

Wie viel Grammatik braucht das Germanistikstudium?

Roland Schäfer und Ulrike Sayatz

roland.schaefer@fu-berlin.de

Deutsche und niederländische Philologie, FU Berlin

Abstract: Typically, a majority of students in BA and MA programs in German language and literature studies (*Germanistik* or *Deutsche Philologie*) at German universities are future school teachers of German, not future linguists. For decades, linguists have been discussing how German grammar and linguistics should be taught to these students. Furthermore, it is often lamented that first-year students have not acquired sufficient skills in descriptive German grammar in their secondary education (*Sekundarstufe*), and that this stands in the way of teaching them German linguistics effectively. Our study makes an empirical contribution to this discussion. We assessed the proficiency of 220 BA students in a test composed of grammar exercises designed for students between the sixth and tenth grade. We show that, contrary to some earlier assessments, they achieve quite good results overall. However, it also becomes apparent from the results that studying linguistics does not significantly improve their performance in the test. Furthermore, exercises which require free analysis combining structural and functional aspects are handled significantly less well than more mechanical one-dimensional exercises. First, we argue that this is actually an ideal point of departure for deeper, more theory-driven linguistic analysis in university courses. Second, we propose that linguistics departments should continually perform similar assessments in order to be able to adapt their teaching methods and content to the previous skills and the intended career of the majority of their students.

Stichwörter Grammatikvermittlung, Linguistik im Studium, Schulgrammatik, Studieneingangstests

1 Grammatik in Schule und Studium

Dass *vom Schulwissen nichts hängen bleibt*, ist ein Gemeinplatz, und wir sind überzeugt, dass es nicht zielführend ist, sich im Kontext der universitären Lehre mit diesem Phänomen zu intensiv oder gar emotional zu befassen. Ein Großteil der deutschen Studienanfängerinnen und Studienanfänger wird Schwierigkeiten mit der Ableitung einer kubischen Funktion, der Beschreibung von Myelinscheiden und Ranvier-Schnürringen oder der Einordnung der Bedeutung der Magna Charta im historischen Kontext haben, genauso wie mit der grammatischen Analyse eines deutschen Satzes (unabhängig von dessen Komplexität). Sind Studierende einmal immatrikuliert, sind jedoch das gewählte Fach und seine Vertreterinnen und Vertreter für eine Ausbildung verantwortlich, die den Inhalten des Faches und den Berufszielen der Studierenden angemessen ist. Jedes germanistische Institut kann und sollte selbst ermitteln, wie diese auszugestalten ist, statt die schlechte schulische Vorbildung zu beklagen. Dies gilt insbesondere angesichts der Tatsache, dass eine grundsätzliche Änderung der Lehrpläne für den Schulunterricht ggf. länger dauert als eine akademische Laufbahn. Linguistinnen und Linguisten in germanistischen Instituten bilden die zukünftigen Lehrerinnen und Lehrer im Fach Deutsch aus, und sie sind damit primär für ihre Ausbildungsqualität verantwortlich. Wir machen in diesem Artikel einen Vorschlag, wo und wie eine empirische Überprüfung des Vorwissens für eine Optimierung der universitären Studieninhalte und Lehrstrategien der germanistischen Linguistik zumindest teilweise ausgestaltet werden kann. So hoffen wir, die seit Jahrzehnten andauernde Diskussion um Grammatik und Lehrerinnen- und Lehrerausbildung im Hochschulstudium (Boettcher 1994; Eisenberg 2004 usw.) mit empirisch gewonnenen Daten bereichern zu können. Wir sind davon überzeugt, dass diese Diskussion im Zentrum der germanistischen Linguistik geführt werden sollte und nicht ausschließlich in der Fachdidaktik – und damit außerhalb des Blickfelds vieler Fachvertreterinnen und Fachvertreter.

Wir steigen – entsprechend unserer Ziele – allerdings nicht in die seit Jahren kontrovers geführte Diskussion um das Für und Wider von Grammatikunterricht in der Schule an sich ein (Gornik 2003; Ossner 2007: 162; Häcker 2009; Bredel 2014: 226–243). Uns geht es um den Stellenwert von Grammatik in der universitären germanistischen Ausbildung unter Berücksichtigung der mehrheitlichen Ausbildungsziele der Studierenden – und dabei unter spezieller Berücksichtigung der zum jetzigen Zeit-

punkt realen Vermittlungsaufgaben, mit denen diese Studierenden später konfrontiert werden. Dass wir mehrheitlich die Lehramtsstudierenden und somit künftigen Deutschlehrerinnen und Deutschlehrer (einschließlich derer im Grundschullehramt) im Blick haben, motiviert sich von selbst, denn über 70% der Studierenden im Fach Deutsche Philologie an der Freien Universität Berlin haben mit genau dieser Option das Fach gewählt (s. Auswertung in Abschnitt 3). An anderen Universitäten mit ähnlichem Profil dürften die Verhältnisse vergleichbar sein.

Dass die universitäre Lehre im Fach fernab von den praktischen beruflichen Anforderungen des Lehramts sei, wird von Studierendenseite oft behauptet und in Seminaren diskutiert. Dabei betreffen die von Studierenden geäußerten Bedenken häufig gar nicht genuin Fragen des linguistischen Fachs, sondern pädagogische, didaktisch-methodische und soziokulturelle Begebenheiten des schulischen Alltags. So werden beispielsweise Schwierigkeiten beim Schriftspracherwerb in der Grundschule und mangelnde Schreibkompetenzen in den weiterführenden Schulstufen häufig unter dem Gesichtspunkt zu großer Schulklassen, eines hohen Anteils von Lernenden mit nicht-deutscher L1 oder auch standardfernen Schreibgebrauchs in den neuen Medien diskutiert.¹ Aus Sicht der linguistischen Ausbildung ist es hingegen wichtig, wie gut die Studierenden z.B. (um eins der genannten Probleme aufzugreifen) darauf vorbereitet werden, die grammatischen Grundlagen der Schreibprinzipien des Deutschen zu verstehen und damit auch vermitteln zu können. Was von dem dafür notwendigen Wissen haben die Studierenden in den germanistischen Seminaren erwerben können? Wie erfolgt der Transfer dieses Wissens in die schulische Lehrtätigkeit? Nach einer bundesweiten Befragung von 1.017 Lehrkräften und Referendaren im Jahr 2013 fühlen sich 48% der Befragten „durch ihre Ausbildung (Studium und Referendariat) nicht hinreichend vorbereitet, Grammatik zu unterrichten“ (Topalovic und Dünschede 2014: 76). Bemerkenswerterweise fiel demgegenüber die Selbsteinschätzung der Befragten hinsichtlich ihrer Kompetenzen in den linguistischen Teildisziplinen Morphologie, Orthographie, Phonologie, Semantik, Pragmatik und Syntax mehrheitlich positiv aus, mit absteigender Tendenz in dieser Reihenfolge.² In der Syntax fühlen sich

¹ Wir berichten dies aus unserem langjährigen Lehralltag, halten unsere Erfahrungen großenteils aber für generalisierbar.

² In unserem Verständnis ist *Orthographie* keine linguistische Teildisziplin. Dies wäre die *Graphematik* (Eisenberg 2013:285–288).

laut Umfrage 18% unsicher und 1% sehr unsicher (Topalovic und Dünschede 2014: 77).

Sieht man die grundlegenden Aufgaben im Deutschunterricht in der Vermittlung, Entwicklung und Beurteilung sprachlicher Fähigkeiten von Schülerinnen und Schülern, so ist zu fragen, welches Wissen über Sprache – und nur dies ist unser Thema – Lehrpersonen dafür haben und benötigen. Mit Eisenberg (2004: 4) ist dies ein Wissen darüber, welche sprachlichen Anforderungen in sprachgebundenen Aufgaben stecken. Schülerinnen und Schüler sollten also im Deutschunterricht lernen, „ihren eigenen Sprachgebrauch in ein Verhältnis [...] zu den Möglichkeiten der Sprache überhaupt zu setzen“ (Ossner 2007: 167). Ein Sprachgebrauch, der die reflexiven Möglichkeiten von Sprache nutzt, erfolgt durch die Einsicht in fundamentale sprachliche Regularitäten und in die Strukturiertheit sprachlicher Phänomene (Portmann-Tselikas 2011: 72). Dieses „durch ‚Reflexion über Sprache‘ und ‚Transfer von explizitem zu implizitem Wissen‘“ geforderte besondere Verhältnis zur Sprache können nach Eisenberg (2004: 23) jedoch „die Schüler nicht entwickeln, wenn es die Lehrer [sic!] nicht haben.“ Damit sind wir wieder bei der Kernfrage nach dem in der universitären Lehre zu vermittelnden Wissen über Sprache angelangt. Folgt man nun der Argumentation Eisenbergs (2004), ist dieses notwendige Wissen zunächst im engeren Sinne ein Wissen um sprachliche Formen und deren Funktionen und damit ein grammatisches Wissen. Die wissenschaftliche Grammatik, die dazu an Universitäten gelehrt werden sollte, ist unabhängig von den im schulischen Kontext relevanten normativen Aspekten gerade nicht präskriptiv, sondern deskriptiv. Es geht also nicht darum, den richtigen Sprachgebrauch zu definieren, sondern historisch gewachsene grammatische Konventionen innerhalb von Sprechergemeinschaften, inklusive möglicher Variationen und Varianten zu beschreiben. Dazu bedarf es in erster Linie empirischer Verfahren (Schäfer 2016: 32–36) sowie einer Methode, um von den sprachlichen Daten zu grammatischen Generalisierungen zu gelangen (Eisenberg 2004: 11).

Ob für den oben beschriebenen Ansatz der Grammatikvermittlung nach Eisenberg nun die schulischen Lehrmaterialien optimal geeignet sind, ist eine Frage, der wir uns nicht widmen, da der Einfluss der germanistischen Institute auf Erstellung und Auswahl dieser Materialien ohnehin viel zu gering ist. Optimal – gemäß Eisenbergs (2004) Sichtweise – ausgebildete Lehrerinnen und Lehrer werden unabhängig von der Ausrichtung und Qualität des real existierenden Unterrichtsmaterials in jedem Fall besser und souveräner in der Lage sein, ihre Lehraufgabe zu erfüllen, als nicht opti-

mal ausgebildete. Dazu müssen sie (1) schlicht und einfach in der Lage sein, die den Schülerinnen und Schülern gestellten Aufgaben selbst zu verstehen und zu lösen. Außerdem müssen sie (2) diese Aufgaben auf größere System- und Funktionszusammenhänge beziehen können. Schließlich sind (3) spezifische methodisch-didaktische Fähigkeiten gefragt, wie das grammatische Wissen je nach Alters- und Leistungsstufe der unterrichteten Schülerinnen und Schüler zu vermitteln ist. Während die unter (3) genannte Aufgabe eindeutig der Fachdidaktik Deutsch zu überlassen ist, sind wir – im Sinn unserer einleitenden Kommentare – nicht der Ansicht, dass die unter (1) genannte Fähigkeit von Studierenden bereits aus ihrer eigenen Schulzeit ins Studium mitgebracht werden muss, und damit etwa nur (2) die Vermittlungsaufgabe der germanistischen Linguistik wäre.³ Die germanistische Linguistik ist unserer Ansicht nach besser beraten, (1) anhand einer konstanten und genauen empirischen Evaluation der Vorkenntnisse der Studierenden diese aktiv in ihren Kompetenzvermittlungsauftrag einzubeziehen und davon ausgehend in einen größeren grammatisch-linguistischen Funktions- und Systemkontext überzuleiten.⁴ Unser Artikel leistet einen empirischen Beitrag zur Systematisierung genau dieser Vorgehensweise. Der dazu entwickelte Fragebogen wird in Abschnitt 2 beschrieben und vor dem Hintergrund der aktuellen Diskussion um sogenannte *Eingangstests* motiviert. Die Auswertung und Diskussion

³ Mit einem solchen Anspruch scheint 2006/2007 ein Eingangstest in Bayern durchgeführt und medienwirksam präsentiert worden zu sein. Laut diversen dramatisierenden Medienberichten (prominent Spiegel Online 2007; SZ Online 2007) wurde damals u.a. festgestellt, dass die Schule nicht ausreichend Grammatikwissen für Studierende der Germanistik vermittele. Zumindest die journalistische Rhetorik ist brachial und demütigend: In einem „Grammatik-Fiasko“ sind Studierende „mit Karacho durchgefallen“ (Spiegel Online), die „Germanistik-Studenten [sic!] haben sich blamiert“, und viele von ihnen sind „Grammatik-Nieten“ (SZ Online). Die „Professoren [sic!] sind erschüttert“ (Spiegel Online), und es muss folglich „in der Schule wieder mehr Grammatik gepaukt“ werden (SZ Online). Das ist – nach allem, was wir gesagt haben – konzeptuell, methodisch und (zumindest in der journalistischen Aufbereitung) rhetorisch ein Ansatz, den wir zu überdenken vorschlagen. Die Universität bildet die Lehrerinnen und Lehrer aus, und ein erhebliches Optimierungspotential liegt daher bei ihr selber. Der Verweis auf die bereits in Boettcher (1994) aufgebrachte Rede vom „Teufelskreis“ zwischen Lehrerinnen- und Lehrerausbildung und Schülerinnen- und Schülerausbildung liegt auch heute noch nah (insbesondere Boettcher 1994:178–181).

⁴ Genau dies ist laut Medienberichten beim bayrischen Test nicht geschehen. In Spiegel Online (2007) heißt es: „Konsequenzen hatte der Test für die Studenten [sic!] nicht. Sie bekamen Literaturhinweise und Lehrmaterialien an die Hand, mit denen sie ihre Wissenslücken stopfen sollten.“

der Ergebnisse erfolgt in Abschnitt 3. Schlussfolgerungen für die universitäre Praxis ziehen wir abschließend in Abschnitt 4.

2 Entwicklung des Tests

In der vorliegenden Studie untersuchen wir das grammatische Basiswissen von Studierenden des Studienganges Deutsche Philologie der Freien Universität Berlin in Form eines Fragebogens bzw. Tests. Über mögliche Funktionen solcher Evaluationen wird unter dem Etikett des *Eingangstests* schon länger diskutiert, rezent zum Beispiel im Jahr 2016 in einem Themenheft von *Didaktik Deutsch*. Nach Schindler (2016: 16) können dabei drei mögliche Funktionen unterschieden werden. In „individual-diagnostischer Funktion“ geben Tests „darüber Auskunft [...], was die Studierenden schon können und was sie bislang noch nicht beherrschen“ und leiten entsprechende Handlungsempfehlungen wie den Besuch bestimmter zusätzlicher Veranstaltungen ab. Weiterhin fungieren Tests im Sinne genereller Zulassungstests (Schindler 2016: 16; Bremerich-Vos 2016: 9–12). Eine dritte Funktion sieht Schindler (2016: 16) in Tests, die auf Studien- und Berufsinhalte bezogen sind und bei Studierenden das Bewusstsein für die fachlichen Inhalte wecken und schärfen. Jede dieser Funktionen ist je nach Situation gut begründbar. Unsere Evaluation bezieht sich vor allem auf die erste und dritte Funktion. Eine wesentliche, von Schindler (2016) nicht genannte Funktion ist allerdings die Optimierung von universitären Schwerpunkten und Vermittlungsmethoden sowie eine Sensibilisierung des universitären Lehrpersonals für seine eigene Vermittlungsaufgabe. Unser Test bzw. unsere Studie ermittelt, wie souverän Studierende im gesamten BA-Studium (also nicht nur im ersten Semester) mit grammatischen Aufgabenstellungen aus schulischen Lehrwerken umgehen. Zweck ist es, Anregungen für die längerfristige Planung der Lehrinhalte in der germanistischen Linguistik zu geben und Vorschläge für die methodische Ausrichtung der universitären Grammatikvermittlung zu unterbreiten. Während der Stoff des Studiums zweifelsohne eine Obermenge des Schulstoffs sein muss, ist die Beherrschung dieses Basiswissens eine Grundvoraussetzung für eine erfolgreiche Tätigkeit im Lehramt (s. Abschnitt 1). Wir würden also einen positiven Effekt des Studiums auf die Leistung der Studierenden im gegebenen Test erwarten.

Da die große Mehrheit unserer Studienteilnehmerinnen und Studienteilnehmer in Berlin und Brandenburg zur Schule gegangen ist (s. Abschnitt 3), wählten wir tatsächliche Aufgaben aus aktuell im Land Berlin zugelassenen und verwendeten Schulbü-

chern der Klassenstufen sechs bis zehn.⁵ Entscheidend war das Anforderungsprofil des tatsächlichen schulischen Grammatikunterrichts, den die Studierenden absolviert haben.⁶ Dabei ist anzumerken, dass nur Aufgabe 5 aus der Grundschule stammt. Alle anderen Aufgaben umfassen ein Spektrum von Klasse 7 bis 10, die jedoch teilweise als Übungs- und Wiederholungsaufgaben auch grammatisches Grundschulwissen abbilden. Die Aufgabenstellungen decken nach unserer Einschätzung eine Schnittmenge von schulischen und universitären Anforderungen ab.

Der Test beinhaltet insgesamt elf Aufgaben,⁷ wobei Aufgabe 8 (Relativpronomina) auch ohne terminologisches Grammatikwissen von Erstsprecherinnen und Erstsprechern problemlos lösbar ist und uns lediglich anzeigen sollte, ob die Probandinnen und Probanden die Aufgaben konzentriert bearbeitet und ernstgenommen haben.⁸ Die Aufgaben wurden in der Anredeform zum *Sie* verändert und waren so nicht vordergründig als Schulbuchaufgaben erkennbar. Inhaltlich und terminologisch gab es keine Veränderungen, obwohl in einigen Fällen (Aufgaben 3, 7 und 11) nach unserer Einschätzung terminologische Ungenauigkeiten in den Aufgabenstellungen Fehler provozieren. In einem Fall (Aufgabe 3) musste die Originalaufgabe gekürzt werden, da in der Quelle Präpositionalphrasen fehlerhaft als Nominalphrasen markiert waren und somit ein Sachfehler vorlag.

Für die Auswahl der zehn Testaufgaben waren die nachfolgend diskutierten Kriterien entscheidend. Die Aufgaben entsprechen zentralen Schwerpunktthemen gemäß dem Bildungsstandard für den Deutschunterricht an Berliner Schulen.⁹ Außerdem thematisieren sie nach Feilke (2012) notwendiges Sprachwissen in Bezug auf bildungssprachliche Kompetenzanforderungen. So ist ein Zusammenhang von sprachlichen

⁵ Die genauen Quellen werden am Ende des Artikels genannt.

⁶ Hier liegt ein wesentlicher Unterschied zum bayrischen Test von 2006/2007, bei dem nach Medienauskunft „[d]ie Aufgaben [...] aus Sicht der Prüfer [sic!] nicht schwer [waren]“ (SZ Online 2007). Unser Test stellt den Studierenden Aufgaben, die ihnen so oder ähnlich sehr wahrscheinlich in ihrer Schulzeit schon gestellt wurden. Wir haben also nicht ohne eigene Erfahrung im Gymnasiallehramt entschieden, was angesichts einer normalen Schulbildung einfach oder schwer sein könnte.

⁷ Der Testfragen selbst sind im Anhang abgedruckt. Eine detailliertere Diskussion einzelner Aufgabenstellungen erfolgt in Abschnitt 3 bei der Ergebnisdarstellung.

⁸ In einem Lückentext mussten hierbei an Stellen Relativpronomen eingesetzt werden, an denen ohnehin nur Relativpronomina stehen konnten. Grammatisches Wissen um Relativpronomina im eigentlichen Sinn ist dafür freilich nicht erforderlich.

⁹ <http://bildungsserver.berlin-brandenburg.de/unterricht/rahmenlehrplaene/bildungsstandards/>

Merkmale der Bildungssprache mit Gebrauchsfunktionen in Bildungskontexten zu formulieren. Insofern handelt es sich um relevantes Sprachwissen, das in sprachgebundenen Aufgaben genutzt werden soll. Nach Feilke (2012: 8–9) sind z.B. sprachliche Mittel wie komplexe Adverbiale (Aufgaben 2, 4), Attribute (Aufgabe 5) und komplexe Sätze (Aufgabe 7) in Gebrauchsfunktionen wie *Explizieren* relevant. Komplementär dazu dienen Aufgaben zu Nominalisierungen (Aufgabe 9), komplexen Nominalphrasen (Aufgabe 3, 11) sowie komplexen Wortbildungen wie Komposita und Derivata (Aufgabe 10) unter dem *Inhaltsaspekt einer Aussage* dem *Verdichten* (Feilke 2012: 8). Sprachliche Mittel zum Ausdruck von Sprecherabsichten wie *Diskutieren* sind nach Feilke (2012: 9) u.a. Sprechaktverben und Konjunktiv (Aufgabe 6), während agensabgewandte Konstruktionen wie z.B. Passiv (Aufgabe 1) eine sprachliche Möglichkeit für sprecherseitiges *Verallgemeinern* darstellen. Die gemeinsame Klammer aller Aufgaben war die Beziehung von Form und Funktion sprachlicher Einheiten und damit sprachliche Mustererkennung und Musterdifferenzierung. Das Anforderungsprofil wurde bei ähnlichem Thema zum Teil bewusst im Schwierigkeitsgrad variiert. So verlangt Aufgabe 2 eine funktionale Analyse bei vorgegebenen Konstituenten, Aufgabe 4 dagegen eine funktionale Analyse vorgegebener Konstituenten (Präpositionalphrasen) mit Unterscheidung von valenzgebundenen Objekten und adjunkten Adverbialen. Am anspruchsvollsten in dieser Aufgabengruppe sind dagegen die Aufgaben 5 und 11 einzustufen. Aufgabe 5 beinhaltet eine funktionale Zuordnung (Attribut) zu sprachlichen Formen bei freier Konstituentenanalyse. Ebenso verlangt Aufgabe 11 neben einer Formanalyse (Kasus) relativ zur Objektfunktion eine Segmentierung der Konstituenten. Die einzige explizit normpositivistische Aufgabe ist Aufgabe 6 zum Konjunktiv, während die graphematische Segmentierung von Wortformen in Aufgabe 9 (ein klassisches Aufgabenformat der Grundschule, hier aus einem Arbeitsheft der 9. Klasse) implizit die orthographische Normbeherrschung morphosyntaktischer Schreibungen im Deutschen (Substantivgroßschreibung sowie Getrennt- und Zusammenschreibung) überprüft. Die Aufgaben zur Mustererkennung und Musterdifferenzierung sind sowohl als Analyseaufgaben (z.B. Aufgabe 1 zur Diathese und Aufgabe 10 zu den Wortfamilien) als auch als Produktionsaufgaben (z.B. Aufgabe 7 zu Nebensätzen) formuliert. Alle Aufgaben können systematisch auf im Studium der Germanistik relevante linguistische Teildisziplinen (Morphologie, Syntax und Graphematik) gemäß grundlegender Modulinhalte bezogen werden.

Jede Aufgabe sollte zudem von den Studierenden auf Vierpunktskalen hinsichtlich ihrer Verständlichkeit (von *sehr gut verständlich* bis *sehr schlecht verständlich*) sowie ihres Schwierigkeitsgrades (von *sehr schwierig* bis *sehr leicht*) evaluiert werden. Am Anfang des Fragebogens wurde von den Studierenden eine Selbstbeurteilung erbeten, in der sie ihre Vorbildung in deutscher Grammatik auf einer Fünfpunktskala (von *sehr gut* bis *sehr schlecht*) bewerten sollten.

Alle Tests wurden in der dritten und vierten Woche des Sommersemesters 2016 im Rahmen linguistischer Lehrveranstaltungen des Bachelorstudienganges Deutsche Philologie der Freien Universität Berlin auf freiwilliger Basis unter Aufsicht der beiden Koautorinnen und Koautoren sowie einer Kollegin und einem Kollegen durchgeführt. Die Testbedingungen wurden weitgehend einheitlich gehalten. Alle teilnehmenden Studierenden hatten zunächst zehn Minuten Zeit, Metadaten einzutragen und sich einen Überblick über die Aufgaben zu verschaffen. Für die Bearbeitung standen dann 20 Minuten zur Verfügung.

3 Auswertung im Detail

In diesem Abschnitt werden die Ergebnisse der Auswertung des Fragebogens präsentiert und diskutiert. Dabei sehen wir unseren Schwerpunkt darin, konkrete Probleme zu benennen und konstruktiv Ursachen für ein eventuelles nicht perfektes Abschneiden der Teilnehmerinnen und Teilnehmer zu suchen. Die Daten werden offen zur Verfügung gestellt, um weitere Analysen durch Dritte zu ermöglichen.¹⁰

In Abschnitt 3.1 wenden wir uns zunächst den Eckdaten der Gruppe der Teilnehmenden zu. In Abschnitt 3.2 diskutieren wir die allgemeinen Ergebnisse und liefern in Abschnitt 3.3 eine Fehleranalyse für die einzelnen Aufgaben. Abschnitt 3.4 widmet sich schließlich den möglichen Effekten, die die universitäre Lehre auf das Abschneiden bei unserer Befragung hat.

¹⁰ Alle abgegebenen Fragebögen stehen eingescannt als PDF zur Verfügung, ebenso unsere Musterlösung, die Tabellen mit den Auswertungen im MS Excel- und LibreOffice-Format, sowie die von uns zur Auswertung verfassten Skripte für das Statistikpaket R (R Core Team 2016) und die rohen R-Ausgaben (Grafiken und Text): <https://github.com/rsling/germanistikstudium>

3.1 Beschreibung der Stichprobe

Insgesamt wurden 220 Fragebögen ausgewertet und 14 Fragebögen aussortiert. Aussortiert wurden vor allem Fragebögen von Austauschstudierenden, die ursprünglich als Referenzgruppe dienen sollten. Die Vergleichs- bzw. Referenzgruppe stellte sich schließlich aber als zu heterogen und klein heraus. Außerdem wurden Fragebögen ausgeschlossen, bei denen nicht sichergestellt werden konnte, dass Deutsch eine der Erstsprachen der Teilnehmerin oder des Teilnehmers war.

Als Geschlecht wurde 173 mal *weiblich* (78,6%), 44 mal *männlich* (20%), zweimal *transgender* (0,9%) und einmal *nichts davon/keine Angabe* (0,5%) angekreuzt. Die Teilnehmerinnen und Teilnehmer wurden außerdem befragt, in welchen Bundesländern sie zur Schule gegangen sind (Mehrfachnennungen möglich), und Tabelle 1 zeigt das deutliche Ergebnis: Die Mehrheit der Studierenden ist in Berlin und Brandenburg zur Schule gegangen. Eine detaillierte Auswertung der Ergebnisse von Studierenden nach einzelnen Bundesländern haben wir wegen der kleinen Gruppengrößen außerhalb von Berlin-Brandenburg nicht durchgeführt.

Bundesland	Anzahl der Nennungen	Prozent
Berlin	136	57,4
Brandenburg	40	16,9
Nordrhein-Westfalen	11	4,6
Niedersachsen	9	3,8
Schleswig-Holstein	6	2,5
Rest (jeweils ≤ 5)	35	14,8

Tabelle 1. Herkunft der Teilnehmerinnen und Teilnehmer nach Bundesland (Mehrfachnennungen möglich).

Weiterhin wurde gefragt, was die Erstsprachen der Teilnehmerinnen und Teilnehmer sind. Insgesamt haben 46 Studierende einen mehrsprachigen und 174 einen einsprachigen L1-Hintergrund. Die am häufigsten genannten L1-Sprachen außer Deutsch waren Türkisch (23), Englisch (7), Arabisch (6) und Russisch (4).

Relevant ist auch die Verteilung der Studienfächer. Hier zeigt sich, dass insgesamt genau 75% der Studierenden einen Lehrberuf anstreben: Grundschullehramt 83

(37,7%), Deutsch Lehramt 82 (37,3%), Deutsch (nicht Lehramt) 51 (23,2%) und andere Fächer 4 (1,8%). Das Fach lebt also (zumindest in der universitären Lehre) von den Lehramtsstudierenden.

3.2 Gesamtbewertung des Ergebnisses

Zur Auswertung wurden die Fragebögen von uns eigenhändig in Tabellen erfasst, wobei die Antworten der Teilnehmerinnen und Teilnehmer nur abgebildet und nicht bewertet wurden. Die eigentliche Bewertung erfolgte nachträglich und automatisch durch einen Abgleich mit einer von uns erstellten Musterlösung in R. Der gesamte Auswertungs- und Bewertungsprozess kann also jederzeit nachvollzogen werden. Bei der Bewertung wurden alle Aufgaben (und innerhalb der Aufgaben die Aufgabenteile) gleich gewertet, da wir keinen Anlass zur Vermutung hatten, dass für einzelne Aufgaben überproportional längere oder kürzere Bearbeitungszeiten zu erwarten wären und die erwartete Bearbeitungszeit als neutralste Gewichtunggrundlage ansahen. Die Ergebnisse der trivialen Aufgabe 8 wurden im Gesamtergebnis allerdings nicht berücksichtigt (s. Abschnitt 2), so dass die zehn verbleibenden Aufgaben mit jeweils 10% gewichtet wurden. Somit kann sehr einfach für jede Aufgabe und den gesamten Fragebogen der Anteil der richtigen Antworten (im Folgenden stets in Prozentwerten angegeben) ermittelt werden.¹¹ Eine Umrechnung des Gesamtergebnisses in akademische Noten (nach der aktuellen Tabelle im Bachelor Deutsche Philologie an der Freien Universität Berlin) von eins bis fünf ermöglicht nun eine erste intuitive Interpretation, die in Abbildung 1 in Form einer Verteilung der Noten gegeben wird.

¹¹ Gar nicht und falsch erledigte Aufgaben wurden einheitlich mit 0 Punkte gewertet. S. auch Tabelle 2.

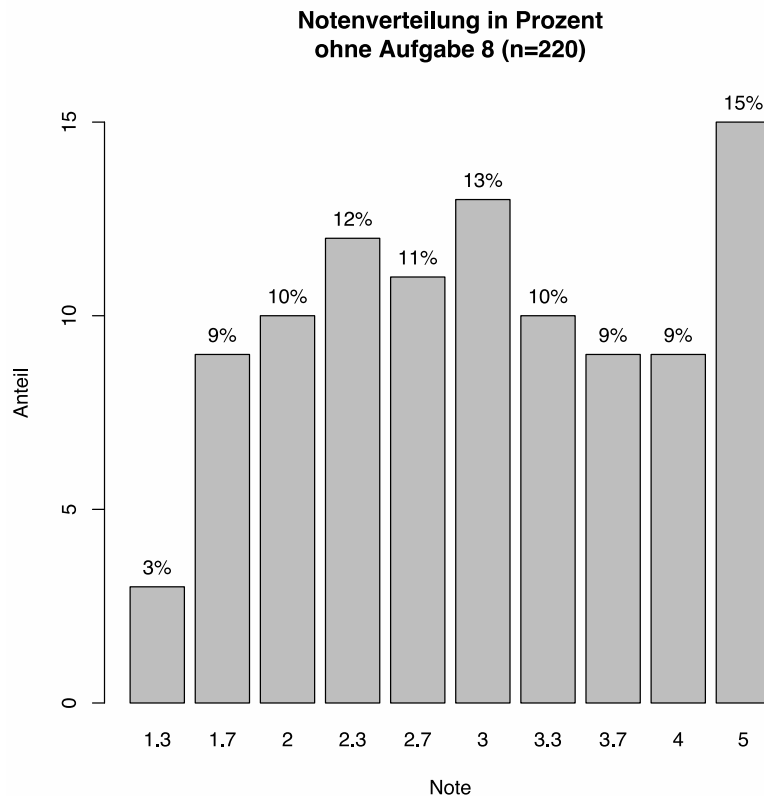


Abbildung 1. Verteilung der Gesamtergebnisse in akademische Noten (1–5).

Wir weisen zunächst darauf hin, dass es bezüglich der erreichten Prozentpunkte keine Hinweise auf Unterschiede zwischen weiblichen (Mittel 67%, Median 67,6%) und männlichen (Mittel 68,7%, Median 73,6%) Teilnehmenden gibt. In einem exakten Wilcoxon-Mann-Whitney-Test kann die H_0 (Lageidentität der Gruppen) nicht zurückgewiesen werden ($p=0,4$).¹² Das Gleiche gilt für die Gruppen mit einsprachigem L1-Hintergrund (Mittel 67,7%, Median 69,5%) und mehrsprachigem L1-Hintergrund (Mittel 65,7%, Median 67,7%): $p=0,31$ in einem exakten Wilcoxon-Mann-Whitney-Test. Als Maß für die Zuverlässigkeit der Erfassung der Fragebögen werten wir die Lagegleichheit der Ergebnisse der Fragebögen, die die beiden Koautorinnen und Koautoren dieses Artikels erfasst haben. Die Mediane liegen bei 69,5% und 69,2%; exakter Wilcoxon-Mann-Whitney-Test ($p=0,3$).

Die Studierenden schneiden also insgesamt (gemäß Abbildung 1 und ohne besonderen Einfluss von Geschlecht oder sprachlichem Hintergrund) recht gut ab, und es kann überhaupt nicht davon gesprochen werden, dass die schulische Ausbildung ohne Effekt geblieben wäre. Ein Viertel der Studierenden liegt bei 2,3 oder besser (erstes

¹² R-Funktion `coin::wilcox_test` mit exakter Verteilung. Am Ende des Artikels findet sich eine kurze Erläuterung der Testverfahren.

Quartil), ein Viertel bei 3,7 oder schlechter (drittes Quartil), der Median liegt bei 3,0. Nur 15% hätten den Test nicht bestanden, wenn es sich um eine Prüfung gehandelt hätte.¹³ Das insgesamt gemischte Ergebnis kann aber nur in Verbindung mit einer detaillierten Betrachtung der erzielten Ergebnisse in den einzelnen Aufgaben konstruktiv interpretiert werden. Diese Betrachtung folgt jetzt.

¹³ Zum bayrischen Test hieß es in Spiegel Online (2007): „[J]eder [sic!] Dritte kassierte eine Fünf oder eine glatte Sechs.“ Dabei ist zusätzlich zu bedenken, dass in unserem Test bestimmte Aufgaben wahrscheinlich wesentlich besser erledigt worden wären, wenn eine simple Begriffsklärung erfolgt wäre. Im Fall von Aufgabe 10 war z.B. der Begriff *Wortfamilie* ein offensichtliches Problem, der im übrigen auch vielen von uns befragten Fachkolleginnen und Fachkollegen nicht geläufig war, die keinen expliziten Schwerpunkt in der Didaktik haben. Grundsätzlich mangelnde Vorkenntnisse sind also auf keinen Fall zu attestieren.

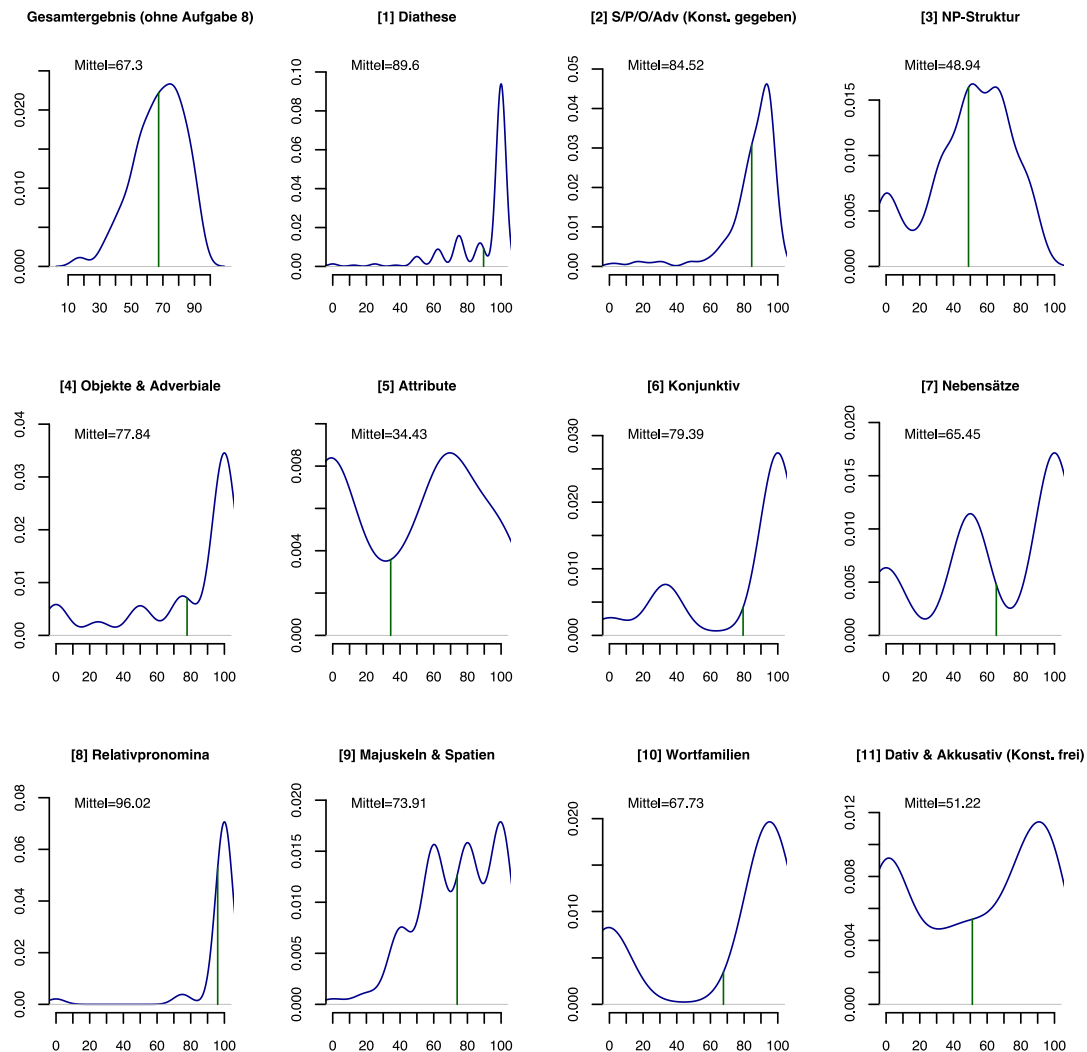


Abbildung 2. Verteilungsdichte (y-Achse) der erreichten Prozentpunkte (x-Achse) insgesamt und pro Aufgabe (Kerndichteschätzungen der empirischen Verteilungen). Die senkrechte Linie markiert den Mittelwert.

Abbildung 2 zeigt zunächst die empirische Verteilung (Kerndichteschätzungen) der erreichten Prozente für das Gesamtergebnis sowie für die einzelnen Fragen. Es ist offensichtlich nicht der Fall, dass die Aufgaben insgesamt mangelhaft oder mittelmäßig gelöst wurden. Allerdings zeigt sich generell die Tendenz zu bi- oder multimodalen Verteilungen, so dass es Gruppen von sehr kompetenten und deutlich weniger kompetenten (aber nicht unbedingt *schlechten*) Studierenden gibt. Aufgabe 1 zur Diathese (Mittel bei 89,6%), Aufgabe 2 zu relationalen syntaktischen Begriffen (84,5%) und Aufgabe 8 (Testaufgabe) zu Relativpronomina (96%) werden überwiegend sehr gut gelöst (Mittelwerte über 80%). Die Lösungen zu Aufgabe 4 zu Adverbialen und Ob-

jekten (77,8%), Aufgabe 6 zum Konjunktiv (79,4%) und Aufgabe 9 zu Majuskeln und Spatien (73,9%) können ebenfalls als überwiegend gut gelten (Mittelwerte zwischen 70% und 80%). Als grenzwertig (Mittelwerte zwischen 60% und 70%) müssen Aufgabe 7 zu Nebensätzen (65,5%) und Aufgabe 10 zu Wortfamilien (67,7%) angesehen werden. Als besonders problematisch (Mittelwerte kleiner als 60%) erweisen sich nur Aufgabe 3 zur Struktur von NPs (48,%), Aufgabe 5 zu Attributen (34,4%) und Aufgabe 11 zu Akkusativ und Dativ bei freier Konstituentenanalyse (51,2%). Zu diesem Ergebnis passt eingeschränkt der Anteil der Aufgaben, die gar nicht erst versucht wurden, s. Tabelle 2.

	Aufgabe	für Klasse	% nicht versucht	mittleres Ergebnis	Problem- Status
1	Diathese	7	0,9	89,6	++
2	S/P/O/adverbiale Bestimmung (Konstituenten vorgegeben)	7	0,5	84,52	++
3	NP-Struktur	9–10	11,8	48,94	–
4	Objekte und Adverbiale	9–10	1,8	77,84	+
5	Attribute	6	25	34,43	–
6	Konjunktiv	9	5	79,3	+
7	Nebensätze	8	3,2	65,45	±
8	Relativpronomina	8	2,3	96,02	++
9	Majuskeln und Spatien	9	0,9	73,91	+
10	Wortfamilien	7	16,8	67,73	±
11	Dativ und Akkusativ (Konstituenten frei)	9–10	20,9	51,22	–

Tabelle 2. Anteil der Aufgaben, die gar nicht erledigt wurden und das mittlere Ergebnis in Prozent für diese Aufgabe. Problemstatus: ++ für 80% oder mehr, + für 70–80%, ± für 60–70% und – für unter 60%.

Wir haben uns also vor allem mit den Aufgaben 3, 5 und 11 sowie in zweiter Annäherung den Aufgaben 7 und 10 zu beschäftigen. Abschnitt 3.3 widmet sich vor allem diesen Aufgaben.

3.3 Detaillierte Analyse der problematischen Aufgaben

In diesem Abschnitt leisten wir eine Fehleranalyse der problematischen Teilaufgaben gemäß Abschnitt 3.2. Die Schlussfolgerungen daraus ziehen wir in Abschnitt 4. In Aufgabe 3 waren die Strukturen der NPs in (1) zu analysieren. Zur Auswahl gestellt wurden die Analysen in (2). Unsere Musterlösung (vor dem Semikolon) und die tatsächlich gegebenen Antworten (nach dem Semikolon) sind in (1) in eckigen Klammern angegeben.

- (1) a. *ein einziger Marktplatz* [b; a: 76, b: 75, d: 37]
- b. *die ganze Welt* [b; a: 60, b: 121, c: 3, d: 3, e: 1]
- c. *die meisten Großunternehmen* [b; a: 115, b: 65, c: 1, d: 5, e: 1]
- d. *ihre Betriebe* [e; a: 1, c: 2, d: 1, e: 188]
- e. *mehr kleine Betriebe* [d; a: 1, b: 1, c: 76, d: 110, e: 1]
- f. *ganz neue Berufsbilder* [c; a: 2, b: 3, c: 87, d: 90, e: 6]
- (2) a. Artikel – Indefinitpronomen – Nomen
- b. Artikel – Adjektiv – Nomen
- c. Adjektiv – Adjektiv – Nomen
- d. Indefinitpronomen – Adjektiv – Nomen
- e. Possessivpronomen – Nomen

Zu (1a) wurde (2a) 76 mal, (2b) 75 mal und (2d) 37 mal zugeordnet. Das Problem ist also vor allem die Differenzierung von Adjektiv und Indefinitpronomen. Da *einzig* nicht als substitutives Pronomen auftreten kann und adjektivisch flektiert, ist die Klassifikation als Adjektiv die plausibelste. Die Ursache für die Unsicherheit ist hier wahrscheinlich in einer Semantisierung des Begriffs *Indefinitpronomen* – im Gegensatz zu einer formbezogenen Betrachtung – zu suchen. Das Gleiche gilt für (1b), für das 60 mal (2a) und 121 mal (2b) angegeben wurde, und (1c) mit 115 mal (2a) und 65 mal (2b). Auch wenn es unsere Überzeugung ist, dass *ihre* in (1d) als Artikel zu klassifizieren ist (Eisenberg 2013: 164, 167–169; Schäfer 2016: 277–278), kommt allein auf Basis der Anzahl der Wörter nur (2e) infrage (188 mal richtig im gegebenen Kontext, nur 4 Fehlantworten). Bei den letzten beiden Strukturen besteht die Unsicherheit wie erwartbar zwischen (2c) und (2d). 76 mal wurde *mehr* als Adjektiv und 110 mal als Indefinitpronomen klassifiziert, *ganz* 87 mal als Adjektiv und 90 mal als Indefinitpronomen. Wie schon bei (1a) und (1c) besteht das Problem also in einer nicht streng

formbezogenen Interpretation von innerhalb der Grammatik rein formal definierten Begriffen wie *Adjektiv* und *Pronomen*.

Bei Aufgabe 5 sollten die Attribute in (3) unterstrichen werden. Wir geben die Musterlösung durch Unterstreichung an.

- (3) Die Inuit, die heute noch auf Jagd gehen, fahren mit schnellen Motorschlitten und kehren in ihre festen Holzhäuser zurück.

Der Prozentsatz derer, die die Aufgabe überhaupt nicht zu lösen versucht haben, ist besonders hoch (25%, s. Tabelle 2). Es ist also davon auszugehen, dass hier häufig ein terminologisches Problem mit dem Begriff *Attribut* bestand. Der Relativsatz wurde 78 mal gar nicht und 30 mal nur teilweise unterstrichen. Nur 53 mal (also von 24,1%) wurde er völlig korrekt unterstrichen. Die Adjektive wurden wesentlich zuverlässiger als Attribute identifiziert (114 und 112 mal vs. 29 und 33 fehlende Unterstreichungen). Zuviel unterstrichen wurden Adverbien von 27, Präpositionen von 26, Substantive von 20, Artikel oder Pronomina von 12 und Verben von 3 Teilnehmerinnen und Teilnehmern. Das Hauptproblem ist also die Nicht-Klassifikation einer komplexen phrasalen Einheit (Relativsatz) als Attribut und damit die Identifikation des funktionalen Attributsbegriffs mit einem kategorialen, nämlich dem des Adjektivs.

In Aufgabe 11 sollten Akkusativobjekte in den (inhaltlich fragwürdigen) Sätzen in (4) einfach und Dativobjekte zweifach unterstrichen werden. Hier wird Teilnehmerinnen und Teilnehmern insgesamt die höchste Freiheit der Analyse gelassen, indem die Konstituentenstruktur, der Objektstatus (relational zum Verb) und der Kasus (relational zum Verb und ggf. morphologisch-formal) erkannt werden müssen.

- (4) a. Leider finden viele nicht sofort einen Ausbildungsplatz.
 b. Ich will den bestmöglichen Schulabschluss erreichen.
 c. Hat mein Wunschberuf eigentlich gute Zukunftsaussichten?
 d. Heutzutage werden den Schulabgängern viel zu wenig Lehrstellen bereitgestellt.
 e. In der Zukunft werden nicht mehr vorwiegend die großen Konzerne die Arbeitsplätze schaffen.
 f. Das wird vielmehr den mittleren und kleinen Betrieben vorbehalten sein.
 g. Kein Industrieland kann sich mehr der Globalisierung der Wirtschaft entziehen.

Tabelle 3 zeigt die Verteilung der Fehler, wobei der Anteil derer, die die Aufgabe gar nicht gelöst haben, mit 20,9% ähnlich hoch lag wie bei Aufgabe 5 mit 25% (vgl. Tabelle 2).

Satz	Erkennung/Segmentierung der NP				Kasusbestimmung	
	komplett	teilweise	gar nicht	zu viel	Akkusativ	Dativ
a	152	10	8	0	159	3
b	137	22	8	2	152	9
c	106	8	50	1	111	4
d	139	7	20	0	22	124
e	110	6	42	6	115	7
f	134	14	16	0	33	115
g	99	37	27	0	23	113

Tabelle 3. Fehlerstatistik für Aufgabe 11.

Die höchste Nichterkennungsquote hat – wegen des Fehlens der kasusmarkierenden Funktion von Determinierern erwartbar – der Akkusativ ohne Determinierer (4c). Ansonsten scheinen die Dativobjekte generell öfter nicht erkannt und öfter falsch als Akkusativ klassifiziert zu werden (4d, 4f, 4g). Falsche Positive sind kein generelles Problem. Phrasen, die (ganz oder teilweise) insgesamt mehr als zehnmal fälschlicherweise markiert wurden, waren in (4c) *mein Wunschberuf* (21 mal Akkusativ, 6 mal Dativ), in (4d) *viel zu wenig Lehrstellen* (39 mal Akkusativ, 3 mal Dativ) und in

(4e) *der Zukunft* (18 mal Akkusativ, 3 mal Dativ). Im Passivsatz (4d) ist sehr wahrscheinlich der Akkusativ des korrespondierenden Aktivs der Distraktor.

In Aufgabe 7 sollten jeweils die beiden unabhängigen Sätze in (5) zu einem Hauptsatz-Nebensatz-Gefüge mit „Konjunktion“ (in der Formulierung der Quelle) umgeformt werden. Hier sind offensichtlich nur Adverbialsätze gemeint.

- (5) a. Es hat geklingelt. Du warst mit dem Referat fertig.
 b. Ich komme morgen in die Schwimmhalle. Ich kann dir beim Training zuschauen.

Bei der Analyse der Fehler ist zu berücksichtigen, dass der etablierte Begriff (*unterordnende*) *Konjunktion* insofern unglücklich gewählt ist, als sich echte Konjunktionen in Koordinationsstrukturen syntaktisch grundsätzlich anders verhalten als Komplementierer (bzw. Subjunktoren), die Nebensätze mit Verb-Letzt-Stellung einleiten (Engel 2009: 61; Schäfer 2016: 186–187, 190). Die Arbeitsanweisung ist also je nach terminologischem Verständnis bereits widersprüchlich. Die Verteilung der gegebenen Antworten wird in Tabelle 4 gegeben.

Satzpaar	Nebensatz (Subjunktoren)	adjunkter Infinitiv	Konnektor/ Konjunktion	Hauptsätze mit Adverb	andere Fehler
(5a)	169	0	26	7	6
(5b)	119	19	24	40	9

Tabelle 4. Verteilung der Antworten für Aufgabe 7.

Bei beiden Satzpaaren überwiegt die Tendenz, eine hypotaktische Struktur zu konstruieren. Dass zu (5b) im Vergleich zu (5a) mit 19 zu 0 Fällen mehr adjunkte *um zu*-Infinitive („um dir beim Training zuschauen zu können“) und mit 40 zu 7 Fällen mehr parataktische Konstruktionen mit konnektivem Adverbial im Vorfeld gewählt wurden („dann kann ich dir beim Training zuschauen“), liegt vermutlich an der Semantik des zweiten Satzpaars, das sich für die Verknüpfung mit üblichen adverbialen Subjunktoren nicht eignet. In beiden Fällen werden parataktische Strukturen mit Konnektoren und Konjunktionen von ca. 10% der Teilnehmerinnen und Teilnehmer gewählt („und ich kann dir beim Training zuschauen“). Generell ist zu schließen, dass die Begriffe *Hauptsatz* und *Nebensatz* (Subjunktoren und Verb-Letzt-Stellung) teilweise

nicht formal aufgefasst wurden, sondern im Sinne einer beliebigen (argumentativen) Satzverknüpfung (z.B. eben auch durch Konstruktionen mit Konnektoren und konnektiven Adverbialen).

In Aufgabe 10 sollten schließlich die Wörter in (6) nach Wortfamilien gruppiert und die Stämme der Wörter unterstrichen werden.

- (6) eröffnen, die Bindung, die Angestellte, das Gestell, anbinden, Offenheit, fröhlich, die Kindheit, kindisch, der Frohsinn

Diejenigen, die die Aufgabe konzeptuell richtig gelöst haben, haben in der Regel die korrekten Wortpaare einander zugeordnet (zwischen 153 und 158 korrekte und zwischen 4 und 5 inkorrekte Zuordnungen). 37 Teilnehmerinnen und Teilnehmer haben die Aufgabe allerdings gar nicht bearbeitet, und 21 haben die Wörter nach Wortklassen geordnet. Es ist also davon auszugehen, dass der Begriff *Wortfamilie* schlicht nicht durchgehend bekannt ist. Die Identifizierung der Stämme haben wir nicht weiter analysiert, bis auf eine systematische Besonderheit bei *Gestell* und *Angestellte*. In diesen (synchron nicht transparent aufeinander bezogenen!) Wörtern haben 73 von 183 (also 40%) der Teilnehmerinnen und Teilnehmer, die die Aufgabe überhaupt beantwortet haben, *gestell* als Stamm unterstrichen. Hier wurde also häufig nach möglichst langen gleichlautenden Wortbestandteilen gesucht, und nicht etwa nach der linguistisch motivierten Kategorie *Stamm*.¹⁴

Die Analyse der wesentlichen bzw. kritischen Fehlerquellen ist damit abgeschlossen. Im Vorgriff auf Abschnitt 4 können wir feststellen, dass Studierenden vor allem in den Bereichen Fehler unterlaufen, die ein geschärftes Verständnis komplexerer Strukturen sowie die Abgrenzung von formaler und funktionaler Grammatik erfordern (Aufgaben 3, 5, 11). Dies sind genau die Bereiche, derer sich die Grammatikausbildung im Germanistikstudium üblicherweise auch annehmen sollte, und die einen idealen Anknüpfungspunkt von Schulgrammatik und weiterführender linguistischer Grammatikbetrachtung böte. Wo terminologische Probleme und einfache Klassifikationsprobleme auftreten (insbesondere Aufgaben 3, 7 und 10), könnte die universitäre Lehre an das vorhandene Wissen anknüpfen und vertiefend diskutieren, wie sich die

¹⁴ Hier muss (entgegen unserer allgemeinen Strategie) im Einzelfall doch kritisiert werden, dass die Feststellung intransparenter etymologischer Beziehungen im schulischen Grammatikunterricht wahrscheinlich fehl am Platz ist.

Schulterminologie begründet, und wo sie scheitert. Das sind ideale Inhalte für das Studium der germanistischen Linguistik im BA für eine Mehrheit, die Lehramt studiert.

In Abschnitt 3.4 analysieren wir die Auswirkung der universitären Lehre auf den Erfolg der Teilnehmerinnen und Teilnehmer. Eine Interpretation und Diskussion der Fehler in einem größeren programmatischen Rahmen erfolgt in Abschnitt 4.

3.4 Effekte des Studienverlaufs

In diesem Abschnitt evaluieren wir den Effekt der universitären Lehre auf die erzielten Ergebnisse. Dazu vergleichen wir in Abbildung 3 die Ergebnisse der Studierenden, die zum Zeitpunkt der Befragung bereits die Einführung in die Linguistik im Rahmen des BA-Studiengangs an der Freien Universität Berlin abgeschlossen hatten, mit den Ergebnissen derer, für die dies nicht der Fall ist.

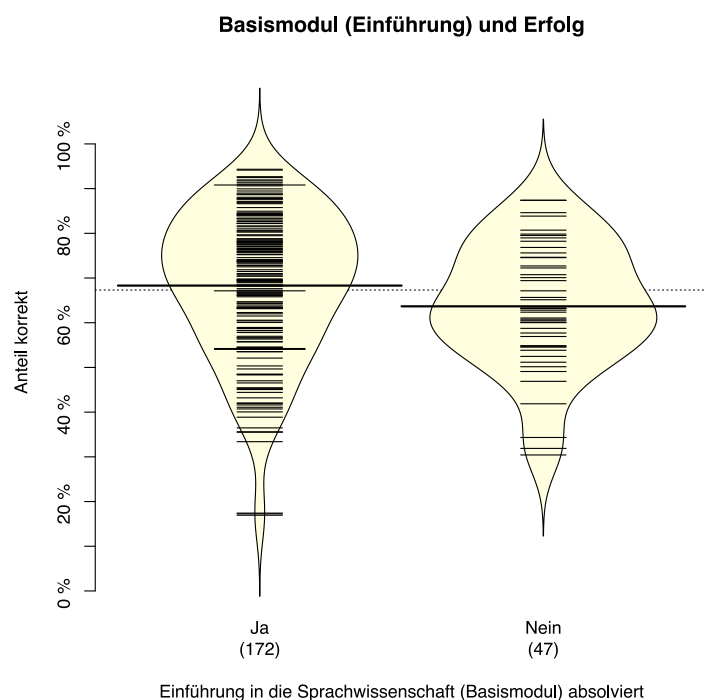


Abbildung 3. Beanplot der Ergebnisse für die Gruppe, die das Einführungsmodul belegt hat und die Gruppe, die es nicht belegt hat.

Deskriptiv ist ein Unterschied zwischen den beiden Gruppen beobachtbar. Das Mittel liegt in der Gruppe mit absolvierter Einführung bei 68,3% (Median 70,1%) und in der Gruppe ohne bei 63,7% (Median 63,2%). Die Differenz von 4,6% (Median 6,9%) ist

bescheiden, und in einem exakten Wilcoxon-Mann-Whitney-Test kann die H_0 (Lageidentität der Gruppen) mit $p=0,035$ zwar zurückgewiesen werden, wenn man einen Schwellenwert von 0,05 zugrunde legt, aber der p-Wert ist nicht sonderlich niedrig. Das zugehörige 95%-Konfidenzintervall für den Lageunterschied ist mit $[0,37..10,42]$ zudem von erheblicher Breite, und insgesamt ist – insbesondere bei einer vorsichtigen Interpretation von Signifikanztest und Konfidenzintervall – marginal davon auszugehen, dass ein (schwacher) Effekt vorliegt. In Abbildung 4 wird daher zusätzlich betrachtet, wie sich die Ergebnisse zwischen dem ersten, zweiten, dritten sowie vierten (oder höheren) Studienjahr unterscheiden.

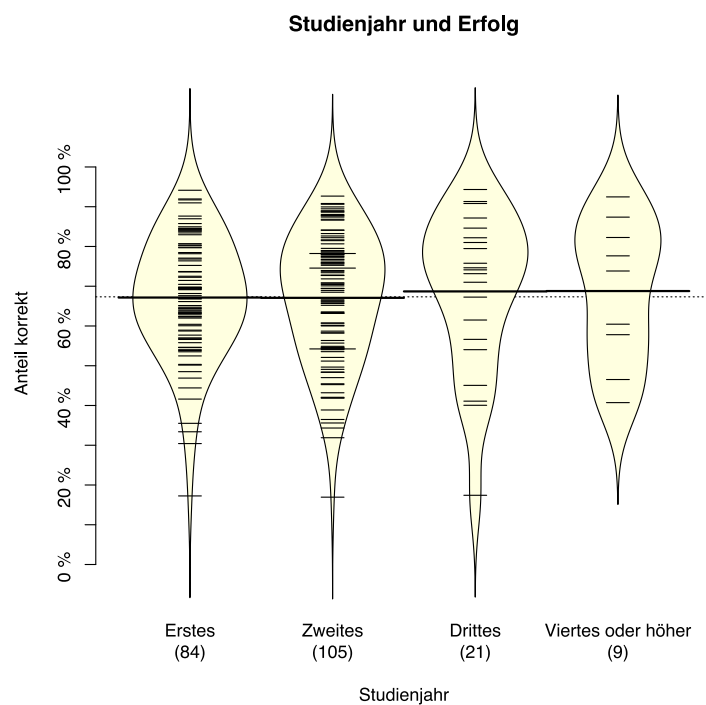


Abbildung 4. Beanplot der Ergebnisse (erzielte Prozent) nach Studienjahren.

In Abbildung 4 wird offensichtlich, dass ein längeres Studium nicht mit einer besseren Beherrschung der Aufgabenstellungen einhergeht. In einem Kruskal-Wallis-Test kann die H_0 (Lageidentität der Gruppen) nicht zurückgewiesen werden ($p=0,475$).¹⁵

¹⁵ R-Funktion `coin::kruskal_test` mit approximativer Monte Carlo-Verteilung (Hothorn et al. 2006). Es ist zu beachten, dass der Test die natürliche Reihenfolge der Klassen (ordinale Skalierung) nicht berücksichtigt. Angesichts dessen, dass schlicht gar kein Unterschied feststellbar ist, ist diese Einschränkung zu vernachlässigen.

Wir haben keine starke Evidenz dafür, dass ein mehrjähriges Studium das Abschneiden im gegebenen Test beeinflusst. Damit wird das ohnehin schwache Ergebnis bezüglich des absolvierten Einführungsmoduls weiter geschwächt. Entweder ist der Effekt nicht existent, oder die Studierenden vergessen langfristig das Gelernte wieder. Für die abschließende Diskussion in Abschnitt 4 halten wir dieses Ergebnis als besonders relevant fest.

4 Diskussion

Die hier vorgestellte erste Untersuchung zeigt, dass die Durchführung grammatischer Tests in der Germanistik einen erheblichen Nutzen für die Planung und Konzeption der universitären Lehre haben könnte – eine Funktion solcher Tests, die bisher wenig diskutiert wurde (vgl. auch Abschnitt 2). Wir haben mit unserem Test ein Anforderungsprofil getestet, das die Studierenden aus ihrer Schulzeit kennen können, und das zugleich zentrale grammatische Inhalte der universitären Lehre berührt. Eine – wie von Schindler (2016) geforderte – abgeleitete Handlungsempfehlung sehen wir aufgrund der erzielten Ergebnisse ausdrücklich nicht in der Etablierung eines Propädeutikums in der Studieneingangsphase. Die Vermittlung von grammatischem Basiswissen stellt eine Kernaufgabe des akademischen Kurrikulums in der Germanistik dar. Dieser Aufgabe muss in der Lehre quantitativ und qualitativ Rechnung getragen werden. Es geht also mit Fuhrhop und Teuber (2016: 15) nicht darum, „Teile der Studienziele zu Voraussetzungen zu erklären“ und diese dann in Propädeutika oder Brückenkursen zu vermitteln sowie in Eingangstests als studienvoraussetzend zu testen.

In Abschnitt 3 wurde festgestellt, dass Studierende (mit erwartbaren Schwankungen) durchaus ein solides Grundwissen in Grammatik und grammatischer Terminologie haben, dass aber das Studium diese Kenntnisse nicht im gewünschten Maß steigert (s. Abbildungen 3 und 4). Suchen wir also nach den Ursachen und fragen zunächst, wie viel an deutschen Universitäten in die sprachwissenschaftliche Lehre investiert wird, so steht die Freie Universität Berlin mit ihrem polyvalenten Bachelor in der Germanistik quantitativ durchaus gut dar. Im Kernfach Deutsche Philologie mit neunzig Leistungspunkten (LP) müssen Studierende mindestens elf Semesterwochenstunden (SWS) mit zwanzig LP absolvieren, davon entfallen zehn LP (fünf SWS) auf die Basisphase und zehn LP (vier SWS) auf die Aufbauphase. Wird in der Vertiefungsphase der Studienbereich Linguistik gewählt, absolvieren die Studierenden im BA nochmals zwanzig LP (acht SWS). Im Maximalfall studieren Studierende der Germanistik an

der Freien Universität Berlin im Kernfach also für vierzig LP (17 SWS) Linguistik. Das gilt für Studierende mit und ohne Lehramtsoption. Ein stichprobenartiger SWS-basierter Vergleich dieser Ausgangslage mit den Linguistik-Anteilen in den Kurrikula anderer deutscher Universitäten zeigt eine insgesamt ausgewogene Gewichtung, wobei deutlich wird, dass die BA-Studiengänge mit Lehramtsoption mitunter deutlich weniger SWS in synchroner germanistischer Linguistik anbieten als polyvalente BA-Studiengänge.¹⁶ Dies betrifft z.B. die Universität Potsdam (Brandenburg) mit maximal sechs SWS im Studiengang *Deutsch Lehramt* (für die Sekundarstufen I und II) und die Johann Wolfgang Goethe-Universität Frankfurt am Main (Hessen) mit maximal zehn SWS im Studiengang *Deutsch Lehramt Gymnasium*. Deutlich mehr investiert hingegen die Ruprecht-Karls-Universität Heidelberg (Baden-Württemberg) mit maximal 22 SWS im Studiengang *Lehramt Gymnasium*. Die Angebote schwanken insgesamt (polyvalent und reines Lehramt) zwischen sechs und 22 SWS. Die Deutsche Philologie der Freien Universität Berlin liegt mit maximal 17 SWS im Mittelfeld zwischen der Christian-Albrechts-Universität zu Kiel (Schleswig-Holstein) im Zweifächer BA mit Profil *Lehramt Gymnasium* mit maximal 16 SWS und der Georg-August-Universität Göttingen (Niedersachsen) im polyvalenten Studiengang *Deutsche Philologie* mit maximal zwanzig SWS.

Für Studierende an der Freien Universität Berlin und anderen deutschen Universitäten ist also prinzipiell genügend Zeit und Raum für die linguistische Ausbildung gegeben. Berücksichtigt man in diesem Zusammenhang die von Topalovic und Dünschede (2014) konstatierte Diskrepanz in der Selbsteinschätzung der Studierenden zwischen erworbener schulgrammatischer Kompetenz und akademischer Kompetenz in den linguistischen Teildisziplinen, so ist zu fragen, warum ein Transfer akademischen Wissens in praktisches Lehrwissen nur unzureichend erfolgt.¹⁷ Wir gehen davon aus, dass die konstante Verbesserung der berufsrelevanten Fähigkeiten der Studierenden also nicht notwendigerweise mehr Ressourcen, sondern eine optimierte Ausrichtung der Inhalte und Lehrmethoden erfordert. Wie diese Ausrichtung aussehen kann, ist keine triviale Frage, aber in jedem Fall sind inhaltliche Fragen betroffen, und es ist

¹⁶ Die genannten Zahlen stammen aus unserer eigenen Auswertung der öffentlich auf den Webseiten der genannten Universitäten zugänglichen Informationen (insbesondere Studienverlaufspläne und Studienordnungen).

¹⁷ Es sei nochmals betont, dass wir hier mit praktischem Lehrwissen nicht die methodisch-didaktischen Kompetenzen meinen, die wir ja auch nicht getestet haben.

außerdem zu überdenken, wie Studierende für die linguistische Ausbildung motiviert und womöglich begeistert werden können. In der linguistisch-germanistischen Lehre geht es sicherlich nicht darum, „wissenschaftlichen Fortschritt unreflektiert auf die Schule abzubilden,“ wie Hennig (2012: 446) bereits richtig im Zusammenhang mit der grammatischen Terminologie für die Schule argumentiert. Zugleich ist aber mit Eisenberg (2004: 22) festzustellen, dass gerade in der Lehrerinnen- und Lehrerausbildung eine falsche Ängstlichkeit gegenüber der Sprachwissenschaft besteht. Nach Eisenberg (2004: 22) hat sich jedoch „die Sprachwissenschaft allgemein und die Grammatik im Besonderen [...] in den vergangenen Jahren so entwickelt, daß wir über viel Wissen verfügen, das unmittelbar von Bedeutung für die Lehrerbildung“ ist. Um dieses grammatische Wissen „selbstbewusst und mit absolut gutem Gewissen zur Wirkung“ (Eisenberg 2004: 23) zu bringen, muss die universitäre Lehre den Gegenstand *Sprache* in den Mittelpunkt stellen und nicht die Linguistik an sich. Eine deskriptive Linguistik ist im besten Fall zunächst theorieneutral und bietet die Basis für einen Erkenntnisprozess von den Daten zu den Generalisierungen und damit auch zu linguistischer Theorie. Die schulischen Aufgaben – zum Beispiel in Form eines Testes wie des unseren – dafür als Ausgangspunkt der Diskussion und Vermittlung zu nehmen, drängt sich nachgerade auf. Durch einen solchen Zugang würde außerdem nicht nur ein wichtiger Anknüpfungspunkt zur Fachdidaktik geschaffen und dem Vorwurf der Praxisferne entgegengewirkt, sondern es würden auch Möglichkeiten eröffnet, Studierende durch diesen Praxisbezug stärker zu motivieren.

Schaut man sich nun die Selbsteinschätzung der Studierenden in der vorliegenden Studie im Verhältnis zu ihrem Erfolg an (vgl. Abbildung 5), so ist festzustellen, dass sie ihr spontan abrufbares grammatisches Wissen im Mittel durchaus richtig und selbstkritisch bewerten, wenn die Streuung auch erheblich ist. Auch hieran kann und sollte die Lehre idealerweise anknüpfen, um Studierenden die Relevanz der linguistischen Ausbildung für ihre konkrete berufliche Zukunft zu vergegenwärtigen, und um daraus letztlich Motivationspotential zu schöpfen.

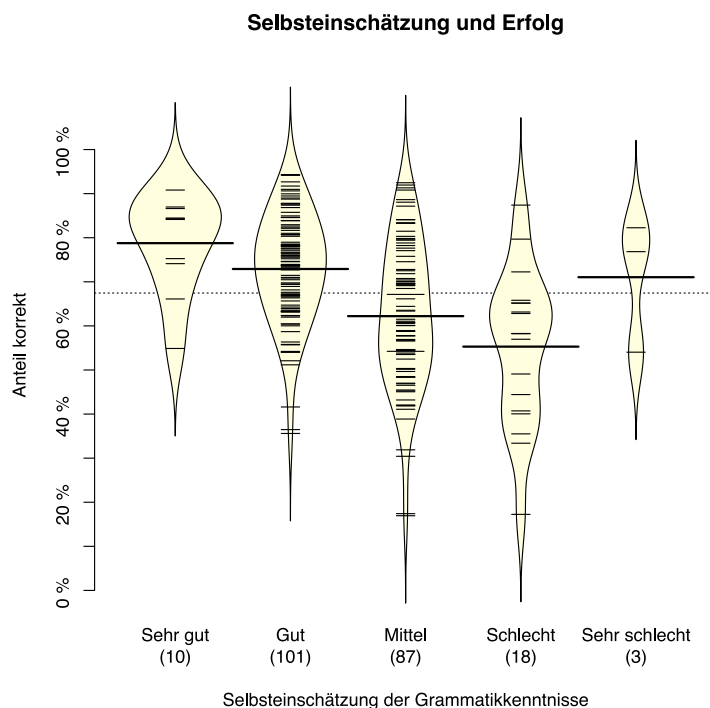


Abbildung 5. Einschätzung der eigenen Grammatikkenntnisse relativ zum Erfolg.

Es ist für den Studienerfolg wichtig, dass die Studierenden ihr erworbenes Wissen permanent und eigenständig reflektieren, und dass dieses Wissen als aktives Wissen kontinuierlich weiterentwickelt wird. Und genau dafür müssen die sprachwissenschaftlichen Lehrveranstaltungen den geeigneten Rahmen bieten. Das wollen wir an einem konkreten Beispiel aus dem Test veranschaulichen. Wir betrachten dazu nochmals die Aufgabenreihe 2, 4, 5 und 11. Alle vier Aufgaben verlangen eine Analyse syntaktischer Konstituenten. Die Schwierigkeit der Aufgaben variiert insofern, als die Aufgaben 2 und 4 im Unterschied zu den Aufgaben 5 und 11 keine Segmentierung verlangen. Die Ergebnisse zeigen, dass eine freie Form-Funktion-Analyse wie in Aufgabe 5 und 11 zu einer erheblich höheren Fehlerquote führt als bei vorgegebener Segmentierung. Während in Aufgabe 5 gerade satzförmig realisierte Attribute nicht erkannt werden, also der funktionale Terminus *Attribut* sehr häufig nur auf die Wortart *Adjektiv* bezogen wird (obwohl er in dieser Bedeutung redundant wäre), zeigen sich in Aufgabe 11 die meisten Fehler in falschen morphologisch-formalen Analysen den Kasus betreffend. Solche Fehleranalysen (siehe auch detaillierte Analysen in Abschnitt 3.3) liefern ideale Ansatzpunkte für die akademische Lehre.

Die in unserem Test festgestellten Probleme in Aufgaben mit freier syntaktischer Konstituentenanalyse bestätigen die Selbsteinschätzung der Studierenden ebenso wie das von Topalovic und Dünschede (2014: 77) genannte starke Unsicherheitsgefühl die Syntax betreffend. Studierende, denen das Handwerkszeug für eine morphologisch-formale Analyse fehlt, werden dann auch den Schritt von den nominalen Komplementen zu den satzförmig realisierten nicht mitgehen können. Solche Zusammenhänge in der Lehre darzustellen und die Studierenden damit dazu zu bringen, im Sinn von Peter Eisenberg „Grammatik zu machen,“ ist die von uns aus dem hier vorgestellten Test gezogene Handlungskonsequenz.¹⁸ In der germanistisch-linguistischen Lehre sind die zentralen Herausforderungen erstens die Weiterentwicklung des Fachs sowie die Heranbildung des akademischen Nachwuchses (und damit eben auch die Theorievermittlung) und zweitens die Auseinandersetzung mit dem erwartbaren Vorwissen der Studierenden und der Vermittlung der für ihre mehrheitliche spätere Tätigkeit wichtigen Fähigkeiten. Dabei ist zu beachten, dass auch für den akademischen Nachwuchs die Fähigkeit, *Grammatik zu machen*, essentiell ist. Wir schlagen vor, dass die germanistisch-linguistische Lehre den richtigen Weg zwischen ihren beiden großen Herausforderungen (unter anderem) mit empirischen Mitteln wie der Durchführung von studienbegleitenden Tests der hier vorgestellten Art sucht, die ihren Ausgangspunkt natürlich beim erwartbaren Schulwissen nehmen.

Nachtrag: Erläuterungen zu den statistischen Verfahren

Im Artikel haben wir verschiedene inferenzstatistische Verfahren verwendet, die hier jetzt konzeptuell erklärt und motiviert werden. Die grundlegende Fragestellung der Inferenzstatistik ist, welche Schlüsse über eine gegebene Stichprobe hinaus auf Basis dieser Stichprobe gezogen werden können. Man möchte (ganz allgemein gesprochen) von den Partikulardaten (Stichprobe) auf das allgemeine Phänomen in der Grundgesamtheit schließen. Es kann z.B. ein Wert in der Grundgesamtheit (Parameter) geschätzt werden oder eine allgemeine Theorie über den Prozess, der die Daten der Grundgesamtheit generiert, anhand der Partikulardaten überprüft werden (Hypothe-

¹⁸ Das Zitat hat Peter Eisenberg im Rahmen einer Vorlesung im Wintersemester 2013/2014 an der Freien Universität Berlin geäußert.

sentest). Bedauerlicherweise herrschen entgegen dem Bild, das Statistik-Einführungen in der Regel zeichnen, bezüglich der prinzipiellen Möglichkeiten und der technischen Umsetzung solcher Ansätze philosophische Uneinigkeit auf Seiten der Statistikerinnen und Statistiker sowie Wissenschaftsphilosophinnen und Wissenschaftsphilosophen. Auf Seiten vieler Praktikerinnen und Praktiker in den Einzelwissenschaften haben sich außerdem viele typische praktische und konzeptuelle Fehler eingeschlichen.¹⁹ Insbesondere werden die Interpretationen statistischer Ergebnisse, die sich in verschiedenen statistischen Philosophien ergeben, nicht genau verstanden und getrennt. Senn (2011) geht besonders pointiert auf die Unterschiede zwischen den vier Traditionen von Ronald A. Fisher (inferentiell frequentistisch), Jerzy Neyman und Egon Pearson (entscheidungssuchend frequentistisch), Harold Jeffreys (subjektiv bayesisch) und Bruno de Finetti (automatisch bayesisch). Zum Unterschied zwischen Fisher und Neyman-Pearson vgl. Perezgonzalez (2015). Wir beschränken uns auf eine Beschreibung der konkret verwendeten Tests und einer Einordnung in eins der vier Systeme, nämlich das von Ronald A. Fisher.

In unserem Artikel werden mehrfach die Testergebnisse von zwei oder mehr Gruppen von Teilnehmenden verglichen, wobei uns interessiert, ob in der Grundgesamtheit (über die Teilnehmenden unseres Experiments hinaus) tatsächliche Unterschiede zwischen diesen Gruppen bezüglich ihrer expliziten Grammatikkenntnisse bestehen. Ein Zwei-Gruppen-Problem ist z.B. der Vergleich der Ergebnisse von Teilnehmenden mit einsprachigem und zweisprachigem L1-Hintergrund. Ein Vier-Gruppen-Problem ist der Vergleich der Ergebnisse von Teilnehmenden aus den vier Studienjahren. Dass in einem gegebenen Experiment die Ergebnisse zwischen diesen Gruppen im Mittel bzw. Median nicht genau gleich sind, ist zunächst einmal erwartbar, da in jedem Experiment zufällige Variation die Ergebnisse beeinflusst.²⁰ Das Ziel der von uns eingesetzten statistischen Tests ist es nun, zu quantifizieren, wie häufig die beobachteten Unterschiede zwischen den Gruppen in einem Experiment auftreten würden, wenn es in der Grundgesamtheit eigentlich keinen Unterschied zwischen ihnen gäbe, wenn also *nur* zufällige Variation die Unterschiede produzieren würde. Wenn es z.B. eigentlich

¹⁹ Diese Situation hat zu teilweise konstruktiven Verbesserungsvorschlägen und Klarstellungen (z.B. Greenland et al. 2016, Senn 2001, Perezgonzalez 2015), aber auch zu extremer bis aggressiver Kritik etablierter Methoden (z.B. Ioannidis 2005, Colquhoun 2014) geführt.

²⁰ Als zufällige Variation gilt hier jeglicher Einfluss auf die quantitativen Ergebnisse, der nicht mit dem untersuchten Phänomen in Verbindung steht.

keinen wesentlichen Unterschied zwischen expliziten Grammatikkenntnissen einsprachiger und mehrsprachiger Studierenden gibt, würde in den meisten Experimenten durchaus ein Unterschied zwischen den Leistungen der beiden Gruppen beobachtet werden. Dieser beobachtete Unterschied wäre aber meistens klein, weil Zufallsstichproben dazu tendieren, die Grundgesamtheit abzubilden. Sehr selten (aber nicht nie) würde unter diesen Bedingungen die rein zufällige Variation eine Stichprobe hervorbringen, in der erhebliche Leistungsunterschiede zwischen den beiden Gruppen gemessen werden. Die wesentlichen Einflussgrößen sind in diesem Zusammenhang einerseits die Größe des tatsächlichen Unterschieds zwischen den Gruppen sowie die Stichprobengröße. Je näher an Null der wahre Unterschied liegt, desto seltener sind extreme Ergebnisse in Experimenten zu erwarten. Je größer die Stichproben sind, desto besser bilden sie prinzipiell die Grundgesamtheit ab, und desto seltener sind folglich extreme Ergebnisse, wenn der wahre Unterschied nahe an Null liegt.²¹ Für die in Abbildung 3 dargestellten Unterschiede wird z.B. mit $p=0,035$ angegeben, dass nur in 3,5% einer großen Zahl von (auf gleiche Weise durchgeführten) Experimenten Unterschiede in der beobachteten Größe oder extremere Unterschiede zwischen dem Erfolg der Teilnehmenden mit und ohne absolvierter Linguistik-Einführung zu erwarten wären, wenn es in der Grundgesamtheit eigentlich keinen Unterschied gäbe. Die Hypothese, dass es keinen Unterschied gibt, wird also von den beobachteten Daten stark geschwächt. Dagegen wird für die in Abbildung 4 dargestellten Daten mit $p=0,475$ angegeben, dass in 47,5% von sehr vielen vergleichbaren Experimenten der beobach-

²¹ Damit folgen wir hier der statistischen Philosophie von Ronald A. Fisher (für eine kompakte Darstellung s. Maxwell & Delaney 2004:34–49; vergleichende Darstellungen in den bereits genannten Perezgonzalez 2015, Senn 2011; weiterführend Mayo & Cox 2011). In diesem Ansatz geht es nur darum, in Form des sogenannten „p-Werts“ zu berechnen, wie häufig ein Experiment ein genauso extremes oder extremeres Ergebnis wie das gemessene Ergebnis hätte, wenn die „Nullhypothese“ (H_0) für die Grundgesamtheit zuträfe. Bei einem angemessen niedrigen p-Wert spricht man von einer „Zurückweisung der H_0 .“ Die Nullhypothese ist für unsere Belange stets: „Es gibt keinen Unterschied zwischen den Gruppen.“ Der Test ist aber auf keinen Fall als konfirmatorisch *für* irgendeine andere Hypothese (z.B. „Es gibt einen Unterschied zwischen den Gruppen.“) zu werten! Der Test liefert auch keinen „statistischen Beweis“ und gibt nicht an wie wahrscheinlich die Nullhypothese oder irgendeine andere Hypothese ist. Zudem hat die ggf. bekannte Rede von der „Alternativhypothese“ (H_1) und dem (präspezifizierten) α -Niveau in Fishers System keinen Platz. Sie sind Teile des Systems von Neyman und Pearson. An genau diesem Punkt treten typischerweise fatale Fehlinterpretationen auf, und wir können hier aus Platzgründen nur vor ihnen warnen, ohne sie im Detail zu diskutieren.

tete oder ein extremerer Unterschied zwischen Teilnehmenden aus verschiedenen Studienjahren gemessen werden würde, wenn es eigentlich keinen wahren Unterschied gäbe. Die Hypothese, dass es keinen Unterschied gibt, wird also von diesem Test nicht nennenswert geschwächt.

Die meisten populären statistischen Tests sind sogenannte „parametrische“ Tests, die die Beobachtungsdaten als Messungen von Zufallsvariablen mit einer eingeschränkten Menge von Parametern betrachten, und die typischerweise starke Annahmen über deren Verteilung voraussetzen. Die Varianz (die Streuung) darf z.B. zwischen den Gruppen nicht variieren, und die Werte müssen homogen normalverteilt (Gaussverteilt) sein. Diese Annahmen sind für unsere Daten i.d.R. verletzt, und wir verwenden daher nur robuste nicht-parametrische Alternativen (implementiert im R-Paket *coin*, Hothorn et al. 2006), womöglich auf Basis einer exakten Verteilung (Zwei-Gruppen-Probleme), sonst auf Basis einer approximativen (Monte Carlo) Verteilung (Mehr-Gruppen-Probleme). Sich scheinbar anbietende parametrische Tests wie der bekannte t-Test und die Varianzanalyse (ANOVA) sind ungünstig für den Vergleich von Anteilswerten bzw. Prozentwerten, da die Varianz bei diesen in der Regel nicht homogen ist: Mit steigender Nähe zu den Anteilswerten 0 und 1 (bzw. 0% und 100%) wird die Varianz intrinsisch eingeschränkt (bei 0 bzw. 1 abgeschnitten), da es keine Anteilswerte von unter 0 oder über 1 gibt.

Danksagungen

Wir danken Luise Reißmann sehr für ihre engagierte Hilfe bei der Erstellung des Fragebogens. Für ihre Hilfe bei der Durchführung der Befragungen danken wir Elena Karagjosova und Andreas Pankau. Für Hinweise zum Text danken wir Felix Bildhauer.

Literatur

- Boettcher, Wolfgang. 1994. Grammatiksozialisation in Schule, Hochschule und Referendarausbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung* 2/1994, 170–186.

- Bredel, Ursula. 2014. *Sprachbetrachtung und Grammatikunterricht*. 2., durchgesehene Auflage. Paderborn etc.: Schöningh.
- Bremerich-Vos, Hans. 2016. Ein Studieneingangstest für Lehramtsstudierende im Fach Germanistik? Zwar mit Bauchschmerzen, aber: Ja! *Didaktik Deutsch* 40/2016, 9–12.
- Colquhoun, David. 2014. An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science* 1(3), 140216.
- Eisenberg, Peter. 2004. Wieviel Grammatik braucht die Schule? *Didaktik Deutsch* 17/2004, 4–25.
- Eisenberg, Peter. 2013. *Grundriss der deutschen Grammatik: Das Wort*. 4. Auflage unter Mitarbeit von Nanna Fuhrhop. Stuttgart: Metzler.
- Engel, Ulrich. 2009. *Syntax der deutschen Gegenwartssprache*. 4., völlig neu bearbeitete Auflage. Berlin: Erich Schmidt Verlag.
- Feilke, Helmut. 2012. Bildungssprachliche Kompetenzen fördern und entwickeln. Basisartikel. *Praxis Deutsch* 233, 4–18.
- Fuhrhop, Nanna und Oliver Teuber. 2016. Orthographisches Wissen und orthographisches Können. Ist ein sinnvoller Studieneingangstest möglich? *Didaktik Deutsch* 40/2016, 13–15.
- Greenland, Sander, Stephn J. Senn, Kenneth J. Rothman, John B. Carlin, Charles Poole, Steven N. Goodman, and Douglas G. Altman. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology* 31, 337–350.
- Hennig, Mathilde. 2012. Grammatische Terminologie in der Schule. Einladung zur Diskussion. *Zeitschrift für germanistische Linguistik* 40(3), 443–450.
- Hothorn, Torsten, Kurt Hornik, Mark A. van de Wiel und Achim Zeileis. 2006. A Lego System for Conditional Inference. *The American Statistician* 60(3), 257–263.
- Gornik, Hildegard. 2003. Methoden des Grammatikunterrichts. In Ursula Bredel, Hartmut Günther, Peter Klotz, Jakob Ossner und Gesa Siebert-Ott (Hrsg.). *Didaktik der deutschen Sprache*. Paderborn etc.: Schöningh. 814–829 (Bd. 2).

- Häcker, Roland. 2009. Wie viel? Wozu? Warum Grammatik in der Schule. In Marek Konopka und Bruno Strecker (Hrsg.). *Deutsche Grammatik – Regeln, Normen. Sprachgebrauch. Jahrbuch des Instituts für Deutsche Sprache 2008*. Berlin, New York: De Gruyter. 309–332.
- Ioannidis, John P. A. 2005. Why most published research findings are false. *PLoS med* 2(8), e124.
- Maxwell, Scott E. und Harold D. Delaney. 2004. *Designing experiments and analyzing data: A model comparison perspective*. Mahwa, London: Taylor & Francis.
- Mayo, Deborah G. und Aris Spanos. 2011. Error Statistics. In Prasanta S. Bandyopadhyay and Malcolm R. Forster (Hrsg.). *Handbook of the Philosophy of Science. Volume 7: Philosophy of Statistics*. Amsterdam: Elsevier. 153–198.
- Ossner, Jakob. 2007. Grammatik in Schulbüchern. In Köpcke, Klaus-Michael und Ziegler, Arne (Hrsg.). *Grammatik in der Universität und für die Schule. Theorie, Empirie und Modellbildung*. Tübingen: Niemeyer. 161–184.
- Perezgonzalez, Jose D. 2015. Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing. *Frontiers in Psychology* 6(223), 1–11.
- Portmann-Tselikas, Paul. 2011. Spracherwerb, grammatische Begriffe und sprachliche Phänomene. Überlegungen zu einem unübersichtlichen Lernfeld. In Köpcke, Klaus-Michael und Ziegler, Arne (Hrsg.). *Grammatik – lehren, lernen, verstehen. Zugänge zur Grammatik des Gegenwartsdeutschen*. De Gruyter: Berlin, Boston. 71 – 90.
- R Core Team. 2016. R: A language and environment for statistical computing. Wien: R Foundation for Statistical Computing. <https://www.r-project.org/>
- Schäfer, Roland. 2016. *Einführung in die grammatische Beschreibung des Deutschen*. Zweite, überarbeitete Auflage. Berlin: Language Science Press.
- Schindler, Kirsten. 2016. Welche Funktion sollten Eingangstests haben? Ein Beitrag aus sprachdidaktischer Perspektive. *Didaktik Deutsch* 40/2016, 16 – 19.
- Senn, Stephen J. 2001. Two cheers for P-values. *Journal of Epidemiology and Biostatistics* 6(2), 193–204.

Senn, Stepehn J. 2011. You may believe you are a Bayesian but you are probably wrong. *Rationality, Markets, and Morals* 2, 48–66.

Spiegel Online. 2007. *Das Studium ist dem Genitiv sein Tod*, Spiegel Online. Erschienen am 18. April 2007, abgerufen am 20. September 2016. <http://www.spiegel.de/lebenundlernen/uni/grammatik-fiasko-das-studium-ist-dem-genitiv-sein-tod-a-477841.html>

SZ Online. 2007. *Germanisten scheitern an Grammatik*, Süddeutsche Zeitung Online. Erschienen am 18. April 2007, abgerufen am 20. September 2016.²² <http://sz.de/1.547946>

Topalovic, Elvira und Susanne Dünschede. 2014. Weil Grammatik im Lehrplan steht. Bundesweite Umfrage zur Grammatik in der Schule. *Der Deutschunterricht* 3/2014, 76–81.

Quellen der Aufgaben

- (1) *Standard Deutsch 7 Basis*. Berlin: Cornelsen Verlag. 1. Auflage, 3. Druck 2012. Aufgabe 2, S. 37. [7. Klasse]
- (2) *Standard Deutsch 7 Basis*. Berlin: Cornelsen Verlag. 1. Auflage, 3. Druck 2012. Aufgabe 1, S. 47. [7. Klasse]
- (3) *Grammatik Alles Klar 3 (9.–10. Klasse)*. Berlin: Cornelsen Verlag. 2. Auflage, 1. Druck 2006. Aufgabe 6, S. 37.
- (4) *Grammatik Alles Klar 3 (9.–10. Klasse)*. Berlin: Cornelsen Verlag. 2. Auflage, 1. Druck 2006. Aufgabe 7, S. 30.
- (5) *Muttersprache plus. Sprach- und Lesebuch 6*. Berlin: Cornelsen Verlag. 2013. Aufgabe 3a, S. 192. [6. Klasse]
- (6) *Arbeitsheft Deutsch 9. Klasse*. Berlin: Cornelsen Verlag. 2007. Aufgabe 2, S. 39. Berlin.

²² Die Datierung auf der aktuellen Version der Webseite von SZ Online (geprüft am 20. September 2016) ist inkorrekt.

- (7) *Doppel-Klick 8, Das Arbeitsheft B.* Berlin: Cornelsen Verlag. 2007. Aufgabe 4, S. 77. [8. Klasse]
- (8) *Doppel-Klick 8, Das Arbeitsheft B.* Berlin: Cornelsen Verlag. 2007. Aufgabe 2, S. 79. [8. Klasse]
- (9) *Arbeitsheft Deutsch 9. Klasse.* Berlin: Cornelsen Verlag. 2007. Aufgabe 1, S. 47. Berlin.
- (10) *Standard Deutsch 7 Basis.* Berlin: Cornelsen Verlag. 1. Auflage, 3. Druck 2012. Aufgabe 1, S. 52. [7. Klasse]
- (11) *Grammatik Alles Klar 3 (9.–10. Klasse).* Berlin: Cornelsen Verlag. 2. Auflage, 1. Druck 2006. Aufgabe 3, S. 40.

Anhang: Testfragen

Hinweis: Die Wiedergabe erfolgt hier rein inhaltlich. Die genaue Präsentation der Fragen kann dem separat veröffentlichten Fragebogen entnommen werden.

1. Aktiv oder Passiv? Bestimmen Sie die folgenden Sätze und kreuzen Sie entsprechend an.

	Aktiv	Passiv
<i>Viele Menschen suchen das große Glück.</i>		
<i>Das Glücksgefühl wird durch Lachen gefördert.</i>		
<i>Auch das Denkvermögen wird dadurch angekurbelt.</i>		
<i>Glücksforscher untersuchen die Wirkung des Lachens.</i>		
<i>Das große Glück wird von vielen Menschen gesucht.</i>		
<i>Die Wirkung des Lachens wird erforscht.</i>		
<i>Ausgiebiges Lachen fördert das Glücksgefühl.</i>		
<i>Häufiges Lachen kurbelt das Denkvermögen an.</i>		

2. Bestimmen Sie alle Satzglieder in den folgenden Sätzen. Kennzeichnen Sie sie so: S für *Subjekt*, P für *Prädikat*, O für *Objekt* und AB für *adverbiale Bestimmung*.

<i>Eine Französin</i>	<i>reiste</i>	<i>mit ihrem Surfbrett</i>	<i>über den indischen Ozean.</i>

<i>Nachts</i>	<i>schief</i>	<i>sie,</i>	<i>tagsüber</i>	<i>surfte</i>	<i>sie</i>

<i>Nach 6300 Kilometern und 60 Tagen</i>	<i>erreichte</i>	<i>sie</i>	<i>Die Insel La Reunion</i>

<i>Im Hafenort Le Port</i>	<i>bereitete</i>	<i>man</i>	<i>ihr</i>	<i>ein großes Willkommensfest.</i>

3. Im unten stehenden Text sind die Nominalgruppen markiert. Wie unterschiedlich sie besetzt sein können, ist in den folgenden Mustern *a–e* angegeben.

- a Artikel – Indefinitpronomen – Nomen
- b Artikel – Adjektiv – Nomen
- c Adjektiv – Adjektiv – Nomen
- d Indefinitpronomen – Adjektiv – Nomen
- e Possessivpronomen – Nomen

Ordnen Sie jeder Nominalgruppe den passenden Buchstaben zu, indem Sie diesen in die eckigen Klammern nach den Nominalgruppen schreiben.

Zuerst wird Europa wie ein einziger Marktplatz [] sein und später die ganze Welt [].

Die meisten Großunternehmen [] werden ihre Betriebe [] über viele Länder verteilen.

*Daneben wird es mehr kleine Betriebe [] und Selbständige geben.
Ganz neue Berufsbilder [] werden entstehen.*

4. Adverbial oder Objekt? Schreiben Sie O bzw. Ad in die Klammern.

Die Rettungsmannschaften sprechen von einer extrem schwierigen Suche [].

Es gebe kaum Hoffnung, in dem unwegsamen Gelände [] Überlebende zu finden.

Sieben Hubschrauber und zwei Transportflugzeuge sind im Landkreis Ismathia [] an der Suchaktion [] beteiligt.

5. Unterstreichen Sie die Attribute in folgendem Satz.

Die Inuit, die heute noch auf Jagd gehen, fahren mit schnellen Motorschlitten und kehren in ihre festen Holzhäuser zurück.

6. Auf welche der unten stehenden Sätze 1–3 beziehen sich die folgenden Aussagen über die Verwendung des Konjunktivs? Setzen Sie jeweils die passende Ziffer ein.

Satz [] ist richtig, weil der Konjunktiv I signalisiert, dass es sich um die Wiedergabe einer fremden Äußerung handelt.

Satz [] ist nicht richtig, weil der Konjunktiv II in der Regel nur dann als Modus für die indirekte Rede gewählt wird, wenn der Konjunktiv I nicht vom Indikativ Präsens zu unterscheiden ist.

Satz [] ist richtig, weil der einleitende Hauptsatz und die Konjunktion „dass“ Signale für die indirekte Rede sind.

Dies sind die zuzuordnenden Sätze:

1. *Die Ministerin sagte, dass wichtige Verhaltensleistungen mit Noten nicht zu erfassen sind.*
2. *Die Ministerin sagte, wichtige Verhaltensleistungen seien nicht mit Noten zu erfassen.*
3. *Die Ministerin meinte, wichtige Verhaltensleistungen wären nicht mit Noten zu erfassen.*

7. Verwandeln Sie jeweils die beiden Hauptsätze in einen Haupt- und einen Nebensatz mit einer Konjunktion.

Es hat geklingelt. Du warst mit dem Referat fertig.

Ich komme morgen in die Schwimmhalle. Ich kann dir beim Training zuschauen.

8. Ergänzen Sie die Relativpronomen in den folgenden Sätzen.

Es gibt einen Fernseher, ___ mit den Zuschauern spricht.

Interessant ist ein Roboter, ___ den Verkehr kontrolliert.

Man kauft Kleidung, ___ sich selbst reinigt.

Du wohnst in einem Haus, ___ unter dem Erdboden liegt.

9. Trennen Sie in den folgenden Sätzen die Wörter voneinander ab. Achten Sie beim Abschreiben auf die richtige Groß- und Kleinschreibung.

*sieversuchtebeimüberquerenderschluc
htnichtindenabgrundhinunterzusehen*

*beimspazierengehenundgeschichtenerzählen
warensichmichaundgabivielnähergekommen*

10. Aus den folgenden Wörtern können Sie insgesamt fünf Wortfamilien bilden. Schreiben Sie sie auf und unterstreichen Sie jeweils den Wortstamm.

eröffnen, das Gestell, anbinden, Offenheit, die Bindung, fröhlich, die Kindheit, kindisch, die Angestellte, der Frohsinn

11. Unterstreichen Sie in den folgenden Sätzen alle Nominalgruppen, die Akkusativobjekte sind, einfach. Die Nominalgruppen, die Dativobjekte sind, unterstreichen Sie bitte doppelt.

Leider finden viele nicht sofort einen Ausbildungsplatz.

Ich will den bestmöglichen Schulabschluss erreichen.

Hat mein Wunschberuf eigentlich gute Zukunftsaussichten?

*Heutzutage werden den Schulabgängern
viel zu wenig Lehrstellen bereitgestellt.*

*In der Zukunft werden nicht mehr vorwiegend
die großen Konzerne die Arbeitsplätze schaffen.*

*Das wird vielmehr den mittleren und kleinen Betrieben
vorbehalten sein.*

Kein Industrieland kann sich mehr der Globalisierung der Wirtschaft entziehen.