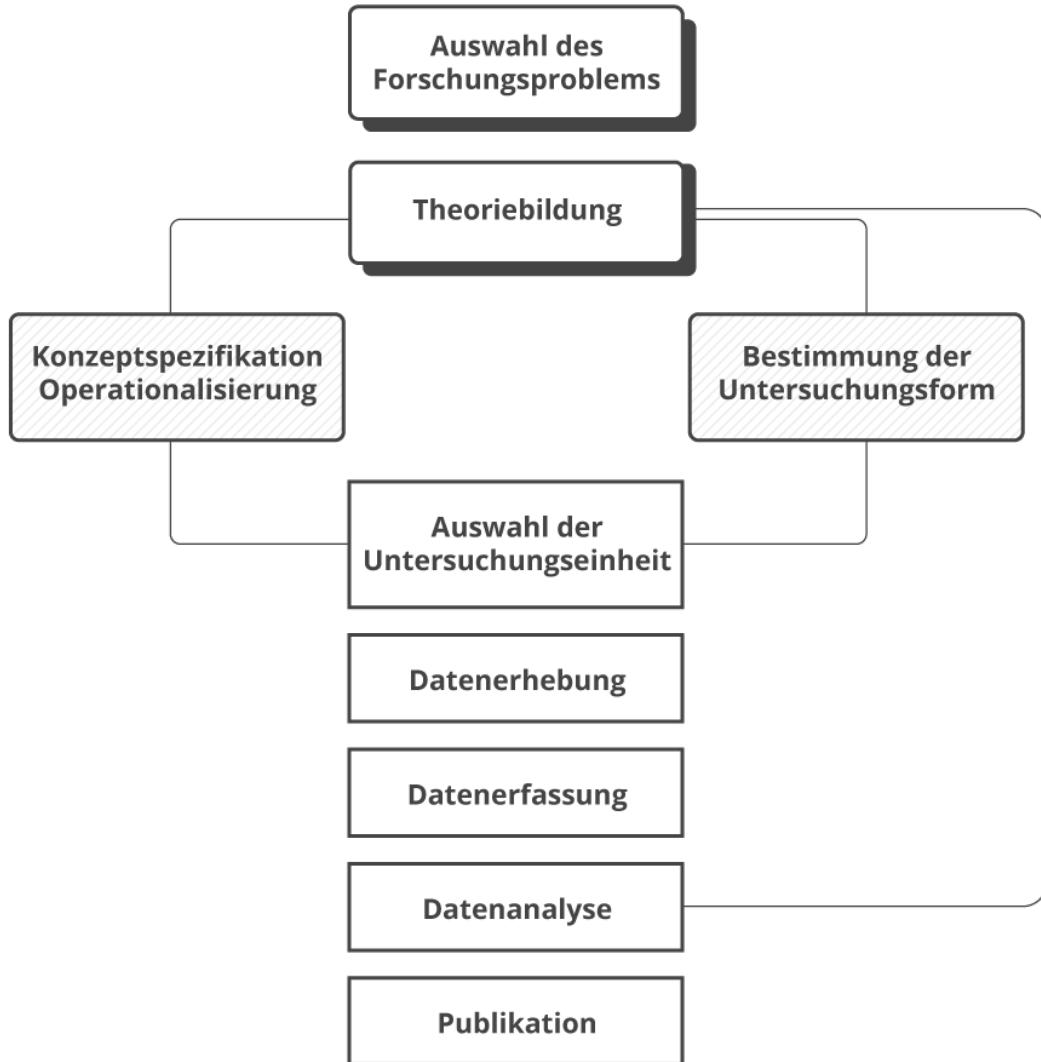


Von Tabellen zu dem Text: Wie kann ich Methodologie und Datenauswertung in der Sprachwissenschaft gründlich gestalten?

Inhalt

1	Forschungsvorgehen	2
1.1	Abhängig und Unabhängige Variable	2
2	Daten und ihre Vielfalt	3
2.1	Was sind meine Daten?	3
2.2	Warum muss ich die Daten vor- und darstellen?	4
3	Methoden, ein Überblick.....	4
3.1	Was sind Methoden (in die Sprachwissenschaften)?.....	4
3.2	Wie kann ich Methoden beschreiben?.....	4
4	Dokumentation und Argumentation.....	5
4.1	Wann sind Daten Argumente?.....	5
4.2	Tabellen und Graphiken, wann sind sie passend?	5
5	Statistik, wann und wie?	7
6	Bibliografie zum wissenschaftlichen Schreiben.....	8
7	References	8

1 Forschungsvorgehen



(Academic Lab, 2025a)

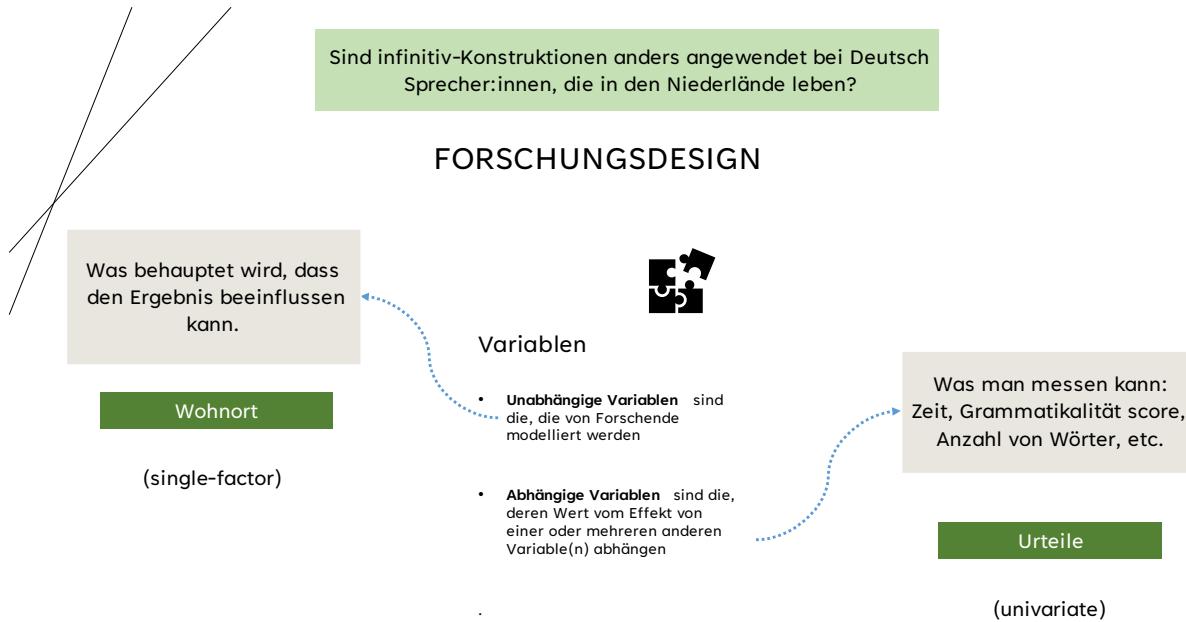
Tipp: Wenn theoretische Hintergründe beschrieben werden, sollten Fachbegriffe und Definitionen markiert werden. Bedeuten die Fachbegriffe immer dasselbe? Kann ich hier ein Beispiel einfügen, um es besser zu erläutern?

1.1 Abhängig und Unabhängige Variable

Unabhängige Variablen sind solche, die von den Forschenden modelliert werden. Abhängige Variablen sind solche, deren Wert vom Effekt einer oder mehrerer anderer Variablen abhängt.

Beispiel 1:

Forschungsfrage: Werden Infinitivkonstruktionen bei Deutsch sprechenden Personen, die in den Niederlanden leben, anders angewendet?



Methoden der Sprachwissenschaft

212

(Ribbert & Kuiken, 2010)

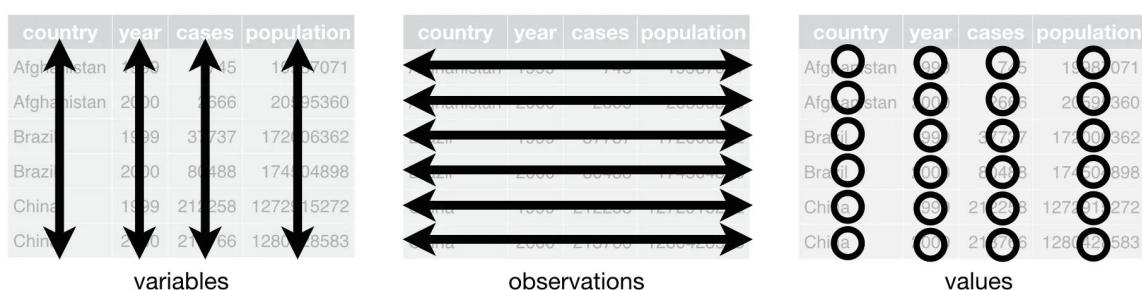
2 Daten und ihre Vielfalt

2.1 Was sind meine Daten?

In der Sprachwissenschaft sind Daten sehr vielfältig. Hier kann man mit Corpora, Interviews, Fragebögen, Narrativen, Ethnografie und anderen Methoden arbeiten. Deshalb muss die Forschungsfrage sehr klar definiert sein. Die Forschungsfrage ermöglicht es, den „Daten-Horizont“ einzugrenzen und das Thema genauer anzugehen.

Wichtig: Die Forschungsfrage begrenzt die Daten, nicht umgekehrt.

Um Daten zu organisieren, gibt es streng genommen keinen „Muss“. Es hilft normalerweise, mit Tabellen zu arbeiten, um jedem Datensatz einen Wert zuzuweisen. Hier bietet sich die „tidy data“-Struktur an: Jede Zeile ist ein Datensatz, jede Spalte eine abhängige Variable.



(Wickham, 2014; Wickham & Grolemund, 2017)

Beispiel 2:

Forschungsfrage: „Kodieren Bewegungsverben im Spanischen primär die Bewegungsrichtung statt der Bewegungsart?“

Methode: Bildbeschreibung

Data	Sprache	Bild	Sprecher	Bewegungsart	Richtung
Das Mädchen ist abgestützt	Deutsch	CAYO_01	ABC_01	-	-
Das Mädchen läuft die Treppe runter	Deutsch	DESCER_01	ABC_02	laufen	runter
La chica se cayó	Spanisch	CAYO_01	ABC_01	Caer	-
La chica bajó las escaleras	Spanisch	DESCER_01	ABC_02	-	bajar

2.2 Warum muss ich die Daten vor- und darstellen?

Daten sind die Grundlage einer jeden Forschung. Wo und wie sie erhoben wurden, ermöglicht es anderen, ihren Wert bzw. ihre Aussagekraft zu bestimmen. Sie sind der Ballast, mit dem sich der Rest erklären lässt.

3 Methoden, ein Überblick

3.1 Was sind Methoden (in die Sprachwissenschaften)?

Allgemein sind Methoden planmäßige Verfahren. Das heißt, man überlegt sich, wie sich eine Forschungsfrage mit einem bestimmten Verfahren beantworten lässt.

Wichtig: Theorie, Forschungsfragen und Methoden müssen zueinander passen, d. h., die Forschungsfrage muss aus einer Lücke in der Theorie entstehen. Aus der Theorie können bestimmte Verfahren übernommen und als Methoden angewendet werden.

„Um wissenschaftliche Argumentationen (also Beweisgründe) von schlichten Behauptungen unterscheiden zu können, sind – neben dem zu bearbeitenden Material (Texte, Gestein, Zellen oder was auch immer) – die jeweiligen Methoden von zentraler Bedeutung. Mit ihnen lässt sich eine Aussage X über einen Gegenstand oder Sachverhalt Y insofern plausibilisieren, als sie zeigen, wie und unter welchen Umständen diese Aussage zustande gekommen ist. Deshalb sollten Methoden nie willkürlich oder unkritisch verwendet werden.“ (Academic Lab, 2025d)

Probleme der Empirie: https://home.uni-leipzig.de/methodenportal/probleme_der_empirie/

Im Rahmen der statistischen Testtheorie wurden Kriterien für eine gute Praxis entwickelt. Diese Kriterien dienen der Qualitätssicherung von Erhebungsverfahren. Die drei Kriterien sind:

- Objektivität: Die Messergebnisse sind unabhängig von der Person, die die Messung durchführt.
- Reliabilität: Messgenauigkeit einer Messung
- Validität: Gültigkeit, d. h. „inwieweit die Messung das misst, was sie messen soll“ (Academic Lab, 2025b)

3.2 Wie kann ich Methoden beschreiben?

Die Beschreibung der Methoden dient gleichzeitig der Dokumentation der Forschungsverfahren und ist eine Art „Versicherung“, dass alle Entscheidungen in der Forschung motiviert sind (und nicht eigenzweckmäßig).

In Bezug auf die Objektivität sollte beschrieben werden, dass unabhängig von der Person, die die Forschung durchführt, das gleiche Ergebnis erzielt wird. Hier sollte genau erklärt werden, warum die Datenzelle #1 als A und die Datenzelle #2 als B ausgewertet wurde. Dann kann dieses Verfahren auch von Personen aus anderen Bereichen nachvollzogen werden. Wenn es hilfreich ist, kann hier ein Flussdiagramm gebaut werden.

Was die Reliabilität angeht, sollte die unabhängige Variable(n) genau erklären, was man von ihr behauptet. Hier sollte genau erklärt werden, warum dies ein guter Wert ist, um eine Hypothese anzunehmen oder zu widerlegen.

Was die Validität angeht, sollte die Konstruktion der Methode inhaltlich valide sein, um die zu messende Merkmale zu erfassen.

Wichtig: Die Methode muss nicht alles können. Im Gegenteil: Es ist wünschenswert, dass die Grenzen der Methodik und der Forschung genau angegeben werden.

Beispiel 3:

Forschungsfrage: Werden Infinitivkonstruktionen bei Deutsch sprechenden Personen, die in den Niederlanden leben, anders angewendet?

“Für die vorliegende Studie wurde den Teilnehmern eine schriftliche Liste mit Satzpaaren vorgelegt, von denen ein Satz in der deutschen Norm und der andere in einer niederländischen Norm stand. Die Studenten sollten angeben, ob beide Sätze richtig waren oder ob einer oder keiner von beiden richtig war. Es wurde angegeben, dass beide Sätze gleich sein sollten.” (Ribbert & Kuiken, 2010, p. 44, angepasst und übersetzt)

Beispiel 4:

Forschungsfrage: Werden Infinitivkonstruktionen bei Deutsch sprechenden Personen, die in den Niederlanden leben, anders angewendet?

“Für die vorliegende Studie wurde den Teilnehmern eine schriftliche Liste von Satzpaaren vorgelegt, die bis auf die Verwendung oder das Weglassen von um identisch waren, wie das folgende Beispiel zeigt:

a. Michael hat probiert, um Daniela anzurufen.

b. Michael hat probiert Daniela anzurufen.

Die Schüler sollten angeben, ob beide Sätze richtig sind oder nur einer oder keiner. Es wurde darauf hingewiesen, dass beide Sätze denselben Sachverhalt bezeichnen sollten. (Siehe Anhang A für weitere Beispiele aus dem Test).” (Ribbert & Kuiken, 2010, p. 44, angepasst und übersetzt)

4 Dokumentation und Argumentation

4.1 Wann sind Daten Argumente?

Daten sind Argumente, wenn man empirisch arbeitet. Dann kann ich meine Aussagen mit Daten belegen und bekräftigen.

4.2 Tabellen und Graphiken, wann sind sie passend?

Tabellen und Graphiken sind nun wichtig, wenn sie was erläutern, erklären oder bildlich was wiedergeben, die den Text komplementieren. Sie sollten immer im Text erwähnt werden und dementsprechend erklärt. Was sollte man in der Tabelle erkennen? Was sollte ich aus dem Bild

mitnehmen? Es ist sehr hilfreich, zum Beispiel, eine Tabelle, um die erhobenen Corpus-Dateien zusammenzufassen.

Checklist (Academic Lab, 2025c):

1. Allgemein

- Hat die Tabelle oder Abbildung tatsächlich einen Mehrwert? Erklärt oder illustriert sie also etwas, das nicht bereits im Text ausführlich entfaltet wurde?
- Wenn ja, was genau ist der Erkenntnisgewinn? Wird er im Text selbst eingeholt? Werden die Abbildungen also im Text ausreichend behandelt und werden Verweise deutlich? Oder sind sie eher außenstehende Bestandteile der Arbeit, die eher im Anhang sinnvoll platziert sind?
- Sind alle Elemente der Grafik oder Tabelle sauber beschriftet (x- und y-Achse, Spalten, Zeilen)?
- Stimmen die Maßeinheiten, sofern es welche gibt?
- Haben alle dargestellten Werte entsprechende Einheiten?
- Enthält die Abbildung oder Tabelle vielleicht irrelevante Daten? Ist sie auf das, was sie leisten soll, zugeschnitten?
- Ist die Tabelle oder Abbildung zu komplex, um sie direkt zu begreifen? Wenn ja, sollten die Informationen auf mehrere Tabellen oder Abbildungen aufgeteilt werden. Mehr als zwei Variablen in einer Abbildung oder Tabelle sind zu vermeiden. Vor allem in Abbildungen ist eine dritte Dimension nicht zu empfehlen.
- Ist die Abbildung auch für sich genommen zu verstehen? (Das sollte sie.)
- Hat die Abbildung einen sprechenden, aussagekräftigen Titel?

2. Design

- seriflose Schrift in Abbildungen und Tabellen für eine bessere Lesbarkeit
- in allen Abbildungen und Tabellen die gleiche Schriftfamilie
- Schriftgröße nicht weniger als 8 pt und nicht mehr als 14 pt
- Kommawerte in Tabellen untereinander am Komma ausrichten für eine schnelle Orientierung
- wenn möglich auf farbige Abbildungen verzichten (Lassen sich die Daten oder Aussagen auch anders darstellen? Farbdruck trägt nicht immer zu einer besseren Übersicht bei und ist zudem deutlich teurer.)
- Druckqualität der Abbildung sowie großer Tabellen prüfen

3. Platzierung

- Taucht die Abbildung/Tabelle dort auf, wo sie im Fließtext verhandelt wird? (Erspart umständliches Blättern.)
- Passt sich die Abbildung/Tabelle in den Satzspiegel ein?
- Hilft vielleicht eine Seite im Querformat, die gesondert in den fortlaufenden Text eingefügt wird?
- Große Darstellungen können in einen Anhang verbannt werden.
- Ist die Beschriftung unmittelbar unter der Abbildung (also nicht aus Versehen umgebrochen)?
- Je nach Größe kann der Fließtext auch umlaufend sein. Wird der Platz jedoch zu schmal, entstehen zu viele Trennungen, was sich schlecht liest.

Beispiel 5:¹

“In summary, it should be noted that the use of Dutch *om* forms a superset of the use of German *um*. Dutch furthermore has a high number of contexts where *om* can be used optionally, and this optionality is driven by the factor “speaker subjectivity”, which is in the first place dependent on the matrix verb being used. In German, on the other hand, the use of *um* is hardly ever optional. It is self-evident that it will be hard for native speakers of German to capture these opaque rules and the subtle preferences for the use of optional *om*.” (Ribbert & Kuiken, 2010, p. 43)

Table 1. *The complementizer in the two languages per category.*

German	Dutch	Example
Obligatory	Obligatory	Piet nimmt die Straßenbahn, um nicht zu spät zu kommen. Piet neemt de tram om niet te laat te komen. Piet takes the tram COMP not too late to come “Piet takes the tram in order to be on time.”
Obligatory	Ungrammatical	Non-existent
Obligatory	Optional	Non-existent
Ungrammatical	Obligatory	*Das ist ein Aufsatz um zu heulen. Dat is een werkstuk om te huilen. this is a paper COMP to cry “This is a paper which could make you cry.”
Ungrammatical	Ungrammatical	Every context where the use of <i>om</i> is ungrammatical in Dutch
Ungrammatical	Optional	*Michael hat versucht um eine Doktorandenstelle zu bekommen. Michael heeft gepoogd om een promotieplaats te krijgen. Michael has tried COMP a Ph.D. stipend to get “Michael has tried to get a Ph.D. grant.”
Optional	Ungrammatical	Non-existent
Optional	Optional	Non-existent

5 Statistik, wann und wie?

Statistik ist nun wichtig, wenn die Forschungsfrage sich quantitativ beantworten lässt. Man spricht dann von einem hypothesenprüfenden Verfahren. Um einem Test auszuwählen, sollte man klar haben:

- Welche sind meine Variable und mein Aufbau?
- Wie werden diese Variable in statistischen Begriffen übersetzt? (un/abhängig Variable, unabhängige Beobachtungen, p-Wert, Signifikanz, etc.)
- Kann ich begründen, warum ich den Test ausgewählt habe?
- Kann ich ausreichend erklären, die Ergebnisse meines Tests?

Eine statische Auswertung ist nicht für jede Forschungsfrage notwendig und kann bei falscher Anwendung irreführend sein (Huff & Geis, 1993). Es ist jedoch wünschenswert, sich daran zu orientieren.

¹ Da es um ein Beispiel für illustrative Zwecke handelt, die Tabelle ist hier als eine Abbildung. Am besten sollte man Tabellen immer selbst konfigurieren und als Tabelle im Dokument einfügen.

Beispiel 6:Table 2. *Number of mistakes per group, German test.*

Group	N	Mean	SD	Standard error mean
Experimental	52	6.1	4.8	.66
Control German	38	1.6	1.8	.28

Table 3. *Number of mistakes per group, Dutch test.*

Group	N	Mean	SD	Standard error mean
Experimental	52	5.1	4.2	.58
Control Dutch	40	3.3	3.2	.51

The experimental group was compared to the German control group by means of an independent samples t-test. Table 2 displays the group statistics. It shows that the group of Germans living in the Netherlands made more mistakes ($M = 6.1$, $SD = 4.8$) than the control group of Germans who had no contact with the Dutch language ($M = 1.59$, $SD = 1.75$). The difference between the groups was significant, $t(68.2) = 6.3$, $p < .001$, two-tailed.⁴

Table 3 shows that the German research participants made more mistakes ($M = 5.1$, $SD = 4.2$) than the Dutch control group ($M = 3.3$, $SD = 3.2$). The difference between the groups was significant, $t(90) = 2.24$, $p < .05$, two-tailed.

(Ribbert & Kuiken, 2010)

Statistische Beratung der Freie Universität Berlin: <https://www.stat.fu-berlin.de/>

6 Bibliografie zum wissenschaftlichen Schreiben

- Becker, H. S. (2021). *Soziologische Tricks: Wie wir über Forschung nachdenken können* (U. Schäfer & E. Heinemann, Trans.). *Positionen*. Hamburger Edition.
- Litosseliti, L. (2010). *Research methods in linguistics. Research methods in linguistics*. Continuum.
- Pinker, S. (2014). *The sense of style: The thinking person's guide to writing in the 21st century!* Viking.
- Podesva, R. J., & Sharma, D. (Eds.). (2018). *Research Methods in Linguistics*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139013734>

7 References

- Academic Lab. (2025a). *Forschungsprozess*. Universität Leipzig. <https://home.uni-leipzig.de/methodenportal/forschungsprozess/>
- Academic Lab. (2025b). *Gütekriterien*. Universität Leipzig. <https://home.uni-leipzig.de/methodenportal/guetekriterien/>
- Academic Lab. (2025c). *Tabellen & Grafiken*. Universität Leipzig. <https://home.uni-leipzig.de/methodenportal/tabellen-grafiken/>
- Academic Lab. (2025d). *Was sind Methoden?* Universität Leipzig. https://home.uni-leipzig.de/methodenportal/was_sind_methoden/
- Huff, D., & Geis, I. (1993). *How to lie with statistics*. Norton.
- Ribbert, A., & Kuiken, F. (2010). L2-induced changes in the L1 of Germans living in the Netherlands. *Bilingualism: Language and Cognition*, 13(1), 41–48. <https://doi.org/10.1017/S1366728909990320>

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, 59(10).

<https://doi.org/10.18637/jss.v059.i10>

Wickham, H., & Grolemund, G. (2017). *R for data science: Import, tidy, transform, visualize, and model data* (First edition). O'Reilly.