. Selbstverlag: Berlin. <im Internet erhältlich>

Dr. Matthias Korn, Postanschrift: Universität Leipzig, Sprachenzentrum, Goethestraße 2, 04109 Leipzig, Deutschland