

Henning Horstmann/Matthias Korn

Cicero als Prüfungsautor im schriftlichen Latinum. Von der Erforderlichkeit einer ganzheitlichen Berechnungsweise des Anforderungsniveaus lateinischer Texte

1. Vorbemerkungen

Auf Bundesebene ist Rechtsgrundlage für den Nachweis des Latinums an Schulen und Universitäten der Beschluss der Kultusministerkonferenz (KMK) vom 22. September 2005. Diese Grundlage wird von den Ländern jeweils auf Verordnungs- oder Erlassebene spezifiziert. Schulrechtlich maßgeblich sind die Dokumente der jeweiligen Landesebene. Die Praxis der schriftlichen Latinumsprüfungen an den Universitäten wurde 2016 von Lichy¹ untersucht. Wir konzentrieren uns daher hier auf den Bereich der Schulen.

2. Kurze Geschichte des Latinums in den Ländern der BR Deutschland

Bis zum Jahr 1979 unterschieden die Länder lediglich zwischen den Abschlüssen „Kleines Latinum“ und „Großes Latinum“. Die Festlegung des Anforderungsniveaus lag allein in der Hoheit des jeweiligen Landes.

Durch den Beschluss der Kultusministerkonferenz vom 26. Oktober 1979 wurden folgende Vereinheitlichungen geschaffen:

¹ Siehe Literaturverzeichnis.

- I. Für den Abschluss wurde die einheitliche Bezeichnung „Latinum“ eingeführt, eine parallellaufende, weitere Verwendung der Bezeichnungen „Kleines Latinum“ und „Großes Latinum“ aber der Hoheit des jeweiligen Landes überlassen.
- II. Als Anforderungsniveau wurde bestimmt: „[...] lateinische Originaltexte im sprachlichen Schwierigkeitsgrad inhaltlich anspruchsvollerer Stellen (bezogen auf Autoren wie Cicero, Sallust, Livius oder auf vergleichbare andere Autoren) [...]“.²
- III. Die Verwendung eines Wörterbuches wurde generell zugelassen.
- IV. Der schulische Erwerb des Latinums ohne Ergänzungsprüfung, das sogenannte Ersitzen, wurde gekoppelt an das Absolvieren aufsteigenden Pflichtunterrichts mit einer Lerndauer von mindestens vier Jahren und das Erreichen der Note „Ausreichend“ (4) an dessen Ende.
- V. Als spezifische Fachleistung der Ergänzungsprüfung wurde die Übersetzung eines lateinischen Textes ins Deutsche festgelegt.

Diese Vereinheitlichungen wurden 26 Jahre später durch den Beschluss der Kultusministerkonferenz vom 22. September 2005 in folgender Weise weiterentwickelt:

- I. Im Hinblick auf das Anforderungsniveau wurden die Autorennamen durch Textgattungen bzw. -sorten ersetzt: „[...] lateinische Originaltexte im sprachlichen Schwierigkeitsgrad inhaltlich anspruchsvollerer Stellen (bezogen auf Bereiche der politischen Rede, der Philosophie und der Historiographie) [...]“.³
- II. Die Verwendung eines Wörterbuches wurde in die Entscheidungshoheit des jeweiligen Landes gelegt.
- III. Als Alternative zur Fachleistung „Übersetzung“ wird die zweiteilige Fachleistung „Übersetzung und Interpretation“ festgelegt.

Bei Lichte betrachtet hat der KMK-Beschluss von 2005 den Beschluss von 1979 nur unwesentlich modifiziert und damit den doch beträchtlichen Entwicklungen im Bereich der Fachdidaktik und der Akzeptanz des Faches kaum Rechnung getragen. Dabei war zudem der Ersatz der Autorennamen aus dem

2 Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 1979, Ziffer 2.

3 Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 2005, Ziffer 2.

Beschluss von 1979 gar nicht erforderlich, weil schon damals „Autoren wie“⁴ und „vergleichbare Autoren“⁵ geschrieben stand und die Fachdidaktik bekanntlich bis heute keine valide Berechnungsweise zur intersubjektiv überprüfbareren Feststellung des Anforderungsniveaus lateinischer Texte entwickelt hat. Befürchtet wurde damals offenbar, dass die Autorennamen Cicero, Sallust und Livius zu wirkmächtig seien. Zu dieser Befürchtung bestand seinerzeit auch aller Anlass, denn die Realität der Ergänzungsprüfungen aus den Jahren zwischen 2000 und 2005 zeigt deutlich, dass weitestgehend Cicero Autor der Texte war, die zur schriftlichen Latinumsprüfung als Aufgaben vorgelegt wurden. Schauen wir also heute, 13 Jahre später, ob sich daran etwas geändert hat.

3. Die Dominanz von Cicero-Texten in der schriftlichen Latinumsprüfung

Unsere Recherchen zu den schriftlichen Ergänzungsprüfungen zum Erwerb des Latinums im Bereich der Schulen ergaben:

Grundlage der Recherchen waren die je 16 offiziellen Landes-CDs mit den schriftlichen Abiturprüfungen 2016 und 2017. Nur vier Länder lieferten aber in diesem Rahmen auch die schriftlichen Ergänzungsprüfungen zum Nachweis des Latinums: Baden-Württemberg, Bayern, Sachsen und Thüringen. Aus diesen vier Ländern lagen somit insgesamt sieben⁶ schriftliche Latinumsprüfungen vor. Fünf von diesen waren aus dem Bereich der Schriften Ciceros. Auf der Grundlage der – zugegebenermaßen spärlichen – öffentlich zugänglichen Dokumente kann also festgestellt werden, dass sich im Bereich der Schulen auch nach der Modifikation von 2005 kaum etwas verändert hat und der seit 1979 bestehende Spielraum weiterhin ungenutzt blieb. Im Bereich Universität, der 2016 von Lichy untersucht wurde, ergibt sich ein durchaus anderes Bild: Von den Lichy vorliegenden 16 Ergänzungsprüfungen an

4 Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 1979, Ziffer 2.

5 Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland 1979, Ziffer 2.

6 Cic. *Catil.* 1,44.; Cic. *Flacc.* 3f.; Cic. *Lael.* 17f.; Cic. *S. Rosc.* 29–34; Cic. *S. Rosc.* 73–76; Lhomond, *De viris*, 358f.: Caesars Tod; Lhomond, *De viris*, 342–344: *Secundum bellum civile*.

Universitäten aus fünf Bundesländern aus den Jahren 2014 und 2015 sind sieben aus dem Bereich der Werke Ciceros, neun aber nicht.⁷

Es ist aber nicht nur die nach wie vor bestehende Dominanz Ciceros, sondern auch die Ungleichartigkeit im Anforderungsniveau der Prüfungsaufgaben, die im Bereich der Ergänzungsprüfung an Schulen Fragen aufwirft (dazu s. u.).

4. Die Auswirkungen der KMK-Beschlüsse zum Latinum

Die Auswirkungen der KMK-Beschlüsse von 1979 und 2005 werden gemeinhin in ihrer Tragweite unterschätzt.

Der KMK-Beschluss von 2005 und die darauf basierenden Spezifizierungen der Länder sind zwar nach wie vor in Kraft, deren Bezugsgröße hingegen, das Latinum, ist als akademische Zulassungsvoraussetzung zugunsten der innerhalb von Universitäten oder Universitätsfakultäten nachgewiesenen Lateinkenntnisse weitestgehend weggebrochen. Damit ist – Status quo im Jahre 2018 – das Latinum ein luxuriöses Bildungsrelikt. Das ist – für sich genommen – nicht weiter schlimm. Als verhängnisvoll hingegen erweist sich die Verflechtung zwischen dem KMK-Beschluss und dem Unterricht im (Wahl-)Pflichtfach Latein. Denn in fast allen Ländern bestimmt der KMK-Beschluss mit seiner Kombination von Anforderungsniveau und Lehrgangsdauer auch nach dem faktischen Ende der Verwertbarkeit des Latinums weiterhin Stofffülle und Lernprogression der Kerncurricula des Faches.

Der Lehrplan Latein Gymnasium des Freistaats Sachsen z. B. sieht in der vierten überarbeiteten Fassung von 2019 beim Lehrgang „Latein als zweite Fremdsprache“ im Rahmen von 16 Jahreswochenstunden drei Schuljahre Lehrbucharbeit (10 JWS) und zwei Lektürejahre (6 JWS) vor. Das prüfungsfreie Ersitzen des Latinums ist bei ausreichenden Leistungen am Ende des zweiten Lektürejahres vorgesehen, in dem der Lehrplan inhaltlich neben den Werken Catulls die Reden Ciceros vorsieht.

Ohne die Orientierung am KMK-Beschluss von 2005 und ohne Bezugnahme auf die Erteilung des Latinums könnten Stofffülle und

⁷ Cic. *Catil.* 4,18f.; Cic. *Mur.* 51f.; Cic. *Phil.* 3,8f.; Cic. *Phil.* 7,25f.; Cic. *rep.* 1,38f.; Cic. *Tusc.* 1,3f.; Cic. *Tusc.* 1,9f.; *Gesta Romanorum* 108; Laktantius, *De ira* 4; Lhomond, *De viris*, 76–80; Marius; Melanchthon, *Declamatio de laude vitae scholasticae*, CR 11, Sp. 302; Ov. *am.* 3,11a; Ov. *met.* 2,150–183; Ov. *rem.* 135–176; Sall. *Cat.* 40; Sall. *Jug.* 97,1 (Stellen nach Lichy).

Lernprogression der Kerncurricula des Lateinunterrichts v.a. im Bereich der zweiten Fremdsprache entschärft werden. Das seinerseits erscheint aus fachdidaktischer Sicht dringend geboten, um die seit zehn Jahren kontinuierlich gesunkene Akzeptanz des Unterrichtsfaches zu erhöhen.

5. Wesentliche Voraussetzung: (k)ein einheitliches Anforderungsniveau

Die nachgewiesene Dominanz von Cicero-Texten im schulischen Latinum zeigt, dass der in den KMK-Vorgaben nun auch explizit angebotene Spielraum bei der Textauswahl kaum ausgenutzt wird. Nun ließe sich einwenden, dass diese relative Zurückhaltung didaktisch überaus sinnvoll ist, weil durch die Konzentration auf einen Autor zugleich ein weitestgehend einheitliches Anforderungsniveau in den bundesweiten Latinumsprüfungen sichergestellt werde. Dass es ein solches Bestreben gibt, kann zunächst einmal nicht verwundern – die gesetzlichen Vorgaben nämlich sind bemerkenswert unklar und offen, fordern sie doch für das Latinum „lateinische Originaltexte im sprachlichen Schwierigkeitsgrad inhaltlich *anspruchsvoller* Stellen.“⁸ Was genau soll das bedeuten? Und wie lassen sich solche Textstellen identifizieren?

Dabei ist nun den Autoren solcher und anderer Vorgaben und Curricula gar kein Vorwurf zu machen, denn eine objektiv bestimmbare Definition des Schwierigkeitsgrades eines lateinischen Textes existiert schlicht nicht. Die Beurteilung von Textschwierigkeit obliegt in der Praxis einzig dem subjektiven Eindruck von Didaktikern, Philologen oder Kommissionsmitgliedern. Handelt es sich beim Beurteilenden etwa um eine Lehrerin bzw. einen Lehrer mit großer fachlicher und didaktischer Expertise, so kann man zumindest darauf bauen, statt einer bloßen „Bauchentscheidung“ tatsächlich eine einigermaßen valide Einschätzung über die Eignung eines bestimmten Textes für eine bestimmte Prüfung zu erhalten. Aufschlussreich jedoch ist, dass meist

⁸ So bspw. der Wortlaut in den aktuellen „Anforderungen der Erweiterungsprüfungen Latinum – Kleines Latinum – Graecum – Hebraicum“, die die Bezirksregierung Köln (NRW) auf ihrer Homepage anführt (vgl. https://www.bezreg-koeln.nrw.de/brk_internet/leistungen/abteilung04/43/erweiterungspruefung/ [letzter Zugriff am 19.04.2019]). Die oben (siehe Kap. 2) angegebene Formulierung des übergeordneten KMK-Beschlusses ist sogar noch offener, fordert sie doch „lateinische Originaltexte im sprachlichen Schwierigkeitsgrad inhaltlich anspruchsvollerer Stellen“ – der relativierende Komparativ ist besonders aufschlussreich, bleibt er doch völlig inhaltsleer.

auch solche Experten einen Text vor allem im Verhältnis zu einem anderen einschätzen können („Text A ist leichter/schwieriger als Text B“), nicht aber absolut. Hinzu kommt: Sobald ein zweiter, dritter oder gar noch mehr Beurteiler hinzugezogen werden, wird in der Regel rasch offenbar, auf wie unsicheren Füßen unser Urteil über einen bestimmten Text steht. Die Autoren dieses Aufsatzes haben beispielsweise den Teilnehmern der diesem Band zugrundeliegenden Tagung jeweils drei Cicero-Latinumsprüfungstexte⁹ aus dem Jahr 2017 (aus drei verschiedenen Bundesländern) vorgelegt und sie darum gebeten, die drei Texte nach ihrem Schwierigkeitsgrad von leicht bis schwer anzuordnen. Beim Abgleich der jeweils gefundenen Reihenfolge ergab sich, dass selbst diese Runde ausgewiesener Didaktik- und Cicero-Experten zu einem völlig heterogenen Gesamtbild kam – alle sechs möglichen Kombinationen wurden mehrfach genannt; ein von einigen als besonders schwer eingeschätzter Text erschien anderen gerade als der leichteste und umgekehrt.

Fest steht, dass in diesem Punkt schon immer eine erstaunlich große formale Unklarheit herrscht. Mangels besserer Alternativen werden in Lehrplänen vage, unkonkrete Formulierungen getroffen, die zumeist lediglich darauf beruhen, bestimmten lateinischen Autoren und Werken ein niedriges, mittleres oder hohes Anspruchsniveau zuzuweisen. Als Beispiel können wieder die nordrhein-westfälischen Vorgaben zum „Kleinen Latinum“ gelten:¹⁰ Dort heißt es, man müsse dafür „lateinische Originaltexte im sprachlichen Schwierigkeitsgrad inhaltlich *mittelschwerer* Stellen“ erfassen und übersetzen können; und weiter: „Als mittelschwere Texte können beispielsweise Texte auf dem Niveau Caesars (*De bello Gallico*), Cornelius Nepos oder leichtere Stellen aus Ciceros Briefen gelten.“ Eine vermeintliche Präzisierung dessen, was als „mittelschwer“ zu gelten hat, ist aus Anwendersicht löblich, aber dennoch wird im Ergebnis so getan, als ob beispielsweise eine Caesar- oder Nepos-Textstelle grundsätzlich „leichter“ sei als z. B. eine Ovid- oder Tacitus-Stelle. Dass es aber natürlich immer auf die konkrete Passage ankommt und dass „leicht“ und „schwierig“ auch nicht näher definiert werden, ist nicht nur didaktisch, sondern etwa auch prüfungsrechtlich gesehen äußerst unbefriedigend.

9 Baden-Württemberg: Ergänzungsprüfung Latinum Schule 2016 auf der Grundlage von Cic. *S. Rosc.* 29–34, Thüringen: Ergänzungsprüfung Latinum Schule 2017 auf der Grundlage von Cic. *Flacc.* 3f., Bayern: Ergänzungsprüfung Latinum Schule 2017 auf der Grundlage von Cic. *Lael.* 17f.

10 Siehe Anm. 8 für die Online-Quelle des Zitats.

6. Die Berechnung von Textschwierigkeit nach Friedrich Maier und Karl Bayer

In der lateinischen Fachdidaktik wurden vor ca. 40 bzw. 20 Jahren zwei Versuche unternommen, hier Abhilfe zu schaffen und den Schwierigkeitsgrad eines Textes messbar zu machen. Sie seien im Folgenden kurz vorgestellt, um vor diesem Hintergrund weitere Überlegungen anstellen zu können:¹¹

(1) *Methode F. Maier:* In seinem groß angelegten und breit rezipierten Grundlagenwerk *Lateinunterricht zwischen Tradition und Fortschritt* (3 Bände, Bamberg 1979–1985) stellte Friedrich Maier auch die Ergebnisse einer breit angelegten empirischen Studie dar. Er untersuchte dabei sowohl Schulautoren (Cäsar, Cicero, Livius, Sallust, Seneca, Ovid, Vergil) als auch Abiturklausuren bzw. Klassenarbeiten (aus den Jahren 1901–1971) und erstellte Frequenzstatistiken zu wichtigen syntaktischen Erscheinungen des Lateinischen (z. B. AcI, Gerundivum, Finalsatz etc.). So ermittelte er (aus insgesamt knapp 90.000 Wörtern) für jedes analysierte syntaktische Phänomen den sogenannten „Teubnerseiten-Koeffizienten“, der angibt, wie oft ein Phänomen auf einer fiktiven Seite einer Teubner-Textausgabe (Durchschnittsumfang: ca. 230 Wörter) auftaucht. Auf der Basis dieser statistischen Mittelwerte konnte Maier postulieren, dass ein Prüfungstext von ca. 110 Wörtern durchschnittlich 16,5 SE (= syntaktische Erscheinungen) enthält – beispielsweise 4–5 Gliedsätze, 3–4 Partizipien, 1–2 *nd*-Formen usw. Jeder neue potentielle Prüfungstext muss demzufolge nur noch syntaktisch untersucht, durchgezählt und mit der Verteilungskurve des statistischen Mittelwerts verglichen werden – kommt man auf weniger als 16,6 SE, ist es ein leichter Text, kommt man auf deutlich mehr, ein schwieriger.

(2) *Methode K. Bayer:* In einem (erstmal) 1998 publizierten Aufsatz demonstrierte Karl Bayer ein eigenes Modell zur Schwierigkeitsgradmessung, das sich als substantielle Weiterentwicklung des Modells von Maier versteht.¹² Er benutzt vier Indikatoren, die für ihn Textschwierigkeit abbilden – durch das (mehr oder weniger arbiträre) Festlegen und Zählen von sogenannten „Gewichtungseinheiten“ (GE)

11 Die jüngst veröffentlichten Arbeiten von Bayer (siehe Literaturverzeichnis), die sich ebenfalls mit den Themen Anspruchsniveau und Schwierigkeit von lateinischen (Lehrbuch-)Texten beschäftigen, können hier unberücksichtigt bleiben, weil sie von ganz anderen Voraussetzungen ausgehen und andere Ziele verfolgen.

12 Vgl. Bayer 1998.

innerhalb jeder Kategorie, deren anschließende Summierung und Division durch die Wortzahl des Textes ermittelt Bayer einen skalierbaren Quotienten, den er wiederum einem bestimmten Anspruchsniveau zuweist. Bayer zieht folgende Indikatoren heran: Wortstellung (Kategorie A/B), d. h. für jede Positionsverschiebung, die zwischen lateinischem Ausgangssatz und deutscher Übersetzung in Gedanken vorzunehmen ist, vergibt er 1 GE; Syntax (Kategorie C), d. h. die syntaktischen Erscheinungen, die schon Maier gezählt hatte, nun aber mit jeweils verschiedenen Punktwerten von 1–4 GE bedacht; sowie sogenannte „weitere Indikatoren“ (Kategorie D), ein Sammelbecken für verschiedene andere Phänomene wie Ellipsen, indirekte Rede, besonders lange Sätze, Homonyme, Verben im Potentialis u. a., die ebenfalls mit unterschiedlich vielen GE versehen sind.

Insgesamt ist es auffällig, dass beide Modelle, insbesondere aber die ausdifferenzierte Berechnungsmethode von Bayer, keinen größeren Niederschlag in der lateinischen Fachdidaktik gefunden haben. Die Verfasser dieses Beitrags sind Mitglieder einer seit 2015/2016 tätigen, sächsisch-thüringisch-niedersächsischen Arbeitsgruppe aus Fachlehrerinnen und -lehrern, Universitätsforschern und ministeriellen Vertretern, die sich im Zuge der Etablierung einer besseren Vergleichbarkeit der jeweiligen Latinumsprüfungen auch mit den Möglichkeiten und Grenzen einer objektiven Einschätzung des Schwierigkeitsgrades von Prüfungstexten befasst hat.¹³ Innerhalb dieser Arbeitsgruppe wurden die Modelle von Maier und Bayer ausprobiert, doch letztlich konnten beide nicht überzeugen. Dies liegt vor allem daran, dass sie nur einen sehr begrenzten Ausschnitt derjenigen Faktoren einbeziehen, die wir für die Konstitution des Schwierigkeitsgrades eines Textes für maßgeblich halten. Obwohl man sich über die vielen Grenzen eines solchen Ansinnens im Klaren sein muss – etwa, dass es keine absolute Textschwierigkeit geben kann, weil jeder Rezipient individuell ist, oder dass sich solch abstrakte Kategorien wie Textverstehen und Schwierigkeitsgrad von vornherein einer mathematischer Berechenbarkeit entziehen –, glauben wir, dass es dennoch möglich ist, mithilfe angepasster Untersuchungsparameter zu einer tiefgründigeren und zumindest valideren Abschätzung kommen zu können.

¹³ Weitere Mitglieder dieser Arbeitsgruppe sind: Anna Philina Burmester (Dresden), Cornelia Eberhardt (Erfurt), Dr. Jörg Freydank (Meißen), Beate Hertel (Saalfeld), Clemens Krause (Gerstungen), Katrin Lucius (Dresden), Gabriele Ott-Großmann (Dresden), Ulrike Rüpke (Erfurt) und Arend Smid (Dresden).

7. Weiterentwicklung

Aus diesen Gründen wurde ein neues Analysemodell für Textschwierigkeit entwickelt, dessen Grundzüge hier im Folgenden vorgestellt werden. Dabei ist gleich zu Beginn darauf hinzuweisen, dass sich das vorliegende Modell noch im Arbeitsstadium befindet und noch breiter empirisch getestet werden muss – wir rufen hier gern zur weiteren Mitarbeit auf. Doch auch im jetzigen Zustand ist das Modell schon benutzbar; dabei berücksichtigt es Aspekte, die bislang nur sporadisch oder gar nicht in den Blick genommen wurden.

Der ursprüngliche Plan bestand in einer Weiterentwicklung der Methode Bayers, die erweitert und vertieft werden sollte. Als grundsätzlich überzeugend schätzten wir nämlich den Versuch ein, aus einer möglichst objektiven Beurteilung eines Textes heraus einen skalierbaren Quotienten zu ermitteln, so dass wir dieses Ziel übernahmen. Auch die Zuweisung von bestimmten Punkten zu bestimmten Erscheinungen griffen wir auf (wobei wir Bayers Begriff „Gewichtungseinheit“ für diese Schwierigkeits-Punkte vorerst auch selbst verwenden). Andererseits erschien die Konzentration auf die Textsyntax (Maier) bzw. zusätzlich auf Wortstellung und ausgesuchte weitere Phänomene (Bayer) als zu begrenzt, um Textschwierigkeit valide zu erfassen. Deshalb wurden auf Basis von Erkenntnissen aus Leseforschung, Sprachwissenschaft, Fremdsprachendidaktik, Lernpsychologie und empirischen Forschungsarbeiten der Lateinidaktik neue übergeordnete Untersuchungsparameter sowie konkrete Items entwickelt, die in der Schwierigkeitsanalyse berücksichtigt werden sollten. Als (bis heute) problematisch erwies sich im Laufe des bisherigen Entwicklungsprozesses nicht nur die komplexe Festlegung auf einen Katalog alter und neuer Kriterien, sondern gerade auch die Beibehaltung einer praktikablen Handhabung. Im Laufe der letzten drei Jahre wurde das Modell deshalb sukzessiv diskutiert, geändert und evaluiert; auch eine wissenschaftliche Einbettung wurde von Beginn an angestrebt: Die Arbeitsgruppe stellte ihr Modell auf Tagungen und Kongressen vor, an der TU Dresden sowie den Universitäten Göttingen und Leipzig war es schon mehrfach Gegenstand fachdidaktischer Seminare und Übungen, und auch Abschluss- und Hausarbeiten dazu wurden vergeben. Weitere Projekte sind bereits geplant; ein Ende dieses Arbeitsprozesses ist auch deshalb, wie gesagt, noch nicht eindeutig abzusehen.

Konkret ruht das neu entwickelte Modell (grafische Abbildung s.u.) auf fünf Säulen, aus denen sich die Textschwierigkeit hinreichend genau ergibt:

- (A) Lexik
- (B) Morphologie
- (C) Syntax
- (D) Textstruktur
- (E) Pragmatik

Innerhalb dieser Kategorien werden über mehrere Items Punkte bzw. Gewichtungseinheiten (GE) verteilt, die am Ende summiert und durch die Gesamtwortzahl geteilt werden. Der so ermittelte Quotient beschreibt als absoluter Wert den Schwierigkeitsgrad des Textes; derzeit ist es aber noch zu früh, nähere Angaben zu einer bestimmten Skalierung zu machen, weil hierfür noch eine längere Testphase nötig ist.

Während Kategorie C (Syntax) recht genau dem entspricht, was Maier und Bayer vorgegeben hatten, handelt es sich bei den übrigen um neue, eigene Indikatoren, die in dieser Form bisher noch nicht zur Messung von Schwierigkeitsgraden herangezogen wurden. Im Rahmen dieses Beitrags kann nun nicht jedes Item innerhalb der fünf Säulen ausführlich vorgestellt und diskutiert werden; stattdessen mögen ein paar übergreifende Hinweise ausreichen: In Kategorie A (Lexik) wird grundsätzlich jedes lateinische Wort mit einem Punkt (GE) für seine spezielle Bedeutung versehen. Zusätzlich aber, und hierauf liegt das eigentliche Augenmerk, wird u. a. evaluiert, wie viele Bedeutungsangaben der Lernende bei ein und demselben Wort berücksichtigen muss (je mehr, desto höhere Verpunktung), und ob die konkret erforderliche Bedeutungsangabe an Position 1 im zugrunde liegenden Verzeichnis steht (falls nicht, mehr GE).¹⁴ In Kategorie B (Morphologie) werden Pronomina, Steigerungsformen und *verba anomala* innerhalb eines Textes verpunktet; besonders wichtig aber ist, dass jede mehrdeutige bzw. homografe Form entsprechende Zuschläge bekommt und dass bei Verben jede Abweichung von einem definierten Standard (= 3. Person Indikativ Präsens Aktiv) gezählt wird. Die schon erwähnte Kategorie C greift die wichtigsten Phänomene der lateinischen Syntax auf und weist ihnen jeweils einen Wert von 1–4 GE zu. In Spalte D (Textstruktur) werden etwa Konnektoren und Prädikate eigens gezählt und verpunktet; auch Ellipsen (vor allem auch solche des Subjekts) sollen hier Berücksichtigung finden, wobei gerade dieses Item

¹⁴ Damit werden die jüngsten empirischen Forschungsergebnisse zum Ablauf des Übersetzungsprozesses bei Schülerinnen und Schülern berücksichtigt (vgl. v.a. Florian 2015).

noch nicht endgültig geklärt ist. Kategorie E (Pragmatik) greift den Umstand auf, dass lateinische Texte oftmals realienkundliches Wissen erfordern, das viele Leser nicht automatisch zur Verfügung haben – je nachdem, ob ein Isomorph oder Allomorph (eigen- bzw. interkulturell übertragbares oder fremdkulturelles Konzept) im Text vorliegt, sollen hier deshalb GE vergeben werden. Zusätzlich sollen alle abstrakten (d. h. nicht-gegenständlichen) Begriffe einen Punkt erhalten, um somit wenigstens grob den Abstraktionsgrad eines Textes messen zu können; auch hier besteht noch weiterer Klärungsbedarf.

Allen diesen fünf Kategorien vorgeschaltet ist die Untersuchung eines Textes hinsichtlich seiner Satzlängen, wobei im Einklang mit lernpsychologischen Forschungsergebnissen grundsätzlich davon ausgegangen wird, dass ab jeweils 6–7 Wörtern die durchschnittliche Aufmerksamkeitsspanne des Arbeitsgedächtnisses überlastet wird; daher erscheint in diesem Fall für den betreffenden Satz ein entsprechender GE-Zuschlag gerechtfertigt.

Noch einige Hinweise zur Nutzungsweise und zum Adressatenkreis: Das Modell ist (zumindest zum jetzigen Zeitpunkt) so ausgelegt, dass nicht zwangsweise alle Kategorien mitsamt allen Items angewendet werden müssen, um einen Text messen. Dies wäre natürlich optimal, aber auch mit großem Zeit- und Notationsaufwand verbunden, der – abgesehen vielleicht von Prüfungskommissionen und Lehrbuchautorenrunden – für eine „normale“ Lehrerin/einen „normalen“ Lehrer im Unterrichtsalltag kaum zu leisten ist. In der vorliegenden Fassung soll und kann das Modell vorrangig dabei helfen, didaktisierte Texte der Sekundarstufe I einzuschätzen, d. h. vom Anfangsunterricht bis etwa zur Latinumsprüfung. Deziert nicht geeignet erscheint es uns (derzeit) für anspruchsvolle Originallektüre, im Besonderen für lateinische Dichtung – hier müssten noch spezielle, eigene Messindikatoren gefunden werden. In Bezug auf den Anfangsunterricht wird umgekehrt unmittelbar deutlich, dass nicht alle fünf Spalten von Anfang an gleiche Bedeutung haben: Morphologie, Syntax, aber auch Textstruktur und Pragmatik werden erst im Laufe der Lehrbucharbeit zu signifikanten Problemfeldern; gleich zu Beginn jedoch – und dann fortwährend – sind die Vokabelbedeutungen relevant. Mit anderen Worten: Das Modell sollte von links nach rechts mit sukzessiver Hinzunahme weiterer Kategorien genutzt werden – es ist ausreichend und sachdienlich, zunächst nur Spalte A zu benutzen, bevor die Folgespalten später hinzukommen. Wir plädieren hier darüber hinaus, obwohl dies noch nicht ausreichend empirisch getestet werden konnte, für ein flexibles Vorgehen: Es macht durchaus Sinn, auch zu Beginn der Spracherwerbsphase ggf. einzelne Items aus den Spalten B–E zur Messung zu A hinzunehmen, um bestimmte, als besonders relevant eingeschätzte Schwierig-

keiten abbilden zu können (z. B. B6: mehrdeutige/homographe Formen; C7: Attribute und Prädikativa, besonders bei nicht direkt erkennbarer KNG-Kongruenz; D2: Prädikate etc.).

Hingewiesen sei noch auf zwei wichtige Aspekte: Erstens, das Modell sieht in der Praxis keinerlei Kompensationen oder Punkterlasse für sich wiederholende Phänomene vor. Unabhängig davon, um welches Item es sich handelt, soll so berücksichtigt werden, dass sich der Kontext immer wieder (ggf. auch nur leicht) ändert; es kann also nicht einfach davon ausgegangen werden, dass eine Form oder ein Phänomen beim zweiten oder dritten Auftauchen grundsätzlich leichter bzw. einfacher verständlich ist als beim ersten Mal. Genauso ist das Modell nicht darauf ausgelegt, in irgendeiner Form das vorherige Unterrichtsgeschehen, einleitende Hinweistexte, unterstützende Visualisierungen oder unter dem Text befindliche Vokabelhilfen in die Berechnung einfließen zu lassen. Selbstverständlich wird ein Text in einem gewissen Sinne „leichter“, wenn bestimmte Phänomene im vorherigen Unterricht drei Wochen lang thematisiert wurden oder wenn unter dem Text bestimmte Phänomene angegeben oder entlastet werden. Das Modell aber ist darauf ausgelegt, möglichst objektiv und valide zu messen und dabei davon auszugehen, dass der Lernende unmittelbar nur mit dem Text allein zu tun hat (Paratexte übrigens werden sehr oft schlicht gar nicht wahrgenommen, selbst wenn man explizit darauf hinweist).

Unter dem Modell (s. Folgeseite) befindet sich zweitens ein Kasten mit einer Auswahl weiterer Items, die potenziell Eingang finden könnten. Damit soll ein weiteres Mal verdeutlicht werden, dass es sich tatsächlich um einen Arbeitsstand handelt – die aufgeführten Punkte (und sicherlich noch andere, nicht genannte) müssen noch diskutiert, präzisiert und entschieden werden.

8. Kriterienkatalog zur Berechnung des Schwierigkeitsgrades von lateinischen Texten

Erarbeitet von und © bei der niedersächsisch-sächsisch-thüringischen Arbeitsgruppe „Weiterentwicklung Latein“ (Stand: April 2019) – Zielperspektive: Rechnerische Ermittlung eines möglichst objektiven Wertes zur Erfassung der Textschwierigkeit, gespeist aus folgenden Parametern (GE = Gewichtungseinheit):

Übergeordnetes Kriterium:

Satzlängenzuschlag: Satzlänge > 6 Wörter: 1 GE; Satzlänge > 12

Wörter: 2 GE;

Satzlänge > 18 Wörter: 3 GE usw.

(ein Satzende wird auch nach Doppelpunkt und Semikolon angesetzt)

A. Lexik	<p>A1) Jedes Wort: grundsätzlich 1 GE</p> <p>A2) Bei mehreren Bedeutungsangaben zu demselben Wort: 2 Bedeutungen: + 1 GE; 3 Bedeutungen: + 2 GE; mehr als 3 Bedeutungen: + 3 GE</p> <p>A3) Bei mehreren Bedeutungsangaben: Steht die im Text erforderliche Bedeutung im zugrunde gelegten Vokabelverzeichnis (Lehrbuchphase) NICHT an der ersten Position: +1 GE</p> <p>A4) Nachgestelltes Wort, das mit dem vorangehenden verschmolzen ist (z. B. <i>-que, -ne, -cum</i>): + 1 GE</p>
-----------------	--

<p>B. Morphologie</p>	<p>B1) Andere Form als 3. Person: 1 GE</p> <p>B2) Anderes Tempus als Präsens (auch bei Infinitiven und Partizipien): 1 GE</p> <p>B3) Anderer Modus als Indikativ: 1 GE</p> <p>B4) Anderes Genus verbi als Aktiv (auch bei Infinitiven und Partizipien): 1 GE</p> <p>B5) Pronomen: 1 GE</p> <p>B6) Mehrdeutige und homographie Form: 1 GE</p> <p>B7) Form eines verbum anomalum (auch Komposita), d. h. konkret: <i>esse</i> (mit Ausnahme der hochfrequenten Formen <i>est, sunt, esse, velle, nolle, malle, fieri, posse, ire, ferre</i>: 1 GE</p> <p>B8) Steigerungsform: 1 GE</p>
<p>C. Syntax</p>	<p>C1) Präpositionslose semantische Kasusfunktion (z. B. abl. instr., gen. qual.): 1 GE</p> <p>C2) Wortstellung – Objekt an erster Stelle im Satz/Kolon: 1 GE</p> <p>C3) Subjekts- und Objektsinfinitiv: 1 GE</p> <p>C4) Nebensatz (Relativsatz, indirekter Fragesatz, Subjunktionalsatz): 2 GE; bei Nebensatz zweiter Ordnung: + 1 GE</p> <p>C5) Relativer Satzanschluss: 2 GE</p> <p>C6) Satzwertige Konstruktion (NcI/AcI, P.C., Abl. abs., nd-Form): 4 GE; mit darin eingeschlossenem Satzglied: 5 GE</p> <p>C7) Attribut und Prädikativum (Apposition und Prädikatsnomen eingeschlossen): 1 GE;</p>

	<p>wenn aber die Kongruenz zum Bezugswort NICHT auf den ersten Blick erkennbar ist (d. h. keine identischen Wortendungen): + 1 GE; bei gesperrtem oder proleptischem Genitivattribut: + 1 GE</p>
<p>D. Textstruktur</p>	<p>D1) Konnektor (d. h. satzverbindende Konjunktion, Partikel, Adverb): 1 GE</p> <p>D2) jedes Prädikat: 1 GE</p> <p>D3) Ellipsen (auch solche des Subjekts; bei mehreren Prädikaten im Satz nur, wenn sie durch ein oder mehrere Satzglieder getrennt sind; Ausnahme: <i>ait, inquit</i>, fehlendes <i>dicit</i> bei der Einleitung von direkter Rede): 1 GE</p> <p>D4) Hyperbaton (bei Nomina): 1 GE</p>
<p>E. Pragmatik</p>	<p>E1) Pro Begriff, Konzept, Hintergrund(wissen) etc., der/das NICHT unmittelbar für SuS in der Gegenwart verständlich bzw. zugänglich ist, werden vergeben: Isomorph (eigenkulturelles bzw. interkulturell übertragbares Konzept): 1 GE Allomorph (fremd-kulturelles, nicht übertragbares Konzept): 2 GE</p> <p>Beispiele für Isomorphie: <i>Africa, Germani, amphora, triclinium, forum, imperator</i></p> <p>Beispiele für Allomorphie: <i>Stoici, Epicurei, religio, virtus, fatum, ara, deus</i></p> <p>E2) Abstraktum: + 1 GE</p>

Fakultative/zu diskutierende Zusätze (Auswahl):

- Signifikanter Stammwechsel bei Verben im Perfekt-Stamm (auch: PPP/PFA; signifikant = alles außer *v-* und *u-*Perfekt): 1 GE
- Veraltete oder ungebräuchliche deutsche Bedeutungsangaben (z. B. *solere*: „pflegen“; *esse*: „gereichen zu“; *servare*: „wahren“): 1 GE
- Kohärenzbrüche und Leerstellen (irritierende inhaltliche Brüche in der Gedankenführung; unerwartetes Auslassen von Informationen): 1 GE

In der Praxis lässt sich das Modell am ehesten so handhaben, dass man den zu untersuchenden Text Wort für Wort in eine Excel-Tabelle einträgt, um dann die einzelnen Spalten durchzugehen und GE/Punkte zu verteilen; am Ende der Tabelle lassen sich über entsprechende Formeln Teilsummen und Quotient dann automatisiert berechnen. Es ist durchaus normal bzw. gewollt, dass einzelne Wörter/Satzteile mehrere Punkte aus unterschiedlichen Spalten auf sich vereinen – in der Gesamtschau lässt sich dann an solchen GE-Kumulationen innerhalb der Tabelle gut erkennen, wo der Text besondere Schwierigkeiten bietet – und ggf. noch verändert werden sollte, bevor er Unterrichts- oder Prüfungsgegenstand wird.

9. Vorläufige Ergebnisse

Die ersten Ergebnisse im Umgang mit dem Schwierigkeits-Modell lassen zwar noch keine pauschalen Einschätzungen zu, sind aber in jedem Fall ermutigend. Hier ist insbesondere eine unveröffentlichte Seminararbeit¹⁵ von Matthias Theißen (Göttingen) zu nennen, der neun mehrfach erprobte mündliche Prüfungstexte für das Kleine Latinum analysierte (6 x Caesar, 3 x Cicero): Dabei zeigte sich, dass der mithilfe unseres Modells ermittelte Quotient in acht von neun Fällen eindeutig mit der Experten-Einschätzung des erfahrenen Prüfers (und Texterstellers) korrelierte, der seinen Texten im Vorfeld die Etiketten „leicht“, „mittel“ und „schwer“ zuweisen sollte. Zusätzlich fiel auf, dass die untersuchten Cicero-Texte im rechnerischen Durchschnitt etwas leichter waren als die Caesar-Texte, wobei sich das Schwierigkeitsprofil

¹⁵ Vgl. Theißen 2018.

allerdings auch markant unterschied: Während die Syntax (Spalte C) bei Caesar etwas anspruchsvoller war, kamen bei Cicero mehr GE aus den Spalten Lexik und Textstruktur (Spalte A/D).

Weiterhin haben sich bereits mehrere kleine Arbeitsgruppen mit den Texten von fünf frühen Lektionen (jeweils L.6–10) aus aktuellen Lehrbüchern (*Prima nova*, *Via mea* und *Roma*) beschäftigt und diese mithilfe des Modells analysiert. Während im Fall von *Roma* hierbei eine signifikante Progression festgestellt werden konnte (L.6: Quotient 3,0; L.7: 3,5; L.8: 3,66; L.9: 3,94; L.10: 4,47), waren die Ergebnisse bei den anderen beiden Lehrbüchern nicht so eindeutig; vielmehr schwankte der Quotient hier teilweise recht stark oder pendelte unverändert um einen Wert von 3,3–3,6.

Im Fall der oben genannten Cicero-Latinumstexte aus dem Jahr 2017 (die hier aus datenschutzrechtlichen Gründen nicht veröffentlicht werden dürfen), dies sei zuletzt noch erwähnt, ergab eine Analyse mittels unseres Modells, dass die drei Texte in den Spalten B (Morphologie) und C (Syntax) interessanterweise praktisch keine signifikanten Unterschiede in den GE-Teilsummen zeigten. Anders sah es hingegen in den anderen Kategorien aus – insbesondere in den Spalten D (Textstruktur) und E (Pragmatik) differierten die Texte statistisch auffällig, was letztlich zu einem recht eindeutigen Ranking führte: Spitzenreiter war der Latinumstext aus Thüringen mit einem Quotienten von 6,66; ein deutlich höherer Wert als die Texte aus Bayern (6,14) und Baden-Württemberg (5,99). Die Gründe für diese Divergenz liegen in mehreren unterschiedlichen Aspekten, die hier nur angerissen werden können: Beispielsweise erwies sich die baden-württembergische Aufgabe in der Textstruktur signifikant weniger komplex als die thüringische mit ihren zahlreichen Ellipsen und ihren langen Sätzen. Gegenüber dem baden-württembergischen Text wiederum sammeln die anderen beiden Texte erkennbar mehr GE in Spalte E (Pragmatik), was v.a. an dem verwendeten politischen Fachvokabular liegen dürfte. In puncto Lexik (Spalte A) schließlich ist der thüringische Text, wie die Analyse ergab, klar komplexer, da hier u. a. besonders viele Vokabelbedeutungen nicht an erster Position im Wörterbuch (verwendet wurde der Neue Stowasser¹⁶) zu finden waren.

Insgesamt scheint sich der mithilfe unseres Modells ermittelte Quotient bislang also in einem Messbereich von ca. 2,5 (Anfangsunterricht) bis ca. 7,0 (Latinumsprüfung) zu bewegen; diese Skalierung muss allerdings noch validiert und präzisiert werden.

16 Vgl. Lošek 2016.

10. Fazit

Abgesehen von konkreten Zahlenwerten sind zwei wichtige vorläufige Ergebnisse festzuhalten: Erstens zeigen sehr viele unserer Analysen, dass es teilweise erstaunlich hohe Unterschiede zwischen Texten gibt, die eigentlich keine haben sollten. Mit anderen Worten, wir sind immer wieder erstaunt, wie unähnlich verschiedene Texte für dieselbe Zielgruppe/Prüfungsform in ihrem Schwierigkeitsgrad sind. Insofern hilft es enorm, und hierin liegt das zweite übergeordnete Ergebnis, dass das vorliegende neue Modell in hohem Maße Sensibilität dafür schafft, was genau an einem Text schwierig ist. Auch wenn die Feinjustierung und Endskalierung noch nicht abgeschlossen ist, hilft es aufgrund seiner systematischen Herangehensweise ungemein dabei, das so unpräzise, subjektive Bauchgefühl in der Einschätzung des Anspruchsniveaus eines Textes vielleicht noch nicht völlig auszuschalten, aber es kriterienorientiert und nachvollziehbar zu hinterfragen.

Nun ist klar, dass in puncto Konzeption, Handhabung und Verwertbarkeit des Modells noch einige Arbeit zu leisten ist. Dennoch haben wir den Eindruck gewonnen, bereits zu diesem Zeitpunkt ein hochinteressantes und nützliches Werkzeug zur didaktischen Analyse lateinischer Texte in der Hand zu haben. Über hilfreichen Input jeglicher Art ist die Arbeitsgruppe deshalb dankbar; gern rufen wir dazu auf, unser Modell zu testen und uns Rückmeldungen zu geben.

Literaturverzeichnis

- Bayer, Karl, Zur Bestimmung des Schwierigkeitsgrades lateinischer Prüfungstexte: *Anregung* 44/4 (1998) 228–241 [erneut veröffentlicht in *Pegasus online* III/2 (2003) 1–19].
- Beyer, Andrea, Lateinische Lehrbuchtexte und ihr Anspruchsniveau: *Latein und Griechisch in Berlin und Brandenburg* 62 (2018) 73–81.
- , *Das Lateinlehrbuch aus fachdidaktischer Perspektive. Theorie, Analyse, Konzeption* (Heidelberg: Universitätsverlag Winter, 2019).
- Florian, Lena, *Heimliche Strategien: Wie übersetzen Schülerinnen und Schüler?* (Göttingen: Vandenhoeck & Ruprecht, 2015).

- Lichy, Nick, *Latinum (un)gleich Latinum? Untersuchungen zur Vergleichbarkeit der schriftlichen Latinumsprüfungen ausgewählter Bundesländer der Bundesrepublik Deutschland* (Masterarbeit TU Dresden, 2016).
- Lošek, Fritz (Hg.), *Stowasser. Lateinisch-deutsches Schulwörterbuch. Begründet von J. M. Stowasser, M. Petschenig, F. Skutsch. Völlige Neubearbeitung* (München: Oldenbourg Schulbuchverlag, 2016).
- Maier, Friedrich, *Lateinunterricht zwischen Tradition und Fortschritt* (3 Bände) (Bamberg: C. C. Buchner, 1979–1984).
- Sächsisches Staatsministerium für Kultus (Hg.), *Lehrplan Gymnasium Latein 2004 / 2007 / 2009 / 2011 / 2019* (Dresden: Saxoprint, im Druck).
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, *Vereinbarung über das Latinum und das Graecum (Beschluss der Kultusministerkonferenz vom 26. Oktober 1979)* (Bonn: Selbstverlag, 1979).
- , *Vereinbarung über das Latinum und das Graecum (Beschluss der Kultusministerkonferenz vom 22. September 2005)* (Berlin: Selbstverlag, 2005).
- Theißen, Matthias, *Die Schwierigkeit mit der Schwierigkeit. Bemessung des Schwierigkeitsgrades mündlicher Prüfungstexte des Kleinen Latinums und Latinums* (Seminararbeit Georg-August-Universität Göttingen, 2018).