

Anwendung der analytischen Bewertungsskala bei der Beurteilung diskursiver und grammatischer Kompetenz fortgeschrittener Deutschlerner

Ćaćić, Ana

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Humanities and Social Sciences / Sveučilište Josipa Jurja Strossmayera u Osijeku, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:142:876210>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom](#).

Download date / Datum preuzimanja: **2025-01-09**



FILOZOFSKI FAKULTET
SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU

Repository / Repozitorij:

[FFOS-repository - Repository of the Faculty of Humanities and Social Sciences Osijek](#)



Sveučilište J.J. Strossmayera u Osijeku
Filozofski fakultet Osijek
Jednopedmetni diplomski studij njemačkog jezika i književnosti

Ana Čačić

**Primjena analitičke ljestvice za vrednovanje diskursne i gramatičke
kompetencije naprednih učenika njemačkog jezika**

Diplomski rad

Mentor: izv. prof. dr. sc. Ivana Jozić

Osijek, 2020.

Sveučilište J.J. Strossmayera u Osijeku
Filozofski fakultet Osijek
Odsjek za njemački jezik i književnost
Jednopredmetni diplomski studij njemačkog jezika i književnosti

Ana Čaćić

**Primjena analitičke ljestvice za vrednovanje diskursne i gramatičke
kompetencije naprednih učenika njemačkog jezika**

Diplomski rad

Humanističke znanosti, filologija, germanistika

Mentor: izv. prof. dr. sc. Ivana Jozić

Osijek, 2020.

J.-J.-Strossmayer-Universität in Osijek
Fakultät für Geistes- und Sozialwissenschaften Osijek
Diplomstudium der deutschen Sprache und Literatur
(Ein-Fach-Studium)

Ana Ćaćić

**Anwendung der analytischen Bewertungsskala bei der Beurteilung
diskursiver und grammatischer Kompetenz fortgeschrittener
Deutschlerner**

Diplomarbeit

Mentor: ao. Univ.-Prof. Dr. Ivana Jozić

Osijek, 2020

J.-J.-Strossmayer-Universität in Osijek
Fakultät für Geistes- und Sozialwissenschaften Osijek
Abteilung für deutsche Sprache und Literatur
Diplomstudium der deutschen Sprache und Literatur
(Ein-Fach-Studium)

Ana Čačić

**Anwendung der analytischen Bewertungsskala bei der Beurteilung
diskursiver und grammatischer Kompetenz fortgeschrittener
Deutschlerner**

Diplomarbeit

Sprachwissenschaft

Mentor: ao. Univ.-Prof. Dr. Ivana Jozić

Osijek, 2020

Erklärung über die eigenständige Erstellung der Arbeit

IZJAVA

Izjavljujem s punom materijalnom i moralnom odgovornošću da sam ovaj rad samostalno napravio te da u njemu nema kopiranih ili prepisanih dijelova teksta tuđih radova, a da nisu označeni kao citati s napisanim izvorom odakle su preneseni.

Svojim vlastoručnim potpisom potvrđujem da sam suglasan da Filozofski fakultet Osijek trajno pohrani i javno objavi ovaj moj rad u internetskoj bazi završnih i diplomskih radova knjižnice Filozofskog fakulteta Osijek, knjižnice Sveučilišta Josipa Jurja Strossmayera u Osijeku i Nacionalne i sveučilišne knjižnice u Zagrebu.

U Osijeku, dana 04.05.2022

utua Cadić, 0122012964
ime i prezime studenta, JMBAG

Inhaltsverzeichnis

1. Einleitung	8
2. Theoretische Grundlagen zur Anwendung der analytischen Bewertungsskala	9
2.1 Was ist Schreiben?	9
2.1.1 Die Schreibfertigkeit.....	10
2.2 Schreibfehler und ihre Korrektur	11
2.3 Das Beurteilen	12
2.3.1 Die Objektivität beim Beurteilen	13
2.3.2 Warum überhaupt Noten vergeben?.....	14
2.4 Verschiedene Bewertungskriterien	15
2.4.1 Die Bewertungsskalen	17
2.4.1.1 Holistische Bewertungsskala.....	18
2.4.1.2 Analytische Bewertungsskala.....	19
2.4.1.3 Die Verbindung von einem holistischen und analytischen Bewertungsverfahren.....	22
3. Anwendung der analytischen Bewertungsskala bei der Beurteilung diskursiver und grammatischer Kompetenz fortgeschrittener Deutschlerner	24
3.1 Ausgangspunkt	24
3.2 Zielstellung	25
3.3 Forschungsdesign	26
3.3.1 Probanden.....	26
3.3.2 Instrument.....	26
3.3.3 Datenerhebung.....	27
3.3.4 Datenanalyse.....	28
3.4 Forschungsergebnisse	30
3.4.1 Deskriptive Statistik	30
3.4.2 T-Test für unabhängige Stichproben.....	34
3.4.3 Korrelationsanalyse.....	34
3.5 Diskussion	38
4. Schlusswort	42
Literaturverzeichnis	44
Anhang	
Zusammenfassung	
Sažetak	

1. Einleitung

Wie könnten wir die Beurteilung von Aufsätzen objektiver gestalten? Das ist eine Frage, die sehr oft von Lehrern, die Deutsch als Fremdsprache (DaF) unterrichten, immer wieder neu aufgeworfen wird. Genau diese Frage ist der Ausgangspunkt für diese Arbeit.

Eine der wichtigsten Aufgaben der DaF-Lehrer ist die Beurteilung von Schüleraufsätzen. Es ist unbestritten, dass die Beurteilung von Aufsätzen immer ein gewisses Maß an Subjektivität in sich trägt. Eine wichtige Frage könnte daher sein, ob uns Bewertungsskalen dabei helfen können, objektiver zu beurteilen. Genau aus diesem Grund haben wir uns dazu entschieden, uns in dieser Arbeit näher mit den analytischen und holistischen Bewertungsskalen zu beschäftigen, beziehungsweise mit den Unterschieden der eben genannten Bewertungsskalen, wobei wir ebenfalls hoffen, auf die oben genannten Fragen eine Antwort geben zu können. Es wird näher untersucht, ob die Anwendung von analytischen Bewertungsskalen zu einer objektiveren, zuverlässigeren Beurteilung führen kann.

Wie der Titel der Arbeit schon sagt, werden wir uns mit der Anwendung der analytischen Bewertungsskala bei der Beurteilung diskursiver und grammatischer Kompetenz fortgeschrittener Deutschlerner befassen. Um aber überhaupt gezielt und konkret etwas über die Bewertungsskalen sagen zu können, müssen wir die Begriffe Schreiben, Schreibfertigkeit, Fehlerkorrektur, Benotung und Bewertungskriterien näher betrachten. Diese Begriffe werden weiter im theoretischen Teil dieser Forschungsarbeit erläutert. Am Schluss des theoretischen Teils befassen wir uns mit den verschiedenen Bewertungsskalen, die uns weiter zum praktischen Teil der Arbeit führen.

Im zweiten Teil der Arbeit, dem praktischen Teil, widmen wir uns der eigentlichen Untersuchung. Es soll festgestellt werden, ob die Anwendung von analytischen Bewertungsskalen zur objektiveren, zuverlässigeren Beurteilung diskursiver und grammatischer Kompetenzen führt. Im Weiteren werden die Unterschiede des holistischen und analytischen Bewertungsverfahrens der Diskurskompetenzen (Kohärenz und Kohäsion) und Grammatikkompetenz (Syntax und Morphologie) anhand von Schüleraufsätzen untersucht, wobei vor allem die entsprechenden Unterschiede näher betrachtet werden. In diesem Teil der Arbeit werden ebenfalls die Korrelationen der Kompetenzen bei den holistischen und analytischen Bewertungsverfahren dargestellt.

Zuerst werden die Datenerhebung und Datenanalyse beschrieben und danach werden die Ergebnisse dargestellt, die zu einer ausführlichen Diskussion über die untersuchten Elemente führen. In der Schlussfolgerung wird dann unter Beachtung der dargestellten Resultate auch ein

Ausblick in Bezug auf die didaktischen Implikationen und weitere ähnliche Untersuchungen gegeben.

2. Theoretische Grundlagen zur Anwendung der analytischen Bewertungsskala

2.1 Was ist Schreiben?

Um über Bewertungsskalen sprechen zu können, müssen wir uns erst mit den Begriffen Schreiben, Schreibfertigkeit, Schreibfehler und Bewertung befassen. Mit der Definition des Schreibens befasst sich Gerd Bräuer, der das Schreiben folgendermaßen definiert:

(...) neben dem Sprechen und der Körpersprache, eines der Hauptmedien menschlicher Kommunikation. Es bewegt sich, im engeren Rahmen gesehen, im Spannungsfeld von Schreibanlaß und AdressatIn und korrespondiert, im weiteren Rahmen, mit dem Verhältnis von Schreibenden, Verteilungsapparat, Aufbewahrungssystem und Lesenden. (Bräuer 1998: 15)

Das Schreiben ist eng mit der Entwicklung des Denkens und der Ordnungsprinzipien verknüpft. Es hilft uns, unsere „chaotischen“ Gedanken zu ordnen und zu strukturieren (vgl. Krumm 1989: 6). Nach Kast (1999: 22) wird im kommunikativen Ansatz die Sprache und damit auch das Schreiben als soziales Handeln aufgefasst. Pragmatisch gesehen definiert man Schreiben als ein Verfahren zur Informationsvermittlung, eine Person teilt einer anderen Person (oder anderen Personen) etwas mit. Die Information, die im Zentrum dieses Vorgangs steht, ist das Produkt des Schreibens, wie aus der Abbildung 1 ersichtlich ist:

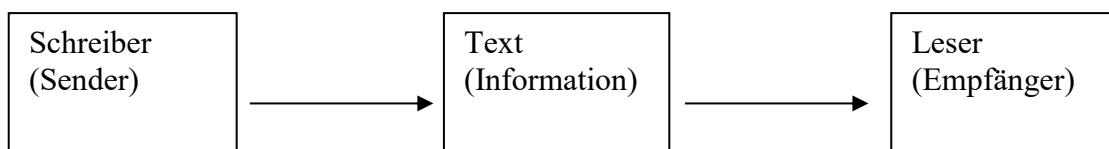


Abbildung 1: Produktorientierte Aspekte des Schreibens (Kast 1999: 22)

Sibylle Bolton (1996: 70) betont ebenfalls die Rolle des Schreibens. Es ermögliche dem Schreibenden über sich selbst, über einen Gegenstand und andere Personen klar zu werden. Fremdsprachenlernende erproben beim Schreiben von Texten ihre Fähigkeit, die gelernten Redemittel und grammatischen Strukturen zu organisieren und sich der grammatischen und diskursiven Kompetenzen schreibend zu vergewissern. Das Verfassen von Texten beeinflusst die sprachlich-geistige Entwicklung, erweitert die kommunikativen Fähigkeiten und ist bedeutend für das Lernen.

Nach Hans-Jürgen Krumm (1993: 24) ermöglicht das Schreiben, Erlebnisse zu verarbeiten, unterhaltende, informative Texte herzustellen, Mitteilungen zu äußern, bietet Gelegenheit zum Nachdenken, Nachgestalten, ermöglicht sprachlich-geistige Überlegungen, unterstützt

Lernprozesse und fördert das Schulleben. Im Vergleich mit dem gesprochenen Wort kann der Schreibende seinen Text mehrfach überprüfen und korrigieren. Der Text kann auch später von anderen gelesen werden und kann für weitere Bearbeitungen dienen. Durch das Lesen des Textes erfahren die Leser etwas über den Schreibenden, da er immer Gefühle, Ansichten und Gedanken in seine Texte einbaut.

Nach Kast (1999) unterscheidet man bei der Art des Schreibens zwischen dem personalen, kreativen und freien Schreiben. Kasts (1999: 29) Meinung nach, ermöglicht das personale Schreiben dem Autor, sich auszudrücken und etwas über sich selbst mitzuteilen. Es geht nicht in erster Linie um den Text als Produkt, sondern um den schreibenden Schüler und um den Text im Prozess des Entstehens. Unter dem Begriff kreatives Schreiben, versteht Kast (1999: 126) das Schreiben von Texten ohne Vorlage, mit Hilfe durch vorbereitende Wortschatzübungen oder strukturierende Hinweise.

Das freie Schreiben bezeichnet Kast (zit. nach Portmann 1999: 126) als ein nicht-vorlagengebundenes Schreiben. Diese Form des Schreibens sei eine primäre Art des Schreibens. Es sei diejenige Form des Schreibens, welche die meisten Menschen benutzen, wenn sie schreiben.

2.1.1 Die Schreibfertigkeit

Mit dem Begriff Schreibfertigkeit befasst sich Sibylle Bolton ausführlich. In ihrer Arbeit behauptet Bolton (1996: 67), dass, im Unterricht zu schreiben, nicht gleich bedeutet, dass jedes Schreiben die Schreibfertigkeit zum Ziel hat. Das Schreiben hat zwei sehr unterschiedliche Funktionen, welche voneinander unterschieden werden müssen. Wir unterscheiden das Schreiben als Zielfertigkeit und das Schreiben als Mittlerfertigkeit:

- a) Es gibt Schreibaktivitäten, die denen das Schreiben das Ziel ist: z.B. wenn ich einen Brief schreibe, ist das Ziel meiner Handlung ein Brief, den ich jemandem schicken möchte.
- b) Es gibt aber auch viele Schreibaktivitäten, bei denen Schreiben ein Mittel für einen anderen Zweck ist: z.B. bei schriftlichen Grammatikübungen; da ist mein Ziel, eine bestimmte Struktur zu üben. (Gerdes, zit. nach Kast 1999: 8)

Laut Valette (1972: 154) folgt der Fremdsprachenlernende zur Entwicklung der Schreibfertigkeit eine Reihe von Schritten. Bevor die Schüler an ihrem Ausdruck und Stil arbeiten können, müssen sie grundlegende Kenntnisse im Bereich Wortschatz, Grammatik und Rechtschreibung beherrschen. Es stellt sich somit die Frage, ob das Schreiben als Zielfertigkeit überhaupt ein legitimes Ziel des Fremdsprachenunterrichts ist und wozu man das Schreiben lehren sollte. Gründe für das Schreiben im Fremdsprachenunterricht bietet Bernd Kast (1999: 20 f.):

- Schreiben aufgrund kommunikativer Bedürfnisse
- formelle Briefe (zur Informationsbeschaffung)

- informelle Briefe/Karten
- Formulare
- Fragen zur Person, Telegramm usw.
- Kurzmitteilungen für Dritte
- Schreiben aufgrund unterrichtspraktischer Bedürfnisse
 - Hausaufgaben, Tests, Notizen, Übungen zu Wortschatz und Grammatik
- Schreiben aufgrund lernpsychologischer Überlegungen
 - die Rolle des Schriftbildes beim Hören, Lesen (und Schreiben)
 - motivationale Überlegungen
 - Integration der Fertigkeiten
- Schreiben als Hilfe bei der Strukturierung geistiger Handlungen

2.2 Schreibfehler und ihre Korrektur

Die Definition der Fehler nach Kast (1999: 170) lautet: „Fehler sind Zwischentappen im Lehr- und Lernprozess, sie verweisen auf wichtige Lehr- und Lernprobleme.“ Fehler wurden somit jahrelang als etwas Negatives (Defizit, Mangel, persönliches Versagen) angesehen. Sie bestimmten den Erfolg oder Misserfolg der Schüler. Heute lautet die Regel „durch Fehler lernt man“. Man betrachtet die Fehler als Notwendigkeit für den Entwicklungsprozess der Lernenden. Lehrer, zum Beispiel, erkennen durch Fehler, was noch einmal oder umfassender erklärt werden sollte (vgl. Hasler, zit. nach Kast 1999: 170). Die Sprachforschung benutzt den Begriff *kreativer Fehler*, der auch bei Kast (1991: 70) zu finden ist. Kast (1991: 170) beschreibt diese als Fehler, bei denen der Lernende richtig gedacht hat. Es sind neue Wortschöpfungen, die als solche in keinem Wörterbuch zu finden sind: „Ein Beispiel aus der Wortbildung: zähmen- Zähmung; kränken- Kränkung, sich grämen- Grämung; dieses Wort finden Sie in keinem Wörterbuch: es existiert nicht“ (ebd.). Fehler können außerdem in verschiedene Kategorien geteilt werden: Ausdruck, Artikel, syntaktischer oder semantischer Bezug, Genus, Inhalt, Kasus, Konjunktion, Modus, morphologische Fehler, Modalverb, Präposition, Pronomen, Rechtschreibung, Satzbau, Satzstellung, Stil, Tempus, Wortwahl, falsche oder fehlende Zeichensetzung, Fehlen von Elementen, überflüssige Elemente, Umstellung (vgl. Kleppin, zit. nach Kast 1999: 171). Kast (1999: 182) unterscheidet nach der Art der Fehler:

- Fehler im Bereich der Orthographie
- kommunikationsbehindernde Fehler
- „hartnäckige“ Fehler
- Flüchtigkeitsfehler
- Fehler in Bezug auf bestimmte Lernziele.

Die Korrektur der Fehler findet nach dem Schreiben statt und stört somit nicht den Schreibprozess. Da die Fehler auf Papier festgehalten werden, lassen sie sich einfach und ohne Störung korrigieren. Bei der Fehlerkorrektur bedienen sich die Lehrer verschiedener Mittel, so z.B. unterstreichen die einen fehlerhafte Wörter und Ausdrücke, andere hingegen bedienen sich verschiedener Zeichen und Symbole bezüglich des Wortschatzes oder der Grammatik (Kast 1999: 168). Laut Macaire (1998: 12) haben Randbemerkungen in Aufsätzen, wenn sie auf Lob, Kritik oder Ratschläge hinweisen eine psychologische Funktion. Durch die Randbemerkungen entsteht ein Kommunikationsprozess zwischen Schüler und Lehrer.

Kast (1999: 172) behauptet, dass man beim Schreiben von Texten nicht nur die Wort- und Satzebene betonen sollte, sondern auch die Textebene. Diese sollte genauso geübt und berücksichtigt werden, wie die anderen Ebenen. Bei der Korrektur werden die Wort- und Satzebene als Einzelphänomene korrigiert, während man bei der Textebene die Gesamtheit der Textintention und Textrealisierung betrachtet. Im Zusammenhang mit Fehlerberichtigungen betont Kast:

Punktuelle Fehlerberichtigungen, wie z.B. bei Rechtschreibung, Artikel und Genus oder bei morphologischen Fehlern, führen zu relativ oberflächlichen, kurzfristigen Bearbeitungsaktivitäten (einen Buchstaben korrigieren, einen Blick ins Wörterbuch oder in eine Grammatiktafel werfen). Fehler in den Bereichen Tempus, Modus, Konnektoren, Textaufbau, Textsorte usw. lösen dagegen satzübergreifende Bearbeitungsaktivitäten aus, da dabei allgemeine Textmerkmale berücksichtigt werden müssen. (Kast 1999: 172)

2.3 Das Beurteilen

Mehrere Autoren sprechen über das Beurteilen. Einer von ihnen, wie schon in dieser Arbeit erwähnt, ist Gerd Bräuer (1998: 123), der Folgendes zum Thema Beurteilen hervorhebt:

Jedes Mal, wenn wir eine Note erteilen, sprechen wir durch die Struktur einer Institution, die sagt, daß unsere Beurteilung eben nicht nur „ein Teil der Information“ ist, sondern die wichtigste, nämlich die Information, welche zählt: sie erscheint auf dem Zeugnis, sie ist das Urteil der Institution, sie steht für das Urteilsvermögen der anderen LehrerInnen dieser Institution. Auf der Grundlage dieser institutionellen und kulturellen Strukturen motiviert Benotung das Verhalten von Studierenden und beeinflusst das Curriculum. (Bräuer 1998: 123)

Beim Beurteilen von Texten unterscheidet man zwischen dem fördernden und dem bewertend-prüfenden Beurteilen. Die fördernde Beurteilung bezeichnet eine schriftliche oder mündliche Rückmeldung des Lesers an den Schreibenden. Die Funktion der fördernden Beurteilung besteht in der Rückmeldung, Orientierung und der Motivation. Die Beurteilung steht nicht am Ende, sondern sie ist ein wichtiger Teil des Schreibprozesses. Die bewertend-prüfende Beurteilung hingegen, befindet sich am Ende eines Schreibprozesses. Durch das Vergeben von

Noten am Ende nennen wir dieses Beurteilen auch Benotung. Die Noten weisen auf den Lernstand und nicht auf die Lernentwicklung der Lernenden hin (vgl. Bauerman, zit. nach Fix 2006:188 f.). Bräuer (1998: 126) übernimmt von Wendy Bishop drei Aspekte des Bewertens:

1. Die Art und Weise der Beurteilung unterliegt dem Stand der Schreibearbeit (d. h. man unterscheidet die Beurteilung von Erstentwürfen und schon abgeschlossenen Arbeiten).
2. Beurteilungskriterien sollten innerhalb der Gruppe entstehen.
3. „Bewerten ist eine Form des Überarbeitens“ (Bräuer 1998: 126)
(Die subjektiven Äußerungen sind unangebracht, nur objektive Vorschläge zur Textveränderung sollten niedergeschrieben werden).

2.3.1 Die Objektivität beim Beurteilen

Altmeyer und Domisch (1998) behaupten, wenn es um das Beurteilen von Aufsätzen geht, geschieht es oft, dass nicht nur verschiedene Lehrer für den gleichen Aufsatz unterschiedliche Noten geben, sondern derselbe Lehrer gibt nach dem ersten und zweiten Durchlesen eines Aufsatzes unterschiedliche Noten. Die Schüler bekommen in verschiedenen Schulen für den gleichen Aufsatz unterschiedliche Noten, d. h. dass verschiedene Lehrer auf unterschiedliche Art und Weise die gleichen Aufsätze bewerten. Laut Bolton (1996: 131) ist es bei geschlossenen Aufgaben leichter objektiv zu beurteilen als bei Aufsätzen oder Prüfungsgesprächen. Um die Beurteilung auch bei Aufsätzen objektiv zu gestalten, hält er es für notwendig eine Bewertungsanleitung zu erstellen. Diese sollte genau definieren, was und wie benotet wird. Die Bewertungsanleitung dient des Weiteren als Maßstab, mit dem alle weiteren Aufsätze und Leistungen der einzelnen Fremdsprachenlernenden gemessen/bewertet werden können. Jeder Schüler wird nach den gleichen Kriterien bewertet und man distanziert sich von den subjektiven Einflüssen, die oftmals das Bewerten bestimmen.

In der Zeitschrift *Fremdsprache Deutsch* spricht Dominique Macaire über Kriterien zur Bewertung von Schüleraufsätzen. Macaires (1998: 13) Meinung nach, sollte man, um objektiv beurteilen zu können, Kriterien, nach denen man beurteilen will, vorher genau bestimmen, z.B. Verständlichkeit und Korrektheit der Sprache, Angemessenheit der Äußerungen usw. Die Lernziele und Beurteilungskriterien sollten übereinstimmen, da Macaire (ebd.) behauptet, die Bewertung von Aufsätzen erfordere funktionale und formale Kriterien. Dabei versteht sie unter dem Begriff *funktionales Kriterium* die Verständlichkeit und unter dem Begriff *formales Kriterium* die Korrektheit des Geschriebenen. Weiterhin behauptet sie, dass die Erwartungen mit den Schülern besprochen werden sollten und eine standardisierende Form der Korrekturzeichen (die

den Lernern bekannt gemacht wurde) benutzt werden sollte. Empfehlenswert wäre ihrer Meinung nach auch, dass die Schüler eine Fehlerstatistik erstellen und Korrekturen als Übungen verwenden würden.

Wenn man über die Objektivität beim Beurteilen spricht, stellt sich immer wieder die Frage, ob es überhaupt möglich ist, Aufsätze objektiv zu beurteilen.

Es gilt als unbestritten, dass die Bewertung von Texten immer ein gewisses Maß an Subjektivität in sich trägt und sich Textqualitäten niemals objektiv messen und benoten lassen. Dies liegt in der Natur eines Textes an sich, der sich durch die Begriffe Verbindung und Verwobenheit definiert und daher nicht in isoliert überprüfbare Einzelteile zerlegbar ist. So hängt jedes Urteil stark vom Geschmack und Empfinden des Lesers ab (...). (Glaboniat 1998: 20)

Laut Glaboniat (1998: 20) ist für das objektive Beurteilen ein einheitliches Bezugs- und Objektivierungssystem¹ erforderlich. Wichtig ist, die Beurteiler und Lernenden für die Arbeit mit festgelegten Kriterien zu schulen. Lado (1964: 256) behauptet, dass „in scoring short or long compositions without dealing separately with the language and graphic problems, it is possible to improve the consistency of the grades by providing the examiner with sample compositions at each level or grade.“

2.3.2 Warum überhaupt Noten vergeben?

Eine der Hauptaufgaben von Lehrerinnen und Lehrern in der Schule ist die Notengebung. Laut Altmeyer und Domisch (1998: 5) haben Noten eine informierende, vergleichende, selektierende und disziplinierende Funktion. Sie geben den Schülern, wie auch den Eltern, Informationen über den Leistungsstand der Lernenden und ermöglichen den Leistungsvergleich verschiedener Schüler. Darüber hinaus weisen Noten die weiterführenden Bildungsanstalten oder den Arbeitgebern auf die Leistungen der Schüler hin. Sie rechtfertigen Versetzungen und Bildungsabschlüsse, belohnen die Leistungen der Schüler und sollen zu mehr Anstrengung motivieren und anspornen. Die schulische Notengebung unterscheidet sich von Land zu Land, so wird z.B. in Deutschland am häufigsten das Notensystem aus sechs Stufen benutzt: 1 (sehr gut), 2 (gut), 3 (befriedigend), 4 (genügend), 5 (mangelhaft), 6 (ungenügend).

Das kroatische Schulnotensystem unterscheidet sich von dem deutschen in zwei Punkten. Es besteht aus fünf Noten, wobei die Note 5 das Äquivalent der Note 1 in Deutschland wäre: 5 (odličan - ausgezeichnet), 4 (vrlo dobar - ziemlich gut), 3 (dobar - gut), 2 (dovoljan - genügend) und 1 (nedovoljan - ungenügend). Sichtlich ist der Unterschied bei der Notenvergabe, dass die

¹ Das Festlegen und Definieren von Kriterien, auf die man sich in der Bewertung bezieht. (Glaboniat, 1998: 20)

Noten andersrum vergeben werden. In Kroatien werden die größeren Zahlen (Noten) für die bessere Leistung und die niedrigeren Zahlen (Noten) für die schlechtere Leistung vergeben.

Altmeyer und Domisch (1998) hinterfragen, ob Noten die ihnen zugeschriebene, informierende, vergleichende und/oder selektierende Rolle erfüllen. Ihrer Meinung nach bleiben jedoch trotz der Notengebung Fragen offen, auf die Noten keine Antworten geben:

Was sagt uns die Note über die Lernumstände, die zu ihr geführt haben? Was sagt sie über die Bewertungskriterien, die an die Leistung des Schülers gelegt wurde? Was erfahren wir über die Prozesse an deren Ende diese Note steht? Auf alle diese Fragen finden wir in den Noten keine Antworten. (Altmeyer und Domisch 1998: 7)

2.4 Verschiedene Bewertungskriterien

Wenn wir über Beurteilung und Notengebung sprechen, können wir Bewertungskriterien nicht auslassen. Nach Bolton (1996: 132), um Aufsätze bewerten zu können, müsse man erst Kriterien festlegen. Diese Bewertungskriterien sind nach Bolton (1996: 132) Beherrschung der Grammatik, der Orthographie, Beherrschung der Interpunktion und des Textaufbaus.

Für diese Kriterien werden Leistungsstufen von sehr guter Leistung bis nicht mehr angemessener Leistung festgelegt, wie aus der Tabelle 1 ersichtlich ist:

Tabelle 1: Bewertungskriterien nach Bolton (Bolton 1996: 132)

	Grammatik	Wortschatz	Orthographie/ Interpunktion
sehr gut	Keine oder nur vereinzelte Fehler	Variationsbereich	Keine oder nur vereinzelte Fehler
gut	Einige Fehler, die jedoch das Verständnis nicht stören.	Der Aufgabe völlig angemessen.	Einige Fehler, die jedoch das Verständnis nicht stören.
befriedigend	Eine Reihe von Fehlern, die das Verständnis aber nur an einigen wenigen Stellen stören.	Gelegentlich nicht ganz angemessene Wortwahl	Eine Reihe von Fehlern, die das Verständnis aber nur an einigen wenigen Stellen stören.
nicht bestanden	Die Fehler stören das das Verständnis erheblich.	Häufig nicht angemessene Wortwahl	Viele und/ oder gravierende Fehler
Oder: Nicht bestanden	Wegen Themenverfehlung (dann werden für die sprachliche Leistung keine Noten vergeben)		

Nach Valette (1972: 182) sieht man anhand von Aufsätzen, ob und wie weit der Lernende in der Lage ist, seine Gedanken zu ordnen, Sätze zu formulieren, eine entsprechende Wortwahl zu

treffen und seine Gedanken niederzuschreiben. Valette schlägt folgende Aspekte vor, die beurteilt werden können (vgl. Valette 1972: 183):

- Der Wortschatzaspekt (der Gebrauch von Substantiven, Adjektiven, Adverbien und Verben)
- Der grammatische Aspekt (Zeitformen der Verben, Ausdrücke für „müssen“ und „sollen“)

Richardson (1998: 18) vertritt die Meinung, dass die Benotung global sein sollte. Sie solle nicht aus dem Zusammenzählen von Einzelfehlern hervorgehen. Seiner Meinung nach sollte man sich die folgenden Fragen bei der Benotung stellen:

- Hat der/die Schreibende die Aufgabe erfüllt?
- Stimmt der „Ton“?
- Gibt es eine sinnvolle Gedankenführung und eine schlüssige Argumentation?
- Wie reich ist das Vokabular? (ebd.)

Laut Glaboniat (1998) sollte man sich vor der Bewertung drei verschiedene Fragen stellen. Zum einen muss hinterfragt werden, inwieweit das Kommunikationsziel erreicht würde, d.h. die Aufgabe erfüllt würde. Darüber hinaus müsse man sich auch mit Mängeln beschäftigen und hinterfragen, wenn es Mängel gibt, wo diese liegen, „auf rein sprachlicher Ebene oder eher im Bereich des Registers oder der Textsortenadäquatheit?“ (Glaboniat 1998: 22) Um Antworten auf die oben genannten Fragen zu bekommen, sollte man Kriterien aufstellen. Diese Beurteilungskriterien schriftlicher Leistungen basieren laut Glaboniat (1998: 22) auf:

- inhaltlicher Bewältigung der Aufgabe,
- kommunikativer Angemessenheit und Textsortenadäquatheit,
- Textaufbau, Strukturierung,
- Ausdruck,
- formaler Richtigkeit.

Näher erklärt Glaboniat (1998), dass die inhaltliche Bewältigung der Aufgabe sich auf die Detaillierung des vorgegebenen Sachverhaltes und auf die Frage, ob das Thema erfasst ist, bezieht. Bei der kommunikativen Angemessenheit und Textsortenadäquatheit steht im Vordergrund das Sprachverhalten und, ob es situations-, zweck- und absichtsgemäß verwendet wird. Die Beurteilungskriterien, die auf Textaufbau und Strukturierung basieren, beurteilen, ob der Text inhaltlich und formal logisch aufgebaut ist und ob er entsprechend der Aufgabenstellung strukturiert ist. Beim Ausdruck wird beurteilt, ob der Wortschatz und Ausdruck der Aufgabe

entsprechend benutzt wird. Und zuletzt werden bei der formalen Richtigkeit die Kenntnisse in der Morphologie, Syntax und Rechtschreibung beurteilt.

2.4.1 Die Bewertungsskalen

Laut Weigle (2002: 108) ist die Benotung eines Aufsatzes das Ergebnis einer Interaktion, die nicht nur die Testteilnehmer und den Text, sondern den Testteilnehmer, die Aufgabenanforderung, den geschriebenen Text, die Bewerter und die Bewertungsskala beinhaltet. Palmer (1996: 213) bietet zwei Definitionen der Bewertungsskalen: „1 the specific features of the language sample to be rated with the scale, and 2 the definition of scale levels in terms of the degree of mastery of these features.“

Weigle (2002) ist der Auffassung, dass eine der ersten Entscheidungen ist, die man bei der Bestimmung der Aufsatzbewertung treffen muss, welches Bewertungssystem benutzt werden soll. Soll ein Aufsatz global beurteilt werden oder sollte man verschiedene Kriterien zur Beurteilung benutzen? Weigle (2002: 109) nennt vier verschiedene Bewertungsskalen: „primary trait scales, holistic scales, and analytic scales. In recent second- language literature, a fourth typ of scale, called a multiple- trait scale (...)“. Die primären Bewertungsskalen sind spezifisch für bestimmte Schreibtests, im Gegensatz dazu sind die holistischen und analytischen Bewertungsskalen in einem breiteren Aufgabenspektrum zu benutzen.

Laut Bailey (1998: 186) werden drei Vorgehensweisen bei der Beurteilung von schriftlichen Texten verwendet, „these three traditional approaches are holistic, analytic, and objective scoring.“ Die holistischen und analytischen Vorgehensweisen werden detaillierter besprochen.

2.4.1.1 Holistische Bewertungsskala

Baileys Meinung nach (1998: 187) ist die Idee der holistischen Bewertungsskala, dass eine einzige Skala verwendet wird, um unterschiedliche Ebenen der Schreibleistungen zu bewerten. Die Bewerter, die die holistische Bewertungsskala benutzen, sind trainiert, nicht über die einzelnen Komponenten der Schreibfertigkeit nachzudenken oder die Anzahl der Fehler, die die Schüler machen, zu zählen. Der Bewerter reagiert auf das Geschriebene als Ganzes, nur eine Note wird für die geschriebene Arbeit vergeben. Normalerweise ist diese Note auf einer Skala von 1 bis 4, 1 bis 6, oder sogar 1 bis 10. Oft wird jede Ebene auf der Skala durch eine verbale Beschreibung der

Leistung begleitet. Um ein Beispiel zu geben, verweist Bailey (1998: 187-188) auf die Skala für den *Test of Written English* (TWE) aus dem *Test of English as a Foreign Language* (TOEFL).

Die Bewertungsskala ist in sechs Ebenen aufgeteilt:

6. Zeigt eindeutige Schreibkompetenzen auf rhetorischer und syntaktischer Ebene, obwohl es gelegentliche Fehler gibt.
Der Aufsatz:
 - ist gut aufgebaut und bearbeitet
 - es wurden entsprechende Details verwendet, die die Thesen und Ideen unterstützen
 - Einheit, Kohärenz, Progression sichtbar
 - Sprachgebrauch dem Niveau entsprechend
 - demonstriert grammatische Vielfalt und angemessene Wortwahl
5. Zeigt Schreibkompetenzen auf rhetorischer und syntaktischer Ebene, obwohl es gelegentliche Fehler gibt
 - allgemein gut aufgebaut und bearbeitet (kann mehr Fehler als bei Punkt 6 beinhalten)
 - es wurden Details verwendet, die die Thesen und Ideen unterstützen
 - Einheit, Kohärenz, Progression sichtbar
 - Sprachgebrauch dem Niveau entsprechend (kann mehr Fehler als bei Punkt 6 beinhalten)
 - demonstriert die grammatische Anwendung und einen angemessenen Wortschatz mit gelegentlichen Fehlern
4. Zeigt minimale Schreibkompetenzen auf rhetorischer und syntaktischer Ebene
 - ausreichend aufgebaut
 - es wurden einige Details verwendet, die die Thesen und Ideen unterstützen
 - adäquater aber mittelmäßiger Gebrauch der Grammatik und des Wortschatzes
 - kann einige gravierende Fehler enthalten, die das Textverständnis stören können
3. Zeigt etwas an Kompetenzentwicklung im Bereich des Schreibens, bleibt aber fehlerhaft entweder auf der rhetorischen oder syntaktischen Ebene (oder auf beiden Ebenen)
 - nicht ausreichend aufgebaut
 - falsche Verwendung der Details, die die Thesen und Ideen unterstützen
 - Anhäufung von Fehlern im Bereich der Grammatik und/oder des Wortschatzes
 - unangemessene Wortwahl
2. Deutet auf Inkompetenz im Bereich des Schreibens hin
 - falsch aufgebaut
 - nur wenige, oder keine Details, irrelevante Fakten
 - schwere und häufige Fehler im Bereich der Grammatik und des Wortschatzes
1. Inkompetenz im Bereich des Schreibens
 - kein Aufbau des Textes
 - inkohärent
 - schwere und häufige Schreibfehler
 - wenn Fragen bestehen - unfähig diese zu verstehen (Bailey 1998: 187-188)

In ihrer Arbeit beschreibt Weigle (2002: 112), dass bei dem holistischen Verfahren jeder Aufsatz schnell gelesen wird und dann anhand einer Bewertungsskala oder Bewertungsrubrik, in der die Bewertungskriterien beschrieben sind, beurteilt wird. Als einen klaren Vorteil des holistischen Bewertens sieht Weigle (2002: 112), dass es schneller ist einen Aufsatz einmal zu lesen und zu benoten, als ihn mehrmals durchzugehen, jedes Mal mit dem Fokus auf ein anderes Bewertungskriterium gerichtet. Ein weiterer Vorteil ist nach Weigle, dass der Lesende sich auf die Stärken des Textes und nicht auf die Mängel beschränkt. Beim holistischen Bewerten spiegelt sich die erste, persönliche Reaktion des Lesers auf den Text, somit bleibt der erste Eindruck des Ganzen erhalten. Weigle (2002: 114) nennt aber auch Nachteile dieser Art des Bewertens, wie z.B., dass

eine einzige Note keine Informationen über die Schreibfähigkeit und über die verschiedenen Aspekte des Schreibens (Grammatik, Wortschatz, Textorganisation) bietet. Holistisch bewertete Aufsätze geben den Schülern keine Rückmeldungen, so weiß der Schüler z. B. nicht, ob er eine 3 wegen der Grammatik oder wegen dem Wortschatz bekommen hat. Weigles Meinung nach schaffen es die breiten Skalen nicht, wichtige Unterschiede bei den schriftlichen Aufgaben zu erfassen. Ohne klare Kriterien könnte die Länge des Textes und die Handschrift der Schüler die Bewertung des Textes beeinflussen.

2.4.1.2 Analytische Bewertungsskala

Bei dem analytischen Bewerten werden die Aufsätze nach verschiedenen Aspekten oder Kriterien des Schreibens bewertet. Verschiedene Aspekte des Aufsatzes werden jeweils separat betrachtet. Johnson, Penny und Gordon (2009: 159) sagen in ihrem Werk, dass im Gegensatz zum holistischen Bewerten, beim analytischen Bewertungssystem „das Ganze“ in Teile getrennt wird. Die verschiedenen Aspekte des Schreibens werden dann einzeln bewertet. Auch Bailey (1998: 189) hebt hervor, dass „(...) in analytic scoring, raters do assess students' performances on a variety of categories. Indeed, an analytic scoring system embodies hypotheses about the underlying constructs that comprise a given skill.“

Bailey (1998: 190) verweist in ihrem Werk auf die weit verbreitete und bekannte analytische Bewertungsskala für *English as a Second Language* (ESL, deut. *Englisch als Zweitsprache*) (Jacobs, Zingraf, Wortmuth, Hartfield und Hughey 1981), die beschreibt, dass wirksames Schreiben aus fünf Aspekten besteht, und dass einige dieser Aspekte wichtiger sind als die anderen. Diesen fünf Aspekten werden je nach Gewichtung Punkte zugewiesen:

The weights given to these components follow: Content: 13 to 30 points possible; organization: 7 to 20 points possible; vocabulary: 7 to 20 points possible; language use: 5 to 25 points possible; and mechanics: 2 to 5 points possible. The points awarded in these five categories would then be added to yield a total score. (Bailey 1998: 190)

Dieser Ansatz wird als analytisches Bewerten bezeichnet, weil die Skalenentwickler die angenommenen Komponenten der Schreibfertigkeit analysiert haben und aus diesen Komponenten die Bewertungskriterien bestimmt haben. Was man wählt, um die Aufsätze mit einer analytischen Bewertungsskala zu bewerten, hängt von der Operationalisierung ihrer Theorie des Schreibens ab (vgl. Bailey 1998: 190).

Tabelle 2 stellt die Bewertungsskala für ESL (Jacobs, Zingraf, Wortmuth, Hartfield und Hughey 1981) dar.

Tabelle 2: Bewertungsskala für ESL nach Jacobs, Zingraf, Wortmuth, Hartfield und Hugheys (1981) (Weigle 2002: 116)

ESL COMPOSITION PROFILE		
STUDENT	DATE	TOPIC
SCORE	LEVEL	CRITERIA
COMMENTS		
CONTENT	30-27	EXCELLENT TO VERY GOOD: knowledgeable · substantive · thorough development of thesis · relevant to assigned topic
	26-22	GOOD TO AVERAGE: some knowledge of subject · adequate range · limited development of thesis · mostly relevant to topic, but lacks detail
	21-17	FAIR TO POOR: limited knowledge of subject · little substance · inadequate development of topic
	16-13	VERY POOR: does not show knowledge of subject · non-substantive · not pertinent · OR not enough to evaluate
ORGANIZATION	20-18	EXCELLENT TO VERY GOOD: fluent expression · ideas clearly stated/ supported · succinct · well-organized · logical sequencing · cohesive
	17-14	GOOD TO AVERAGE: somewhat choppy · loosely organized but main ideas stand out · limited support · logical but incomplete sequencing
	13-10	FAIR TO POOR: non-fluent · ideas confused or disconnected · lacks logical sequencing and development
	9-7	VERY POOR: does not communicate · no organization · OR not enough to evaluate
VOCABULARY	20-18	EXCELLENT TO VERY GOOD: sophisticated range · effective word/ idiom choice and usage · word form mastery · appropriate register
	17-14	GOOD TO AVERAGE: adequate range · occasional errors of word/idiom form, choice, usage <i>but meaning not obscured</i>
	13-10	FAIR TO POOR: limited range · frequent errors of word/idiom form, choice, usage · <i>meaning confused or obscured</i>
	9-7	VERY POOR: essentially translation · little knowledge of English vocabulary, idioms, word form · OR not enough to evaluate
LANGUAGE USE	25-22	EXCELLENT TO VERY GOOD: effective complex constructions · few errors of agreement, tense, number, word order/function, articles, pronouns, prepositions
	21-18	GOOD TO AVERAGE: effective but simple constructions · minor problems in complex constructions · several errors of agreement, tense, number, word order/function, articles, pronouns, prepositions <i>but meaning seldom obscured</i>
	17-11	FAIR TO POOR: major problems in simple/complex constructions · frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragment, run-ons, deletions · <i>meaning confused or obscured</i>
	10-5	VERY POOR: virtually no mastery of sentence construction rules · dominated by errors · does not communicate · OR not enough to evaluate

MECHANICS	5	EXCELLENT TO VERY GOOD: demonstrates mastery of conventions · few errors of spelling, punctuation, capitalization, paragraphing	
	4	GOOD TO AVERAGE: occasional errors of spelling, punctuation, capitalization, paragraphing <i>but meaning not obscured</i>	
	3	FAIR TO POOR: frequent errors of spelling, punctuation, capitalization, paragraphing · poor handwriting · <i>meaning confused or obscured</i>	
	2	VERY POOR: no mastery of conventions · dominated by errors of spelling, punctuation, capitalization, paragraphing · handwriting illegible · OR not enough to evaluate	
TOTAL SCORE		READER	COMMENTS

Bailey (1998: 192) behauptet, dass die Bewertungskriterien sich ändern, oder besser gesagt, neu entwickelt werden sollten, so wie sich unser theoretisches und praktisches Wissen über die Kriterien weiterentwickelt. Weigle (2002: 114) betont diesbezüglich, dass ein analytisches Bewertungssystem mehr Informationen über die Leistung in verschiedenen Schreibaspekten bietet und aus diesem Grund bei der Bewertung bevorzugt wird. Laut Weigle (2002:120) bietet das analytische Bewertungsverfahren nützliche, diagnostische Informationen über die schriftlichen Fähigkeiten der Schüler. Ihrer Meinung nach ist das analytische Bewertungssystem besonders nützlich beim Training der Beurteiler, weil unerfahrene Beurteiler die Kriterien in separaten Skalen besser verstehen und anwenden können. Die verschiedenen Aspekte des Schreibens werden einzeln betrachtet und bewertet. Ein Aufsatz kann zum Beispiel ganz gut erarbeitet sein, aber zahlreiche grammatische Fehler beinhalten, oder ein Aufsatz kann im Bereich der Morphologie und Syntax als gut bewertet werden, aber im Bereich der Kohäsion und Kohärenz als schlecht bewertet werden. Weigle behauptet darüber hinaus (2020), dass das analytische Bewerten zuverlässiger ist. Ein Notensystem, in dem mehrere Noten (für verschiedene schriftliche Aspekte) für einen Text vorhanden sind, erhöhe die Zuverlässigkeit der Endnote. Am Ende bekommen die Schüler eine Rückmeldung über die verschiedenen Aspekte des Schreibens. Dieser Aspekt fehlt beim holistischen Bewertungsverfahren.

Wie auch das holistische Bewertungssystem hat auch das analytische einige Nachteile. Die Bewertung dauert länger als bei der holistischen Bewertung, da der Bewerter mehrmals den Text durchlesen muss und für verschiedene Aspekte separate Noten vergeben muss. Wenn die Noten aus den unterschiedlichen Skalen kombiniert werden, um eine Endnote zu bilden, geht ein großer Teil der Informationen, die die analytische Bewertungsskala bietet, verloren. Erfahrene Bewerter, die analytische Bewertungsskala benutzen, können im Nachhinein holistisch bewerten. Sie vergeben eine Note, die sie vom Aufsatz erwarten und vergeben dementsprechend die Noten für die verschiedenen Schreibaspekte der analytischen Bewertungsskala (vgl. Weigle 2020: 121).

Als besonders herausfordernd hat sich die Bewertung der Diskurskompetenz gezeigt, wobei sich immer mehr Wissenschaftler (z. B. Chiang 2003; vgl. auch Bagarić Medve und Čelebić 2011) für ein analytisches Bewertungsverfahren in Bezug auf die Diskurskompetenz ausgesprochen haben. Es wird angenommen, dass das analytische Bewertungsverfahren besser zur Erfassung der einzelnen Aspekte bezüglich der Qualität einer schriftlichen Arbeit dient. Darüber hinaus wird durch das analytische Bewertungsverfahren ebenfalls gesichert, dass unter der Kategorie Diskurskompetenz auch tatsächlich nur diesbezügliche Aspekte bewertet werden und nicht z. B. grammatische Aspekte, die zur Grammatikkompetenz gehören. Laut Bagarić Medve und Čelebić (2011: 35) gibt es in Bezug auf die Bewertung von Aufsätzen von Zweit- und Fremdsprachenlernenden eine „Korrelation zwischen den Ergebnissen, die durch die Anwendung unterschiedlicher Skalen gewonnen wurden, aber sie wiesen auch auf Variationen bei den Bewertungsergebnissen hin (vgl. Cohen und Manion 1994, Bacha 2001, Shoonen 2005, Xi 2007, Koch 2009).“

Bagarić Medve und Čelebić (2011) haben in ihrer Untersuchung 155 Aufsätze kroatischer DaF-Lernender holistisch und analytisch in Hinsicht auf zwei Komponenten der Diskurskompetenz (Kohäsion und Kohärenz) und auf zwei Komponenten der Grammatikkompetenz (Syntax und Morphologie) bewertet. Das Ziel war, festzustellen, welches der erwähnten Verfahren eine höhere Bewertungszuverlässigkeit und die Möglichkeit eines tieferen Verständnisses der Diskurs- und Grammatikkompetenz bietet. Ihre Ergebnisse verweisen darauf, dass „der Gebrauch der analytischen Bewertungsweise bei der Bewertung der Diskurs- und Grammatikkompetenz eine bessere Unterscheidung derselben Kompetenzen ermöglicht“ (ebd.: 43).

2.4.1.3 Die Verbindung von einem holistischen und analytischen Bewertungsverfahren

In einer empirischen Untersuchung stellte man fest, „dass die Übereinstimmung mit anderen Beurteilern umso größer ist, je globaler der Zugriff ist“ (Grzeski und Fischer, zit. nach Fix 2006: 199). Im Folgenden wird die Verbindung des holistischen und analytischen Bewertungsverfahrens (Grzeski und Fischer, zit. nach Fix 2006: 199 ff.) erläutert. Zu viele Kriterien erwiesen sich für die Bewertung der Aufsätze im Vergleich mit einer kleineren Anzahl von Kriterien als weniger valide und zuverlässig. Grzeski und Fischer setzen sich für die Kombination von holistischen und analytischen Bewertungsverfahren ein. Laut Fix (2006: 199) ist ein globales Bewertungssystem bei erfahrenen Lehrern legitim. Fix betont jedoch, dass Schüler

und Eltern keine Rückmeldung durch eine globale Bewertung bekommen und daher auf differenziertere Raster angewiesen sind. Die Bewertungskriterien müssen am Ende für alle Beteiligten transparent sein.

Die Verbindung von einem holistischen und analytischen Bewertungsverfahren wird in zwei Durchgängen vorgenommen. Im ersten Durchgang wird die Angemessenheit des Textes durch die Globalbeurteilung (holistisch) bewertet. Darauf folgt der zweite Durchlauf nach klassischen Kriterien (analytisch). Als erstes wird eine Schätznote für den Ausdruck vergeben. Diese wird nicht durch Kriterien differenziert, sondern ergibt sich aus dem ersten Eindruck. Die Textsorte entscheidet dabei, ob es um Originalität, Authentizität oder Poetizität geht. Nachdem die erste Note (Schätznote) vergeben wurde, wird beim zweiten Durchlesen eine Note anhand von Kriterien beschlossen. Zum Schluss der Bewertung werden die zwei Noten verglichen. Wenn Abweichungen bezüglich der Noten bestehen, werden die oben genannten Durchgänge wiederholt und die Note wird durch das neue Urteil bestimmt (vgl. Fix 2006: 201). In seinem Buch beschreibt Fix (2006) dieses Bewertungsverfahren. Im ersten Durchgang wird eine Schätznote von 1 bis 6 vergeben. Im zweiten Durchgang werden Punkte nach Kriterien vergeben. Fix (2006) definiert die Kriterien und Punktevergabe folgendermaßen:

- a) Inhalt und Aufbau (7 bis 8 Punkte),
- b) Formulierungen und Stil (4 bis 5 Punkte),
- c) sprachliche Richtigkeit und Darstellung (2 bis 3 Punkte).

Die Punkte werden in Noten umgerechnet. Drei Punkte ergeben eine Notenstufe. Am Ende wird die Schätznote mit der Kriteriennote verglichen, die Abweichungen werden bearbeitet und die Endnote wird vergeben.

3. Anwendung der analytischen Bewertungsskala bei der Beurteilung diskursiver und grammatischer Kompetenz fortgeschrittener Deutschlerner

Nachdem im theoretischen Teil der Arbeit die Grundlagen für die analytische Bewertung und Bewertungsskala beschrieben wurden, wird im praktischen Teil der Arbeit die Anwendung der analytischen Bewertungsskala bei der Beurteilung diskursiver und grammatischer Kompetenz fortgeschrittener Deutschlerner analysiert. Zu diesem Zweck wurde eine Untersuchung durchgeführt, die aus mehreren Teilen besteht. Alle Untersuchungsergebnisse wurden im Anschluss quantitativ und qualitativ dargestellt und diskutiert.

3.1 Ausgangspunkt

Der Ausgangspunkt der Forschung ist die Tatsache, dass beim Beurteilen von Aufsätzen verschiedene Lehrer für den gleichen Aufsatz unterschiedliche Noten vergeben, oder sogar derselbe Lehrer nach dem ersten und zweiten Durchlesen eines Aufsatzes unterschiedliche Noten vergibt. Wenn man einen Aufsatz in zwei unterschiedlichen Schulen beurteilen lässt, kann es geschehen, dass sich daraus zwei unterschiedliche, unobjektive Noten ergeben. Durch diese Unobjektivität werden die Schüler in ihrer Sprech- und Schreiblust durch Noten demotiviert. Um die diskursiven und grammatischen Kompetenzen der Fremdsprachenlernenden beurteilen zu können, ist eine Bewertungsskala notwendig. Oft werden bei der Beurteilung von Aufsätzen holistische Bewertungsskalen benutzt, da man mit diesen die Aufsätze schneller beurteilen kann. Außer der holistischen Bewertungsskala dient die analytische Bewertungsskala, die im theoretischen Teil der Arbeit näher beschrieben wurde, als Maßstab zur Beurteilung von Aufsätzen.

Es stellt sich nun die Frage, ob die Anwendung von analytischen Bewertungsskalen zur objektiveren, zuverlässigeren Beurteilung diskursiver und grammatischer Kompetenz führt. Die Analyse der Anwendung analytischer Bewertungsskalen bei der Beurteilung diskursiver und grammatischer Kompetenz fortgeschrittener Deutschlerner soll uns helfen, die Antwort auf die oben gestellte Frage zu finden.

3.2 Zielstellung

In dieser Arbeit analysieren wir die Anwendung der analytischen Bewertungsskala bei der Beurteilung diskursiver (Kohäsion und Kohärenz) und grammatischer (Syntax und Morphologie) Kompetenz. Die Forschungsfragen, auf die wir eine Antwort suchen, und die dazugehörigen Hypothesen lauten:

- Forschungsfrage 1: Bestehen signifikante Unterschiede zwischen den Ergebnissen der Schüleraufsätze in Bezug auf die oben angeführten Aspekte der Diskurs- und Grammatikkompetenz beim analytischen und beim holistischen Bewertungsverfahren? Führt die Anwendung von analytischen Bewertungsskalen zur objektiveren, zuverlässigeren Beurteilung diskursiver und grammatischer Kompetenzen?

Hypothese 1: Man geht davon aus, dass es Unterschiede zwischen den Ergebnissen der Schüleraufsätze in Bezug auf die oben angeführten Aspekte der Diskurs- und Grammatikkompetenz beim analytischen und beim holistischen Bewertungsverfahren geben wird. Die statistische Analyse wird zeigen, dass der Gebrauch der analytischen Bewertungsskala bei der Bewertung der Diskurs- und Grammatikkompetenz eine bessere Unterscheidung dieser Kompetenzen ermöglicht. Auf diese Weise bekommt man zuverlässigere Daten über das Niveau der Diskurs- und Grammatikkompetenz.

- Forschungsfrage 2: In welcher Beziehung stehen die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim analytischen und die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim holistischen Bewertungsverfahren?

Hypothese 2: Man geht davon aus, dass zwischen den Aspekten der Diskurskompetenz (Kohärenz und Kohäsion) und den Aspekten der Grammatikkompetenz (Morphologie und Syntax) bei dem analytischen Bewertungsverfahren und dem Bewertungsverfahren des Probestaatsabitur (eher holistisch) statistisch betrachtet teilweise eine Korrelation bestehen wird. Die Ergebnisse der einzelnen hier untersuchten Aspekte der Grammatikkompetenz bei dem analytischen und dem holistischen Bewertungsverfahren werden stärker miteinander korrelieren als die Ergebnisse bezüglich der Diskurskompetenz.

3.3 Forschungsdesign

3.3.1 Probanden

Die Untersuchung umfasst 58 Aufsätze fortgeschrittener Deutschlerner aus kroatischen Mittelschulen. Die Aufsätze entstanden beim Probestaatsabitur und beim ersten Staatsabitur als schriftliche Prüfung aus dem Unterrichtsfach Deutsch auf der höheren Stufe. Die Probanden besuchten die vierte Klasse der Mittelschule und lernten seit mindestens acht Jahren Deutsch in der Schule. Der Unterricht der höheren Stufe fand nach dem Unterrichtsplan und -programm des Gymnasiums und laut dem Gemeinsamen europäischen Referenzrahmen auf dem B2 Sprachniveau statt.

Da die Beurteilung beim Probe- und Staatsabitur nach objektiven Kriterien vollzogen werden sollte, handelte es sich bei der Auswahl der Aufsätze um ein Zufallsauswahlverfahren. Die Beurteiler hatten keine Informationen über die Schreiber. Die Aufsätze wurden mit Ziffern gekennzeichnet, um die Identität der Schüler zu wahren. Die Tatsache, dass die Aufsätze durch ein Zufallsauswahlverfahren ausgewählt wurden, verleiht der Untersuchung mehr Repräsentativität, da die Aufsätze, die in der Untersuchung benutzt worden sind, den Durchschnitt der Deutschkenntnisse fortgeschrittener Deutschlerner an kroatischen Mittelschulen repräsentieren.

Prof. Dr. Vesna Bagarić Medve reichte mir die Aufsätze weiter, die ihr durch das Nationale Zentrum zur externen Evaluation des Bildungswesens (kroatisch: *Nacionalni centar za vanjsko vrednovanje obrazovanja*, NCVVO) zu Forschungszwecken überlassen wurden. Nur ein Teilkorpus der Arbeiten wurde für unsere Forschungszwecke benutzt, wie bereits oben schon erwähnt, insgesamt 58 Aufsätze.

3.3.2 Instrument

Beim Probe- und Staatsabitur bestand der schriftliche Teil der Aufgabe darin, einen Aufsatz zu einem vorgeschriebenen Thema zu schreiben. Die Schüler bekamen Leitfragen, die ihnen helfen sollten, den Text zu erstellen. Der Aufsatz musste 200 bis 250 Wörter umfassen.

Die Aufsätze wurden danach von Bewertern nach einer holistischen Bewertungsskala beurteilt. Diese Bewertungsskala enthielt vier Kriterien: Erfüllung der Aufgabe, Kohärenz und Kohäsion, Wortschatz, Grammatik. Diese Kriterien wurden jeweils durch zwei bis drei

Deskriptoren beschrieben. Für jede der vier Kriterien vergaben die Bewerter Punkte, die am Ende zusammengezählt und in eine Note umgewandelt wurden.

Als Forschungsinstrument diente für die Beurteilung der Aufsätze eine analytische Bewertungsskala, die nach Chiangs (2003) Bewertungsskala von Merkmalen der Diskurs- und Grammatikkompetenz angepasst wurde. Chiang entwickelte die Bewertungsskala für schriftliche Arbeiten in englischer Sprache. Da aber in unserer Untersuchung Aufsätze in deutscher Sprache geschrieben wurden, wurde Chiangs Bewertungsskala den Umständen entsprechend angepasst. Die analytische Bewertungsskala (siehe Anhang 1), die in unserer Untersuchung benutzt wurde, besteht aus 4 Kriterien (Kohärenz, Kohäsion, Morphologie und Syntax) und jeweils 5 Deskriptoren. Für die einzelnen Deskriptoren werden die Punkte 0 bis 5 (5 = volle Übereinstimmung; 4 = Übereinstimmung; 3 = Unentschlossenheit; 2 = ich stimme nicht überein; 1 = ich stimme überhaupt nicht überein; 0 = kann nicht bewertet werden) vergeben. Am Ende werden die Punkte, genauso wie bei dem holistischen Bewertungsverfahren, zusammengezählt und daraus ergibt sich eine Endnote.

3.3.3 Datenerhebung

Die Datenerhebung, bzw. die Durchführung der Untersuchung, war ein Schritt der Forschungsarbeit, auf den ich keinen Einfluss hatte, da dieser schon vor dieser Untersuchung vorgenommen wurde. Die Daten, die in dieser Untersuchung benutzt werden, stammen von der schriftlichen Prüfung der Deutschlerner beim Probestaats- und Staatsabitur in Kroatien, bei denen die Schüler für den schriftlichen Teil der Prüfung einen Aufsatz schreiben mussten. Wie schon erwähnt, wurden die Aufsätze Prof. Dr. Bagarić Medve zu Untersuchungszwecken zur Verfügung gestellt, die diese dann an mich für diese Forschungsarbeit weiterreichte.

Die schriftliche Aufgabe bestand darin, einen Aufsatz zum vorgegebenen Thema zu schreiben. Für die Aufgabe hatten die Schüler 75 Minuten zur Verfügung. Beim Schreiben durften die Schüler kein Wörterbuch oder andere Hilfsmittel benutzen.

Die Gliederung der Aufgabe sollte in Form von Einleitung, Hauptteil und Schlussteil gestaltet werden. Der Text musste 200 bis 250 Wörter umfassen. Die maximale Anzahl von Wörtern wurde jedoch nicht begrenzt. Es wurde darauf hingewiesen, dass eine geringe Anzahl von Wörtern zu Punktabzügen führen wird. Als Schreibhilfe befanden sich auf der Aufgabenstellung Leitfragen zum vorgegebenen Thema, auf die die Schüler im Hauptteil ihres Aufsatzes eingehen sollten.

Die Aufsätze, die im Rahmen dieser Untersuchung benutzt wurden, stammen aus dem Probestaatsabitur aus den Jahren 2009 und 2010. Im Jahr 2009 lautete die Aufgabe im schriftlichen Teil der Prüfung, einen Aufsatz mit dem Thema *Jugendliche und Kleidung* zu schreiben. Die Leitfragen zum Hauptteil, die die Schüler als Hilfe bekamen, lauteten:

- Welche Rolle spielt Kleidung im Leben der Jugendlichen?
- Welche Vor- und Nachteile haben teure Klamotten im Alltag?
- Wie wichtig sind Kleidung und Outfit für Sie und warum?

Im Jahr 2010 lautete das Aufsatzthema *Medien*. Die Leitfragen für den Hauptteil lauteten:

- Wie beeinflussen Medien das Leben von Jugendlichen?
- Welche Vorteile und welche Nachteile haben Mobiltelefone als Kommunikationsmittel in der heutigen Welt?
- Welche Medien bevorzugen Sie und warum?

3.3.4 Datenanalyse

Nach der Datenerhebung folgte die Datenanalyse. Als erster Schritt wurden vor dem Lesen der Aufsätze die Kriterien für die analytische Beurteilungsskala festgelegt. Die folgende Beurteilungsskala besteht aus vier Kriterien (Kohärenz, Kohäsion, Morphologie und Syntax) mit jeweils fünf Deskriptoren. Für die einzelnen Deskriptoren wurden die Punkte 0 bis 5 vergeben. Die Bewertungsskala wurde von Chiang (2003) abgeleitet.

Kohärenz

A Die Reihenfolge der Ideen ist klar, logisch und kohärent (zusammenhängend), der rote Faden zieht sich durch den Text, was dazu führt, dass der Text verständlich ist.

B Neue Ideen (Informationen) werden an der richtigen Stelle und in angemessener Weise eingeführt.

C Die Ideen, die diskutiert/genannt werden, sind erarbeitet.

D Auszüge sind logisch und gerechtfertigt im Hinblick auf die Themen und die vorgeschriebenen Punkte (Relevanz).

E Der Übergang zwischen den Absätzen ist fließend.

Kohäsion

- A Die Verwendung von grammatischen Kohäsionsmitteln ist vielfältig.
- B Grammatische Kohäsionsmittel werden angemessen verwendet.
- C Lexikalische Äquivalente, Paraphrasen und sinnvolle Wiederholung (lexikalische Kohäsionsmittel) werden verwendet.
- D Die Interpunktion wird bei der Trennung von Ideen und Sätze verwendet.
- E An den Stellen, wo keine Kohäsionsmittel verwendet werden, ist der Übergang zwischen den Sätzen fließend.

Syntax

- A Die Reihenfolge der Wörter in einem Satz ist richtig – Genauigkeit.
- B Kongruenz (Übereinstimmung von Subjekt - Prädikat, Artikel und Substantiven) ist richtig hergestellt.
- C Komplexere Verbformen und Phrasen (das Passiv, der Partizip, Modalverben, der Konjunktiv, die Komparation) sind richtig vertreten.
- D Satzgefüge sind vertreten und vielfältig.
- E Angemessenheit der Verwendung aller Satzarten im Text

Morphologie

- A Alle Artikel (bestimmter, unbestimmter und der Nullartikel) werden richtig verwendet.
- B Deklination und Komparation werden richtig verwendet.
- C Konjugation und Zeitformen werden richtig verwendet. Zeitformen werden richtig gebildet.
- D Präpositionen und die Rektion werden richtig verwendet.
- E Die Wortbildung ist richtig (Wörter waren richtig gebildet).

Nach der Festlegung der Bewertungsskala wurden die Aufsätze gelesen und nach der analytischen Bewertungsskala wurden für jeden Aufsatz Punkte von 0 bis 5 für die jeweiligen Deskriptoren vergeben (5 = volle Übereinstimmung; 4 = Übereinstimmung; 3 = Unentschlossenheit; 2 = ich stimme nicht überein; 1 = ich stimme überhaupt nicht überein; 0 = kann nicht beurteilt werden). Die Punkte wurden zur besseren Übersicht in eine Tabelle eingetragen (siehe Anhang 2). Danach wurden die Punkte für jeden Aufsatz zusammengezählt und eine Gesamtnote wurde ermittelt. Zuletzt wurde die Datenanalyse der Forschungsergebnisse, in der Verfahren der deskriptiven Statistik, des t-Testes und der Korrelationsanalyse verwendet wurden, mit Hilfe des Softwareprogramms SPSS für Windows durchgeführt.

3.4 Forschungsergebnisse

Um Antworten auf unsere Forschungsfragen zu bekommen, untersuchten wir verschiedene Aspekte und Unterschiede zwischen den Mittelwerten der Ergebnisse bezüglich der Diskurs- und der grammatischen Kompetenz beim analytischen und holistischen Bewertungsverfahren, sowie die Korrelationen der angeführten Kompetenzen bei den eben genannten Bewertungsverfahren. Wie schon erwähnt, um die Ergebnisse zu bekommen, wurden quantitative Verfahren der deskriptiven Statistik, des t-Testes und der Korrelationsanalyse verwendet. Im Anschluss werden die quantitativen Ergebnisse (Tabelle 4 bis 11) interpretiert und in Kapitel 3.5 weiter diskutiert.

3.4.1 Deskriptive Statistik

Das Verfahren der deskriptiven Statistik wurde durchgeführt, um die gewonnenen Daten zu beschreiben und auszuwerten, bzw. um den Mittelwert der Daten zu ermitteln.

Tabelle 4 zeigt die arithmetischen Mittelwerte (Durchschnittswerte) bei der Beurteilung der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenz (Syntax und Morphologie) beim Probestaatsabitur (eher holistisches Beurteilungsverfahren²) und beim erneuten analytischen Bewertungsverfahren.

Tabelle 4: Deskriptive Statistik für das Probestaatsabitur (eher holistisch) und die wiederholte analytische Bewertung (WAB)

	N	Min	Max	AM	SA
Probestaatsabitur: Kohärenz, Kohäsion	67	,00	5,00	3,3657	1,41844
Probestaatsabitur: Morphologie, Syntax	67	,00	5,00	3,2388	1,50594
WAB: Kohärenz, Kohäsion	68	,00	4,90	3,2338	1,21761
WAB: Morphologie, Syntax	68	,00	5,00	3,3221	1,31147

Anhand der Mittelwerte ist aus der Tabelle sichtbar, dass die Durchschnittswerte beim Probestaatsabitur und der erneuten analytischen Beurteilung ziemlich ähnlich sind. Der Durchschnittswert beim Probestaatsabitur liegt für die Diskurskompetenz (Kohärenz, Kohäsion) bei 3,37 und für die Grammatikkompetenz (Syntax und Morphologie) bei 3,24 und bei der

² Beim Probestaatsabitur wurde das holistische Bewertungsverfahren bevorzugt.

analytischen Beurteilung für die Diskurskompetenz (Kohärenz, Kohäsion) bei 3,23 und für die Grammatikkompetenz (Syntax und Morphologie) bei 3,32. Die Durchschnittswerte aller Kriterien zeigen daher Ähnlichkeiten auf und liegen um den Wert 3,2 oder 3,3.

Anhand der Ergebnisse ist ersichtlich, dass die Grammatikkompetenz (Syntax und Morphologie) beim Probestaatsabitur strenger benotet wurde. Im Gegensatz dazu wurde beim erneuten analytischen Bewerten die Diskurskompetenz strenger bewertet. Es gibt keine relevanten Abweichungen bei den oben genannten Beurteilungsverfahren oder Kriterien, die sich auf die Beurteilungsverfahren beziehen. Die Ergebnisse zeigen, dass die Mehrzahl der Aufsätze als mittelmäßig bewertet wurde.

Tabelle 5 zeigt den Mittelwert der erzielten Punkte bei der analytischen Beurteilung der verschiedenen Deskriptoren für die Diskurskompetenz (Kohärenz, Kohäsion) und Grammatikkompetenz (Morphologie und Syntax).

Tabelle 5: Deskriptive Statistik der einzelnen Kriterien beim analytischen Beurteilungsverfahren der Diskurskompetenz (Kohärenz und Kohäsion, Punkte 1 bis 10) und Grammatikkompetenz (Morphologie und Syntax, Punkte 11 bis 20)

		N	Min	Max	AM	SA
1.	Die Reihenfolge der Ideen ist klar, logisch und zusammenhängend	68	,00	5,00	3,1765	1,53507
2.	Neue Ideen werden an der richtigen Stelle und in angemessener Weise eingeführt	68	,00	5,00	2,9265	1,51919
3.	Die Ideen, die diskutiert/genannt werden sind erarbeitet	68	,00	5,00	3,5588	1,38631
4.	Die Auszüge sind logisch und gerechtfertigt im Hinblick auf die Themen und die Relevanz	68	,00	5,00	3,1029	1,40510
5.	Der Übergang zwischen den Absätzen ist fließend	68	,00	5,00	3,0000	1,55496
6.	Die Verwendung von grammatischen Kohäsionsmitteln ist vielfältig	68	,00	5,00	3,3824	1,24624

7.	Die grammatischen Kohäsionsmittel werden angemessen verwendet	68	,00	5,00	3,5882	1,26060
8.	Es werden lexikalische Äquivalente, Paraphrasen und sinnvolle Wiederholung verwendet	68	,00	5,00	3,2059	1,31079
9.	Die Interpunktion wird bei der Trennung von Ideen und Sätzen verwendet	68	,00	5,00	2,9559	1,23909
10.	An den Stellen wo keine Kohäsionsmittel verwendet werden, ist der Übergang zwischen den Sätzen fließend	68	,00	5,00	3,4412	1,46992
11.	Die Reihenfolge der Wörter im Satz	68	,00	5,00	3,3676	1,53472
12.	Kongruenz – Übereinstimmung von Subjekt – Prädikat, Artikel und Substantiven	68	,00	5,00	3,8088	1,28432
13.	Komplexere Verbformen und Phrasen sind vertreten	68	,00	5,00	3,0147	1,51108
14.	Satzgefüge - sind vertreten und vielfältig	68	,00	5,00	3,1765	1,37071
15.	Angemessenheit der Verwendung aller Satzarten im Text	68	,00	5,00	3,5147	1,29859
16.	Verwendung der Artikel	68	,00	5,00	3,5735	1,52899
17.	Deklination und Komparation	68	,00	5,00	3,0735	1,60518
18.	Konjugation und das Bilden der Verbformen	68	,00	5,00	3,4265	1,47937
19.	Präpositionen und Rektion	68	,00	5,00	3,2941	1,52647
20.	Wortbildung	68	,00	5,00	2,9706	1,42442
	Valid N (listwise)	68				

Die Punkte, die vergeben wurden, reichten von 0 (*kann nicht beurteilt werden*) bis 5 (*völlige Übereinstimmung*). Beim näheren Betrachten der Tabelle sieht man, dass die Mittelwerte der vergebenen Punkte zwischen 2,9 und 3,8 variieren. Die wenigsten Punkte wurden für die Deskriptoren 2 (*Neue Ideen werden an der richtigen Stelle und in angemessener Weise eingeführt*), 9 (*Die Interpunktion wird bei der Trennung von Ideen und Sätzen verwendet*) und 20 (*Wortbildung*) vergeben. Die eben genannten Deskriptoren liegen bei 2,9.

Für die Deskriptoren 3 (*Die Ideen, die diskutiert/genannt werden sind erarbeitet*), 7 (*Die grammatischen Kohäsionsmittel werden angemessen verwendet*), 15 (*Angemessenheit der Verwendung aller Satzarten im Text*) und 16 (*Verwendung der Artikel*) wurden die meisten Punkte vergeben. Der Mittelwert bei diesen Deskriptoren liegt im Bereich 3,5.

Nur ein einziger Deskriptor lag über dem Mittelwert 3,5 und zwar der Deskriptor 12 (*Kongruenz – Übereinstimmung von Subjekt – Prädikat, Artikel und Substantiven*), dessen Mittelwert bei 3,81 lag.

Wenn man sich die Ergebnisse der deskriptiven Statistik anschaut, kann festgestellt werden, dass die Schüler beim Verfassen des Aufsatzes die größten Schwierigkeiten beim Einführen neuer Ideen, der Interpunktion und der Wortbildung hatten, und die wenigsten Schwierigkeiten bei der Deklination und Konjugation.

Vergleicht man nun den Mittelwert der verschiedenen Deskriptoren nach den vier Kriterien, lassen sich folgende Ergebnisse ablesen:

Tabelle 6: Deskriptive Statistik der analytischen Beurteilung nach den Kriterien der Diskurskompetenz (Kohärenz, Kohäsion) und Grammatikkompetenz (Morphologie und Syntax)

	N	Min	Max	AM	SA
Kohärenz	68	,00	5,00	3,1529	1,41722
Kohäsion	68	,00	5,00	3,3147	1,12200
Morphologie	68	,00	5,00	3,2676	1,37899
Syntax	68	,00	5,00	3,3765	1,28993

Bei allen vier Kriterien liegt der arithmetische Mittelwert über dem Wert 3 und unter dem Wert 3,38, also bei einem Mittelwert von 3,2779. Es sind keine größeren Abweichungen feststellbar. Die meisten Punkte wurden im Bereich der Syntax vergeben und die wenigsten im Bereich der Kohärenz. Den Ergebnissen der deskriptiven Statistik zufolge, zeigen die Schüler Schwächen im Bereich der Kohärenz und die wenigsten (mit nur 0,2366 Vorsprung) im Bereich der Syntax.

3.4.2 T-Test für unabhängige Stichproben

Durch den t-Test für unabhängige Stichproben sollte überprüft werden, ob es einen bedeutenden Unterschied zwischen dem holistischen und dem analytischen Bewertungsverfahren in Bezug auf die Aspekte Kohäsion und Kohärenz aus dem Bereich der Diskurskompetenz und in Bezug auf die Aspekte der Syntax und Morphologie aus dem Bereich der Grammatikkompetenz gibt. Der t-Test sollte ebenfalls überprüfen, welches der zwei Bewertungsverfahren zuverlässigere Daten über die Grammatik- und Diskurskompetenz liefert.

Tabelle 7: Ergebnis des T-Tests für unabhängige Stichproben für die Variablen Diskurskompetenz (DK) vs. Grammatikkompetenz (GK): holistische (hol) und analytische (anl) Bewertungsweise

Variable Diskurskompetenz (DK) vs. Grammatikkompetenz (GK): holistische (hol) und analytische (anl) Bewertungsweise		hol	anl	t	Sig.
DKhol: DKanl	AM	3,3657	3,2119	1,766	,082**
	SD	1,41844	1,21325		
GKhol: GKanl	AM	3,2388	3,300	-,727	,470**
	SD	1,50594	1,30860		

**p< .01

Die Mittelwerte der Bewertungsergebnisse des analytischen und holistischen Bewertungsverfahrens der Diskurs- und Grammatikkompetenz sind sehr nah beieinander. Wie aus der Tabelle 7 erkennbar ist, besteht kein statistisch signifikanter Unterschied zwischen den Ergebnissen der Diskurskompetenz (Kohärenz und Kohäsion) beim holistischen und analytischen Bewertungsverfahren [T=1,766; p=,082], sowie der Grammatikkompetenz (Syntax und Morphologie) bei den zwei Bewertungsverfahren [T=-,727, p=,470].

3.4.3 Korrelationsanalyse

Um die Beziehung zwischen den Ergebnissen der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim analytischen Bewertungsverfahren (Probestaatsabitur) und die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim holistischen Bewertungsverfahren zu untersuchen, wurde eine

Korrelationsanalyse (Korrelation nach Pearson) durchgeführt. Die Ergebnisse dieser Analyse sind anhand der Tabellen 8 und 9 zu deuten.

Tabelle 8: Korrelationsanalyse zwischen der analytisch beurteilten Kohärenz und dem Probestaatsabitur (eher holistisch)

Korrelation (Pearson)	Die Reihenfolge der Ideen ist klar, logisch und zusammenhängend	Neue Ideen werden an der richtigen Stelle und in angemessener Weise eingeführt	Ideen, die genannt werden sind erarbeitet	Die Auszüge sind logisch und gerechtfertigt im Hinblick auf die Themen und die Relevanz	Der Übergang zwischen den Absätzen ist fließend
Probestaatsabitur (Kohäsion u. Kohärenz)	,798(**)	,800(**)	,793(**)	,786(**)	,775(**)

Tabelle 9: Korrelationsanalyse zwischen der analytisch beurteilten Kohäsion und dem Probestaatsabitur (eher holistisch)

Korrelation (Pearson)	Verwendung von gramm. KM ist vielfältig	Gramm. KM werden angemessen verwendet	Lexi. Äquivalente, Paraphrasen und sinnvolle Wiederholungen werden verwendet	Interpunktion wird bei der Trennung von Ideen und Sätzen verwendet	An den Stellen wo keine KM verwendet werden, ist der Übergang zwischen den Sätzen fließend
Probestaatsabitur (Kohäsion u. Kohärenz)	,786(**)	,722(**)	,747(**)	,510(**)	,795(**)

**Korrelation ist signifikant bei dem 0.01 Niveau (2-tailed).

Aus den Tabellen 8 und 9 geht hervor, dass zwischen den untersuchten Aspekten der Diskurskompetenz beim analytischen Bewertungsverfahren und dem Bewertungsverfahren des Probestaatsabitur (eher holistisch) eine statistisch signifikante Korrelation festgestellt werden konnte.

Der höchste Korrelationsgrad wurde zwischen dem Probestaatsabitur (ein eher holistisches Beurteilungsverfahren) und zwischen den ersten drei Aspekten der Kohärenz (*Die Reihenfolge der Ideen ist klar, logisch und zusammenhängend* ($r=,798^{**}$); *neue Ideen werden an der richtigen Stelle und in angemessener Weise eingeführt* ($r=,800^{**}$); *die Ideen, die genannt werden sind erarbeitet* ($r=,793^{**}$)) und des letzten Aspektes der Kohäsion (*An den Stellen wo keine Kohäsionsmittel verwendet werden, ist der Übergang zwischen den Sätzen fließend*) des analytischen Beurteilungsverfahrens festgestellt ($r=,795^{**}$).

Bei dem Aspekt der Kohärenz (*Die Auszüge sind logisch und gerechtfertigt im Hinblick auf die Themen und die Relevanz*) und dem Aspekt der Kohäsion (*Verwendung von grammatischen Kohäsionsmitteln ist vielfältig*) wurde ebenfalls eine statistisch signifikante Korrelation festgestellt ($r=,786^{**}$).

Der etwas niedrigere Korrelationsgrad wurde beim Aspekt der Kohäsion (*Interpunktion wird bei der Trennung von Ideen und Sätzen verwendet*) festgestellt ($r=,510^{**}$), obwohl festgehalten werden muss, dass es sich noch immer um einen hohen Korrelationsgrad handelt.

Da fast alle Korrelationsgrade der verschiedenen Aspekte nah beieinanderliegen, sticht der Aspekt der Interpunktion mit dem höchsten Korrelationsunterschied hervor.

Betrachtet man nun beide Tabellen im Unterschied zu einander, wird ersichtlich, dass die höchste Korrelation beim Probestaatsabitur und dem erneuten analytischen Bewertungsverfahren im Bereich der Kohärenz liegt.

Um den grammatischen Aspekt (Syntax und Morphologie) nicht außer Acht zu lassen, wird in den Tabellen 10 und 11 die statistisch signifikante Korrelation im Bereich der holistischen und analytischen Grammatikkompetenzbewertung dargestellt.

Tabelle 10: Korrelationsanalyse der holistischen und analytischen Beurteilung der Syntax

	Reihenfolge der Wörter im Satz	Kongruenz – Übereinstimmung von Subjekt – Prädikat, Artikel und Substantiven	Komplexe Verbformen und Phrasen sind vertreten	Das Vorkommen und die Vielfalt der Satzgefüge	Angemessenheit der Verwendung aller Satzarten im Text
Probestaatsabitur Grammatik	,830(**)	,797(**)	,822(**)	,825(**)	,859(**)

**Korrelation ist signifikant bei dem 0.01 Niveau (2-tailed).

Tabelle 11: Korrelationsanalyse der holistischen und analytischen Beurteilung der Morphologie

	Verwendung der Artikel	Deklination und Komparation	Konjugation und das Bilden der Verbformen	Präpositionen und Rektion	Wortbildung
Probestaatsabitur Grammatik	,744(**)	,760(**)	,817(**)	,744(**)	,814(**)

**Korrelation ist signifikant bei dem 0.01 Niveau (2-tailed).

Genau wie bei der Diskurskompetenz (siehe Tabelle 8 und 9) wurde auch zwischen den Aspekten der Grammatikkompetenz (Syntax und Morphologie) des analytischen Bewertungsverfahrens und des Bewertungsverfahrens beim Probestaatsabitur (eher holistisch) eine statistisch signifikante Korrelation festgestellt.

Der höchste Korrelationsgrad wurde zwischen den Aspekten der Syntax (*Angemessenheit der Verwendung aller Satzarten im Text* ($r=,859^{**}$); *Reihenfolge der Wörter im Satz* ($r=,830^{**}$)) bei dem analytischen Bewertungsverfahren und dem holistischen Beurteilungsverfahren des Probestaatsabiturs festgestellt. Der etwas niedrigere Korrelationsgrad ist bei den Aspekten der Morphologie *Verwendung der Artikel* ($r=,744^{**}$) und *Präpositionen und Rektion* ($r=,744^{**}$) erkennbar.

Betrachtet man nun die Korrelationen aus Tabelle 10 und 11, bzw. die Korrelationsanalysen der holistischen und analytischen Beurteilung bezüglich der Syntax und Morphologie, kommt man zu interessanten Ergebnissen. Im Bereich der Syntax (siehe Tabelle 10) sind fast alle Korrelationsgrade höher als im Bereich der Morphologie. Obwohl im Bereich der Syntax und Morphologie Korrelationen zwischen den zwei Bewertungsverfahren festgestellt wurden, ist durch die Tabellen ersichtlich, dass im Bereich der Syntax (Tabelle 10) fast alle Korrelationsgrade höher als im Bereich der Morphologie (Tabelle 11) sind. Die Korrelation zwischen dem holistischen und dem analytischen Bewertungsverfahren ist im Bereich der Syntax etwas höher als sie es im Bereich der Morphologie ist.

3.5 Diskussion

Nachdem die zwei Forschungsfragen quantitativ bearbeitet wurden, wird nun versucht, die gewonnenen Resultate zu interpretieren und die unterschiedlichen Resultate in Bezug auf alle relevanten Faktoren zu beleuchten.

In der ersten Frage haben wir uns damit beschäftigt, ob signifikante Unterschiede zwischen den Ergebnissen der Schüleraufsätze in Bezug auf die Diskurskompetenz (Kohärenz und Kohäsion) und auch in Bezug auf die grammatische Kompetenz (Morphologie und Syntax) sowohl beim analytischen als auch beim holistischen Bewertungsverfahren bestehen. Die Ergebnisse der statistischen Analyse (siehe Tabelle 4) haben keinen signifikanten Unterschied bei der Bewertung der Diskurs- und Grammatikkompetenz beim Probestaatsabitur (eher holistisches Beurteilungsverfahren) und beim erneuten analytischen Bewertungsverfahren gezeigt. Dies bedeutet, dass unsere erste Hypothese, die davon ausging, dass der Gebrauch der analytischen Bewertungsskala bei der Bewertung der Diskurs- und Grammatikkompetenz eine bessere Unterscheidung derselben Kompetenzen ermöglicht, nicht bestätigt werden konnte. Die Durchschnittswerte in den untersuchten Bereichen der Diskurs- und Grammatikkompetenz liegen sehr nah beieinander. Da in dieser Untersuchung 58 Aufsätze fortgeschrittener Deutschlerner aus kroatischen Mittelschulen untersucht wurden, könnte das einer der möglichen Gründe für die gewonnenen Ergebnisse sein.

Die weiteren Resultate zeigten (siehe Tabelle 4), dass die Grammatikkompetenz beim Probestaatsabitur strenger benotet wurde. Im Gegensatz dazu wurde beim erneuten analytischen Bewerten die Diskurskompetenz strenger bewertet. Es ist anzunehmen, dass sich die Lehrer beim Probestaatsabitur mehr auf die Grammatikkompetenzen und Sprachfehler konzentriert haben als auf die Diskurskompetenzen.

Den Ergebnissen nach wurden die meisten Punkte im Bereich der Syntax und die wenigsten im Bereich der Kohärenz vergeben (siehe Tabelle 6). Das legt die Folgerung nahe, dass die Schüler die meisten Schwächen im Bereich der Kohärenz und die wenigsten im Bereich der Syntax zeigten. Ein möglicher Grund dafür könnte sein, dass man den Grammatikkompetenzen, insbesondere der Syntax, mehr Aufmerksamkeit schenkt als den Diskurskompetenzen. Den quantitativen Resultaten zufolge hatten die Schüler die meisten Schwierigkeiten beim Einführen neuer Ideen, der Interpunktion und der Wortbildung, und die wenigsten Schwierigkeiten bei der Deklination und Konjugation (vgl. Tabelle 5). Das könnte darauf hinweisen, dass sich der Unterricht mehr auf die Grammatikkompetenzen richtet und weniger auf die Diskurskompetenzen. So könnte auch erklärt werden, warum die Lerner häufig falsch, zu wenig oder zu viel Diskursmittel verwenden.

Zuletzt haben wir auch im zweiten Teil der ersten Forschungsfrage versucht, eine Antwort auf die Frage zu bekommen, ob die Anwendung von analytischen Bewertungsskalen zur objektiveren, zuverlässigeren Beurteilung diskursiver und grammatischer Kompetenzen führt. Hierzu haben wir den t-Test für unabhängige Stichproben und das quantitative Verfahren der deskriptiven Statistik verwendet. Die Bewertungsergebnisse deuten darauf hin, dass bei dem analytischen und holistischen Bewertungsverfahren der Diskurs- und Grammatikkompetenz keine signifikanten Unterschiede bestehen.

Bagarić Medve und Čelebić (2011) (siehe auch Kapitel 2.4.1.2) haben sich in ihrer Arbeit mit einer ähnlichen Frage beschäftigt, jedoch haben sie einen statistisch bedeutenden Unterschied feststellen können. In ihrer Untersuchung kamen sie zum Ergebnis, dass die Mittelwerte der Bewertungsergebnisse von Diskurs- und Grammatikkompetenz beim holistischen Bewertungsverfahren nah beieinander liegen. Beim erneuten, analytischen Bewertungsverfahren von Diskurs- und Grammatikkompetenz haben sie einen statistisch bedeutenden Unterschied festgestellt.

Die Ergebnisse der statistischen Analyse verweisen darauf, dass der Gebrauch der analytischen Bewertungsweise bei der Bewertung der Diskurs- und Grammatikkompetenz eine bessere Unterscheidung derselben Kompetenzen ermöglicht. Dadurch bekommt man zuverlässigere Daten über das Niveau der Diskurskompetenz in Bezug auf die Grammatikkompetenz und womöglich auch in Bezug auf andere Kompetenzen der Lerner. (Bagarić Medve und Čelebić 2011: 43)

Man darf jedoch nicht außer Acht lassen, dass unsere Untersuchung nur 58 Schüleraufsätze umfasste, also eine kleinere Anzahl von Stichproben. Für zuverlässigere Ergebnisse, die eine klarere Antwort auf die oben gestellte Frage bieten könnten, sollte die Untersuchung mehr als 150 Stichproben enthalten (vgl. Bagarić Medve und Čelebić 2011).

In der zweiten Forschungsfrage wurde untersucht, in welcher Beziehung die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim analytischen und die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim holistischen Bewertungsverfahren stehen. Man ging hier hypothetisch von einer teilweisen Korrelation aus, wobei man annahm, dass die Ergebnisse der Grammatikkompetenz (Morphologie und Syntax) beim analytischen und holistischen Bewertungsverfahren stärker miteinander korrelieren werden als die Ergebnisse bezüglich der Diskurskompetenz (Kohärenz und Kohäsion). Diese Hypothese wurde teilweise bestätigt, wobei betont werden muss, dass sowohl eine sehr starke Korrelation für die Aspekte der Grammatik- als auch für die Aspekte der Diskurskompetenz beobachtet werden kann.

Aus den Tabellen 8 und 9 geht hervor, dass zwischen den Aspekten der Diskurskompetenz (Kohärenz und Kohäsion) beim analytischen Bewertungsverfahren und dem Bewertungsverfahren

des Probestaatsabiturs (eher holistisch) eine statistisch signifikante Korrelation festgestellt werden konnte.

Die Bewertungsergebnisse zeigen, dass die Punkte zwischen dem Probestaatsabitur (ein eher holistisches Beurteilungsverfahren) am meisten mit den ersten drei Aspekten der Kohärenz (A *Die Reihenfolge der Ideen ist klar, logisch und zusammenhängend* ($r=,798^{**}$); B *Neue Ideen werden an der richtigen Stelle und in angemessener Weise eingeführt* ($r=,800^{**}$); C *Ideen, die genannt werden sind erarbeitet* ($r=,793^{**}$)) und mit dem Aspekt E der Kohäsion (*An den Stellen wo keine Kohäsionsmittel verwendet werden, ist der Übergang zwischen den Sätzen fließend*) ($r=,795^{**}$) bei dem erneuten analytischen Bewertungsverfahren korrelieren. Eine statistisch signifikante Korrelation wurde bei dem Aspekt D der Kohärenz (*Die Auszüge sind logisch und gerechtfertigt im Hinblick auf die Themen und die Relevanz* ($r=,786^{**}$)) und dem Aspekt A der Kohäsion (*Verwendung von grammatischen Kohäsionsmitteln ist vielfältig* ($r=,786^{**}$)) beim Probestaatsabitur und dem erneuten analytischen Bewertungsverfahren festgestellt.

Ein möglicher Grund für die oben genannte Korrelation wäre, dass sich die Prüfer des Probestaatsabiturs (eher holistisches Bewertungsverfahren) bei der Punktevergabe für die Diskurskompetenzen (Kohärenz und Kohäsion) am meisten an den oben genannten Aspekten der Kohärenz orientiert haben.

Darüber hinaus ist aus den Ergebnissen ersichtlich, dass die höchste Korrelation beim Probestaatsabitur und dem erneuten analytischen Bewertungsverfahren im Bereich der Kohärenz liegt. Gründe dafür könnten darin liegen, dass je weniger Fehler beim Aspekt der Kohärenz zu finden sind, desto lesbarer und verständlicher die Schüleraufsätze gewirkt haben. Der Text scheint logisch und gut erarbeitet, und ist deshalb fließender und einfacher zu verfolgen. Diese Gründe könnten die Prüfer, egal bei welchem Bewertungsverfahren, bewusst oder unbewusst bei der Punktevergabe beeinflusst haben.

Ein etwas niedrigerer Korrelationsgrad zwischen dem holistischen und analytischen Bewertungsverfahren wurde lediglich beim Aspekt der Kohäsion (*Interpunktion wird bei der Trennung von Ideen und Sätzen verwendet*) ($r=,510^{**}$) festgestellt. Nach Krüger (2012) versteht man unter der Interpunktion die Gliederung von Texten durch Satzzeichensetzung. Die Satzzeichensetzung folgt einem Regelapparat. Durch die festgelegten Regeln ist es einfacher diesen Aspekt zu bewerten, egal welches Bewertungsverfahren benutzt wird. Krüger (zit. nach Nerius 2012: 4) sagt, dass sich die Interpunktion von einem intonatorischen zu einem syntaktischen Prinzip entwickelt hat.

Dieser Punkt bringt uns zur Bewertung der Grammatikkompetenz (Syntax und Morphologie). Aus den Tabellen 10 und 11 wird deutlich, dass zwischen den Aspekten der

Grammatikkompetenz (Syntax und Morphologie) beim Probestaatsabitur (eher holistisches Bewertungsverfahren) und dem erneuten analytischen Bewertungsverfahren statistisch signifikante Korrelationen bestehen.

Der höchste Korrelationsgrad wurde zwischen den Aspekten E der Syntax (*Angemessenheit der Verwendung aller Satzarten im Text*) ($r=,859^{**}$) und A (*Reihenfolge der Wörter im Satz*) ($r=,830^{**}$) bei dem analytischen Bewertungsverfahren und dem holistischen Beurteilungsverfahren des Probestaatsabiturs festgestellt. Der niedrigste Korrelationsgrad ist bei den Aspekten der Morphologie A (*Verwendung der Artikel*) ($r=,744^{**}$) und D (*Präpositionen und Rektion*) ($r=,744^{**}$) festzustellen.

Betrachtet man die Bewertungsergebnisse genauer (siehe Tabelle 10), kommt man zur Schlussfolgerung, dass die Korrelation zwischen dem holistischen und analytischen Bewertungsverfahren im Bereich der Syntax höher ist als im Bereich der Morphologie.

Darüber hinaus ist bei der Grammatikkompetenz auch erkennbar, dass die morphologischen Aspekte weniger beachtet wurden als die syntaktischen Aspekte.

Wenn wir uns nun die Aspekte der Syntax aus der analytischen Bewertungsskala (siehe Anhang 2) ansehen und die Ergebnisse der Untersuchung, beispielsweise die hohe Korrelation zwischen den Aspekten A (*Reihenfolge der Wörter im Satz*) ($r=,830^{**}$) und E (*Angemessenheit der Verwendung aller Satzarten im Text*) ($r=,859^{**}$) noch einmal unter die Lupe nehmen, wird klar, dass diese Aspekte das Lesen von Texten verständlicher und einfacher machen. Die Texte wirken somit lesbarer und verständlicher, für den Leser scheinen die Aufsätze mit diesen Hilfsmitteln leichter verfolgbar zu sein. Diese Tatsache könnte den hohen Korrelationsgrad im Bereich der Syntax erklären und warum die Aspekte der Syntax mehr beachtet wurden.

Aus dieser Untersuchung geht hervor, dass beide Bewertungsverfahren (analytisches und holistisches) beim Bewerten von Schüleraufsätzen gleich gut sind, d.h. objektiv betrachtet, scheint es keinen Unterschied bezüglich der Bewertungsverfahren zu geben. Es scheint aber dennoch so zu sein, dass das analytische Bewertungsverfahren bei der Benotung bestimmter Teilkompetenzen didaktisch hilfreicher ist. Auf diese Weise können der Prüfer und der Schreiber, bzw. Lehrer und Schüler, nach der Bewertung der Aufsätze, Informationen darüber bekommen, an welchen (Teil-)Aspekten noch gearbeitet werden sollte.

Abschließend muss nochmals betont werden, dass eine Notwendigkeit besteht, weitere Untersuchungen durchzuführen, um zuverlässigere Ergebnisse zu bekommen.

4. Schlusswort

In dieser Forschungsarbeit wurde die Anwendung der analytischen Bewertungsskalen bei der Beurteilung der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenzen (Syntax und Morphologie) fortgeschrittener Deutschlerner untersucht. Der Schwerpunkt der Untersuchung lag darauf, herauszufinden, ob signifikante Unterschiede beim analytischen und einem eher holistischen Bewertungsverfahren bei der Bewertung der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenzen (Syntax und Morphologie) bestehen. Untersucht wurde auch in welcher Beziehung die Ergebnisse der einzelnen Aspekte der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenz (Syntax und Morphologie) beim analytischen und die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim holistischen Bewertungsverfahren stehen.

Den Ergebnissen zufolge bestehen keine signifikanten Unterschiede bei der Bewertung der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenzen (Morphologie und Syntax) beim Probestaatsabitur (eher holistisches Beurteilungsverfahren) und beim erneuten analytischen Bewertungsverfahren. Es ist wichtig zu betonen, dass diese Untersuchung 58 Schüleraufsätze umfasst. Für zuverlässigere Ergebnisse sollte die Untersuchung an einer größeren Anzahl von Aufsätzen wiederholt werden.

Aus der Untersuchung ging auch hervor, dass die Grammatikkompetenz (Syntax und Morphologie) beim Probestaatsabitur im Gegensatz zu der Diskurskompetenz (Kohärenz und Kohäsion) strenger benotet wurde. Beim Probestaatsabitur (ein eher holistisches Bewertungsverfahren) haben die Ergebnisse gezeigt, dass die syntaktischen Aspekte mehr beachtet wurden als die morphologischen Aspekte.

Beim näheren Betrachten der Ergebnisse im Bereich der Diskurskompetenz (Kohärenz und Kohäsion) beim Probestaatsabitur (eher holistisches Bewertungsverfahren) und dem erneuten analytischen Bewertungsverfahren wird deutlich, dass beim Aspekt der Kohärenz die wenigsten Punkte vergeben wurden. Die Bewertungsergebnisse dieser Untersuchung haben gezeigt, dass die Schüler die meisten Schwächen im Bereich der Kohärenz und die wenigsten im Bereich der Syntax zeigen.

Zum Schluss kann aus didaktischer Sicht aufgrund der Forschungsergebnisse geschlossen werden, dass die Schüler die meisten Schwächen im Bereich der Kohärenz und die wenigsten im Bereich der Syntax zeigen. Demzufolge kann angenommen werden, dass den Grammatikkompetenzen, insbesondere der Syntax, im Unterricht wahrscheinlich mehr Aufmerksamkeit geschenkt wird als den Diskurskompetenzen. Es scheint so, als richte sich der

Unterricht mehr auf die Grammatikkompetenzen und weniger auf die Diskurskompetenzen. Dies könnte erklären, warum die Lerner häufig falsch, zu wenig oder zu viel Diskursmittel verwenden. Die Ergebnisse weisen ebenfalls darauf hin, dass sich der Unterricht mehr auf die Aspekte der Kohäsion und Kohärenz richten sollte, ohne selbstverständlich die Aspekte der Grammatikkompetenz auszuschließen. Diese aus der Untersuchung gewonnenen Resultate könnten ein nützlicher Anhaltspunkt für die weitere Ausarbeitung der Kurrikula in Mittelschulen darstellen.

Literaturverzeichnis

Altmeyer, Axel; **Domisch**, Rainer (1998): „Benoten und Bewerten im Deutschunterricht“. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*. Heft 19: Benoten und Bewerten, 5-9.

Bachmann, Lyle F.; **Palmer**, Adrian S. (1996): *Language Testing and Practice*. Oxford: University Press.

Bagarić Medve, Vesna; **Čelebić**, Ivan (2011): Bewertung der Diskurskompetenz: Holistischer vs. Analytischer Ansatz. In: Pon, Leonard; Karablić, Vladimir; Cimer, Sanja (Hrsg.): *Applied Linguistics Today: Research and Perspectives – Angewandte Linguistik heute: Forschung und Perspektiven*. Frankfurt am/Main etc.: Peter Lang, 33-49.

Bailey, Kathleen M. (1998): *Learning about Language Assessment: Dilemmas, Decisions, and Directions*. Boston: Heinle & Heinle Publishers.

Bolton, Sibylle (1996): *Probleme der Leistungsmessung Lernfortgeschrittstetsts in der Grundstufe*. München: Goethe Institut.

Bräuer, Gerd (1998): *Schreibend lernen: Grundlage einer theoretischen und praktischen Schreibpädagogik*. Innsbruck: Studien-Verlag.

Canale, Mihael (1983): From communicative competence to communicative language pedagogy. In: Richards, J. C.; Schmidt, R.W. (Hrsg.): *Language and Communication*, London: Longman, 2-27.

Chiang, Steve (2003): The importance of cohesive conditions to perceptions of writing quality at the early stages of foreign language learning. *System* 31, 471-484.

Fix, Martin (2006): *Texte schreiben: Schreibprozesse im Deutschunterricht*. Paderborn: Schöningh.

Glaboniat, Maunela (1998): Auf der Suche nach Objektivität. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*. Heft 19: Benoten und Bewerten, 20-23.

Harden, Theo (2006): *Angewandte Linguistik und Fremdsprachendidaktik*. Tübingen: Narr Francke Attempto Verlag.

Harris, David P. (1969): *Testing English as a Second Language*. New York: McGraw-Hill Book Company

Johnson, Robert L., **Penny**, James A., **Gordon**, Belita (2009): *Assessing performance: Designing, scoring, and validating performance tasks*. New York: Guilford Press.

Kast, Bernd (1999): *Fertigkeit Schreiben*. Berlin: Langenscheidt.

Krumm, Hans-Jürgen (1993): Kulturgeprägte Einflüsse auf das Schreiben und ihre möglichen Konsequenzen für den Schreibunterricht Deutsch als Fremdsprache. In: Tütken, Gisela; Neuf-

Müinkel, Gabriele, (Hrsg): *Schreiben im DaF-Unterricht an Hochschulen und Studienkollegs I. Forschungsergebnisse- Didaktische Konzeption- Übungsformen*, Regensburg, 23 - 33.

Krumm, Hans-Jürgen (1989): Thema: Schreiben. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*. Heft 1, 5-8.

Krüger, Alexandra (2010): *Interpunktion – Ihre Funktion und Anwendung. Mit einer Analyse der Interpunktion in Schülertexten*. München: Grin Verlag.

Kruse, Otto; **Jakobs**, Eva-Maria; **Ruhmann**, Gabriela (1999): *Schlüsselkompetenz Schreiben: Konzepte, Methoden, Projekte für Schreibberatung und Schreibdidaktik an der Hochschule*. Krefeld: Luchterhand.

Lado, Robert (1964): *Language Testing: The Construction and Use of Foreign Language Tests: a Teacher's Book*. New York: McGraw-Hill.

Macaire, Dominique (1998): Die Note der Noten. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*. Heft 19: Benoten und Bewerten, 10-13.

Richardson, Peter (1998): Von der „Erstfassung“ zur „Ausgabe letzter Hand. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*. Heft 19: Benoten und Bewerten, 14-19.

Valette, Rebecca M. (1972): *Tests im Fremdsprachenunterricht*. Berlin: Cornelsen-Velhagen & Klasing.

Weigle, Sara Cushing (2002): *Assessing Writing*. Cambridge: Cambridge University Press.

Internetquellen

https://books.google.hr/books?id=IXmNswmcIoAC&pg=PA4&dq=interpunktion&hl=en&sa=X&ved=2ahUKEwiV2qyB8s_tAhXIqaQKHXnbDmoQ6AEwCXoECAkQA#v=onepage&q=interpunktion&f=false, abgerufen am 28.12.2020

Anhang 1

Analytische Bewertungsskala – abgeleitet von Chiang (2003)

5 = volle Übereinstimmung

4 = Übereinstimmung

3 = Unentschlossenheit

2 = ich stimme nicht überein

1 = ich stimme überhaupt nicht überein

0 = kann nicht bewertet werden

Kohärenz

A Die Reihenfolge der Ideen ist klar, logisch und kohärent (zusammenhängend), der rote Faden zieht sich durch den Text, was dazu führt, dass der Text verständlich ist.

B Neue Ideen (Informationen) werden an der richtigen Stelle und in angemessener Weise eingeführt.

C Die Ideen, die diskutiert/genannt werden, sind erarbeitet.

D Auszüge sind logisch und gerechtfertigt im Hinblick auf die Themen und die vorgeschriebenen Punkte (Relevanz).

E Der Übergang zwischen den Absätzen ist fließend.

Kohäsion

A Die Verwendung von grammatischen Kohäsionsmitteln ist vielfältig.

B Grammatischen Kohäsionsmittel werden angemessen verwendet

C Lexikalische Äquivalente, Paraphrasen und sinnvolle Wiederholung (lexikalische Kohäsionsmittel) werden verwendet.

D Die Interpunktion wird bei der Trennung von Ideen und Sätzen verwendet.

E An den Stellen, wo keine Kohäsionsmittel verwendet werden, ist der Übergang zwischen den Sätzen fließend.

Syntax

A Die Reihenfolge der Wörter in einem Satz ist richtig - Genauigkeit

B Kongruenz (Übereinstimmung von Subjekt - Prädikat, Artikel und Substantiven) ist richtig hergestellt.

- C Komplexere Verbformen und Phrasen (das Passiv, der Partizip, Modalverben, der Konjunktiv, die Komparation) sind vertreten.
- D Satzgefüge sind vertreten und vielfältig.
- E Angemessenheit der Verwendung aller Satzarten im Text

Morphologie

- A Artikel (bestimmter, unbestimmter und der Nullartikel) werden richtig verwendet.
- B Deklination und Komparation werden richtig verwendet.
- C Konjugation und Zeitformen werden richtig verwendet. Zeitformen werden richtig gebildet.
- D Präpositionen und die Rektion werden richtig verwendet.
- E Die Wortbildung ist richtig (Wörter waren richtig gebildet).

Anhang 2

Tabelle für die Punktevergabe bei dem analytischen Bewertungsverfahren

Punkte:

5 = volle Übereinstimmung

4 = Übereinstimmung

3 = Unentschlossenheit

2 = ich stimme nicht überein

1 = ich stimme überhaupt nicht überein

0 = kann nicht beurteilt werden

Kohärenz	Punkte
Die Reihenfolge der Ideen ist klar, logisch und kohärent (zusammenhängend), der rote Faden zieht sich durch den Text.	
Neue Ideen (Informationen) werden an der richtigen Stelle und in angemessener Weise eingeführt.	
Die Ideen, die diskutiert/genannt werden, sind erarbeitet.	
Auszüge sind logisch und gerechtfertigt im Hinblick auf die Themen und die vorgeschriebenen Punkte (Relevanz).	
Der Übergang zwischen den Absätzen ist fließend.	

Kohäsion	Punkte
Die Verwendung von grammatischen Kohäsionsmitteln ist vielfältig.	
Grammatische Kohäsionsmittel werden angemessen verwendet.	
Lexikalische Äquivalente, Paraphrasen und sinnvolle Wiederholung (Lexikalische Kohäsionsmittel) werden verwendet.	
Die Interpunktion wird bei der Trennung von Ideen und Sätze verwendet.	

An den Stellen, wo keine Kohäsionsmittel verwendet werden, ist der Übergang zwischen den Sätzen fließend.	
---	--

Syntax	Punkte
Genauigkeit der Wortstellung im Satz (Reihenfolge der Wörter im Satz).	
Kongruenz (Übereinstimmung von Subjekt - Prädikat, Artikel und Substantiven) ist richtig hergestellt.	
Komplexere Verbformen und Phrasen (das Passiv, das Partizip, Modalverben, der Konjunktiv, die Komparation) sind vertreten.	
Satzgefüge sind vertreten und vielfältig.	
Angemessenheit der Verwendung aller Satzarten im Text.	

Morphologie	Punkte
Artikel (bestimmter, unbestimmter und der Nullartikel) werden richtig verwendet.	
Deklination und Komparation werden richtig verwendet.	
Konjugation und Zeitformen werden richtig verwendet. Zeitformen werden richtig gebildet.	
Präpositionen und die Rektion werden richtig verwendet.	
Die Wortbildung ist richtig (Wörter werden richtig gebildet).	

Anhang 3

Aufgabenstellung und zwei Beispiele aus dem Korpus

Im Jahr 2009 lautete das Thema, zu dem ein Aufsatz verfasst werden sollte, „Jugendliche und Kleidung“. Die Punkte, auf die die Schüler im Hauptteil eingehen sollten, lauteten:

1. Welche Rolle spielt Kleidung im Leben der Jugendlichen?
2. Welche Vor- und Nachteile haben teure Klamotten im Alltag?
3. Wie wichtig sind Kleidung und Outfit für Sie und warum?

Zwei Beispiele aus dem Korpus:

Aufsatz 263 UNGENÜGEND

Der Jugendliche im Leben ist spielt nicht Kleidung muss können Fußball spielen. Sie muss kann schreiben und Leben Texte. Vorteile und Nachteile teuer Klamotten im Alltag ist nicht Computer und nicht spielen Basketball. Sind Kleidung und Outfit für sie ist muss nicht ein zusammen habt. Kleidung im leben Rolle Kleidung. Ich nicht kann der Jugendlichen spielt Fußball Es ist nicht spielt. Ja, warum du nicht spielt. Bruder und ich muss Computer, und spielen Basketball. Warum Jugendliche trinkt Pivo? Ich muss trinkt Cola. Muss du trink Pivo?

Meine Mutter nicht trinkt Cola, es trinkt Pivo

Du trinkst Cola. Vorteile teuer Klamotten ist muss trinken Cola. Haben du Thelephone.

Ich muss verstanden trinkt Pivo, ich muss trinken Cola. Was du trinkt. Kleidung nicht fahren dem Auto. Es ist schön, aber Jugendlichen muss Auto habe. Meine mutter nicht fahren dem Auto. Es habt vierundfunzig jahre alt. Mein Vater ist trinkt Pivo. Bruder und ich trinken und essen in schönen Restaurant. Mutter, Vater, Schwester, Bruder, und Oma mit essen und trinken nicht Restaurant es trinkt und isst zu Hause. Meine Schwester und Bruder ist Jugendlichen und ich. Nachteile ist ja nicht Computer und meine freundin Computer habt.

Outfit für sind nicht haben.

Im folgenden Aufsatz möchte ich Ihnen erklären, welche Rolle die Kleidung im Leben der Jugendlichen spielt. Dazu gehören auch die Vor- und Nachteile im Alltag und wie wichtig die Kleifung bzw. das Outfit ist.

Im Leben der Jugendlichen spielt die Kleidung eine sehr bedeutende Rolle. Heute ist es nicht mehr so wichtig wie es aussieht, das was man trägt, sondern welcher Modeartikel bzw. welche Marke es ist. Oft werden Schüler, d.h. Jugendliche ausgegrenzt oder verspottet, weil sie nicht in so guter finanziellen Lage bez. deren Eltern, sind, um sich Kleidung dieser Art zu kaufen. Solche Artickel haben einen sehr hohen Preis. Nach dem Outfit werden Jugendliche beurteilt, weil man sie nicht kennt, oder sie seien „nicht würdig“ mit den besserstehenden zu kommunizieren. Die obengenannten Nachteile sind auch meistens der Hauptgrund des Gemeinschaftsausschlusses. Im Alltag haben teure Klamotten den Vorteil, dass man zwar immer schick angezogen ist und auch toll aussieht, aber den Nachteil, dass solche Kleidung auch sehr empfindlich ist und das sie nicht für jede Arbeit geeignet ist. Für mich ist mein Outifit sehr wichtig. Meistens Sorge ich dafür, dass es eine Marke ist, oder dass es einfach gut aussieht. Da meine Eltern nicht in der finanziellen Möglichkeit sind mir Markensachen zu kaufen, gehe ich in den Ferien arbeiten und kaufe sie mir dann selbst.

Man sollte die Menschen nicht nach ihren Aussehen beurteilen, sondern nach ihren Eigenschaften und Fähigkeiten, denn nur so schafft man sich Freunde fürs Leben.

Zusammenfassung

Diese Diplomarbeit beschäftigt sich mit der Frage, ob signifikante Unterschiede beim analytischen und einem eher holistischen Bewertungsverfahren der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenzen (Syntax und Morphologie) bestehen. Untersucht wird auch in welcher Beziehung die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim analytischen und die Ergebnisse der einzelnen hier untersuchten Aspekte der Diskurs- und Grammatikkompetenz beim holistischen Bewertungsverfahren stehen. 58 Aufsätze fortgeschrittener Deutschlerner aus kroatischen Mittelschulen, die in der schriftlichen Prüfung beim Probestaatsabitur und beim ersten Staatsabitur geschrieben worden waren, wurden analytisch und holistisch in Hinsicht auf die Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenzen (Syntax und Morphologie) bewertet.

Um Antworten auf unsere Forschungsfragen zu bekommen, untersuchten wir verschiedene Aspekte der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenz (Syntax und Morphologie), Unterschiede zwischen den Mittelwerten der Diskurs- und Grammatikkompetenz beim analytischen und holistischen Bewertungsverfahren, sowie die Korrelationen der Kompetenzen bei den eben genannten Bewertungsverfahren. Bei der Untersuchung wurden quantitative Verfahren der deskriptiven Statistik, des t-Testes und der Korrelationsanalyse verwendet.

Aus den Ergebnissen ging hervor, dass keine signifikanten Unterschiede bei der Bewertung der Diskurs- (Kohärenz und Kohäsion) und Grammatikkompetenzen (Morphologie und Syntax) beim Probestaatsabitur (eher holistisches Beurteilungsverfahren) und beim erneuten analytischen Bewertungsverfahren bestehen. Zudem geht aus den Ergebnissen hervor, dass den Grammatikkompetenzen, insbesondere der Syntax, mehr Aufmerksamkeit geschenkt wurde, als den Diskurskompetenzen (Kohärenz und Kohäsion).

Schlüsselwörter: analytisches Bewertungsverfahren, holistisches Bewertungsverfahren, Grammatikkompetenz, Morphologie, Syntax, Diskurskompetenz, Kohärenz, Kohäsion.

Sažetak

U ovome se radu pokušava utvrditi postoje li značajne razlike pri analitičkom i holističkom načinu vrednovanja diskursne (koherentnost i kohezija) i gramatičke kompetencije (sintaksa i morfologija). Istražuje se i odnos između rezultata pojedinih aspekata diskursne i gramatičke kompetencije, spomenutih u ovome radu, primjenom analitičkog načina ocjenjivanja kao i rezultati pojedinih aspekata diskursne i gramatičke kompetencije pri holističkom načinu ocjenjivanja.

58 eseja naprednih učenika njemačkog jezika iz hrvatskih srednjih škola, koji su pisani na pismenom ispitu na probnoj državnoj maturi i prvoj državnoj maturi, ocijenjeno je analitičkim i holističkim načinom s obzirom na diskursne (koherencija i kohezija) i gramatičke kompetencije (sintaksa i morfologija).

Kako bismo dobili odgovore na naša istraživačka pitanja, ispitali smo kvantitativno različite aspekte diskursne (koherencija i kohezija) i gramatičke kompetencije (sintaksa i morfologija), razlike između srednjih vrijednosti diskursnih i gramatičkih kompetencija pri analitičkom i holističkom načinu ocjenjivanja, kao i korelacije kompetencija u gore spomenutim načinima ocjenjivanja. U istraživanju su korištene kvantitativne metode opisne statistike, t-test i korelacijska analiza.

Rezultati su pokazali kako nije bilo značajnih razlika u ocjenjivanju diskursnih (koherentnost i kohezija) i gramatičkih (morfologija i sintaksa) kompetencija na probnoj državnoj maturi (više holistički način ocjenjivanja) i pri ponovljenom analitičkom načinu ocjenjivanja. Rezultati također pokazuju kako se gramatičkim kompetencijama, posebno u aspektu sintakse, pridaje više pažnje nego diskursnim kompetencijama (koherencija i kohezija).

Ključne riječi: analitički način ocjenjivanja, holistički način ocjenjivanja, gramatička kompetencija, morfologija, sintaksa, diskursna kompetencija, koherencija, kohezija