

Grundsätze des Korrektursystems zur schriftlichen Reifeprüfung in Griechisch und Latein¹

Walter FREINBICHLER – Peter GLATZ – Florian SCHAFFENRATH

Mit Beschluss des Österreichischen Nationalrates² sollten ab dem Schuljahr 2013/14 auch in den Fächern Griechisch und Latein³ die Aufgaben für die schriftliche Reifeprüfung österreichweit durch das Bundesministerium für Frauen und Unterricht vorgegeben werden. Bis dahin wurden die Maturaaufgaben von der jeweiligen Lehrperson einer Klasse nach allgemeinen Vorgaben zusammengestellt: Maßgeblich waren die einschlägigen Passagen des Schulunterrichtsgesetzes, der Reifeprüfungsverordnung⁴ sowie *mutatis mutandis* des „Consensus“, eines Leitfadens zur Leistungsbeurteilung im Griechisch- und Lateinunterricht im Allgemeinen.⁵ Die Einführung der neuen Reifeprüfung, der sogenannte Zentralmatura, wurde um ein Jahr verschoben. (Das Bundesgesetz zur Änderung des Schulunterrichtsgesetzes wurde am 02. August 2012 im Bundesgesetzblatt I Nr. 73/2012 veröffentlicht.)

Mit der Vorbereitung der standardisierten Reifeprüfung in Griechisch und Latein, die auch der an alle Fächer gestellten Anforderung der „Kompetenzorientierung“⁶ genügen sollte (daher der Name „standardisierte kompetenzorientierte schriftliche

¹ Die folgenden Gedanken beruhen auf einer Idee von Walter Freinbichler und Peter Glatz, die Florian Schaffenrath nach der intensiven Diskussion mit der Arbeitsgruppe zur Vorbereitung der standardisierten kompetenzorientierten schriftlichen Reifeprüfung in Latein und Griechisch (vgl. www.bifie.at/node/79) ausformuliert hat.

² Vgl. BGBl. Nr. 112/2009 vom 17. November 2009.

³ In § 42e des in Anm. 1 zitierten Gesetzes heißt es unter (2): „Die Aufgabenstellungen sind wie folgt zu bestimmen: [...] für die Prüfungsgebiete Deutsch, [...], (Lebende) Fremdsprache (Englisch, Französisch, Spanisch, Italienisch, Latein, Griechisch; in weiteren Sprachen nach Maßgabe einer Verordnung des zuständigen Bundesministers) und Mathematik (unter Berücksichtigung der jeweiligen lehrplanmäßigen Anforderungen) der Klausurprüfung (Klausurarbeiten und mündliche Kompensationsprüfungen) durch den zuständigen Bundesminister [...]“.

⁴ Die gesetzlichen Grundlagen (SchUG und RPVO) finden sich, soweit fachspezifisch relevant, mit Kommentar in: Rechtsgrundlagen und Leitlinien zur kompetenzorientierten Leistungsfeststellung und Leistungsbeurteilung in den klassischen Sprachen Latein und Griechisch, Stand: Juli 2011, 7–13 (einzusehen unter: www.bifie.at/node/529).

⁵ „Consensus. Leistungsbeurteilung in Latein und Griechisch. Eine Initiative der Sodalitas, Bundesarbeitsgemeinschaft für klassische Philologen in Österreich, in Zusammenarbeit mit den Landesarbeitsgemeinschaften und dem Landesschulrat für Niederösterreich“, Letztfassung: 19.5.2006. In diesem „Consensus“ finden sich keine konkreten Angaben zur schriftlichen Reifeprüfung, diese sind jedoch in Analogie aus den Angaben zu den Schularbeiten in der Lektürephase zu erschließen.

⁶ Zur Kompetenzorientierung in den klassischen Sprachen vgl. Scholz, Ingvalde / Weber, Karl-Christian: Denn sie wissen, was sie können. Kompetenzorientierte und differenzierte Leistungsbeurteilung im Lateinunterricht, Göttingen 2010; Kuhlmann, Peter: Kompetenzorientierung und Lateinunterricht in der Oberstufe, in: Forum Classicum 54 (2011), 114–123.

Reifeprüfung – SKR“),⁷ beauftragte das Bundesministerium für Frauen und Unterricht das Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (BIFIE).⁸ Dieses Institut übertrug die konkrete Entwicklung von Kompetenzmodellen, Klausuraufgaben und eines Korrektursystems, die Pilotierung und Feldtestung der vorläufigen Ergebnisse⁹ sowie die Vorbereitung eines österreichweiten Schulversuches¹⁰ im Jahr 2012/13 (d.h. ein Jahr vor der flächendeckenden Einführung des neuen Systems, nach damaligem Stand) an eine Arbeitsgruppe, die zunächst von Friedrich Lošek und seit September 2009 von Klassischen Philologen der Universität Innsbruck geleitet wird.

Die Arbeitsgruppe hat in einem ersten Schritt Kompetenzmodelle für Griechisch, für das vierjährige und für das sechsjährige Latein ausgearbeitet.¹¹ Grundsätzlich und übergreifend wurden in diesen Modellen einerseits das Übersetzen griechischer bzw. lateinischer Texte in die Unterrichtssprache und andererseits das Lösen von Arbeitsaufgaben zu griechischen bzw. lateinischen Texten als die Grundkompetenzen des Altsprachlichen Unterrichts festgeschrieben, die im Rahmen der schriftlichen Reifeprüfung standardisiert abgeprüft werden können. Der gesamte kulturkundliche Aspekt der Fächer musste ausgeblendet bleiben, da der seit September 2004¹² geltende modulare Lehrplan keine spezifischen Inhalte, seien sie kulturkundlicher, seien sie literaturgeschichtlicher Art, österreichweit zwingend vorschreibt. Die Behandlung dieser Aspekte des altsprachlichen Unterrichts muss im Rahmen der Reifeprüfung der mündlichen Prüfung¹³ oder der vorwissenschaftlichen Arbeit¹⁴ vorbehalten bleiben.¹⁵

⁷ Zur allgemeinen Entwicklung des Projektes vgl. Pinter, Anna: Standardisierung und Kompetenzorientierung in Österreich. Die neue schriftliche Reifeprüfung in den klassischen Sprachen, in: AU 54, 4+5 (2011), 116–121.

⁸ Auf der Homepage des BIFIE finden sich alle einschlägigen Dokumente zur Vorbereitung der standardisierten schriftlichen Reifeprüfung in Griechisch und Latein unter <https://www.bifie.at/node/79> (2.1.2012).

⁹ Vgl. Niedermayr, Hermann: Standardisierung und Kompetenzorientierung im österreichischen Lateinunterricht. Erste Erfahrungen und mögliche didaktische Folgerungen, in: Latein Forum 72 (2010), 1–19.

¹⁰ Zum Schulversuch vgl. Lošek, Fritz: Die neue schriftliche Reifeprüfung aus Latein und Griechisch im Schulversuch 2013, in: Circulare 2/2011, 3.

¹¹ Abgedruckt in: Rechtsgrundlagen und Leitlinien (wie Anm. 4), 14–28.

¹² Zum neuen Oberstufenlehrplan von 2004 vgl. Lošek, Fritz: Latein im 21. Jahrhundert. Ein Grenzgang zwischen ‚toter Sprache‘ und lebendigem Trendfach. Bilanz der Entwicklung in Österreich (mit Fokus auf den fächerverbindenden Unterricht), in: Rauscher, Erwin (Hg.): Unterricht als Dialog. Von der Verbindung der Fächer zur Verbindung der Menschen, Baden 2011, 213–238, bes. 218–224.

¹³ Vgl. Die kompetenzorientierte Reifeprüfung. Latein und Griechisch. Richtlinien und Beispiele für Themenpool und Prüfungsaufgaben (www.bmukk.gv.at/medienpool/21679/reifepruefung_ahs_lflatgr.pdf)

¹⁴ Vgl. www.ahs-vwa.at/.

¹⁵ Zur neuen mündlichen Reifeprüfung in Griechisch und Latein vgl. Lošek, Fritz / Schaffenrath, Florian: Die neue mündliche Reifeprüfung aus Latein und Griechisch. Ein Arbeitsbericht, in: Circulare 4/2011, 1–2.

Im Folgenden soll nach einer kurzen Vorstellung der neuen Aufgabenstellungen, die sich von der bisherigen Praxis augenfällig dadurch unterscheiden, dass nunmehr zwei Texte zur Bearbeitung vorgelegt werden (dazu gleich mehr), auch ein neues Korrektursystem¹⁶ präsentiert werden. Derzeit sieht es so aus, dass die Lehrperson einer Klasse die Aufgabenstellungen der schriftlichen Reifeprüfung zwar von zentraler Stelle vorgegeben bekommt, die Performanzen ihrer Kandidaten und Kandidatinnen dann jedoch selbst korrigieren muss. Da eines der Motive für die Einführung einer standardisierten Reifeprüfung in Österreich der Wunsch nach besserer Vergleichbarkeit der Leistungen der Kandidaten und Kandidatinnen war, wäre es nicht zielführend, die Aufgabenstellung zwar für alle einheitlich zu machen, bei der Korrektur aber dem einzelnen Lehrer/der einzelnen Lehrerin gänzlich freie Hand zu lassen.

1. Grundsätzliches

Die Leistungsmessung und Leistungsbeurteilung stellt einen wichtigen Bereich der Fachdidaktik zum altsprachlichen Unterricht dar.¹⁷

Derzeit werden in Österreich parallel verschiedene Korrektursysteme je nach Vorliebe des jeweiligen Lehrers/der jeweiligen Lehrerin verwendet: Zu nennen sind dabei die reine Negativkorrektur (sog. Fehlerzählmodell),¹⁸ die Negativkorrektur mit Vergabe von Bonuspunkten für besonders gelungene Leistungen,¹⁹ die Positivkorrektur²⁰ oder auch eine Kombination von Negativ- und Positivkorrektur bezogen auf die Übersetzungs- bzw. Interpretationsaufgabe²¹. Versuche haben gezeigt,²² dass die Anwendung verschiedener Korrekturmodelle auf dieselbe Performanz und v.a. die uneinheitliche Gewichtung von Fehlern zu ganz unterschiedlichen Bewertungen und Benotungen führen können. Dieser Unschärfe in der Vergleichbarkeit wird durch die Einführung eines neuen Korrektursystems, das

¹⁶ Eine erste auf Schularbeiten bezogene Vorstellung des Systems bietet Pinter, Anna / Schaffenrath, Florian: Schularbeiten und Korrektur nach den Prinzipien der neuen Reifeprüfung, in: *Circulare* 3/2010, 1–9.

¹⁷ Vgl. Fink, Gerhard / Maier, Friedrich: *Konkrete Fachdidaktik Latein – L2*, München 1996, 166–188; Kuhlmann, Peter: *Fachdidaktik Latein kompakt*, Göttingen 2009, 143–156.

¹⁸ Vgl. Bayer, Karl: Objektivierung der Leistungsmessung, dargestellt an der Negativ-Korrektur, in: *Anregung* 21 (1975), 95–101; Seitz, Wolfgang: Neues Punktesystem für Latein-Schularbeiten. Ein ARGE-Vorschlag zur Diskussion, in: *Apis Latina* 2 (1992), 2–3.

¹⁹ Vgl.: Mader, Walter: *Corrigo ergo sum?*, in: *Latein Forum* 19 (1993), 1–11, wieder abgedruckt in: *IANUS* 15 (1994), 42–49; Schwab, Helge: Problemlose Leistungsbeurteilung bei schriftlichen Arbeiten im Lateinunterricht, in: Pietsch, Wolfgang J. / Prochaska, Roman A. (Hg.): *Lateinische Texte verstehen und erleben*, Graz 1994, 114–116.

²⁰ Vgl. Consensus (wie Anm. 5), 9–13; Widhalm-Kupferschmid, Wilhelmine: Ein Modell der Positivkorrektur, in: *Circulare* 4/2004, 8–9.

²¹ Kritisch zur getrennten Beurteilung von Übersetzungs- und Interpretationsteil Siehs, Gottfried: Beurteilung von Schularbeiten, in: *Circulare* 4/2004, 7–8.

²² Vgl. Schwab (wie Anm. 20), 115; Oswald, Renate: Umfrage und Statistik zur Korrekturpraxis, in: *IANUS* 19 (1998), 53–58.

auf alle Performanzen der schriftlichen Matura in gleicher Weise anzuwenden ist, begegnet. Es erhebt nicht den Anspruch, objektiv zu sein oder das persönliche Ermessen des einzelnen Korrektors/der einzelnen Korrektorin ganz auszuschalten. Es erhebt aber sehr wohl den Anspruch, ein höheres Maß an *interrater*-Reliabilität (s.u.) als die bisher bekannten Systeme zu besitzen.²³

2. Aufgabenformate

Den Kandidaten und Kandidatinnen, die ab dem Schuljahr 2014/15 zur schriftlichen Reifeprüfung in Griechisch oder Latein antreten, werden zwei Texte vorgelegt, ein Übersetzungstext (ÜT), bei dem die Übersetzung bewertet wird, und ein Interpretationstext (IT), auf dessen Basis zehn Arbeitsaufgaben zu lösen sind, die bewertet werden. Im sechsjährigen Latein und in Griechisch umfasst der ÜT mindestens 120 Wörter, der IT mindestens 80 Wörter und beide Texte zusammen maximal 220 Wörter. Im vierjährigen Latein umfasst der ÜT mindestens 110 Wörter, der IT mindestens 80 Wörter und beide Texte zusammen maximal 210 Wörter. Die Gesamtarbeitszeit beträgt 270 Minuten. Die Benutzung eines (elektronischen) Wörterbuches ist erlaubt.²⁴ Modellbeispiele für schriftliche Reifeprüfungsaufgaben im neuen Format wurden auf der Homepage des BIFIE veröffentlicht.²⁵

3. Abgrenzung von anderen Systemen

Die Zielsetzung des neuen Korrektursystems ist nicht eindimensional. Sein erster Zweck besteht darin, eine Benotung der Maturaperformanzen auf der Grundlage der derzeit gültigen Notendefinitionen herbeizuführen, die transparent ist und die vom einzelnen Korrektor/der einzelnen Korrektorin und den Umständen seiner Korrektur möglichst unabhängig ist. In der Testtheorie spricht man bei diesem Phänomen von *interrater*-Reliabilität.²⁶

Die Art des Beurteilungssystems beeinflusst andererseits auch die Performanzen: Wenn ein Korrektursystem ausschließlich überprüft, ob Fehler gemacht wurden, wie es bei der klassischen Form der Negativkorrektur geschieht, wird ein Kandidat im Zweifelsfall zögern und eine wörtlich-abbildende Übersetzung, die ihm

²³ Zum grundsätzlichen Problem der Subjektivität des mit dem neuen Korrektursystem verbundenen *rating*-Verfahrens vgl. McNamara, Tim.: *Language Testing*, Oxford 2000, 37–38.

²⁴ Grundsätzliche Informationen zur Aufgabenstellung bei der schriftlichen Reifeprüfung in Griechisch und Latein wurden in der Broschüre „Reifeprüfung neu. Die standardisierte kompetenzorientierte Reifeprüfung in den Klausurfächern Griechisch und Latein“ [BIFIE Wien 2011] niedergelegt.

²⁵ Vgl. <https://www.bifie.at/node/520> (vierjähriges Latein), <https://www.bifie.at/node/519> (sechsjähriges Latein) bzw. <https://www.bifie.at/node/1509> (Griechisch).

²⁶ Vgl. McNamara (wie Anm. 24), 56–59; Hughes, Arthur: *Testing for Language Teachers*, Cambridge 2003, 53–52.

grammatikalisch richtig erscheint, einer freieren Übersetzung vorziehen, auch wenn diese den Sinn des ausgangssprachlichen Textes viel treffender wiedergeben würde.

Schon früh wurde an solchen Korrekturformen, die im deutschsprachigen Raum lange Zeit üblich waren,²⁷ Kritik geübt: Bereits 1846 fordert Karl Friedrich von Nögelsbach in der „Vorrede“ zur ersten Auflage seiner „Lateinischen Stilistik“ die Berücksichtigung sprachlicher Eigenschaften über die Elementargrammatik hinaus und kommt dabei auch auf die Korrekturmethode seiner Zeit zu sprechen:

Ich kann mir in der Tat nichts langweiligeres und nutzloseres [sic] denken, als wenn jenes Durchgehen [sc. einer Schülerübersetzung] lediglich darin besteht, daß man dem einzelnen Schüler sein Heft mit der Bemerkung zurückgibt: hier ist und dort ist ein Fehler gemacht worden. Nicht nur wird von den übrigen, jetzt gerade nicht beteiligten Schülern höchst selten einer auf die Berichtigung der Fehler des anderen merken, sondern der angeredete Schüler selbst wird in der Meinung stehen, er habe eben nur hier oder dort gefehlt, alles übrige sei richtig. So wird sich denn überhaupt die Ansicht festsetzen, man könne nur im einzelnen fehlen, und der Wert oder Unwert der gelieferten Arbeit hänge lediglich von der Zahl der Fehler ab. Eine nur zu gewöhnliche, aber höchst verderbliche Ansicht, welche im Schüler alle stilistische Bemühung von vorneherein untergräbt!²⁸

Nögelsbach erkennt an dieser Stelle, die er seinem Hauptanliegen entsprechend stark auf die Stilistik zuspitzt, ein Grundproblem der Negativkorrektur: Nur weil in einer Passage einer Performanz kein Fehler angestrichen wurde, bedeutet dies noch lange nicht, dass der Schüler/die Schülerin den Inhalt der Stelle verstanden und alles richtig gemacht hat.

Doch nicht nur der gestrenge Stilist des 19. Jhs., sondern auch Fachdidaktik unserer Tage machen sich Gedanken über die Sinnhaftigkeit der Übersetzungen, deren Aussehen und Qualität auch durch die Art der Korrektur beeinflusst wird. Karl-Wilhelm Weeber etwa bricht in seinem Buch „Mit dem Latein am Ende?“ eine Lanze für die holistische Betrachtung von Texten, die eben nicht als bloße Ansammlung von Einzelsätzen zu verstehen sind:

Dass manches Produkt, das so [sc. durch die perspektivisch eingeeengte Übersetzung von Einzelsätzen] zustande gekommen ist, spätestens dann,

²⁷ Vgl. Bauder, Manfred: Leistung, Kontrolle und altsprachlicher Unterricht im Wandel der Zeiten, in: Forum Classicum 49 (2006), 122–134, bes. 128–129; Engelbrecht, Helmut: Wozu Schulnoten? Leistungsbeurteilung an höheren Schulen vom Mittelalter bis heute, in: Kremser Humanistische Blätter 12 (2008), 11–46.

²⁸ Zitiert nach: Nögelsbach, Karl Friedrich von: Lateinische Stilistik, Nürnberg 1905 (Nachdr. Darmstadt 1980), XVI–XVII.

wenn man es als Ganzes durchliest, als hanebüchener Unsinn noch freundlich umschrieben ist, ist allgemein bekannt – und wird erstaunlich gottergeben akzeptiert. Als wenn es dann besonders lateinisch zuginge, wenn der offensichtliche Nonsens blüht und gedeiht!²⁹

Wenn man diese Kritik ernst nimmt und die Gottergebenheit, von der Weeber in diesem Zitat spricht, nicht hinnehmen will, muss man sich zunächst überlegen, welches Produkt eigentlich erwartet wird, wenn die Handlungsanweisung an die Schüler/die Schülerinnen „Übersetze den folgenden Text!“ lautet, und wie die gewünschten Eigenschaften des Produktes auch überprüft werden können.

Mit Blick auf die in der Übersetzungswissenschaft gängige Äquivalenz-Diskussion³⁰ wurde in den Kompetenzmodellen für die Dimension „Übersetzen“ folgende Grunddefinition festgeschrieben: „Ziel der Übersetzung ist nicht primär die formale Übereinstimmung zwischen Ausgangs- und Zielsprache, sondern die Produktion eines in Inhalt, Sinn und Funktion äquivalenten Textes, der die Textnormen der Zielsprache berücksichtigt.“³¹

Auf dieser Grundlage muss ein Korrektursystem also in erster Linie überprüfen, ob Ausgangs- und Zieltext in Inhalt, Sinn und Funktion äquivalent sind und übereinstimmen. Ein System, das kontrolliert, ob jedes einzelne griechische oder lateinische Wort übersetzt wurde, und das ein Nachahmen der (Satz-)Strukturen des Ausgangstextes verlangt, wäre unangemessen.³² Außerdem muss das System berücksichtigen, ob die sprachlichen Normen der Zielsprache eingehalten wurden.

4. Wesen

Grundsätzlich werden bei Korrektursystemen zählende von abwägenden Systemen unterschieden.³³ Bei ersteren gibt es bestimmte Kriterien, nach denen vorher festgesetzte Punkte (nicht) erreicht werden können. Am Ende ergibt sich eine Gesamtzahl, die entsprechend einer festen Skala zu einer bestimmten Note führt. Abwägende Systeme legen zunächst für jeden zu erreichenden Notenwert bestimmte Qualitätskriterien fest und formulieren Musterlösungen für die jeweilige Stufe (sog.

²⁹ Weeber, Karl-Wilhelm: Mit dem Latein am Ende? Tradition mit Perspektiven, Göttingen 1998, 45.

³⁰ Zur allgemeinen Darstellung der Äquivalenzdiskussion vgl. Stolze, Radegundis: Übersetzungstheorien. Eine Einführung, Tübingen 2011, 87–104.

³¹ Rechtsgrundlagen und Leitlinien (wie Anm. 4), 14.

³² Ein solches System erscheint vielmehr im Elementarunterricht angemessen, da dort die Hauptaufgabe der Übersetzung für den Schüler darin besteht zu demonstrieren, dass er in der Lage ist, mit den aktuell vorliegenden sprachlichen Phänomenen umzugehen.

³³ Im angelsächsischen Raum unterscheidet man zwischen scoring und rating. Zählende Systeme zielen durch geschlossene Testformate meist auf rezeptive Fertigkeiten ab, abwägende Systeme hingegen durch offene Testformate v.a. auf produktive Fertigkeiten. Die Übersetzung liegt zwischen diesen beiden Bereichen.

benchmarks). Die zu beurteilende Performanz des Schülers/der Schülerin wird dann mit den Kriterien und den *benchmarks* verglichen und kann so einer bestimmten Notenstufe zugeordnet werden.

Das neue Korrektursystem in Griechisch und Latein ist eine Mischung aus beiden Formen, wobei die zählende Variante deutlich im Vordergrund steht, während die abwägende Form auf einen kleinen Aspekt im Bereich des ÜT beschränkt bleibt: die Bewertung der sprachlichen Qualität in der Zielsprache. Letztlich wird auch dieser Bereich in Punkte umgerechnet, sodass ein Maximum von 60 Punkten erreicht werden kann. Die Kandidaten haben die Möglichkeit, im Bereich des ÜT 36 Punkte, im Bereich des IT 24 Punkte zu erreichen. Diese prozentuelle Verteilung ist der Grund dafür, dass von den Korrektoren/den Korrektorinnen keine zusätzlichen Bonuspunkte (oder Plus-Punkte) vergeben werden dürfen. Die Vergabe von halben Punkten ist ebenfalls nicht vorgesehen (s.u.).

Schließlich wird die Gesamtheit der erreichten Punkte in eine Note umgerechnet: 53–60 Punkte bedeuten „Sehr gut“, 45–52 Punkte „Gut“, 37–44 Punkte „Befriedigend“, 30–36 Punkte „Genügend“. Voraussetzung für eine positive Beurteilung ist das Erreichen von insgesamt zumindest 30 Punkten (von 60), wobei im Kompetenzbereich Übersetzung die Untergrenze von 18 Punkten und im Kompetenzbereich Interpretation die Untergrenze von 12 Punkten nicht unterschritten werden darf. Wenn eine Kandidatin/ein Kandidat in einem der beiden Kompetenzbereiche unter den Mindestanforderungen liegt, ist die Arbeit unabhängig von der Leistung im anderen Kompetenzbereich mit „Nicht genügend“ zu beurteilen.

5. Beurteilung des ÜT

Im Bereich des ÜT wird (im Gegensatz zum IT, s.u.) ausschließlich die Übersetzungsleistung des Kandidaten/der Kandidatin beurteilt. Relevant für die Korrektur sind folgende drei Kriterien : (1.) Sinn, (2.) Lexik, Morphologie und Syntax, (3.) sprachliche Qualität in der Zielsprache. Alle drei Dimensionen werden im Folgenden näher vorgestellt.

In allen genannten Bereichen ist eine vorgegebene Zahl von Punkten zu erreichen: zwölf im Bereich des Sinns, sechs im Bereich der Lexik, sechs im Bereich der Morphologie, sechs im Bereich der Syntax und sechs im Bereich der sprachlichen Korrektheit. Auch hier soll der einleitend angesprochene Steuerungseffekt (sog. *wash back*) für die Performanzen zum Tragen kommen: Ein hohes Kontingent an Punkten ist für die denotative Äquivalenz (Sinneinheiten) vorgesehen, und eine immer noch beträchtliche Zahl für sprachliche Korrektheit in der Unterrichtssprache.

5.1. Sinn

Die erste Dimension, die bei einer Performanz überprüft wird, betrifft die denotative Äquivalenz zwischen dem ausgangssprachlichen Originaltext und der Schülerperformanz. Zu diesem Zweck wird der Text in Sinneinheiten gegliedert, was sich im Griechischen und Lateinischen aufgrund der Prinzipien des Prosarhythmus,³⁴ der beliebten Klammerstellung satzwertiger Ausdrücke, der häufigen Schlussstellung des finiten Verbums und anderer Eigenschaften besonders leicht bewerkstelligen lässt. Um zu einem ausgewogenen Verhältnis zwischen dieser und den folgenden Dimensionen zu gelangen, ist die Zahl der Sinneinheiten mit zwölf verbindlich festgesetzt.

Die Grundidee besteht nun darin, bei jeder einzelnen Sinneinheit von der rhetorischen Ausgestaltung abzusehen und zu bestimmen, was der Inhalt der Passage ist. Cicero hat diesen Unterschied in seinem *Orator* mit der Spannung zwischen *sententia* (Inhalt) und *ornatus* (rhetorische Ausgestaltung) zu beschreiben versucht, z.B. *Conlocata autem verba habent ornatum, si aliquid concinnitatis efficiunt, quod verbis mutatis non maneat manente sententia* (Cic. or. 81).³⁵

Hier kommt einer der ersten subjektiven Bereiche dieses Korrektursystems zum Vorschein: Bereits durch die Art und Weise der Abteilung³⁶ der einzelnen zwölf Sinneinheiten kann es für die Kandidaten und Kandidatinnen leichter oder schwieriger werden, Punkte zu bekommen. Zudem empfiehlt es sich für den Korrektor/die Korrektorin, für sich selbst in Form einer Paraphrase festzuhalten, was er/sie für den Sinn oder den Inhalt (*sententia*) der Sinneinheit hält, um die konkrete Performanz damit vergleichen zu können. Je nachdem, wie detailliert diese Paraphrase ausfällt, wieviele Aspekte aufgenommen bzw. ausgeklammert werden, wird es für den Schüler einfacher bzw. schwieriger, einen Punkte zu bekommen. Bei der Reifeprüfung wird im Rahmen der Aufgabenstellung auch das Korrektursystem mit abgeteilten Sinneinheiten und Paraphrasen mitgeliefert.

Der Beginn der *Res gestae divi Augusti* diene im Folgenden als Beispiel für die konkrete Anwendung dieser Dimension:

³⁴ Zur Einteilung lateinischer Prosa in Kommata und Kola vgl. Schmid, Walter: Über die klassische Theorie und Praxis des antiken Prosarhythmus, Wiesbaden 1959.

³⁵ Konkret geht es an der zitierten Stelle um den Unterschied zwischen Gedanken- und Wortfiguren, vgl. Sandys, John Edwin (Hg.): Marcus Tullius Cicero. Ad Marcum Brutum Orator, Cambridge 1885 (Nachdruck Hildesheim / New York 1973), 92. Dennoch eignet sie sich gut zur Illustration des Unterschiedes zwischen Inhalt und sprachlicher Form.

³⁶ Zur Arbitrarität der Abtrennung formuliert Helfried Gschwandtner treffend: „Was eine Sinneinheit ist, bleibt weitgehend Ihrem persönlichen Geschmack überlassen.“ (Gschwandtner, Helfried: Ökonomie des Lateinunterrichts. Leistungsbewertung und Lektüre, in: Latein Forum 27 (1995), 43–53, hier: 45).

Annos undeviginti natus exercitum privato consilio et privata impensa comparavi, per quem rem publicam a dominatione factionis oppressam in libertatem vindicavi.

Das durch die Schlussstellung des Verbuns (*comparavi*) markierte Ende des Hauptsatzes und der Beginn des Relativsatzes (*per quem*) bieten sich an, um hier eine Grenze zu ziehen und so zwei Sinnheiten voneinander abzugrenzen:

SE 1: *Annos undeviginti natus exercitum privato consilio et privata impensa comparavi*
SE 2: *per quem rem publicam a dominatione factionis oppressam in libertatem vindicavi.*

Als nächstes muss bedacht werden, welche Inhalte in diesen Sinneinheiten transportiert werden. Eine mögliche Paraphrase könnte lauten: „Ich war 19 Jahre alt. Ich habe ein Heer aufgestellt. Ich hatte keinen Auftrag von staatlicher Seite. Ich habe es mit eigenen Mitteln ausgeführt.“ Eine Paraphrase von SE 2 könnte lauten: „Der Staat hatte unter einer Gewaltherrschaft zu leiden. Ich habe ihn von dieser befreit.“ Bei der Formulierung solcher Paraphrasen sollen sprachlich-stilistische Erscheinungen keine Rolle spielen. So bleibt etwa das anaphorische *privato – privata* aus SE 1 hier unberücksichtigt, und auch die eigentliche Satzstruktur (z.B. der Relativsatz in SE 2) sollte nicht relevant sein (außer Phänomene wie kausallogische Verknüpfungen u.dgl.).

Wie könnten nun mögliche Performanzen aussehen, die mit diesen Paraphrasen zu vergleichen sind? Nehmen wir folgende beiden Musterbeispiele (P1 und P2) an:

P1: Ohne offiziellen Auftrag nahm ich mit 19 mein Geld und hob Truppen aus.
P2: 19 Jahre geboren kaufte ich das Heer mit privatem Plan und mit privatem Aufwand.

Obwohl P2 die sprachliche Struktur des Ausgangstextes getreulich abbildet, jedes Wort übersetzt und an seiner nämlichen Stelle platziert, wird doch deutlich, dass der Inhalt der Sinneinheit nicht verstanden wurde: Obwohl noch von keiner Armee die Rede war, wird der bestimmte Artikel „das“ vor „Heer“ gesetzt. Dass mit dem lateinischen Partizip *natus* Altersangaben ausgedrückt werden, wurde nicht erkannt bzw. adäquat ins Deutsche übertragen. Die Übersetzungen „Plan“ und „Aufwand“ für *consilio* und *impensa* machen den deutschen Text für einen Leser, der mit dem Original nicht vertraut ist, nur schwer (wenn überhaupt) verständlich.

P1 hingegen hat stark in die Satzarchitektur eingegriffen, die Altersangabe von der Spitzenstellung im Lateinischen mitten in den Satz verschoben und aus dem ersten

Ablativ einen Präpositionalausdruck, aus dem zweiten einen eigenen Hauptsatz gemacht. Dennoch sind alle Elemente, die für den Inhalt als wesentlich erkannt wurden, vorhanden.

Wenn der Inhalt der Sinneinheit, wie er in der Paraphrase umschrieben worden ist, in der Performanz wiedergegeben wird, so erhält der Kandidat an dieser Stelle einen Punkt. Wenn nicht, werden null Punkte vergeben. Die Vergabe von halben Punkten oder von Zusatzpunkten ist nicht vorgesehen. Maximal sind in dieser Dimension also zwölf Punkte zu erreichen. In unseren Musterbeispielen würde P1 einen Punkt, P2 null Punkte erhalten.

5.2. Checkpoints

Um zu überprüfen, ob ein Kandidat/eine Kandidatin den Inhalt eines Originaltextes verstanden hat, reicht die Kontrolle der denotativen Äquivalenz in den zwölf Sinneinheiten. Weil der altsprachliche Unterricht jedoch in erster Linie zum genauen Übersetzen und nur in geringerem Maß zum Paraphrasieren erziehen will, müssen auch Elemente in das Beurteilungssystem eingebaut werden, die diesem Umstand Rechnung tragen.

Es werden daher drei Dimensionen bestimmt, in denen die Kandidaten und Kandidatinnen besondere Fähigkeiten besitzen müssen, wenn sie eine korrekte Übersetzung anfertigen wollen: Im Bereich der „Lexik“ müssen sie in der Lage sein, die im speziellen Kontext passende Bedeutung eines Wortes herauszufinden (Monosemierung). Im Bereich der „Morphologie“ besteht die Grundüberlegung darin, dass zur korrekten Übersetzung eines Wortes morphologische Kenntnisse entscheidend sind und eingesetzt werden müssen. Dasselbe gilt auf Satzebene für die „Syntax“, wobei die Grenze zwischen morphologischen und syntaktischen Phänomenen (v.a. in der Kasuslehre) nicht immer ganz klar zu bestimmen ist. Viele Grammatiken behelfen sich, indem sie von „Morphosyntax“ sprechen.

5.2.1. Lexik

Über den gesamten Text verstreut werden sechs Lexik-Checkpoints (LEX) stichprobenartig verteilt. Geeignet sind besonders solche Wörter, denen im Deutschen eine gewisse Fülle von möglichen Übersetzungen gegenübersteht. Eher ungeeignet sind daher Wörter wie *octo* (acht) oder *gingiva* (Zahnfleisch). Wie schon bei den Sinneinheiten empfiehlt es sich auch hier für den Korrektor/die Korrektorin, die eigene Erwartungshaltung im Vorfeld festzuhalten. Aufgrund der potentiell unüberschaubaren Zahl richtiger Übersetzungen wird man sich mit wenigen Beispielen für richtige bzw. falsche Lösungen begnügen.

Bleiben wir zur exemplarischen Demonstration bei der oben zitierten Eingangspassage der *Res gestae* und legen einen LEX-Checkpoint auf *consilio*. Zunächst wird man Beispiele für mögliche korrekte Übersetzungen anführen. Diese könnten z.B. lauten: Beschluss, Entschluss, Überlegung. Dann wird man Bedeutungen anführen, die nicht richtig sind, wie z.B. Sitzung, Versammlung. So ergibt sich zwischen den beiden Polen einer optimalen bzw. absolut unakzeptablen Übersetzung ein breites Feld von mehr oder weniger passenden Lösungen.

Bei der nun folgenden Überprüfung der Performanzen können für die einzelnen Checkpoints wieder ein oder null Punkte vergeben werden. Die Überlegung dabei darf ausschließlich sein, ob der Checkpoint aus lexikalischer Sicht korrekt wiedergegeben wurde. Nicht berücksichtigt werden dürfen hier morphologische oder syntaktische Aspekte.

Nicht immer wird in einer Performanz ein lateinisches Wort mit genau einem deutschen Wort wiedergegeben. Dies ist kein Zeichen von Schwäche, sondern zeichnet im Gegenteil den guten Übersetzer/die gute Übersetzerin mitunter aus.³⁷ In diesem Fall muss der Blick u.U. etwas erweitert und gefragt werden, ob das betroffene Satzglied korrekt wiedergegeben wurde. Wenn in P1 *privato consilio* mit „ohne offiziellen Auftrag“ übersetzt wurde, trifft dies die Sache im Kern und darf daher auch bei den LEX-Checkpoints nicht mit Punkteabzügen geahndet werden. P2 übersetzt *consilio* mit „Plan“. Der Verdacht liegt zwar nahe, dass der Verfasser dieser Performanz nicht verstanden hat, worin der Witz der Stelle liegt – Octavian hat im Jahr 44 v. Chr. bekanntlich ohne öffentliches Mandat mit der Aushebung von Truppen gegen Marcus Antonius begonnen und wurde erst nachträglich durch den Einsatz Ciceros legitimiert – aber die Frage darf hier nur lauten, ob eine unter lexikalischem Gesichtspunkt korrekte Übersetzung für *consilium* gefunden wurde, und da dies der Fall ist, ist auch hier der Punkt zu geben.

5.2.2. Morphologie

In der ersten praktischen Erprobung dieses Systems hat sich herausgestellt, dass die im Folgenden zu besprechenden Morphologie-Checkpoints (MORPH) mit den vergleichsweise größten Problemen verbunden sind.

Die Grundidee besteht darin, dass für eine korrekte Übersetzung bestimmter Stellen, auf die ein MORPH-Checkpoint gelegt wird, besonders morphologische Kenntnisse ausschlaggebend sind. Wie bei allen bisher besprochenen Dimensionen gilt auch hier, dass sich der Korrektor/die Korrektorin im Vorfeld überlegen muss, welche morphologischen Aspekte für ihn/sie entscheidend sind. Wenn wir beispielsweise

³⁷ Als Beispiel diene die Behandlung von lateinischen Doppelausdrücken, die im Deutschen mit einem Einzelwort, das um ein intensivierendes Adverb erweitert wurde, wiedergegeben werden können, z.B. *fundi fugarique* – völlig aus dem Feld geschlagen werden, vgl. Nägelsbach (wie Anm. 29), 341.

auf *comparavi* in unserem Mustertext einen MORPH-Checkpoint legen, werden aus morphologischer Perspektive v.a. zwei Aspekte entscheidend sein: Tempus (die Handlung fand in der Vergangenheit statt) und Person (es wird nicht über jemanden gesprochen, sondern ein Ich spricht über sich selbst).

Wenn Tempus und Person als wesentliche Aspekte bestimmt wurden, dürfen bei der Korrektur ausschließlich diese beiden Aspekte berücksichtigt werden. In beiden Musterperformanzen P1 und P2 ist sowohl die 1. Person als auch die Vergangenheit wiedergegeben. Beide würden an dieser Stelle einen Punkt erhalten. Wie wäre es aber, wenn es in einer Performanz lautete „ich habe (das Heer) verglichen“? Auch in diesem Fall muss der Punkt gegeben werden, da sowohl die 1. Person als auch die Vergangenheit erkannt und wiedergegeben wurden. Der Fehler ist lexikalischer Natur, und dieser Aspekt ist bei den MORPH-Checkpoints irrelevant.

Bei der Auswahl der MORPH-Checkpoints muss darauf Wert gelegt werden, dass das morphologische Kriterium in einer Übersetzung eindeutig zum Ausdruck kommen muss und nicht durch eine mögliche freiere Übersetzung vollkommen verdeckt wird. Würde man im Satz *Canis mihi est* einen MORPH-Checkpoint auf *mihi* legen und als Kriterium „Kasus“ angeben, erwartet man sich in der Übersetzung ein „mir“. Für die völlig korrekte Übersetzung „Ich besitze einen Hund“ dürfte der MORPH-Punkt nicht vergeben werden. Dieser MORPH-Checkpoint wäre daher schlecht gewählt. Diese Passage könnte im Bereich der im Folgenden zu besprechenden Syntax-Checkpoints besser überprüft werden.

5.2.3. Syntax

Im Bereich der „Syntax“ (SYN) ist nun die Satz- und Textebene erreicht, weswegen auch der Begriff Check-„point“ nicht mehr ganz glücklich gewählt ist, aber in Analogie zu den vorhergehenden Dimensionen Lexik und Morphologie weiterhin benutzt wird.

Diese Checkpoints dienen der Überprüfung, ob satzwertige Konstruktionen, Partizipialausdrücke, Gliedsätze u.dgl. möglichst korrekt übersetzt wurden. Auch hier empfiehlt es sich, nach der Festsetzung eines SYN-Checkpoints eine gewisse Erwartungshaltung zu formulieren.

In unserem Musterbeispiel könnte ein SYN-Checkpoint auf *natus* liegen. Mögliche korrekte Übersetzungen sind „mit (19 Jahren)“ oder „im Alter von (19 Jahren)“. Mögliche falsche Übersetzungen sind „Sohn“ oder „geboren“. Von unseren Musterperformanzen würde P1 den Punkte für diesen SYN-Checkpoint erhalten, P2 hingegen nicht.

5.3. Qualität in der Zielsprache

Bisher konnten die Kandidaten und Kandidatinnen 30 Punkte für ihre Leistungen zum ÜT erreichen: zwölf Punkte für die einzelnen Sinneinheiten, jeweils sechs für die LEX-, MORPH- und SYN-Checkpoints. Weitere sechs Punkte sind für die Dimension „Qualität in der Zielsprache“ vorgesehen.

Wie bereits einleitend ausgeführt, handelt es sich bei dieser Dimension nicht um einen Bereich, in dem gezählt wird, sondern in dem die Performanz in ihrer Gesamtheit eingeschätzt wird. Es geht nicht um einzelne Stellen, wie bei den LEX/MORPH/SYN-Checkpoints, sondern um die Arbeit in ihrer Gesamtheit, um ihren sprachlichen Gesamteindruck. Folgende drei Niveaustufen werden unterschieden:

Niveaustufe 3: Bei der Formulierung der Übersetzung werden die Normen der Zielsprache, vor allem in den Bereichen Wortstellung, Textkohärenz und Idiomatik, eingehalten, sodass ein gut verständlicher und grammatikalisch korrekter Text entsteht. Die Übersetzung zeichnet sich durch viele gelungene Formulierungen aus.

Niveaustufe 2: Bei der Formulierung der Übersetzung werden die Normen der Zielsprache, vor allem in den Bereichen Wortstellung, Textkohärenz und Idiomatik, überwiegend eingehalten, sodass ein mit einiger Anstrengung verständlicher und grammatikalisch weitgehend korrekter Text entsteht. Die Übersetzung zeichnet sich durch einige gelungene Formulierungen aus.

Niveaustufe 1: Bei der Formulierung der Übersetzung werden die Normen der Zielsprache, vor allem in den Bereichen Wortstellung, Textkohärenz und Idiomatik nur ansatzweise eingehalten, sodass ein kaum verständlicher Text entsteht. Die Übersetzung weist kaum gelungene Formulierungen auf.

Entspricht eine Performanz der Beschreibung der Niveaustufe 3, so werden in dieser Dimension sechs Punkte vergeben. Entsprechend werden für Stufe 2 drei Punkte und für Stufe 1 null Punkte vergeben.

Auch wenn in dieser Dimension ausschließlich die Qualität der Zielsprache in den Blick genommen werden soll, muss der zu beurteilende Text doch in einem Zusammenhang mit dem ausgangssprachlichen Text stehen. Auf diese Verbindung zielt die Angabe „bei der Formulierung der Übersetzung“. Für einen „gut verständlichen Text“ gibt es nur dann sechs Punkte, wenn es sich tatsächlich um eine Übersetzung handelt oder handeln soll.

Wenn Teile des ÜT nicht bearbeitet wurden, kann im Bereich der sprachlichen Qualität dennoch die maximale Punktezahl vergeben werden, gewisse Mindestanforderungen im Bearbeitungsstand sind jedoch einzuhalten: Die maximale Punktzahl kann nur dann vergeben werden, wenn mindestens neun (der insgesamt zwölf) Sinneinheiten bearbeitet wurden. Wenn aber weniger als sechs Sinneinheiten bearbeitet wurden, ist die sprachliche Qualität mit null Punkten zu bewerten.

Bei den angeführten Musterperformanzen müsste man natürlich alle zwölf Sinneinheiten in Betracht ziehen, um über die sprachliche Qualität zu urteilen. Betrachtet man nur SE 1, dann ist P1 ein gut lesbarer Text gelungen (sechs Punkte), während P2 gegen zahlreiche Regeln der deutschen Sprache verstößt: Die Altersangabe ist nicht korrekt, die Verwendung des bestimmten Artikels bei einem nicht eingeführten Objekt ist unpassend. Es entsteht ein kaum verständlicher Text, weswegen hier null Punkte zu vergeben sind.

6. Beurteilung des IT

Zum IT werden zehn Arbeitsaufgaben formuliert, wobei etwa zwei Drittel davon geschlossene oder halboffene Formate sind, während ein Drittel aus offenen Aufgaben besteht. Eine genaue Darstellung und Beschreibung der für die Reifeprüfung vorgesehenen Formate erfolgte an anderer Stelle.³⁸ Bereits im Rahmen der Aufgabenstellung wird angegeben, wie viele Punkte für die einzelne Aufgabe maximal zu erreichen sind. Bei der Reifeprüfung können pro Aufgabe maximal vier Punkte, bei allen Aufgaben zum IT zusammen maximal 24 Punkte erreicht werden.

Bereits im Rahmen der Aufgabenstellung ist zu überlegen, für welche Leistung wie viele Punkte zu vergeben sind. Allgemeine Angaben hängen hier in erster Linie vom Format ab: Wenn bei *multiple choice*-Aufgaben aus insgesamt zwei Alternativen zu wählen ist (z.B. richtig/falsch-Aufgaben), dann kann die Hälfte der maximal bei der entsprechenden Aufgabe zu erreichenden Punkte erst dann vergeben werden, wenn mehr als die Hälfte der Items, d.h. der Unteraufgaben, aus denen sich die Aufgabe zusammensetzt, richtig gelöst wurde. So soll verhindert werden, dass ein Kandidat/eine Kandidatin durch Raten Punkte erhält.

Bei halboffenen Aufgaben hingegen, bei denen der Kandidat/die Kandidatin etwa kleinere Lücken füllen muss, ist es durchaus möglich, die Hälfte der maximal zu erreichenden Punkte zu bekommen, wenn auch die Hälfte der Items richtig gelöst wurde. Hierzu ein Beispiel: Wenn die Aufgabe darin besteht, im IT vier Begriffe aus einem bestimmten Sachfeld zu finden und maximal zwei Punkte erreicht werden können, dann kann für zwei korrekt gefundene Begriffe bereits ein Punkt vergeben werden.

³⁸ Vgl. Pinter, Anna / Widhalm-Kupferschmid, Wilhelmine: Bausteine zum Erstellen von Schularbeiten (ÜT+IT): <https://www.bifie.at/node/1387> (18. 9.2014)

Etwas komplexer gestaltet sich die Punktevergabe bei den offenen Arbeitsaufgaben: Hier ist zunächst darauf zu achten, im Rahmen der Aufgabenstellung auch eine maximale Wortanzahl für die Lösung vorzugeben. Wird diese maximale Wortanzahl um mehr als zehn Prozent überschritten, wird von der erreichten Punktezahl ein Punkt abgezogen.

Im Rahmen der Aufgabenstellung ist weiters ein Erwartungshorizont zu formulieren, verbunden mit einer klaren Vorstellung darüber, für welche Leistung wie viele Punkte vergeben werden. Besteht die Aufgabe z.B. in einer Paraphrasierung oder Zusammenfassung des IT und wird mit vier Punkten versehen, so muss sich der Korrektor/die Korrektorin in der Vorbereitungsphase vier wesentliche Elemente überlegen, deren Vorhandensein in der konkreten Performanz er/sie dann überprüfen kann. Bei anderen Fragestellungen empfiehlt es sich, komplexere Fragen durch Leitfragen zu gliedern und diesen Leitfragen dann einzelne Punkte zuzuweisen.