

## Kapitel 8: Webergänzung

### Experimentformen der Linguistik und Psycholinguistik

Bei psycholinguistischen Experimenten werden üblicherweise Reaktionszeiten und Fehlerquoten verglichen. Die zu prüfenden Hypothesen sind also fast immer der Art, dass gesagt wird, Faktor A sorgt für längere Reaktionszeiten (und/oder mehr Fehler) als Faktor B oder umgekehrt. Bei psycholinguistischen Experimenten gibt es zwar auch eine Vielzahl von möglicherweise intervenierenden Variablen, aber sie sind doch relativ leicht zu kontrollieren, indem man mögliche einflussnehmende Faktoren jeweils unter der „A“- und der „B“-Bedingung gleich hält. Ein ganz einfaches Beispiel für ein solches Experiment wäre, dass wir annehmen, bei der Sprachverarbeitung werde auf Morpheme zurückgegriffen. Diese Hypothese würde die Voraussage machen, dass das Erkennen mehrmorphemiger Wörter länger dauert als das Erkennen von Simplicia.<sup>1</sup> Eine Möglichkeit des Überprüfens dieser Hypothese wäre, dass wir Reaktionszeiten beim Erkennen von Wörtern messen, die aus mehreren Morphemen bestehen, und beim Erkennen von gleich langen Wörtern derselben Buchstaben- und Silbenanzahl, die nur aus einem Morphem bestehen (z.B. *Eisbär/Gibbon*). Nach unserer Hypothese müsste also die Reaktionszeit bei *Eisbär* länger sein, denn beim Erkennen müsste man auf zwei Morpheme zurückgreifen, als die bei *Gibbon*, denn das besteht nur aus einem Morphem. Dies gilt natürlich nur für den Fall, dass alle anderen Bedingungen, die die Reaktionszeit beeinflussen können, konstant gehalten werden. In diesem Fall wäre zwar die Länge in Buchstaben und die Silbenzahl gleich, jedoch ist *Eisbär* das häufiger gebrauchte Wort, das deswegen schneller erkannt werden müsste. Weil es nie gelingen wird, die zu kontrastierenden Elemente in jeder Hinsicht gleich zu halten, müssten diese Effekte in einem Experiment mit ca. 40 Wörtern in jeder der Gruppen so ausgeglichen werden, dass manchmal das einmorphemige und manchmal das mehrmorphemige Wort das häufiger gebrauchte ist und dass für die beiden Gruppen jeweils die durchschnittliche Frequenz konstant gehalten wird. Im also den Unterschied zwischen *Eisbär* und *Gibbon* auszugleichen, wird man z.B. auch ein Paar wie *Schwein* und *Seehund*, bei dem das einmorphemige Wort frequenter ist, aufnehmen. Die Häufigkeit von Wörtern ermittelt man über Häufigkeitswörterbücher oder -datenbanken; die gebräuchlichste dieser Datenbanken ist CELEX.<sup>2</sup>

Es ist kaum möglich, in einer Einführung alle Arten von linguistischen Experimenten vorzustellen. Immerhin wollen wir einen kurzen Überblick über einige verbreitete Techniken geben.

#### 1 Verfahren und Erhebungsinstrumente in Experimenten zur Sprachproduktion

Wenn es um Sprachproduktion geht, sind u.a. folgende Verfahren möglich:

**Texte/Sätze laut lesen lassen:** Dieses Verfahren hat nur Sinn, wenn kaum mehr als die Aussprache, Satzintonation u.ä. interessiert. Man kann Texte vorgeben, in denen die Aussprachephänomene enthalten sind, die einen interessieren, man muss jedoch bedenken, dass keineswegs sicher ist, dass die Sprecher beim freien Sprechen genauso artikulieren würden. Beim Vorlesen gibt es allerhand hyperkorrekte Aussprachen, die beim freien Sprechen nicht vorkommen.

---

<sup>1</sup> Das Erkennen von Wörtern misst man meist in einer Experimentform, die „lexical decision“ bzw. „lexikalische Entscheidungsaufgabe“ genannt wird. Dabei sitzen die Versuchsteilnehmer vor einem Bildschirm, auf dem in zufälliger Abfolge tatsächliche Wörter und bedeutungslose Buchstabenfolgen, die von der phonologischen Form her Wörter der betreffenden Sprache sein könnten (*glompf* z.B. für das Deutsche), erscheinen. Sie müssen dann immer so schnell wie möglich auf eine Taste für „ja“ und eine andere für „nein“ drücken, wenn sie etwas als Wort oder Nichtwort (Pseudowort) erkannt haben.

<sup>2</sup> Zum psycholinguistischen Experimentieren brauchen Sie neben den an Ihrer Universität zugänglichen Möglichkeiten zur Frequenzermittlung der benutzten Wörter auch meist ein Programm, das am Computer das Stimulusmaterial präsentiert und gleichzeitig die Reaktionszeiten misst, etwa NES<sup>2</sup> oder ERTS. Dabei sind Sie auf das Programm angewiesen, das an Ihrer Universität vorhanden ist, und für die vorhandenen Programme wird es Anleitungen geben. Insofern haben wir darauf verzichtet, hier eine genauere Einführung zu geben.

**Einzelwörter vom Computerbildschirm ablesen lassen:** Hier besteht die Möglichkeit, Reaktionszeiten zu messen, d.h. für verschiedene Gruppen von Wörtern zu überprüfen, wie lange es dauert, bis die Versuchsteilnehmer anfangen das Wort auszusprechen, oder wie viele Fehler sie beim Lesen machen.

**Sätze/Äußerungen vervollständigen lassen:** Bei diesem Verfahren bekommt man eine einigermaßen freie Sprachproduktion, aber man kann trotzdem bis zu einem gewissen Grad durch die Vorgabe steuern, was produziert wird. Die Möglichkeiten sind sehr variabel. Man kann Äußerungen so vorgeben, dass nur noch eine einzige Antwort möglich ist, z.B. *Fritz bringt mich mit seinen dummen Witzen auf die .....*, aber auch so, dass völlig verschiedene Reaktionen möglich sind, z.B. *Wenn meine Eltern anderer Meinung sind als ich, dann .....*

Häufig wird das Verfahren eingesetzt, wenn man Fehler oder Reaktionszeiten bei bestimmten grammatischen Formen messen will. Das kann man z.B. so arrangieren, dass man einen Satz in Einzelwörtern auf dem Computerbildschirm zeigt und dann das letzte Wort in einer Zitierform vorgibt und bittet, es in der im Satz korrekten Form auszusprechen. Dabei wird die Zeit gemessen, die die Versuchsteilnehmer brauchen, bis sie ihre Reaktion beginnen, und eventuell auch die Anzahl der Fehler, die sie dabei machen. So kann man z.B. Unterschiede in der Reaktionszeit oder in der Fehleranzahl bei regelmäßiger im Vergleich zu unregelmäßiger Flexion feststellen.

Bei dieser Experimentform kann man auch mit maskierten Primes arbeiten. Damit ist gemeint, dass man einen Reiz kürzer als 60 Millisekunden einblendet, sodass er nicht bewusst wahrgenommen, aber trotzdem verarbeitet wird. So kann man zum Beispiel die Einflüsse einer Sprache auf die andere bei Mehrsprachigen überprüfen. Unmittelbar bevor das Wort in der gewünschten Sprache ausgesprochen werden soll, wird ein ähnliches Wort aus der anderen Sprache eingeblendet. Dadurch wird sichergestellt, dass dieses Wort aus der anderen Sprache auch aktiviert ist. Sein Einfluss muss sich dann in Reaktionszeiten oder Fehleranzahlen nachweisen lassen.

**Imitieren lassen:** Man kann ganz verschiedene Imitationsaufgaben stellen; das geht von „einzelne Laute oder Wörter nachsprechen lassen“ bis „Sätze/Texte reproduzieren“. Meist wird dieses Verfahren angewendet, wenn es darum geht, festzustellen, welche (oft auch unerwarteten) Merkmale der zu reproduzierenden Einheiten die Versuchsteilnehmer bemerken, ob sie also z.B. einen Laut korrekt nachsprechen oder ihn an einen Laut ihrer L1 angleichen. Eine besondere Form dieser gelenkten Sprachproduktion nennt man *shadowing*. Dabei werden die Versuchsteilnehmer gebeten, alles, was sie hören, so schnell wie möglich nachzusprechen. Dabei kann man überprüfen, inwieweit sie fehlerhaft präsentierte Äußerungen „reparieren“, d.h. also z.B. nachsprechen: „Ich begrüße Sie“, obwohl ihnen „Ich beglücke Sie“ vorgesprochen wurde. Man kann dabei beobachten, dass die Versuchsteilnehmer ganz offensichtlich nicht papageienhaft reproduzieren, was ihnen vorgesprochen wurde, sondern dabei ihr Sprachverarbeitungssystem benutzen. Als unabhängige Variable kann man z.B. die Position des Fehlers im Wort/in der Äußerung variieren und überprüfen, wie sich das auf die Anzahl der Reparaturen auswirkt.

**Wortassoziationen nennen oder aufschreiben:** Hier werden Versuchsteilnehmer gebeten, so schnell wie möglich und ohne Reflexion die Wörter zu nennen/aufzuschreiben, die ihnen zu einem Stimuluswort einfallen. Dabei wird Zeitdruck erzeugt, um ein bewusstes Aussuchen der genannten Wörter möglichst auszuschließen.

Von diesem Verfahren hat man zunächst sogar angenommen, es könne dazu dienen, die Struktur des mentalen Lexikons wiederzugeben. Es zeigt sich tatsächlich bei dieser Art von Experimenten, dass bei vielen Wörtern eine hohe Übereinstimmung zwischen Sprechern besteht in Bezug darauf, was sie als erste Assoziation nennen (z.B. zu *Tisch Stuhl*). Die als erste Assoziation genannten Wörter stehen oft in einer Oberbegriff/Unterbegriff- oder Schwesterbegriff-Relation zu dem Stimulusbegriff. Gele-

**Ergänzendes Internetmaterial zu: Albert, R., & Marx, N. (2025). *Empirisches Arbeiten in Linguistik und Sprachlehrforschung* (4., überarbeitete Auflage). Narr Studienbücher. Narr Francke Attempto.**

gentlich werden auch typische Fortsetzungen (*Hund – bellen*) oder Kollokationen (*Maßnahmen – ergreifen*) genannt. Bei späteren Assoziationen werden sehr häufig Elemente von vorgestellten Situationen genannt (*König – Königin – Schloss – Schlosscafé – Erdbeereis mit Sahne*). Spätestens bei der letzten genannten Assoziation muss man bezweifeln, dass tatsächlich die Struktur des mentalen Lexikons durch diese Assoziationen gezeigt wird.

Assoziationsexperimente wurden auch häufig mit Mehrsprachigen vorgenommen, teils um herauszufinden, welche Einflüsse vorgegebene Sprachwechsel haben oder bei welchen Gelegenheiten spontane Sprachwechsel auftreten, teils um herauszufinden, ob übersetzungsäquivalente Wörter auch übersetzungsäquivalente Assoziationen hervorrufen.

Man hat die Tatsache, dass die Assoziationen der Sprecher derselben Sprache recht stark übereinstimmen, auch benutzt, um für einzelne Sprachen sogenannte „Assoziationsnormen“ zusammenzustellen; darin kann man nachschlagen, welche Wörter besonders häufig zusammen genannt wurden (u.a. in: Hasselhorn/Hager 1994. *Handbuch deutschsprachiger Wortnormen*. Göttingen: Hogrefe). Diese Listen von Assoziationsnormen werden bei Experimenten benutzt, bei denen „assoziierte Begriffe“ in Kontrast zu anderen, z.B. phonologisch oder semantisch ähnlichen, gesetzt werden.

**Bilder benennen:** Das Benennen von Zeichnungen oder Fotos ist eine klassische Sprachproduktionsaufgabe, die man am Computer durchführen lassen muss, wenn man die Reaktionszeiten messen will, die man aber auch einfach anhand von vorgelegten Bildern durchführen kann, wenn es darum geht, überhaupt eine Äußerung zu produzieren. Der Nachteil des Verfahrens ist, dass man nur Dinge verwenden kann, die eindeutig zu zeichnen oder zu fotografieren sind, und das sind im Wesentlichen konkrete Substantive.

**Bilder beschreiben / Geschichten erzählen anhand von Bildern / Unterschiede zwischen zwei Bildern benennen:** Mit diesen Aufgaben kann man auch längere Texte erzeugen, allerdings per Computer nur die Reaktionszeit bis zur Produktion des ersten Wortes messen. Gerade das Benennen von Unterschieden zwischen zwei Bildern ist geeignet, auch bestimmte grammatische Formen zu elizitieren. Wenn man z.B. bei zwei geeigneten Bildern fragt „Was ist zwischendurch passiert?“, kann man Partizipien II elizitieren („Das Haus ist eingestürzt“).

**Nacherzählungen (anhand von gehörten/gelesenen Erzählungen, Videos u.a.):** Auch hier kann man seine Vorgaben so wählen, dass bestimmte Wörter, bestimmte grammatische Phänomene vorkommen müssten, die man in der freien Sprachproduktion erhalten will. Bei derartigen Vorgaben ist es ganz besonders wichtig, mit mehreren Personen aus dem Bekanntenkreis auszuprobieren, ob die gewünschten Phänomene tatsächlich eliziert werden. Man kann sich gar nicht vorstellen, wie viele Möglichkeiten Versuchsteilnehmer finden, um die gewünschten sprachlichen Phänomene zu vermeiden.

**Stroop-Tests:** Dieses Verfahren gibt über die Sprache eine andere Information als über ein gleichzeitig gegebenes Bild, damit untersucht werden kann, wie sich die damit verbundene Störung des Bildbenennens auf die Reaktionszeiten und/oder die Fehleranzahl auswirkt. In der einfachsten Form des Tests werden die Versuchsteilnehmer z.B. aufgefordert, Tintenfarben zu benennen von geschriebenen, damit nicht übereinstimmenden Farbbezeichnungen, also z.B. steht dort *rot* mit grüner Tinte geschrieben. Dabei kann man z.B. die Sprachen variieren und feststellen, dass die inkongruenten Farbbezeichnungen um so mehr stören, je präsenter die Sprache dem Versuchsteilnehmer ist. Ein anderes gebräuchliches Stroop-Verfahren blendet in zu benennende Bilder Wörter als Schrift ein, wobei man z.B. variieren kann, ob die eingeblendeten Wörter dem als Bild dargestellten Wort semantisch oder phonologisch ähnlich sind.

Ergänzendes Internetmaterial zu: Albert, R., & Marx, N. (2025). *Empirisches Arbeiten in Linguistik und Sprachlehrforschung* (4., überarbeitete Auflage). Narr Studienbücher. Narr Francke Attempto.

**Kunstwörter morphologisch verändern:** Hierzu gehört das bereits erwähnte *Wug*, mit dem die Pluralformen erhoben wurden. Den Versuchsteilnehmern wird eine Zeichnung dieses Tiers vorgelegt und gesagt: „Hier ist ein *Wug*.“ Dann kommt ein zweites Kärtchen mit noch einem solchen Tier, und es wird gesagt: „Und da kommt noch eins. Jetzt haben wir hier zwei .....“.

## 2 Verfahren und Erhebungsinstrumente in Experimenten zur Sprachrezeption

Wenn es um Sprachrezeption geht, ist die Anzahl der Experimentformen noch größer.

**Lexikalische Entscheidungstests:** Am verbreitetsten sind die bereits erwähnten lexikalischen Entscheidungsaufgaben, wobei ein wichtiges Verfahren das sogenannte *Priming* (dt. Bahnung; meist wird aber der englische Terminus verwendet) ist. Das Verfahren misst Einflüsse eines vorher gezeigten Reizes auf die Reaktionszeit für den gemessenen Reiz und schließt daraus auf Verbindungen in unserer mentalen Repräsentation der Sprache. Z.B. wird die Reaktionszeit für das Erkennen von *König* als einem Wort des Deutschen verkürzt, wenn vorher *Krone* (ein assoziativ verbundenes Wort), *Königin* (ein morphologisch verbundenes Wort) oder *king* (ein über ein gemeinsames Konzept verbundenes Wort aus einer anderen Sprache) gezeigt wird. So kann man mit dem Priming-Verfahren Hypothesen über bestehende Verbindungen testen.

**Phonemmonitoring:** Ein weiteres verbreitetes Verfahren ist, dass man den Versuchsteilnehmern die Aufgabe stellt, jedes Mal einen Knopf zu drücken, wenn sie einen bestimmten Laut hören (*phoneme monitoring*). Abhängig davon, wie komplex das vorher Präsentierte war, oder von der Umgebung, in der dieser Laut präsentiert wurde, verändert sich die Reaktionszeit. Für dieses Verfahren gibt es auch eine Papier- und Bleistift-Variante, die unkompliziert zu handhaben ist und sich also auch für linguistische Hausarbeiten eignet: Man lässt seine Versuchsteilnehmer unter Zeitdruck jeweils einen bestimmten Buchstaben in einem schriftlich präsentierten Text markieren. Dabei wird nicht die Reaktionszeit gemessen, sondern die Zahl der Fehler, d.h. der nicht markierten Buchstaben. In bestimmten syntaktischen oder morphologischen Umgebungen werden nämlich mehr Fehler dieser Art gemacht als in anderen, d.h. man schließt aus der Zahl der gemachten Fehler auf die Schwierigkeit der Verarbeitung der jeweiligen Textstelle.

**Gating:** Ein anderes Verfahren präsentiert unterschiedlich lange Segmente von zu erkennendem Sprachmaterial und überprüft, ab wann es richtig ergänzt wird, also eindeutig erkannt wurde. So wird zum Beispiel das Wort *Elefant* auditiv wie folgt in Teilen präsentiert, wobei die Versuchspersonen nach jedem Teil angeben müssen, welches Wort gemeint ist:

E  
El  
Ele  
Elef  
Elefa  
Elefan  
Elefant

Dieses Verfahren ist zum Beispiel benutzt worden, um festzustellen, wo der „Erkennungspunkt“ eines Wortes liegt, d.h. wann das Wort als „Elefant“ erkannt wird. Das wäre hier wohl, wenn man das „f“ gehört hat, denn bis dahin wäre auch noch die Fortsetzung „Element“ möglich. Wenn das Wort „Elefant“ ganz präsentiert ist, sind natürlich noch (längere) flektierte oder abgeleitete Wörter mit *Elefant* als Bestandteil möglich, z.B. „Elefanten“, „elefantös“.

**Segmentverschiebung:** Wenn es um komplexe Wörter geht, wird häufig mit einem Verfahren gearbeitet, bei dem einzelne Elemente aus einem Wort in ein anderes verschoben werden sollen (*segment shifting*), also soll etwa zu einem auf dem Computerbildschirm mit Markierung der ersten Silbe präsentierten **GEBURT** ein Wort gebildet werden, das das ebenfalls auf dem Bildschirm präsentierte **BET** enthält, wobei die markierte Silbe zu verschieben ist (die Versuchsteilnehmer sollen also *Gebet* sagen). Man kann dabei z.B. überprüfen, ob Reaktionszeiten verschieden sind, wenn es sich bei dem zu verschiebenden Element um ein Morphem oder eine in diesem Kontext nicht bedeutungstragende Silbe handelt.

**Fehlererkennung:** Weitere Verfahren messen das Erkennen von Fehlern (die Zeit, die dazu benötigt wird, oder – auch das ist wieder ohne komplizierte Computerprogramme möglich – die Anzahl der übersehenen Fehler in zu variierenden Kontexten), wobei verschiedene Möglichkeiten vorgegeben werden können, unter denen die richtige auszusuchen ist, oder wobei einfach fehlerhafte und fehlerlose Sätze präsentiert werden und die Versuchsteilnehmer unter Zeitdruck die Fehler finden müssen.<sup>3</sup>

### 3 Auswertung der Experimentergebnisse

Wenn wir unser Experiment durchgeführt haben, stehen wir vor dem Problem, was wir mit den gesammelten Daten anfangen. Zunächst einmal geht es darum, welche Daten wir überhaupt verwenden können.

Versuchsteilnehmer, für die wir nicht für jede der Bedingungen Ergebnisse haben, nehmen wir aus der Auswertung heraus. Das geschieht recht häufig, wenn wir im realen Unterricht experimentieren. Dann kann es sein, dass manche Schüler z.B. nur bei der einen erprobten Methode da waren, aber bei der anderen gefehlt haben. Deren Daten nimmt man oft ganz aus der Auswertung heraus, wenn man ein Messwiederholungsdesign hatte.<sup>4</sup>

Bei vielen psycholinguistischen Experimenten ist es üblich, die Reaktionszeiten von Personen ganz aus der Auswertung herauszunehmen, die mehr als 10 % Fehler gemacht haben, das hängt jedoch auch von der Art des Experiments ab. Auf jeden Fall nimmt man die Reaktionszeiten für fehlerhafte Reaktionen aus der Auswertung heraus. Die übrig gebliebenen Daten werden dann mit den in den nächsten Kapiteln beschriebenen statistischen Verfahren aufbereitet.

### Aufgabe

Stellen Sie sich vor, Sie wollen experimentell überprüfen, ob die Verarbeitung der deutschen Plurale von Substantiven auf -s (sie gelten nach manchen Abhandlungen als die regelmäßige Pluralbildung) schneller geht als die Verarbeitung anderer Plurale. Sie haben sich dazu ein Experiment ausgedacht, in dem Versuchspersonen so schnell wie möglich die Singularform eines Substantivs nennen müssen, wenn der Computer das entsprechende Wort im Plural zeigt (also bei „Autos“ müssen sie „Auto“ sagen, bei „Mütter“ „Mutter“ usw.). Der Computer misst die Reaktionszeiten. Mit welchen intervenierenden Variablen müssen Sie rechnen, und wie können Sie sie ausschalten?

---

<sup>3</sup> Der Zeitdruck, der in den o.a. Experimentbeschreibungen fast immer auftaucht, dient dazu, die für die Sprachverarbeitung schwierigen Aufgaben von den für die Sprachverarbeitung leichteren Aufgaben zu trennen. Mit genügend Zeit werden natürlich alle Fehler gefunden, aber dann hätte man keine auswertbaren Daten mehr.

<sup>4</sup> Es gibt aber auch Möglichkeiten, mit solchen „fehlenden Werten“ umzugehen, z.B. indem man eine so genannte „multiple Imputation“ durchführt. Wie dies zu machen ist, erfragen Sie am besten bei einem Statistikexperten.

## Aufgabenlösung

Eine deutlich intervenierende Variable ist die Worthäufigkeit. Bei jeder Benennaufgabe sind Versuchspersonen deutlich schneller bei häufig vorkommenden Wörtern. Sie müssten also darauf achten, dass die Wörter in allen Gruppen dieselbe durchschnittliche Häufigkeit haben (anhand von Korpora zur Vorkommenshäufigkeit, z.B. der CELEX-Datenbank).

Eine weitere intervenierende Variable ist die Wortlänge. Je länger das Wort ist, umso länger dauert das Lesen (abhängig von der Zahl der Buchstaben) und dauert die Ausspracheplanung (wohl abhängig von der Zahl der Silben). Auch Silbenzahl und Anzahl der Buchstaben müssten also über die Gruppen konstant gehalten werden.

Zu befürchten ist auch, dass der Grad der vorzunehmenden Veränderungen eine Rolle spielt. Insofern wäre es weise, Fälle mit *m*-lauten getrennt zu betrachten, denn beim -s-Plural gibt es keine *m*-laute.

Ein weiteres Problem ist, dass Sie gar nicht wissen, ob die Wörter mit dem -s-Suffix von den Versuchspersonen als Plural oder als Genitiv wahrgenommen werden (*des Autos*), das könnten Sie nur dadurch lösen, dass Sie sich in allen Gruppen auf Feminina beschränken, die im Genitiv kein Suffix haben (*der Oma, der Sahne*).

Und als letztes Problem ergibt sich, dass für die meisten Wörter mit anderen Pluralen von ihrer morphologischen Struktur her mehrere Plurale infrage kommen, bei den Wörtern mit -s-Plural (Abkürzungswörter, Wörter, die auf Vollvokal enden) jedoch nur einer. Dieses Problem haben Sie vielleicht umschifft, indem Sie nicht den Plural, sondern den Singular bilden lassen, aber es ist natürlich ungeklärt, was es ausmacht, ob die Beziehung zwischen Singular- und Pluralform eindeutig ist oder nicht. Insofern sollten Sie diese ungeklärte Frage im Forschungsbericht erwähnen.