

# EIN *CLOZE* TEST ZUR SPRACHSTANDSERHEBUNG PORTUGIESISCHER L2 LERNER DES DEUTSCHEN

Clara Setas<sup>1</sup>, Inês Sousa<sup>1</sup>, João Antunes<sup>1</sup>, Cristina Flores<sup>2</sup>

Escola de Letras, Artes e Ciências Humanas, Universidade do Minho

## Abstract

This study aims to test the feasibility of a research tool, a *Cloze* Test, developed to assess the proficiency of German as a second language (L2) in native speakers of European Portuguese (EP). The *Cloze* Test is a fill-in-the-gap test used as a tool to measure the participants' proficiency of linguistic competence, taking into account their lexical, morphosyntactic and discursive knowledge of varying complexity levels. The research conducted presents a *Cloze* Test in German that was tested on a group of 37 adult native Portuguese speakers who have acquired German as L2 (mainly in a classroom environment). The results of this study indicate that this type of instrument allows to assess language proficiency as it presents a clear complexity hierarchy, with the potential to identify the different proficiency levels of the participants. Moreover, statistical correlations between test scores, self-assessments and extralinguistic input factors show that test scores correlate with both self-assessment and the type of L2 contact. The results are further discussed concerning the role of various input variables. The article ends with a discussion of

---

<sup>1</sup> Email: [csetas@gmail.com](mailto:csetas@gmail.com); [ines.msousaa@gmail.com](mailto:ines.msousaa@gmail.com); [joao.v.antunes11@gmail.com](mailto:joao.v.antunes11@gmail.com)

<sup>2</sup> ORCID: <https://orcid.org/0000-0001-5629-9556>; Email: [cflores@elach.uminho.pt](mailto:cflores@elach.uminho.pt)

some limitations of this study, while also highlighting future research to be conducted in this field.

**Keywords:** German as a second language, L2 acquisition, *Cloze* Test, language proficiency assessment, self-assessment, input factors

### **Zusammenfassung**

Der *Cloze* Test ist ein Lückentest zur Sprachstandserhebung, der lexikalisches, morphosyntaktisches und diskursives Wissen unterschiedlicher Komplexität berücksichtigt. Die vorliegende Studie stellt einen *Cloze* Test in deutscher Sprache vor, der an einer Gruppe von 37 erwachsenen portugiesischen Muttersprachlern getestet wurde, die Deutsch als Zweitsprache erworben haben. Die Ergebnisse zeigen, dass der *Cloze* Test eine klare Komplexitätshierarchie aufweist, die das Potential hat, die unterschiedlichen Sprachniveaus der Teilnehmer zu erfassen. Zudem zeigen statistische Korrelationen zwischen den Ergebnissen des Tests, den Selbsteinschätzungen und den extralinguistischen Inputfaktoren, dass die Testergebnisse sowohl mit der Selbsteinschätzung korrelieren als auch mit der Art des L2-Kontakts. Die Ergebnisse werden im Hinblick auf die Rolle der Inputvariablen im L2-Erwerb erörtert. Der Artikel weist abschließend auf mögliche Limitationen des Tests und adressiert zukünftige Forschungsfragen.

**Schlüsselwörter:** Deutsch als Zweitsprache, Zweitspracherwerb, *Cloze* Test, Sprachstandserhebung, Selbstbewertung von Sprachkompetenzen, Inputfaktoren

## 1. Einleitung

Ein wichtiger Teilbereich des Forschungsfeldes zum Zweitspracherwerb beschäftigt sich mit den Faktoren, die Einfluss auf die Leistung und die Kenntnisse des Lernenden haben. Wir konzentrieren uns in diesem Artikel auf eine *späte L2*. Diese wird im Jugend- oder Erwachsenenalter nach der Muttersprache erlernt, üblicherweise in einer formalen Klassenzimmerumgebung. Der Leistungsstand von L2-Lernern ist in der Regel sehr variabel, oft auch wenn man Lerner vergleicht, die eine ähnliche Dauer an formalem Unterricht aufweisen. Diese Diversität ist auf eine große Anzahl von Faktoren zurückzuführen, wie zum Beispiel das Alter zu Erwerbsbeginn, der Umfang des Inputs und Outputs, die Lehr- und Lernstrategien von Lehrenden und Lernern und individuelle Variablen wie Sprachbegabung, Motivation oder sogar das soziale Umfeld (Duff, 2017).

Um sowohl den Prozess des späten Zweitspracherwerbs als auch die Rolle der einzelnen Faktoren, die diesen Prozess beeinflussen, genau beschreiben zu können, müssen die Instrumente zur Erhebung des Lernniveaus die Fähigkeiten der Lernenden zuverlässig und sachlich unabhängig beurteilen können. Es besteht jedoch ein Mangel an Konsens unter den Forschern in Bezug auf die Wahl der zuverlässigsten Methoden und Instrumente der Kompetenzbewertung. Die häufigsten Methoden bzw. Instrumente für diese Beurteilung sind standardisierte Tests oder bestehende Einstufungsprüfungen, sogenannte *Cloze* Tests oder *C*-Tests, und schließlich auch mündliche Interviews oder Akzentbewertungen.

Im Folgenden werden wir uns mit einem dieser Instrumente beschäftigen, dem *Cloze* Test. Ziel dieses Artikels ist es, ein Instrument vorzustellen, das für portugiesische L2-Lerner des Deutschen entwickelt und an einer Gruppe von 37 Teilnehmern getestet wurde. Die Ergebnisse zeigen, dass der entwickelte Test Unterschiede im Sprachstand der Teilnehmer gut erfasst und dementsprechend nach seiner Validierung an einer größeren Gruppe von Teilnehmern, im

Deutschunterricht einfach und praktisch von Sprachlehrern zur Sprachstandserhebung ihrer Schüler eingesetzt werden könnte. Zudem wird gezeigt, dass die Ergebnisse des angewandten *Cloze* Tests positiv mit den Eigenbewertungen der Sprecher und auch mit der Art des Sprachinputs außerhalb des Klassenzimmers korrelieren.

## 2. Der *Cloze Test* als Instrument der Sprachstandserhebung

Der *Cloze* Test ist ein Lückentest, der sprachliche Strukturen in verschiedener Hinsicht auswertet: lexikalisch, morphosyntaktisch und diskursiv. Die Lücken können aus ganzen Wörtern oder einem Wortteil bestehen. Normalerweise sollten der erste und der letzte Satz unverändert bleiben. Wie von Chung und Ahn (2019) beschrieben, wird der „Cloze Test als integrativer Test angesehen, da die Lernenden auf ihr gesamtes Sprachwissen zurückgreifen und die Passage rekonstruieren müssen, indem sie die Beziehungen zwischen Sätzen oder innerhalb eines Satzes beobachten und die Lücken mit geeigneten Antworten ausfüllen.“ (S.627)<sup>3</sup>. Die Effizienz dieses Tests liegt in der Fähigkeit, Sprecher in verschiedene Gruppen je nach ihren Kompetenzen zu unterteilen. Das ermöglicht auch, externe Einflussfaktoren zu identifizieren. Darüber hinaus weist der *Cloze* Test eine große Wirksamkeit, Zuverlässigkeit und Praktikabilität auf (Tremblay, 2011).

Der C-Test unterscheidet sich vom *Cloze* Test in mehreren strukturellen Aspekten. So basiert der C-Test nicht auf der Auslassung bestimmter grammatischer Strukturen, sondern zielt darauf, dass der Lernende eine feste Anzahl von Lücken unabhängig von deren Komplexität (z.B. jedes siebte Wort) ausfüllt. Die Vor- und Nachteile beider Methoden sind Gegenstand

---

<sup>3</sup> Aus dem Englischen übersetzt: “The cloze test is considered to be an integrative test because examinees must draw on their overall language knowledge and reconstruct the passage by observing relationships between sentences or within a sentence and filling in the blanks with appropriate answers.”

großer Debatten in der Literatur (Brown, 1984, 1988), auf die wir hier jedoch nicht weiter eingehen werden.

Der *Cloze* Test erschien erstmals in Taylors (1953) Forschungsarbeit über die Fähigkeit von Lückentests, die Lesefähigkeiten von Schulkindern in den USA zu messen (Taylor, 1953). Seit den siebziger Jahren werden *Cloze* Tests in vielen Studien als Instrument zur Beurteilung der L2-Leistung benutzt (Hinofotis & Snow, 1980; Jonz, 1976; Oller & Inal, 1971). Einige Forscher sind der Ansicht, dass *Cloze* Tests ausschließlich für die Bewertung von Sprechern mit geringer lexikalischer und grammatischer Kompetenz hilfreich sind (Alderson, 1979; Markham, 1985), während andere argumentieren, dass der Test auch ein hohes Niveau an diskursiver Kompetenz bewerten kann (Bachman, 1982; Brown, 1983; Fotos, 1991). Die Ergebnisse von Lückentests korrelieren in der Regel mit den Ergebnissen von standardisierten Kompetenztests (Bachman, 1985; Fotos, 1991; Hanania & Shikhani, 1986).

Bei der Erstellung eines *Cloze* Tests müssen verschiedene Elemente berücksichtigt werden: die Auswahl und Länge des Textes, die grammatischen Zielstrukturen (Inhalts- und Funktionswörter), die Antworttypen (Mehrfachauswahl, ganz oder teilweise fehlendes Wort), die Bewertungskriterien und die Anzahl der Lücken im Text (empfohlen sind zwischen 30 und 45 Lücken bei einer angemessenen Verteilung der linguistischen Strukturen).

Der Test sollte weder zu einfach noch zu schwierig sein, sondern über die gesamten Textsegmente hinweg einen niedrigen bis hohen Schwierigkeitsgrad aufweisen. Nur auf diese Weise können Unterschiede der Kompetenzniveaus innerhalb der getesteten Lernergruppen nachgewiesen werden. Infolge dessen zeigen Tremblay und Garrison (2010, S. 84) in den Ergebnissen ihrer Studie, dass der Lückentest die Fähigkeit hat, L2-Lernern mit einem weiten Spektrum von Leistungsniveaus voneinander zu unterscheiden und eine breite Verteilung von

Punktzahlen sowohl hinsichtlich der Testteilnehmer als auch in Hinblick auf die getesteten Strukturen zu gewährleisten.

Die Nachteile der Verwendung von *Cloze* Tests für die L2-Kompetenzforschung sind vor allem mit der Schwierigkeit verbunden, geeignete Tests zur Beurteilung von kindlichen Lernern und von erwachsenen Lernenden mit sehr geringen Lese- und Schreibfähigkeiten zu entwickeln, denn das Ausfüllen der Lücken aktiviert nicht nur grammatisches und lexikalisches Sprachwissen, sondern auch Lese- und Verständnisfähigkeiten sowie Diskurs- und Textkohäsionskompetenzen. Ein weiterer Nachteil des *Cloze* Tests besteht darin, dass er nicht in der Lage ist, mündliche Fähigkeiten zu testen. In Bezug auf letzteres schlagen Tremblay und Garrison (2010) vor, neben dem *Cloze* Test eine mündliche Aufgabe (z.B. eine Hörverständnisaufgabe) zu verwenden.

Die Methode des *Cloze* Tests ist einfach zu handhaben und auszuwerten. Der größte Vorteil liegt darin, dass der Forscher die Struktur und Gestaltung des *Cloze* Tests den spezifischen Zwecken und den Zielen der Untersuchung anpassen kann.

### **3. Die Studie**

Die vorliegende Studie wurde an der Universidade do Minho im Rahmen des Masters *Deutsch-Portugiesische Studien / Estudos Luso-alemães* entwickelt. Im Folgenden beschreiben wir die Struktur des Tests und des Fragebogens, die Datenerhebungsmethode, die Teilnehmer und die Forschungsfragen.

#### **3.1. Der entwickelte *Cloze* Test**

Die Entwicklung des *Cloze* Tests erfolgte in verschiedenen Phasen. Zunächst wurde über die Art des auszuwählenden Textes diskutiert. Unter Berücksichtigung der Zielgruppe unserer

Studie – erwachsene Hochschulstudenten – entschieden wir uns, eine Kurzgeschichte zu schreiben, die ein Familienereignis schildert. Den Text selber zu schreiben hatte den Vorteil, dass wir gezielt über das verwendete Vokabular und die Sprachstrukturen entscheiden konnten. Die Wahl des Themas der Kurzgeschichte wurde von unseren eigenen Erfahrungen als portugiesische Studierende, die die deutsche Sprache als Zweitsprache gelernt haben, bestimmt.

Der Text erzählt die Geschichte eines Paares, das heiraten wird. Die Handlung spielt sich am Tag der Hochzeit ab, wobei auch der Ort und andere anwesende Personen wie die Eltern des Bräutigams oder auch die Gäste, eine Rolle spielen. Im Laufe der Handlung kommt es jedoch zu einem unerwarteten Ereignis.

Nach der Gliederung der Geschichte wurden die unterschiedlichen, auszulassenden Strukturen mit Hinblick auf deren Schwierigkeitsgrad ausgesucht. Insgesamt hat der *Cloze* Test 39 Lücken<sup>4</sup>, die Verben, Substantive, Adjektive, Adverbien, Präposition und Konjunktionen erfassen (siehe Tabelle 1).

Im Text können diese Lücken auf zwei Arten erscheinen: (1) durch das vollständige Weglassen von Wörtern (z.B. *und* [\_ \_ \_]) oder (2) durch das Weglassen von Wortteilen (z.B. *Zeremonie* [Zerem \_ \_ \_]). Die Lücken testen lexikalisches Wissen, wie z.B. in der Struktur 16 *Nachricht*, oder/und grammatisches Wissen wie Kasusdeklination (z.B. Struktur 3), trennbare Partikel (z.B. Struktur 28) oder Verbflexion (z.B. Struktur 23).

Die Durchführung dieser Studie erfolgte in zwei Formen. Da die aktuellen Umstände der Covid-19-Pandemie es nicht erlaubten, die Testauswertung gleichzeitig in Anwesenheit aller Teilnehmer zu erhalten, hat ein Teil der Teilnehmer den *Cloze* Test im Deutschunterricht (in einer Papierversion) gemacht und ein anderer Teil hat *online* teilgenommen. Die verwendete

---

<sup>4</sup> Es waren ursprünglich 40 Lücken, eine musste leider wegen eines technischen Fehlers ignoriert werden.

Online-Plattform ist *cognition.run*, eine Plattform zur Erstellung von Instrumenten für linguistische Forschung, wie z.B. *Cloze* Tests unter Verwendung der Programmiersprache *JavaScript*. Auf diese Weise konnten wir eine größere Anzahl von Teilnehmern auf eine sicherere und effektivere Weise erfassen.

**Tabelle 1**

*Getestete Strukturen*

	<b>Lücke</b>	<b>Struktur</b>		<b>Lücke</b>	<b>Struktur</b>
<b>1</b>	[auf]	Präposition	<b>21</b>	i[hn]	Personalpronomen, Akkusativ
<b>2</b>	gewart[et]	Verb, Partizip II	<b>22</b>	[war]	Verb, Präteritum, Indikativ
<b>3</b>	eine[m]	Artikel, Deklination	<b>23</b>	klingel[te]	Verb, Präteritum, Indikativ
<b>4</b>	sonnig[en]	Adjektiv, Deklination	<b>24</b>	i[m]	Präposition + Dativ
<b>5</b>	plötz[lich]	Adverb	<b>25</b>	ha[tten]	Verb, Plusquamperfekt, Indikativ
<b>6</b>	w[ürde]	Verb (Konjunktiv II)	<b>26</b>	N[ach]	Präposition
<b>7</b>	schwie[rigen]	Adjektiv, Deklination	<b>27</b>	d[ie]	bestimmter Artikel, Akkusativ
<b>8</b>	Luka[s]	Genitivmarkierung	<b>28</b>	[ab]	trennbare Partikel
<b>9</b>	Zeit[punkt]	Substantiv	<b>29</b>	verzweifelt[en]	Adjektiv, Deklination
<b>10</b>	w[ie]	Adverb, Relativsatz	<b>30</b>	hätt[en]	Verb, Konjunktiv II



11	[sie]	Personalpronomen, Akkusativ	31	n[icht]	Negationsadverb
12	Des[halb]	Konjunktion	32	[zu]	Infinitivpartikel
13	d[ass]	Konjunktion	33	d[es]	bestimmter Artikel, Genitiv
14	A[ls]	Konjunktion	34	s[ich]	Reflexivpronomen
15	Zerem[onie]	Substantiv	35	Oh[ne]	Präposition
16	Nachr[icht]	Substantiv	36	z[um]	Präposition + Dativ
17	aufger[egt]	Adjektiv	37	[um]	Konjunktion, Finalsatz
18	z[ur]	Präposition + Dativ	38	pass[iert]	Verb, Perfekt, Indikativ
19	i[hm]	Personalpronomen (Dativ)	39	gebl[ieben]	Verb, Partizip II
20	helf[en]	Verb (Infinitiv + ‚zu‘)			

### 3.2. Fragebogen

Zusätzlich zu dem *Cloze* Test wurden die Teilnehmer gebeten, auch einen kurzen soziolinguistischen Fragebogen auszufüllen, der uns Informationen über ihren Erwerb der deutschen Sprache liefert.

Der Fragebogen besteht aus zwei Teilen: Im ersten wurden persönliche Daten der Teilnehmer wie Alter, Geschlecht, Geburtsort, Stadt und Studiengebiet abgefragt.<sup>5</sup> Der zweite Teil des Fragebogens behandelt Fragen, die sich auf den Kontakt der Teilnehmer mit der deutschen Sprache und deren Erwerb beziehen, wie z.B. wie lange diese Teilnehmer schon Deutsch lernen, wie sie ihre Sprachkenntnisse einschätzen, unter welchen Umständen sie die Sprache verwenden, wie viel Zeit sie dafür aufwenden und ob sie außerhalb des Klassenzimmers

<sup>5</sup> Alle Teilnehmer blieben anonym, da nicht nach ihrem Namen gefragt wurde.

Kontakt mit dem Deutschen haben oder hatten (z.B. durch Erasmus oder nahe Verwandte, die in einem deutschsprachigen Land leben). Dieser zweite Teil des Fragebogens ist etwas umfangreicher und detaillierter und ermöglicht so die Beschreibung des Sprach-Inputs und -Outputs der Teilnehmer.

### **3.3. Teilnehmer**

Die Teilnahme an der Studie setzte voraus, dass die Informanten zwischen 18 und 30 Jahre alt sind, Europäisches Portugiesisch als Muttersprache sprechen und mindestens seit 2 Jahren Kontakt mit der deutschen Sprache im Klassenraum haben. Es nahmen 40 Informanten an der Studie teil, von denen drei ausgeschlossen wurden, weil sie den Teilnahmekriterien nicht entsprachen.

Die Teilnehmer sind zwischen 19 und 27 Jahre alt (Durchschnitt: 21,7). Die meisten Teilnehmer sind Bachelor-Studierende des Niveaus B1 der Universidade do Minho (23 Teilnehmer). Es nahmen zusätzlich 14 Masterstudierende aus verschiedenen geisteswissenschaftlichen Instituten verschiedener Universitäten teil, die längeren Kontakt mit der deutschen Sprache hatten. Insgesamt hatten die Teilnehmer zwischen 2 und 16 Jahre Kontakt mit dem Deutschen (Median: 3,00; Durchschnitt: 4,2 Jahre; StdAbw.: 2,7).

Es konnte festgestellt werden, dass die meisten Teilnehmer ihren ersten Kontakt mit der deutschen Sprache im Universitätsstudium hatten und hauptsächlich in diesem Kontext die deutsche Sprache erlernt haben. Der wöchentliche Kontakt mit dem Deutschen ist relativ gering (1,8 Stunden pro Woche). Dennoch gibt es eine beträchtliche Anzahl von Teilnehmern, die an einem Erasmus-Austauschprogramm in Deutschland teilgenommen haben oder regelmäßigen Kontakt zu deutschsprachigen Freunden, Verwandten bzw. Bekannten pflegen (13). 24

Teilnehmer haben keinen Kontakt mit der deutschen Sprache außerhalb des Klassenzimmers (1 fehlende Antwort).

In der Selbsteinschätzung (in einer Skala von 1- sehr gut bis 3- schlecht) hat die Mehrheit der Teilnehmer ihre Sprachkompetenz als mittelmäßig eingestuft (sehr gut: 5; mittelmäßig: 26; schlecht: 5; ohne Antwort: 1).

### **3.4. Forschungsfragen**

Das Ziel der vorliegenden Studie ist es, den entwickelten *Cloze* Test an einer Gruppe von portugiesischen L2-Sprechern des Deutschen zu testen und damit die folgenden Forschungsfragen zu beantworten:

1. Ermöglicht dieses Erhebungsinstrument, unterschiedliche Sprachniveaus zu erfassen?
2. Werden bestimmte Lücken von wenig Teilnehmern und andere Lücken von der Mehrheit der Teilnehmer ausgefüllt, und lässt sich anhand der Ergebnisse der Teilnehmer eine Komplexitätshierarchie definieren?
3. Korrelieren die Ergebnisse des *Cloze* Tests mit der Selbstbewertung der Teilnehmer?
4. Korrelieren die Ergebnisse des *Cloze* Tests mit extralinguistischen Inputfaktoren, wie z.B. die Anzahl der Unterrichtsjahre in der L2, der wöchentliche Kontakt mit der L2 und die Art des Kontakts außerhalb des Klassenzimmers?

## **4. Ergebnisse**

### **4.1. Allgemeine Ergebnisse und Komplexitätshierarchie**

Beginnen wir zunächst mit den allgemeinen Ergebnissen der Teilnehmer am *Cloze* Test. Die Tabelle 2 zeigt die Mindest- und Maximalwerte, den Mittelwert und den Median der richtigen

Antworten sowie den der Auslassungen. Für die korrekten Antworten wurden die grundsätzlich erwarteten und die nicht erwarteten, aber möglichen Antworten addiert.

## Tabelle 2

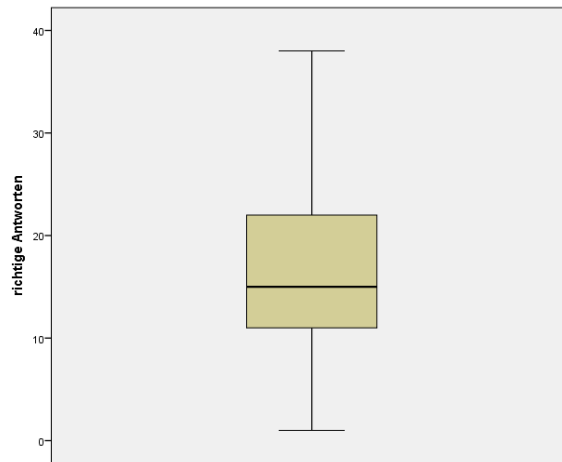
### *Globale Ergebnisse*

	<b>richtige Antworten</b>	<b>Auslassungen</b>
Mittelwert	16,26	10,55
Median	15,00	9,00
Standardabweichung (StdAbw)	9,09	9,28
Minimum - Maximum	1 – 38	0 – 31

Die Ergebnisse zeigen in einer Breitspanne von 1 bis zu 38 (von 39) richtigen Antworten eine große Variation zwischen den Sprechern. Der Mittelwert liegt bei 16.26 (StdAbw: 9,09) und der Median bei 15,00, d.h. bei rund der Hälfte der Lücken (der Text hatte 39 Lücken) – siehe auch die Grafik 1. zur Verteilung der richtigen Antworten. Auch die Auslassungen sind aussagekräftig: Der Mittelwert beträgt 10,55 (StdAbw: 9,28), der Median liegt bei 9,00. Die Anzahl der Auslassungen geht von 0 bis 31, d.h. einige Teilnehmer haben alle Lücken ausgefüllt (wenn auch nicht alle korrekt), was bei anderen in einer großen Anzahl an Lücken nicht der Fall war.

## Grafik 1

*Verteilung der richtigen Antworten in der Stichprobe*

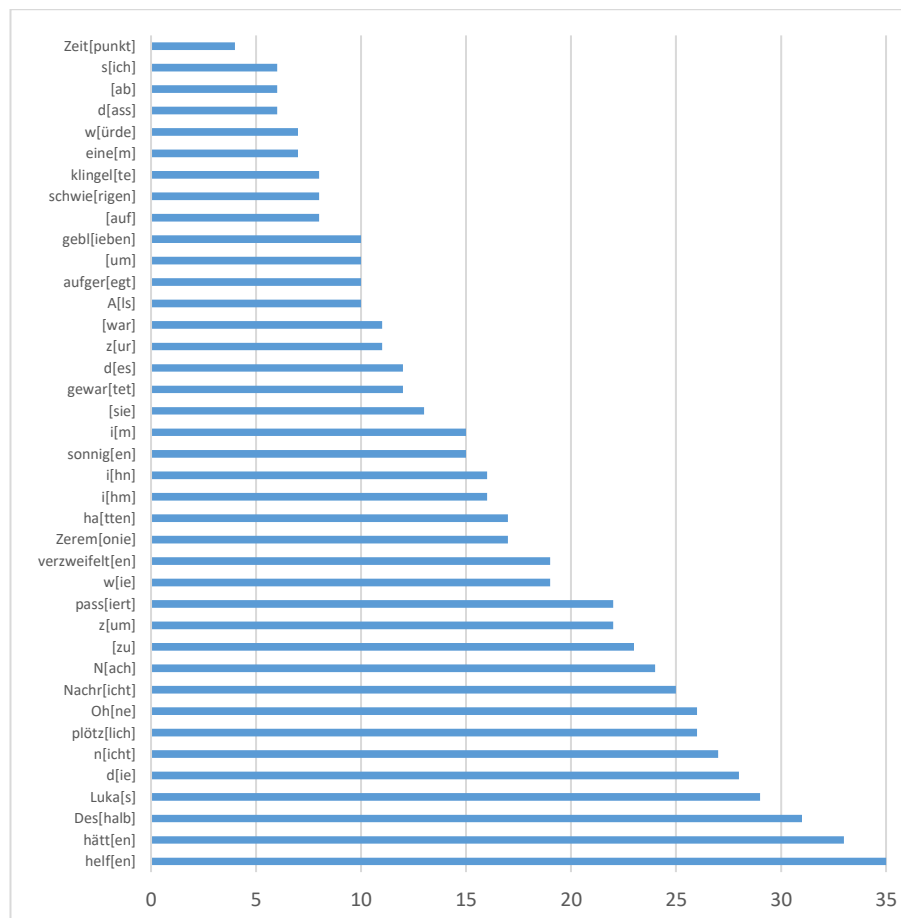


Die Ergebnisse zeigen klar, dass der *Cloze* Test eine weite Spanne an Sprachkompetenzen erfasst, die von minimalen Deutschkenntnissen bis zu sehr hohen Sprachleistungen im Deutschen geht. Inwiefern diese Ergebnisse mit extralinguistischen Faktoren wie der Dauer des Spracherwerbs und der Quantität und Art des Kontakts mit dem Deutschen korreliert werden wir im Abschnitt 4.2. erörtern.

Zunächst werfen wir einen Blick auf die Sprachstrukturen, die durch die Lücken erfragt wurden. Wie im Abschnitt 3.1 dargestellt, erfassten die Auslassungen sowohl lexikalisches Wissen als auch diverse grammatische Aspekte wie die Verbal- und Nominalflexion sowie die Verwendung von Präpositionen und Konjunktionen. Die Grafik 2 zeigt die Schwierigkeitshierarchie der Lücken.

## Grafik 2

### Schwierigkeitshierarchie des Cloze Tests



Die Ergebnisse zeigen eine klare Hierarchie: Sie reicht von Lücken, die aufgrund deren hohen Schwierigkeitsgrads von nur sehr wenig Sprechern ausgefüllt wurden, bis zu Lücken, die aufgrund ihrer Leichtigkeit von relativ vielen Teilnehmern ausgefüllt wurden. Zu den leichten Strukturen gehören die Infinitivflexion  $\{-en\}$  der Verbform *helfen* und das Morphem der 3. Person Plural  $\{-en\}$  in *hätten*, die je 35 und 33 Teilnehmer richtig ausgefüllt haben. An dritter Stelle stehen mit 31 richtigen Antworten die Adverbien *deshalb/deswegen*. Es folgen mit 29 bis 26 richtigen Antworten der Genitiv mit  $\{-s\}$  in Eigennamen, der definite feminine Artikel im Akkusativ *die*, die Negationspartikel *nicht*, das Adverb *plötzlich* und die Präposition *ohne*. Zu den besonders

schwierigen Strukturen zählen das Kompositum *Zeitpunkt/Zeitraum*, das nur 4 Teilnehmer kannten. Es folgen mit nur 6 richtigen Antworten das Reflexivpronomen *sich*, das sich auf das Verb *aufhalten* bezieht; die trennbare Partikel *ab-* des Verbes *absagen* sowie die Konjunktion *dass* in der Nebensatzverbindung mit *ohne dass*. Es lässt sich somit definieren, welche Strukturen in niedrigen Sprachstandsniveaus beherrscht werden und welche nur besonders kompetente Sprecher kennen.

#### 4.2. Zum Einfluss extralinguistischer Faktoren

In dieser Sektion werden die Ergebnisse des *Cloze* Tests mit den erfragten extralinguistischen Variablen statistisch korreliert. Die statistischen Analysen wurden im Programm SPSS Version 23 durchgeführt. Da die Daten keine Normalverteilung aufweisen, werden nicht-parametrische Assoziationstests durchgeführt.

Im Fragebogen wurde eine Selbstbewertung der Sprachkompetenz der Teilnehmer in einer Skala von 1 (gut) bis 3 (schlecht) erfragt. Es sollte gezielt getestet werden, ob die Ergebnisse im *Cloze* Test mit der Selbsteinschätzung der Sprecher übereinstimmt. In der Literatur wird von vielen Autoren die Einstellung vertreten, dass die Selbsteinschätzung von L2-Sprechern meistens mit den Ergebnissen von Sprachstandserhebungen übereinstimmt und somit eine sehr gültige Methode ist, die Sprecher nach Sprachniveaus zu klassifizieren (Dolovic, 2018; LeBlanc & Painchaud, 1985; Ross, 1998). Um eine potentielle Korrelation zwischen den Ergebnissen des Tests und der Selbsteinschätzung der Teilnehmer zu errechnen, wurde ein Spearman-Assoziationstest durchgeführt, wobei als Referenzwert für die Sprachkompetenz die Anzahl der richtigen Antworten als Variable eingegeben wurde. Das Ergebnis zeigt eine signifikant negative Korrelation zwischen beiden Variablen ( $r_s = -,466$ ,  $p = ,004$ ), d.h. je höher der Wert der korrekten Antworten, desto niedriger der Selbsteinschätzungswert (wobei 1 der Bewertung 'sehr gut' und

3 der Bewertung 'schlecht' entspricht). Das bedeutet, dass die Sprecher mit der höchsten Zahl der korrekten Antworten haben sich selbst auch als sehr kompetent eingeschätzt.

Als Inputfaktoren wurden im Fragebogen folgende Informationen erfragt: (1) die Dauer (Anzahl der Jahre) des Deutschunterrichts; (2) die Anzahl der Wochenstunden, in der die Teilnehmer mit der deutschen Sprache Kontakt hatten sowie (3) der Kontakt außerhalb des Klassenzimmers.

Zunächst wurde ein Spearman-Assoziationstest durchgeführt, um die Korrelation zwischen den Testergebnissen (korrekte Antworten) und der Dauer des Deutschunterrichts (in Jahren) zu messen. Das Ergebnis zeigt, dass die zwei Variablen nicht miteinander korrelieren ( $r_s = -,268$ ,  $p = ,119$ ), d.h. der Sprachstand der Teilnehmer, gemessen an den *Cloze* Testwerten, wird nicht durch die Dauer des Kontakts mit der Sprache bestimmt. Dieses Ergebnis scheint auf den ersten Blick der allgemeinen Auffassung zu widersprechen, dass die Anzahl der Jahre, während der ein Lerner eine Sprache lernt, den Erwerb dieser Sprache beeinflusst. Ein genauerer Blick auf die extralinguistischen Daten der Teilnehmer verdeutlicht jedoch, dass die meisten Sprecher nur drei Jahre Deutschunterricht haben. Der Medianwert liegt bei 3 Jahren (siehe Abschnitt 3.3), was nicht ausreichend scheint, um einen Effekt auf den Sprachstand der Gruppe zu haben.

Die zweite getestete Variable ist die Anzahl der Wochenstunden, in denen die Teilnehmer Kontakt mit dem Deutschen hatten. Wie in 3.3. dargestellt, liegt der Durchschnitt hier bei 1,8 Wochenstunden. Das entspricht im Grunde der Dauer des Deutschunterrichts pro Woche. Auch in diesem Fall zeigt ein Spearman-Assoziationstest keine Korrelation zwischen den Testergebnissen und den Wochenstunden, die der Lerner dem Lernen der deutschen Sprache widmet ( $r_s = -,014$ ,  $p = ,937$ ). Auch diese Variable zeigt keine Korrelation mit den Ergebnissen und trägt somit wenig zur Erklärung der Variation zwischen den Sprechern bei.



Anders ist das Ergebnis, wenn wir die Art des Kontakts mit dem Deutschen analysieren. Wie in 3.3. beschrieben, haben 24 Studierende keinen Kontakt mit der deutschen Sprache außerhalb des Klassenzimmers; 13 Studierende haben entweder eine Erasmuserfahrung oder Kontakt mit deutschsprachigen Freunden, Verwandten bzw. Bekannten. Um eine Korrelation zwischen der Art des Kontakts und den Ergebnissen des *Cloze* Tests zu berechnen, wurde ein Point-Biserial-Assoziationstest durchgeführt (die Kontaktvariable ist nominal; Kontakt außerhalb des Klassenzimmers: ja/nein). Das Ergebnis weist eine signifikant positive Korrelation zwischen den Testergebnissen und dem Kontakt außerhalb des Klassenzimmers auf ( $r_{pb} = ,590$ ,  $p < ,001$ ), d.h. die Teilnehmer, die eine Erasmuserfahrung hatten und/oder mit Freunden, Verwandten bzw. Bekannten außerhalb des Unterrichts auf Deutsch kommunizieren, haben bessere Testergebnisse und weisen somit einen höheren Sprachstand auf. Es ist somit die Variable der Art des Sprachkontakts, die die Sprachleistung der Teilnehmer stark beeinflusst. Dieses Ergebnis wird in der nächsten Sektion genauer diskutiert.

## 5. Diskussion und Schlussfolgerung

In der vorliegenden Studie wurde ein Instrument vorgestellt, das zur Überprüfung des Sprachstands von portugiesischen L2-Sprechern des Deutschen entwickelt wurde. Der *Cloze* Test wurde mit einer Gruppe von 37 Hochschulstudierenden durchgeführt.

Die erste grundlegende Frage beschäftigte sich mit der Validität des entwickelten Instruments. In der Literatur werden *Cloze* Tests als effiziente und wirksame Instrumente beschrieben, um L2-Sprecher nach ihren Sprachkompetenzen zu unterteilen (Tremblay, 2011). Um wirksam zu sein, müssen die Lückentexte jedoch bestimmte Voraussetzungen erfüllen. In erster Linie müssen sie unterschiedliche grammatische, lexikalische und diskursive Strukturen mit unterschiedlichem Schwierigkeitsgrad erfassen. Nur eine Gradierung der Komplexität der

ausgelassenen Strukturen ermöglicht es, unterschiedliche Sprachniveaus zu differenzieren und sowohl Lerner mit niedrigen als auch mit sehr hohen Sprachkenntnissen zu testen.

Die Komplexitätshierarchie, die sich aus der Auswertung des vorliegenden *Cloze* Tests ergeben hat, lässt uns unsere erste Forschungsfrage positiv beantworten: In der Tat ermöglicht dieses Erhebungsinstrument, unterschiedliche Sprachniveaus der Teilnehmer zu erfassen. Wie in Tabelle 2 dargestellt, reicht die Spanne der richtigen Antworten von 1 bis 38 (bei einer Höchstpunktzahl von 38) und der Auslassungen von 0 bis 31.

Es konnte zudem eine klare Komplexitätshierarchie definiert werden, was unsere zweite Forschungsfrage positiv beantwortet. Der Schwierigkeitsgrad reicht von 4 bis 35 ausgefüllte Lücken. Dabei haben sich bestimmte Strukturen wie das Infinitivmorphem *{-en}* oder die Adverbien *deshalb/deswegen* als relativ einfach erwiesen, wohingegen lexikalisches Wissen, z.B. Komposita wie *Zeitraum/Zeitpunkt* betreffend, oder die Erkenntnis, dass bestimmte Verben reflexiv sind (*sich aufhalten*) oder eine trennbare Partikel haben (*ab in absagen*), nur von wenig Teilnehmern erfolgreich gemeistert wurden.

Eine weitere Frage erkundet, ob die Ergebnisse des *Cloze* Tests mit der Selbstbewertung der Teilnehmer korrelieren. In der Literatur zu diesem Thema wird trotz variierender Ergebnisse von vielen Autoren der Standpunkt vertreten, dass die Selbsteinschätzung der Sprecher sehr zuverlässig ist und häufig mit den Ergebnissen aus Sprachtests korreliert (Dolovic, 2018; LeBlanc & Painchaud, 1985; Ross, 1998). Zu diesem Schluss kommen wir auch in der vorliegenden Studie. Der Assoziationstest zwischen der Selbsteinschätzung der Teilnehmer und der Anzahl der richtig ausgefüllten Lücken zeigt in der Tat eine signifikante Korrelation, d.h. die Sprecher, die ihre Sprachkompetenz als gut eingeschätzt haben, hatten in der Tat die besten Ergebnisse; die Teilnehmer mit den schlechtesten Ergebnissen im *Cloze* Test hatten ihre Sprachleistung als schlecht bewertet.

Die letzte Frage beschäftigt sich mit dem Einfluss der extralinguistischen Inputfaktoren auf den gemessenen Sprachstand der Teilnehmer. Um den Einfluss dieser Variablen zu verstehen, müssen wir die Ergebnisse der drei Korrelationen zusammen interpretieren. Es wurde gezeigt, dass die Anzahl der Jahre des Deutschunterrichts und die Anzahl der Wochenstunden, die die Sprecher der deutschen Sprache widmen, keinen Einfluss auf die Ergebnisse haben. Jedoch ist zu beachten, dass die große Mehrheit der Sprecher weniger als vier Jahre Deutsch lernt und zudem sehr wenig wöchentlichen Kontakt mit dem Deutschen hat (im Durchschnitt weniger als zwei Stunden pro Woche). Schaut man jedoch auf die dritte Variable, d.h. auf den Kontakt außerhalb des Klassenzimmers, dann kommen wir zu dem Schluss, dass dies eine ausschlaggebende Variable ist. Die Teilnehmer, die auch außerhalb des Unterrichts Kontakt mit dem Deutschen haben oder hatten, wenn auch relativ reduziert, sind die Sprecher, die die besten Leistungen im *Cloze* Test vorweisen. Es wird somit klar, dass der reduzierte Kontakt mit dem Deutschen, der sich nur auf den Klassenzimmerkontext beschränkt, nicht ausreichend für die Entwicklung hoher Sprachkompetenzen in einer Zweitsprache ist. Der Kontakt außerhalb des Klassenzimmers ist ausschlaggebend. Die Ergebnisse bestätigen somit, was in der Forschung schon lange über den Erwerb einer zweiten Sprache bekannt ist: Um erfolgreich zu sein, erfordert der Erwerbsprozess langen, kontinuierlichen und vielfältigen Kontakt mit der Zielsprache (Ellis, 1997).

## **6. Limitationen der Studie**

Die vorliegende Studie wurde im Rahmen eines Projekts des Masters *Deutsch-Portugiesische Studien* entwickelt. Sie entspricht somit den ersten Schritten der Autoren dieser Studie in der wissenschaftlichen Forschung. In diesem Sinne ist die vorliegende Studie durch mehrere Faktoren beeinträchtigt. Eine erste Limitation ist mit der ungleichartigen, durch die aktuelle

Covid-19-Pandemie bedingten Erhebungsmethode verbunden. Der Test wurde teils in Papierversion, teils als *Online*-Instrument angewandt. Ein Mann-Whitney-Test hat zwar keine signifikanten Unterschiede zwischen beiden Methoden gezeigt hat ( $U = 96.5, p = \text{n.s.}$ ). Trotzdem wäre eine einheitliche Erhebungsmethode angebracht gewesen.

Wie die Datenanalyse gezeigt hat, wäre auch eine größere Vielfalt im Lernerprofil der Teilnehmer wünschenswert, v.a. in der Dauer des Kontakts mit der deutschen Sprache. So haben Tremblay und Garrison (2010) in ihrer Studie zur Bewertung der L2-Französischkenntnisse erwachsener Lerner gezeigt, dass die Anzahl der Unterrichtsjahre in der L2 das aussagekräftigste Indiz für die Leistung im *Cloze* Test ist. Daher ist davon auszugehen, dass auch der vorliegende *Cloze* Test bei einer weiteren Zeitspanne an Unterrichtsjahren der Teilnehmergruppe eine Korrelation zwischen diesen Variablen aufzeigen könnte.

Ausgehend von den vorläufigen Ergebnissen der vorliegenden Studie liegt ein nächster Schritt nun in der Validierung dieses Instruments mit einer größeren Stichprobe. Zudem trägt auch eine Korrelation der Testergebnisse mit anderen Sprachstandsmessungen, wie etwa Einstufungsprüfungen, zur Validierung des *Cloze* Tests bei und soll in zukünftiger Arbeit hinzugefügt werden. Eine größere Stichprobe und weitere Korrelationen mit anderen Instrumenten der Sprachstandserhebung ermöglichen es letztendlich, konkrete Sprachstandsniveaus für den vorliegenden Test zu definieren.

### **Literaturverzeichnis**

- Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly*, 13, 219-227. DOI: 10.2307/3586211
- Bachman, L. F. (1982). The trait structure of close test scores. *TESOL Quarterly*, 16, 61-70.  
DOI: 10.2307/3586563

- Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, 16, 61-70. DOI: 10.2307/3586277
- Brown, J. D. (1983). A closer look at cloze: Validity and reliability. In J. W. Oiler, Jr. (Ed.), *Issues in language testing research* (pp. 237-250). Rowley, MA: Newbury House.
- Brown, J. D. (1984). A cloze is a cloze is a cloze? In J. Handscombe, R. A. Orem, & B. P. Taylor (Eds.), *On TESOL '83* (pp. 109-119). Washington, DC: TESOL.
- Brown, J. D. (1988). Tailored cloze: Improved with classical item analysis techniques. *Language Testing*, 5, 19-31. DOI: 10.1177/026553228800500102
- Chung, E. S., & Ahn, S. (2019). Examining Cloze Tests as a Measure of Linguistic Complexity in L2 Writing. *Language Research*, 55(3), 627-649. <https://doi.org/10.30961>
- Dolovic, H. (2018). An examination of self-assessment and interconnected facets of second language reading. *Reading in a Foreign Language*, 30(2), 189-208.
- Duff, P. (2017). Social dimensions and differences in instructed SLA. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 379-395). New York: Routledge. DOI: 10.4324/9781315676968
- Ellis, R. (1997). *Second Language Acquisition*. Oxford: Oxford University Press.
- Fotos, S. (1991). The cloze test as an integrative measure of EFL proficiency: A substitute for essays on college entrance examinations? *Language Learning*, 41, 313-336. DOI: 10.1111/j.1467-1770.1991.tb00609.x
- Hanania, E., & Shikhani, M. (1986). Interrelationships among three tests of language proficiency: Standardized ESL, cloze and writing. *TESOL Quarterly*, 20, 97-110.
- Hinofotis, F. B., & Snow, B. G. (1980). An alternative cloze testing procedure: Multiple-choice format. In J. W. Oller, Jr. & K. Perkins (Eds.), *Research in language testing* (pp. 129-133). Rowley, MA: Newbury House.
- Jonz, J. (1976). Improving on the basic egg: The M-C cloze. *Language Learning*, 26, 255-256. DOI: 10.1111/j.1467-1770.1976.tb00276.x

- LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly*, 19, 673–687. DOI: 10.2307/3586670
- Markham, P. L. (1985). The rational deletion cloze and global comprehension in German. *Language Learning*, 35(3), 423-430. DOI: 10.1111/j.1467-1770.1985.tb01085.x
- Oller, J. W. Jr., & Inal, N. (1971). A cloze test of English prepositions. *TESOL Quarterly*, 5, 315-326.
- Ross, S. (1998). Self-assessment in second language testing: a meta-analysis and analysis of experiential factors. *Language Testing*, 15(1), 1–20. DOI: 10.1177/026553229801500101
- Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30, 414-438. DOI: 10.1177/107769905303000401
- Tremblay, A., & Garrison, M. D. (2010). Cloze tests: A tool for proficiency assessment in research on L2 French. In M. T. Prior, Y. Watanabe, & S.-K. Lee (Eds.), *Selected proceedings of the Second Language Research Forum 2008* (pp. 73-88). Somerville, MA: Cascadilla Press.
- Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: "Clozing" the Gap. *Studies in Second Language Acquisition*, 33(3), 339-372. DOI: 10.1017/S0272263111000015

## **Anhang A**

### **Der Cloze Test**

Heute ist der Hochzeitstag von Lukas, dem jüngsten Mitglied der Familie Fischer. Lukas und Lena, Braut und Bräutigam, haben sehnsüchtig [ \_ \_ \_ ] diesen Tag gewar[ \_ \_ \_ ], und nach monatelanger Planung werden sie endlich an eine[ \_ ] sonnig[ \_ \_ ] Sommertag heiraten.

Die Familie Fischer hatte nicht erwartet, dass Lukas so plötz[ \_ \_ \_ ] um Lenas Hand anhalten w[ \_ \_ \_ ]. Angesichts der schwie[ \_ \_ \_ \_ ] finanziellen Verhältnisse der Familie fanden Luka[ \_ ] Eltern den Zeit[ \_ \_ \_ \_ ] nicht besonders angemessen.

Es war jedoch jedem klar, w[ \_ \_ ] sehr er [ \_ \_ ] liebte. Des[ \_ \_ \_ ] würde er keinen weiteren Monat vergehen lassen, ohne d[ \_ \_ ] diese Hochzeit stattfinden würde.

A[ \_ ] die Gäste bereits Platz genommen haben und sich auf den Beginn der Zerem[ \_ \_ \_ ] vorbereiteten, erhielt Lukas eine beunruhigende Nachr[ \_ \_ \_ ]. Nervös und aufger[ \_ \_ ] rief er seinen Vater z[ \_ \_ ] Seite und sagte i[ \_ \_ ], dass niemand wisse, wo Lena sei. Also beschlossen alle, ihm bei der Suche nach ihr zu helf[ \_ \_ ].

Er konnte es nicht fassen, dass sie i[ \_ \_ ] nicht heiraten wollte. Er entschied, sie anzurufen. Aber es [ \_ \_ ] umsonst. Ihr Handy klingel[ \_ \_ ] auf dem Nachttisch i[ \_ ] Hotelzimmer, wo sie übernachtet ha[ \_ \_ \_ ].

N[ \_ \_ ] mehrstündiger Suche nach der Braut und unter Zeitdruck sagte Lukas d[ \_ \_ ] Zeremonie schließlich [ \_ \_ ]. Die verzweifelt[ \_ \_ ] Mitglieder beider Familien hätt[ \_ \_ ] nie erwartet, dass Lena doch n[ \_ \_ \_ ] bereit war, Lukas [ \_ \_ ] heiraten.

Im Morgengrauen d[ \_ \_ ] nächsten Tages erfuhren alle, wo s[ \_ \_ ] die fehlende Braut aufhielt. Oh[ \_ \_ ] ihren Verlobten Bescheid zu sagen, war sie allein z[ \_ \_ ] Veranstaltungsraum gelaufen, [ \_ \_ ] die Ausstattung des Raums zu überprüfen. Als sie die Toilette benutzen musste, ist es dann pass[ \_ \_ \_ ] : Sie ist darin gefangen gebl[ \_ \_ \_ \_ ]. Und so verlief der Hochzeitstag von Lukas und Lena.