

Mit welchem Wort soll ich's benennen?

Eine lexikalische und korpusbasierte Analyse deutscher Verben



Matthias Spörrle, Kerstin Barth, Beatrice Gerber-Braun,
Jennifer Stich, Friedrich Försterling

Ludwig-Maximilians-Universität München

Kurzzusammenfassung

Verben beschreiben entweder sachbezogene Vorgänge (z.B. anpflanzen) oder interpersonale Ereignisse (z.B. loben). Diese interpersonalen Verben können in Handlungs- (z. B. schlagen) und Zustandsverben (z.B. bewundern) unterteilt werden. Sowohl für Handlungs- als auch für Zustandsverben kann unterschieden werden, ob sie Kausalattributionen auf das Subjekt (z.B. schlagen, ängstigen) oder auf das Objekt (bestrafen, bewundern) der Interaktion auslösen. Bisherige sprachpsychologische Untersuchungen haben überwiegend identische Vokabeln aus einem kleinen Ausschnitt des Gesamtvokabulars verwendet. Auch fehlt eine umfassende Systematisierung deutscher Verben hinsichtlich der dargestellten Klassifikation. Um erstmals Häufigkeitsangaben der Verbverteilungen der gesamten Sprache zu erhalten, wird eine Analyse unterschiedlicher deutscher Verb- und Textkorpora durchgeführt, aus der ein Datensatz mit insgesamt über 10 000 Verben resultiert. Häufigkeitsanalysen belegen eine Dominanz sachbezogener im Vergleich zu interpersonalen Verben und innerhalb interpersonalen Verben eine deutliche Überzahl von Handlungsverben. Der erstellte Korpus erlaubt sprachlich repräsentative Analysen interpersonalen Verben, Aussagen zur sprachlichen Repräsentation interpersonalen Ereignisse und attributionstheoretische Analysen sprachlicher Daten.

1. Einleitung

Ein Verb bezeichnet eine Handlung (ein Geschehen oder einen Vorgang wie beispielsweise „trinken“) oder einen Zustand (wie beispielsweise „triefen“). Dementsprechend bezeichnen interpersonale Verben Geschehnisse (schlagen) oder Zustände (überraschen) zwischen Personen. Sprachpsychologische Forschung konnte aufzeigen, dass hierbei vom Rezipienten verbsspezifisch tendenziell eher das Subjekt oder das Objekt für diese Interaktion kausal verantwortlich gemacht wird.

Beispiel:

Sandra quält Lisa, weil sie böse ist.

Frage: Wer ist böse?

Alternativfrage: An wem liegt es, dass dies (das Quälen) passiert?

Überwiegende Antwort: Sandra (Subjekt des Satzes)

Sandra verurteilt Lisa, weil sie böse ist.

Frage: Wer ist böse?

Alternativfrage: An wem liegt es, dass dies (das Verurteilen) passiert?

Überwiegende Antwort: Lisa (Objekt des Satzes)

Somit kann zwischen Zustands- und Handlungsverben differenziert werden, die entweder Attributionen auf das Subjekt oder das Objekt nahelegen. Hieraus resultiert ein Schema mit $2 \times 2 = 4$ Verbklassen (siehe Tabelle 1).

Tabelle 1: Die vier Verbklassen interpersonalen Verben

Verbart	Zustandsverben		Handlungsverben	
Umschreibung	fühlen, erfahren		machen, tun	
Kriterium	Mentale Interaktion		Interaktives Verhalten	
Verbklasse	SE Stimulus- Experiencer	ES Experiencer- Stimulus	AP Agent- Patient	AE Agent- Evocator
Attribution	Subjekt	Objekt	Subjekt	Objekt
Beispiel	überraschen	bewundern	schlagen	beweinen

Die sprachpsychologische Forschung verwendet zur Analyse dieser impliziten Kausalität in Sprache überwiegend eine weitgehend unveränderte kleine Auswahl (ca. 250) interpersonalen Verben. Ein zentrales Ziel dieser Arbeit war es daher, einen möglichst umfassenden Textkorpus an deutschsprachigen Verben bereitzustellen. Ein solcher Korpus ermöglicht nicht nur die Untersuchung sprachpsychologischer Phänomene auf einer breiteren Stimulusgrundlage als bisher, sondern erlaubt auch die sprachliche Analyse von bestehenden Textdokumenten (z.B. zur Untersuchung der möglichen Intentionen des Urhebers).

Zudem versucht die vorliegende Arbeit mittels integrierter lexikalischer und textkorpusbasierter Analysen erstmals, empirisch fundierte Aussagen zu den folgenden in der Literatur diskutierten Fragen zu treffen:

- Wie viele interpersonale Verben existieren in der deutschen Sprache absolut und in Relation zu nicht-interpersonalen Verben?
- Wie viele Verben gehören zu den jeweiligen Verbklassen?
- Wie häufig werden die jeweiligen Verben in der deutschen Sprache verwendet?
- Führen ein lexikalischer und ein textkorpusbasierter Forschungsansatz zu vergleichbaren Ergebnissen?

Es sei an dieser Stelle darauf hingewiesen, dass die Erhebung noch nicht abgeschlossen ist und die Ergebnisse noch vorläufigen Charakter besitzen.

Wir danken an dieser Stelle herzlich:

Herrn Dipl.- Inf. Christian Biemann vom Institut für Informatik an der Universität Leipzig
Herrn Dipl.-Ing. Peter Hinterseer vom Lehrstuhl für Kommunikationsnetze an der Technischen Universität München
für ihre freundliche Unterstützung bei der Datengewinnung.

2. Methode

Insgesamt wurden 6 Verbquellen in der Untersuchung berücksichtigt:

- Ein Lexikon deutscher Verben aus dem Hueber Verlag [etwa 4200 Verben] (Griesbach & Uhlig, 2000)
- Ein Lexikon deutscher Verben aus dem Langenscheidt Verlag [etwa 2600 Verben] (Wendt & Thurmair, 1999)
- Eine im Internet verfügbare Verbliste aus einer Textverarbeitungssoftware [etwa 5800 Verben] (Schulz, 1994)
- Das Valenzlexikon deutscher Verben des Instituts für deutsche Sprache [etwa 600 Verben] (Engel & Schumacher, 2003)
- Eine Liste interpersonalen Verben aus einer Habilitationsschrift zum Thema Verbkausalität [etwa 1000 Verben] (Rudolph, 1997)
- Eine vom Institut für Informatik (Abteilung: Automatische Sprachverarbeitung) in Leipzig übersandte Liste deutscher Verben [etwa 5300 Verben] (Universität Leipzig, 2004)

Für jedes dabei gefundene Verb wurde codiert, in wie vielen dieser Quellen es verzeichnet ist. Zudem wurden für jedes Verb zwei Häufigkeitsindices (die absolute Häufigkeit und die Häufigkeitsklasse) aus dem Wortschatz Leipzig übernommen. Jedes Verb wurde als nicht-interpersonales oder interpersonales Verb klassifiziert, sowie - falls es sich um ein interpersonales Verb handelte - einer der vier Verbklassen zugeordnet.

3. Ergebnisse

Aus den 6 Verbquellen ergaben sich insgesamt 10176 deutsche Verben. Anzahl interpersonalen Verben: Insgesamt können 3761 (37%) als interpersonale klassifiziert werden ($p < .001$, $g = .13$) und damit deutlich mehr als die in der Literatur vermuteten 2000 (Hoffman & Tchir, 1990).

Verbklassen interpersonalen Verben: Es werden 3479 Verben (92,5%) als Handlungsverben und 282 Verben (7,5%) als Zustandsverben klassifiziert. Für die vier Verbklassen ergeben sich folgende Verteilungen ($\chi^2(3, N=3761) = 3871.8$, $p < .001$):

AE: 26,7 % (Schätzung in der Literatur: 12%)

AP: 65,8 % (Schätzung in der Literatur: 50%)

ES: 4,8 % (Schätzung in der Literatur: 18%)

SE: 2,7 % (Schätzung in der Literatur: 15%)

Interkorrelationen der Häufigkeitsindices: Die bei beiden Häufigkeitsindices sowie die Häufigkeit des Auftretens in den 6 unterschiedlichen Quellen korrelieren signifikant untereinander ($.42 < r_s < .74$, $p < .001$) und belegen damit, dass die Auftretenshäufigkeit in verschiedenen Lexika und Systematisierungen mit der Auftretenshäufigkeit in einem Sprachkorpus zusammenhängt.

Analyse der Auftretenshäufigkeiten: Mann-Whitney-U-Tests belegen für alle drei Häufigkeitsindices, dass ein durchschnittliches Handlungsverb signifikant ($p < .001$) seltener verwendet wird als ein durchschnittliches Zustandsverb. So wird beispielsweise ein durchschnittliches Handlungsverb 4700 mal im Textkorpus erfasst, ein Zustandsverb hingegen etwa 11000 mal. Die insgesamt gesehen selteneren Zustandsverben konzentrieren somit eine häufigere durchschnittliche Verwendungshäufigkeit auf sich.

Kruskal-Wallis-Rangvarianzanalyse belegen signifikante ($p < .001$)

Unterschiede zwischen den vier Verbklassen für alle vier Häufigkeitsindices:

AE: 2400 mal im Textkorpus vorhanden

AP: 5600 mal im Textkorpus vorhanden

ES: 14000 mal im Textkorpus vorhanden

SE: 7800 mal im Textkorpus vorhanden

4. Diskussion und Fazit

Die lexikalische und korpusbasierte Analyse deutscher Verben erweist sich als fruchtbar und sinnvoll:

- Korrelationen belegen den Zusammenhang beider Datenspeicher.
- Auf einer Makroebene können über lexikalische Analysen Schätzungen zur Anzahl der Repräsentanten einzelner Verbklassen gegeben werden.
- Auf einer Mikroebene liefert der Korpus Hinweise zur Verwendungshäufigkeit dieser einzelnen Verben.
- In der Sprache überwiegen nicht-interpersonale Verben.
- Bei interpersonalen Verben wiederum überwiegen Handlungsverben deutlich gegenüber Zustandsverben.
- Ein Zustandsverb wird dementsprechend durchschnittlich häufiger verwendet.

Engel, U. & Schumacher, H. (2003). *Kleines Valenzlexikon deutscher Verben*. Mannheim: Institut für Deutsche Sprache,

Griesbach, H. & Uhlig, G. (2000). *4000 deutsche Verben und ihr Gebrauch*. Ismaning: Max Huber.

Hoffman, C. & Tchir, M. A. (1990). Interpersonal verbs and dispositional adjectives: The psychology of causality embodied in language. *Journal of Personality and Social Psychology*, 58, 765-778.

Rudolph, U. (1997). *Zur impliziten Kausalität in Sprache*. Unveröffentlichte Habilitationsschrift, Ludwig-Maximilians-Universität München.

Schulz, M. (1994). *Deutsches Wörterbuch für Ispell* [nur: verben.txt]. Online verfügbar: <http://unipen.nici.kun.nl/scrawls/dictionaries/udi/german>.

Universität Leipzig (2004). *Projekt deutscher Wortschatz*. Online verfügbar: <http://wortschatz.uni-leipzig.de>.

Wendt, H. F. & Thurmair, M. (1999). *Verb-Tabellen*. München: Langenscheidt.