

# **Experimente**

Medienlinguistische Methodik

Arne Rubehn

arne.rubehn@uni-passau.de

04.11.2025

## Zusammenfassung

In dieser Sitzung beleuchten wir die **Erhebung** von wissenschaftlichen Daten durch **Experimente**. Hierbei sollen folgende Leitfragen thematisiert werden: Welche experimentellen Verfahren gibt es in der Linguistik? Wie funktionieren diese Verfahren im Wesentlichen? Was ist bei der Analyse von experimentellen Daten zu beachten?

*Dieses Handout orientiert sich maßgeblich an Kleineberg und Henneke (2023, 227-233).*

## 1 Einführung

Experimente finden unter **kontrollierten Bedingungen** statt und untersuchen das Verhalten von Proband:innen. Die Situation wird hierbei bewusst manipuliert, um eine **Hypothese** zu überprüfen. Die Kriterien der Objektivität, Reliabilität und Repräsentativität müssen auch bei Experimenten gegeben sein. Für die Durchführung von Experimenten mit Versuchspersonen ist in der Regel die Zustimmung einer Ethikkommission notwendig, die potenzielle Risiken für Proband:innen einschätzt.

Experimente untersuchen grundsätzlich den Zusammenhang zwischen verschiedenen **Variablen**, also prinzipiell veränderbaren Faktoren einer Untersuchung. Wir unterscheiden hierbei zwischen drei Arten von Variablen:

- **unabhängige Variable:** Sie wird im Experiment bewusst verändert, um mögliche Auswirkungen auf andere Variablen zu untersuchen.
- **abhängige Variable:** Sie wird gemessen, um festzustellen, ob sie sich durch die Änderung der unabhängigen Variable ändert.
- **Störvariablen:** Variablen, die einen Effekt auf die abhängige Variable haben können, den man „herausrechnen“ (*kontrollieren*) möchte.

Illustrieren wir diese verschiedenen Variablen mal anhand einer Aussage, die man des Öfteren mal auf WG-Partys hört:

*„Ich treffe beim Beerpong besser, wenn ich getrunken habe.“*

Will man diese Aussage nun in einem wissenschaftlichen Experiment überprüfen, muss man die verschiedenen Variablen zunächst einmal klar benennen und definieren. In diesem Fall möchte man herausfinden, ob der Alkoholpegel des Werfers oder der Werferin einen Einfluss auf die Trefferquote hat. Dementsprechend wäre der Alkoholpegel die **unabhängige Variable**, die Forschende auch relativ einfach manipulieren könnten. Die

Trefferquote hingegen wäre in diesem Setup die **abhängige Variable**. Je nach genauem Experimentdesign treten allerdings auch verschiedene potenzielle **Störvariablen** auf, die berücksichtigt werden müssen: Untersucht man die gleiche Gruppe an Proband:innen in einer Sitzung, könnte auch die Wurfoutine einen Einfluss haben. In diesem Fall könnte es sein, dass Proband:innen nach einem Getränk tatsächlich besser werfen – aber nicht, weil der Alkohol einen förderlichen Effekt hat, sondern weil sie besser „eingeworfen“ sind. Untersucht man die gleiche Gruppe in unterschiedlichen Sitzungen, kann man zwar diesen Effekt verhindern, allerdings könnte hier die Tagesform die Ergebnisse verfälschen; bei der Untersuchung verschiedener Proband:innengruppen spielen wiederum die unterschiedlichen Fähigkeiten der einzelnen Testpersonen eine Rolle. Dieses Beispiel zeigt: Es ist in den meisten Fällen unmöglich, den Einfluss von Störvariablen komplett zu verhindern, aber durch sorgfältiges Experimentdesign und ausreichend große Stichproben (*Repräsentativität!*) können diese Effekte zumindest einmal minimiert werden.

Das zweite Standbein des formellen Rahmens für Experimente sind die **Hypothesen**, die überprüft werden. Das typische Verfahren hierbei ist es, zunächst von einer **Nullhypothese** auszugehen, die keinen Zusammenhang zwischen den untersuchten Variablen annimmt. Dem gegenüber stehen **Gegenhypothesen**, die einen bestimmten, signifikanten Zusammenhang zwischen den untersuchten Variablen annehmen. Die Nullhypothese ist erst dann zu verwerfen, wenn es solide Evidenz für eine der formulierten Gegenhypothesen gibt! Für unser obiges Beispiel wäre also als Nullhypothese anzunehmen, dass es keinen Zusammenhang zwischen dem Alkoholpegel und der Trefferquote eines Werfers oder einer Werferin beim Beerpong gibt. Dem gegenüber steht die Gegenhypothese, dass doch ein solcher Zusammenhang besteht – ganz streng genommen lassen sich hier sogar zwei Gegenhypothesen formulieren: Die eine besagt, dass ein höherer Alkoholpegel mit einer *höheren* Trefferquote einhergeht; die andere, dass ein höherer Alkoholpegel mit einer *niedrigeren* Trefferquote einhergeht.

## 2 Psycholinguistische Experimente

Psycholinguistische Experimente spielen – wie der Name schon sagt – eine große Rolle in der **Psycholinguistik**, die sich mit den kognitiven Prozessen der Sprachverarbeitung und -produktion beschäftigt. Auch angrenzende Felder wie z.B. die Phonetik profitieren von solchen experimentellen Daten. Im Gegensatz zu neurolinguistischen Verfahren (s. nächster Abschnitt) werden kognitive Abläufe nicht direkt gemessen, stattdessen werden andere externe Messgrößen erfasst, die Rückschlüsse auf kognitive Prozesse erlauben. Diese indirekten Verfahren haben den Vorteil, dass sie deutlich einfacher und kostengünstiger umzusetzen sind, da sie ohne größere technische Hilfsmittel auskommen.

## 2.1 Reaktionszeitmessungen

Eine der einfachsten Möglichkeiten, kognitive Prozesse indirekt zu messen, ist die **Reaktionszeitmessung**. Hierbei wird die Zeit erfasst, die zwischen der Präsentation eines Reizes und der Reaktion der Versuchsperson verstreicht. Um solche Messungen möglichst aussagekräftig zu machen, werden Proband:innen in der Regel dazu angehalten, die ihnen gestellte Aufgabe möglichst schnell und akkurat zu beantworten. Die Idee dahinter, Reaktionszeiten zu messen, beruht auf der simplen Annahme, dass die Reaktionszeit in direktem Zusammenhang mit kognitiven Abläufen steht. Das Gehirn braucht also länger, um anspruchsvollere Reize zu verarbeiten, als einfachere.

Das lässt sich mit einem simplen Experiment, dem **Stroop-Test**, illustrieren (Stroop, 1935). Hierbei werden den Proband:innen verschiedene Farbwörter präsentiert, die in unterschiedlichen Farben gedruckt sind. Die Aufgabe ist es nun, die Druckfarbe zu benennen. Hierbei kann die Druckfarbe mit dem Farbwort übereinstimmen, oder sich von ihm unterscheiden. Die unabhängige Variable des Experiments ist also die Kongruenz zwischen Druckfarbe und Farbwort:

- **kongruente Bedingung:** rot blau
- **inkongruente Bedingung:** blau rot

Abhängige Variablen in diesem Experiment sind folglich die Reaktionszeit, aber auch die Fehlerrate. Beide abhängige Variablen zeigen einen stabilen Effekt: Die inkongruente Bedingung führt zuverlässig zu einer höheren Reaktionszeit und einer höheren Fehlerrate. Ein solcher inkongruenter Stimulus (z.B. blau) ist also für das Gehirn offenbar schwieriger zu verarbeiten als ein kongruenter.

Der Stroop-Test stellt eine klassische **Benennungsaufgabe** dar – die Aufgabe der Proband:innen ist es also, einen Reiz zu benennen. Bei Experimenten zur Sprachproduktion sind neben solchen Benennungen auch **Kategorisierungen** üblich, bei denen es also darum geht, Reizen einer bestimmten Kategorie zuzuweisen. In beiden Fällen wird gemessen, wie schnell Proband:innen eine sprachliche Antwort auf einen Reiz generieren. Konkret wird hierbei die **Speech Onset Time** gemessen, also die verstrichene Zeit zwischen Präsentation des Reizes und Beginn der Antwort. Eine weitere typische Methode ist der Einsatz von **Distraktoren**, also ablenkende Reize, die zusätzlich präsentiert werden. Hier findet sich ein stabiler Effekt wieder, dass solche Distraktoren typischerweise einen Einfluss auf die Reaktionszeit haben, selbst wenn Versuchspersonen dazu angewiesen sind, diese Distraktoren zu ignorieren.

Experimente zur Sprachverarbeitung haben logischerweise ein umgekehrtes Setting: Versuchspersonen bekommen einen sprachlichen Reiz präsentiert und müssen auf diesen Reiz in irgendeiner Form reagieren. Diese Reaktion erfolgt typischerweise nicht-sprachlich,

um die Prozesse der Sprachverarbeitung und Sprachproduktion auseinanderhalten zu können – typischerweise erfolgt die Reaktion mechanisch, etwa durch das Drücken von Knöpfen. Ein häufiges Paradigma ist hierbei die **lexikalische Entscheidungsaufgabe**, bei der Proband:innen echte Wörter von Pseudo-Wörtern unterscheiden sollen. Die Messung der Reaktionszeit – genau genommen zwischen der Präsentation des Reizes und dem Knopfdruck – erfolgt hierbei meistens automatisiert und ist daher sehr effizient und verlässlich.

## 2.2 Eye-Tracking

Ein etwas aufwändigeres Verfahren zur Untersuchung der Sprachverarbeitung ist das **Eye-Tracking**. Im Wesentlichen wird hierbei mit einer Infrarotkamera die Bewegung bzw. Position der Pupillen gemessen. Kalibriert mit einer unbeweglichen Oberfläche (wie z.B. einem Bildschirm) kann somit akkurat aufgezeichnet werden, wann die Versuchspersonen wohin schauen.

Eye-Tracking kommt logischerweise zum Einsatz, um die visuelle Informationsaufnahme zu untersuchen. Im Kontext von Sprache kann es hierbei darum gehen, wie Menschen einen Text lesen, aber auch, wie sie Personen wahrnehmen, die mit ihnen sprechen (z.B. die Wahrnehmung von Gestik, Mundbewegungen, Gesichtsausdrücken, ...).

Bei der Analyse von Eye-Tracking-Daten fokussiert man sich auf zwei Prozesse. **Fixationen** bedeuten, dass das Auge auf einem Punkt ruht und eine Informationsaufnahme stattfindet. Sie geben uns Aufschlüsse über den zeitlichen Ablauf der sprachlichen Wahrnehmung: Wörter oder Strukturen, die beim Verständnis Probleme bereiten, werden häufiger und/oder länger fixiert. Das Gegenstück hierzu sind **Sakkaden**, die Sprünge zwischen Fixationen bezeichnen. Bei Sakkaden ist vor Allem die Richtung relevant. Bei Texten unterscheidet man entsprechend zwischen progressiven Sakkaden, also Sprüngen in die Leserichtung, und regressiven Sakkaden, Sprüngen gegen die Leserichtung.

## 2.3 Typische Effekte

Bezüglich der Verarbeitung von Wörtern gibt es einige stabile Effekte, die für die beschriebenen experimentellen Setups relevant sind. Sie können entweder bewusst als unabhängige Variable integriert werden, oder können als potenzielle Störvariablen auftreten, für die kontrolliert werden muss.

Ein sehr häufiger und faszinierender Effekt, der gerne gezielt untersucht wird, ist der **Primingeffekt**. Priming besagt, dass das Gehirn Reize schneller verarbeiten kann, wenn vorher bereits ähnliche Reize präsentiert wurden. In der ursprünglichen Version der oben beschriebenen lexikalischen Entscheidungsaufgabe (Meyer & Schvaneveldt, 1971) ging

es tatsächlich darum, genau solche Primingeffekte zu untersuchen – ein Teil der echten Worte wurde in semantisch verwandten Paaren präsentiert (z.B. *nurse* - *doctor*), andere in zufälligen Paaren ohne semantischen Zusammenhang. Die Pseudo-Wörter waren in diesem Experiment lediglich eine Ablenkung – das tatsächliche Interesse der Studie bestand darin, herauszufinden, ob echte Wörter schneller erkannt werden, wenn sie bereits durch ein semantisch verwandtes Wort *geprimet* wurden. Ein solcher Primingeffekt konnte in der Tat nachgewiesen werden.

Primingeffekte finden sich nicht nur in der Semantik, sondern auch auf allen anderen linguistischen Ebenen wieder. Als Beispiele seien hier genannt:

- **semantisches Priming:** *Katze* - *Hund*
- **phonologisches Priming:** *Ball* - *Fall*
- **morphologisches Priming:** *haben* - *hatte*
- **Übersetzungspriming:** *Maus* - *mouse*

Andere stabile Effekte, die tendenziell eher als Störvariablen auftreten, sind **Frequenz-, Längen-** und **Nachbarschaftseffekte**. Hochfrequente Wörter werden schneller erkannt und verarbeitet; längere Wörter benötigen mehr Zeit zur Verarbeitung; und Wörter mit vielen „Nachbarn“ (also semantisch oder phonologisch ähnlichen Wörtern) werden langsamer erkannt. Die Reize *Haus*, *Yacht* und *Atomwaffensperrvertrag* direkt miteinander zu vergleichen, wäre daher nicht ratsam, da sich sich im Bezug auf all diese Effekte stark voneinander unterscheiden.

### 3 Neurolinguistische Experimente

Neurolinguistische Verfahren werden eingesetzt, um Hirnaktivitäten **direkt** zu messen. Hierbei wird zwischen **elektrophysiologischen** und **hämodynamischen Verfahren** entschieden (Mayer & Dogil, 2025).

Elektrophysiologische Verfahren wie die Elektroenzephalographie (EEG) messen direkt das Aktionspotenzial, also die direkte physiologische Entsprechung von Hirnaktivität. Interessant ist hierbei, dass solche Potenziale andere Spannungskurven aufweisen, wenn die Erwartung des Gehirns verletzt wird. Sehen wir uns die beiden folgenden Beispiele an:

- (1) Nina takes her coffee with sugar and **cream**.
- (2) Nina takes her coffee with sugar and **socks**.

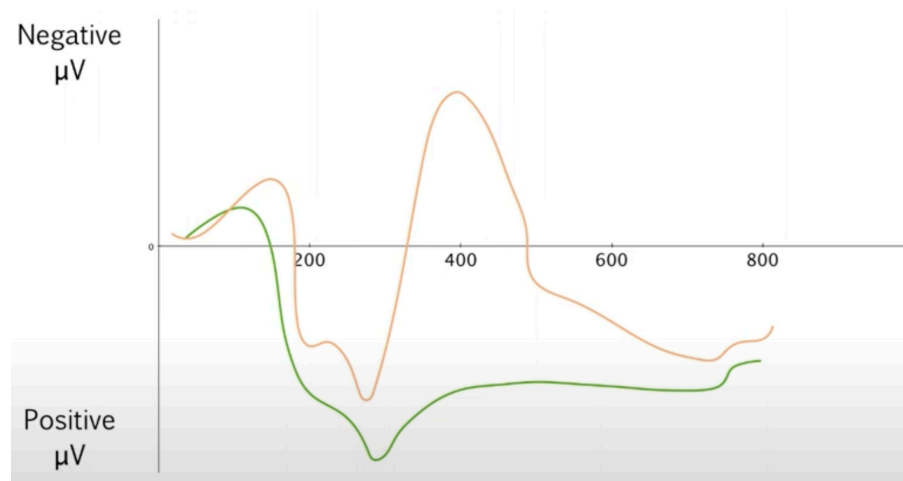


Abbildung 1: Aktionspotenziale bei erwartetem (grün) und unerwartetem (beige) Stimulus.

Im Fokus steht jeweils das letzte Wort des Satzes. Die Spannungskurven der Aktionspotenziale sind in Abb. 1 dargestellt. Die untere, grüne Kurve ist die physiologische Reaktion auf (1) – der Satz wird erwartungsgemäß und sinnvoll beendet; die Kurve zeigt ein normales Aktionspotenzial an. Wird in (2) hingegen das letzte Wort ersetzt, sodass die Aussage keinen Sinn mehr ergibt, wird die Erwartung des Gehirns verletzt – als direkte physiologische Reaktion zeigt sich die obere, beige Kurve. Dieses Verhalten können sich Forschende zu nutzen machen, um etwa Grammatikalität oder Sinnhaftigkeit von Aussagen zu untersuchen. Grundlegend gilt, dass solche elektrophysiologischen Verfahren eine hohe zeitliche Auflösung haben, aber eine geringe räumliche. Sie eignen sich also gut, um den zeitlichen Ablauf der Verarbeitung eines Signals zu untersuchen, aber weniger gut, um genau zu bestimmen, wo ein Signal verarbeitet wird.

Ist tatsächlich die genaue Hirnregion von Interesse, eignen sich hämodynamische Verfahren wie die funktionale Magnetresonanztomographie (fMRT). Diese Verfahren messen nicht die elektrischen Aktionspotenziale selbst, sondern den sekundären Effekt, dass aktivierte Nervenzellen kurzzeitig einen erhöhten Stoffwechsel aufweisen. Dementsprechend sind hämodynamische Verfahren zwar zeitlich weniger genau, haben aber eine deutlich bessere räumliche Auflösung. Daher werden diese Verfahren typischerweise verwendet, um zu erforschen, welche Hirnareale bei welchen Reizen aktiviert werden. Hämodynamische Verfahren spielen auch bei der Diagnose und Erforschung von Sprachstörungen wie Aphasie eine Rolle (Mayer & Dogil, 2025).

## 4 Datenauswertung

Hat man nun durch die Durchführung eines Experimentes seine Forschungsdaten erhoben, folgt als nächster Schritt natürlich die Auswertung dieser Daten. Daten lassen sich grundsätzlich auf drei Ebenen analysieren:

- **Deskriptiv:** Welche (oberflächlichen) Unterschiede oder Muster sind zu erkennen? z.B. *Rauchen Frauen mehr als Männer?*
- **Korrelativ:** Welche Variablen hängen systematisch zusammen? z.B. *Inwiefern unterscheidet sich der Tabakkonsum abhängig von sozialen Faktoren?*
- **Kausal:** Welche Variablen bedingen andere Variablen direkt? z.B. *Inwiefern bedingen psychische Krankheitsbilder den Tabakkonsum?*

Rein deskriptive Analysen haben eine vergleichsweise geringe Aussagekraft. Sie dienen eher der Hypothesenbildung für spätere Analysen. **Korrelationen** lassen sich durch verschiedene statistische Verfahren berechnen, wobei Programmiersprachen wie R oder Python viele praktische Module hierzu liefern. Korrelationen gelten als signifikant, wenn sie einen niedrigen p-Wert erreichen. Hier wird häufig  $p < 0.05$  als Schwellwert angesetzt; allerdings ist dieser Wert relativ willkürlich gewählt (McElreath, 2015). Vereinfacht gesagt stellt der p-Wert eine Wahrscheinlichkeit dar, dass die vorliegenden Daten durch die Nullhypothese zufällig generiert werden können, also unter der Annahme, dass die beteiligten Variablen komplett unabhängig voneinander sind.

Durch die Korrelation von Variablen sollen **kausale Zusammenhänge** ergründet werden. Allerdings ist hier ein extrem wichtiger, statistischer Grundsatz zu beachten: **Korrelation impliziert keine Kausalität!**

Die Korrelation ist eine deutlich schwächere Beziehung als die Kausalität. Sie beschreibt lediglich, dass zwei Variablen (über den Zufall hinaus) häufig miteinander auftreten. Daraus ergibt sich auch, dass Korrelation zwangsläufig *bilateral* ist: Wenn X mit Y korreliert, korreliert entsprechend auch Y mit X.

Die Korrelation alleine begründet noch keinen kausalen Zusammenhang. Der Beweis für einen solchen muss darüber hinaus erbracht werden. Nehmen wir als Beispiel die folgende Beobachtung: *Wenn die Sonne scheint, trinken die Menschen mehr Bier*. Der Bierkonsum korreliert also mit dem Sonnenschein (und andersherum korreliert auch der Sonnenschein mit dem Bierkonsum). Die Kausalität – wenn es denn eine gibt – kann allerdings nur in eine Richtung erfolgen: Die Menschen trinken mehr Bier, *weil* die Sonne scheint. Andersherum funktioniert das nicht: Die Sonne wird nicht anfangen zu scheinen, weil viele Leute Bier trinken. Kausalität ist also zwangsläufig eine *unilaterale*, also eine einseitige Beziehung.

Besonders relevant ist diese Unterscheidung, wenn die kausalen Zusammenhänge komplexer werden. Fügen wir unserem simplen Beispiel also noch eine dritte Variable hinzu: Wenn die Sonne scheint, bekommen die Menschen auch mehr Sonnenbrände. Der kausale Zusammenhang liegt hier auf der Hand. Da aber sowohl der Sonnenbrand, als auch der Bierkonsum kausal vom Sonnenschein abhängt, korrelieren auch Bierkonsum und Sonnenbrand miteinander! Hier wäre es fatal, einen direkten kausalen Zusammenhang hinter der Korrelation zu vermuten: Weder bekommt man einen Sonnenbrand, weil man Bier

trinkt, noch trinkt man Bier, weil man einen Sonnenbrand bekommt.

Dieses Beispiel ist bewusst so gewählt, dass die kausalen Zusammenhänge auf der Hand liegen, um den Unterschied zwischen Korrelation und Kausalität zu verdeutlichen. Allerdings sind diese kausalen Zusammenhänge in der „echten Welt“ meistens nicht so klar. Korrelationen lassen sich leicht finden, allerdings muss man gründlich hinterfragen, ob es tatsächlich auch einen direkten kausalen Zusammenhang gibt (der ja oft auch suggeriert wird).

Experimente stellen tatsächlich noch die beste Möglichkeit dar, potenzielle kausale Zusammenhänge systematisch zu untersuchen, da wir ja Variablen gut kontrollieren und systematisch manipulieren können. In anderen Settings müssen solche Analysen *post-hoc* erfolgen, also anhand von bereits erhobenen, unveränderlichen Daten. Dennoch muss auch in experimentellen Verfahren eine Korrelation nicht zwangsläufig eine Kausalität implizieren.

## 5 Zusammenfassung

In der vergangenen und dieser Sitzung haben wir eine Methoden kennengelernt, um wissenschaftliche Daten für die linguistische Forschung zu erheben.

**Befragungen** sind hierbei die kostengünstigste Art der Datenerhebung, insbesondere in Form von Fragebögen, die von Gewährspersonen selbstständig ausgefüllt werden können. Sie liefern vergleichbare Daten, bei denen jedoch Details und relevante Zusatzinformationen fehlen können.

**Beobachtungen** sind zeitintensivere Datenerhebung, die reichhaltige Daten von wenigen Personen generieren. Durch die große Bandbreite, aber relativ geringe Vergleichbarkeit der Daten eignet sich diese Erhebungsform vor allem für qualitative Analysen.

**Experimente** generieren sehr genaue und vergleichbare Daten, sind in ihrer Umsetzung allerdings ziemlich aufwändig und teuer. Wie wir gesehen haben, sind die erhobenen Datenpunkte zumeist nicht sprachlich, sondern sprachliche „Begleiterscheinungen“. Experimente beleuchten daher meistens die kognitiven oder physiologischen Prozesse, die mit Sprachverarbeitung und -produktion zusammenhängen.

## Literatur

- Kleineberg, D. & Henneke, I. (2023). Empirische Linguistik. In R. Klabunde & W. Mihatsch (Hrsg.), *Linguistik*. Berlin & Heidelberg: J.B. Metzler.
- Mayer, J. & Dogil, G. (2025). *Sprache und Gehirn: Ein neurolinguistisches Tutorial*. Stuttgart: Universität Stuttgart. Zugriff auf <https://www2.ims.uni-stuttgart.de/sgtutorial/index.html>
- McElreath, R. (2015). *Statistical rethinking: A Bayesian course with examples in R and STAN*. Chapman and Hall / CRC.
- Meyer, D. & Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90 (2), 227-234.
- Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18 (6), 643-662.