

PSYCHOLOGIE | SPRACHANALYSE

Du bist, was du sprichst

Die Statistik lügt nicht, meint der Psychologe James W. Pennebaker. Mit seinem Textanalyseprogramm will er die Persönlichkeit von Schreibern enttarnen – ein höchst umstrittener Ansatz.

VON JAN DÖNGES

VERRÄTERISCHE UNTERHALTUNG
Sprache vermittelt nicht nur Inhalte, sondern offenbart auch etwas über unsere Wesensart.



AUF EINEN BLICK

Sag mir, wer du bist

1 Das Sprachanalyseprogramm LIWC (Linguistic Inquiry and Word Count) des Psychologen James W. Pennebaker soll aus beliebigen Texten Aussagen über den Charakter des Verfassers ableiten können.

2 Laut Pennebaker sind nicht die Inhalte eines Textes der Schlüssel zur Persönlichkeit des Autors, sondern unscheinbare Worte wie Pronomen, Artikel oder Konjunktionen. Diese wurden nämlich weniger bewusst verwendet.

3 Befürworter computergestützter Verfahren halten diese Form der Textauswertung für objektiver als menschliche Gutachter. Kritiker bemängeln, dass sie größere inhaltliche Zusammenhänge vernachlässigen.

Die Arbeiten des Psychologen James W. Pennebaker spalten die Zunft der Sprachwissenschaftler – denn er schert sich keinen Deut um die klassischen Forschungsansätze der Linguisten. Seit knapp zwei Jahrzehnten sucht Pennebaker in Texten nach der Persönlichkeit ihrer Autoren, ohne auch nur einmal die Syntax zu analysieren oder Bedeutungsnuancen zu erfassen – er liest das Geschriebene nicht einmal. Der Psychologe verlegt sich stattdessen aufs Zählen von Wörtern.

Dass sich die Persönlichkeit eines Menschen auch in seiner Sprache niederschlägt, ist unumstritten. Schriftsteller suchen seit eh und je nach der passenden Ausdrucksweise für ihre fiktiven Charaktere. Sigmund Freud (1856–1939) sah in Versprechern das Unterbewusstsein zu Tage treten. Wer ein ausgeprägtes Sprachgefühl hat, meint zwischen so manchen Zeilen auch die Person des Verfassers zu errahnen. Aber mittels blinder Statistik die Oberfläche eines Texts vermessen, um daraus ein Charakterprofil des Autors zu gewinnen – geht das?

Sein wichtigstes Instrument entwickelte der seit 1997 an der University of Texas in Austin forschende Sozialpsychologe Mitte der 1990er Jahre: LIWC (Linguistic Inquiry and Word Count), »Luke« ausgesprochen. Das computergestützte Auszählverfahren durchforstete seitdem ungeheure Textmassen. Seine jüngste Analyse präsentierte der 58-jährige Pennebaker während der US-Präsidentschaftswahl 2008: Auf der Website www.wordwatchers.wordpress.com veröffentlichte er ein sprachliches Psychogramm der Kandidaten. Sowohl in den Debatten als auch in Interviews zeigten sich klare Tendenzen: John McCain sprach seine Wähler sehr direkt und persönlich an, seine Wortwahl erwies sich als emotionsgeladen und impulsiv. Barack Obama hingegen verwendete häufig Kausalbeziehungen, was auf komplexe Gedankengänge hindeute. Tendenziell äußerte er sich auch vager als sein republikanischer Konkurrent (siehe Kasten S. 27).

Zur maschinellen Textverarbeitung kam Pennebaker wie die Jungfrau zum Kind. In den frühen 1990er Jahren galt sein Hauptinteresse dem so genannten expressiven Schreiben. Bei dieser Methode lernen traumatisierte Personen, das Erlebte durch Tagebucheinträge und das Verfassen von Essays zu verarbeiten (siehe G&G 4/2007, S. 53). Der Psychologe wollte zeigen, dass nicht nur die Reflexion dessen, was die Betroffenen zu Papier bringen, sondern bereits die Aufzeichnungen selbst Fortschritte der Patienten anzeigen. Um dies zu beweisen, arbeiteten

sich Doktoranden durch Hunderte von Aufsätzen mit größtenteils belastendem Inhalt. »Die emotionale Inanspruchnahme war enorm«, erinnert sich Pennebaker, die Bewertungen des Geschriebenen wurden unzuverlässig. »Ich suchte einen objektiveren Weg, die Texte zu analysieren. Deshalb haben wir LIWC entwickelt.«

Die Software sucht nach Worten, die Signalwirkung haben: Wie optimistisch ist der Verfasser? Beschäftigen ihn Gedanken an die Vergangenheit oder dominieren Zukunftsthemen? Verwendet er Ausdrücke, die auf komplexe Argumentationen schließen lassen? All das leistet – neben dem eigentlichen Zählprogramm – ein Wörterbuch mit derzeit 4500 Einträgen, aufgeteilt in mehr als 70 verschiedene Wortfelder. Gezählt wird, wie oft Wörter der einzelnen Kategorien auftreten.

Von A wie Affekt bis Z wie Zahlwort

»Einsamkeit«, »weinen« oder »glücklich« fallen beispielsweise unter »Emotionswörter«, die wiederum in Untergruppen für positive und negative Affekte aufgeteilt sind. In diesen Subkategorien werden zudem Begriffe für einzelne Gefühle wie Angst oder Trauer genauer unterschieden. Andere Kategorien sind unspezifische »Wahrnehmungswörter« (»sehen«, »hören«, »fühlen«) oder der Komplex »kognitive Prozesse« (»müsste«, »denken« oder »weil«). Hinzu kommen 22 allgemeine Kategorien zur Wortart wie Pronomen, Hilfsverben, Artikel, Zahlwörter oder Satzverknüpfungen – so genannte Funktionswörter.

Die elektronische Sprachanalyse ist keine Erfindung Pennebakers. Um die mühsame Auswertung per Hand abzukürzen, erfand der Harvard-Psychologe Philip Stone (1936–2006) bereits im Jahr 1966 das erste computergestützte Zählverfahren: den »General Inquirer«. Fast parallel dazu trieben 1969 Louis Gottschalk von der University of California in Irvine sowie seine Kollegin Goldine Gleser (1915–2004) die Entwicklung eines vergleichbaren Programms voran. Mittels der nach ihnen benannten Gottschalk-Gleser-Skalen wollten die Psychologen automatisch einordnen, wie sehr ein bestimmter Affekt, etwa Angst oder Aggressivität, die Mitteilungen eines Patienten prägt.

1987 sorgte eine Veröffentlichung Gottschalks für Aufsehen: Er behauptete, mit seinem Verfahren Hinweise dafür gefunden zu haben, dass der damals amtierende US-Präsident Ronald Reagan (1911–2004) ab ungefähr 1980 an mentalen Beeinträchtigungen gelitten habe.



ENFANT TERRIBLE
DER SPRACHFORSCHUNG:
James W. Pennebaker – Erfinder
des Textanalyseprogramms
LIWC (Linguistic Inquiry and
Word Count)

Erst 1994 wurde öffentlich, dass Reagan an Alzheimer-Demenz erkrankt war.

Richtig populär wurden derartige Methoden mit Beginn der 1990er Jahre, als ausreichend leistungsfähige Computer erschwinglich wurden. Textverarbeitungssoftware und schließlich das Internet sorgten außerdem für immer mehr Datenmaterial in digitalisierter Form. Die Zahl der Analysetools schoss in die Höhe.

Dabei steht normalerweise im Fokus, was gesagt wird – die Inhaltsebene der Sprache. »Es ist von der Psychologie praktisch unbemerkt geblieben, dass vor allem die ›kleinen‹ Wörter wie ›ich‹, ›wir‹, ›ein‹ oder ›alle‹ Rückschlüsse auf das Verhalten zulassen«, kommentiert Pennebaker. Substantive, Verben und Adjektive mögen die Bedeutung eines Satzes tragen, sie hingen aber auch stark von Zufällen ab – etwa vom vorgegebenen Aufsatzthema.

Das Wie hingegen offenbart Pennebaker zufolge die Persönlichkeit. Der Stil eines Textes, der sich unter anderem in den »kleinen« Wörtern ausdrückt, sei einer bewussten Manipulation weniger zugänglich. Wer sich in ein schmeichelhafteres Licht rücken will, wählt die passenden Ausdrücke, kontrolliert aber kaum die Anzahl seiner Artikel oder Pronomen, so der Psychologe.

Die Sprache der Terroristen

Das Prinzip wird in Pennebakers »Al-Kaida-Studie« deutlich: Er analysierte 2007 zusammen mit einigen Kollegen im Auftrag der amerikanischen Bundespolizei (FBI) insgesamt 58 Texte von Osama Bin Laden und Aiman Al-Zawahiri, der als Nummer zwei in der Hierarchie der Terrororganisation gilt. Als Vergleichsmaßstab zogen sie 17 Proben anderer extremistischer Pamphlete heran.

Drei der LIWC-Kategorien stachen besonders heraus: Erstens verwendeten Bin Laden und Zawahiri signifikant mehr Emotionswörter, darunter überdurchschnittlich viele feindselige. Zweitens traten Verben selten in ihrer Vergangenheitsform auf. Die Wissenschaftler deuteten dies als vermindertes Interesse an zurückliegenden Ereignissen. Drittens fiel den Forschern

insbesondere der hohe Anteil von Pronomen der 3. Person Plural (»sie«, »ihnen«, et cetera) auf. Al-Kaida definierte sich stärker als andere Terroristengruppen über die Abgrenzung zu einer Gegnerschaft, so die Schlussfolgerung – nach dem Motto: »Wir gegen die im Westen.« Schon im Vorfeld der Studie hatten die Forscher LIWC-Befunde mit Einschätzungen unabhängiger Beobachter abgeglichen und ermittelt, dass sich der Grad des Extremismus einer Gruppe am besten auf Grund der Pronomen der 3. Person Plural vorhersagen lasse.

Was Pronomen alles verraten können, zeigte sich auch im Vergleich der beiden Islamisten selbst. So verdreifachte sich zwischen 2004 und 2006 die Häufigkeit, mit der Al-Zawahiri das Wort »ich« gebrauchte, während sie bei Bin Laden etwa konstant blieb. »Ein hoher Prozentsatz von Ich-Wörtern deutet üblicherweise auf starke Selbstbezogenheit, aber auch auf Unsicherheit und eine Verteidigungshaltung hin. Der konkrete inhaltliche Kontext bestätigt das im Großen und Ganzen auch«, so Pennebaker.

Andere Studien hatten gezeigt, dass Wörter, die für Abwägungen oder differenzierte Aussagen verwendet werden (»außer«, »aber«, »ausgenommen« et cetera), mit einer höheren kognitiven Komplexität, mit besseren Schulnoten und sogar mit der wahrheitsgemäßen Wiedergabe eines Sachverhalts zusammenhängen. Für Bin Laden ergab die Zählung, dass die Komplexität der Gedankengänge in seinen Texten im Lauf der Jahre ein hohes Niveau erreichte, während sie bei seinem Vize auf einer niedrigen Stufe stagnierte.

So plausibel Pennebakers Auslegungen erscheinen, ein bisschen erinnert seine Methode doch an Kaffeesatzleserei. Beispielsweise fallen in den Texten Bin Ladens 1,49 Prozent aller Wörter in die Kategorie »Hass und Zorn«, in den zur Kontrolle zufällig ausgewählten Vergleichstexten sind es 0,89 Prozent. Rein statistisch gesehen ist der Unterschied signifikant – aber wie aussagekräftig ist das Resultat wirklich?

»Diese Abweichungen scheinen vielleicht gering, sind es aber nicht«, sagt Markus Wolf von der Universität Heidelberg. Mit einem Team erarbeitete der Psychologe eine deutsche Übersetzung des LIWC-Wörterbuchs, das in Zukunft bei einer e-mail-gestützten Therapienachsorge zum Einsatz kommen soll. »Um höhere Abweichungen zu erzielen, müsste der Textumfang größer sein oder aber die menschliche Ausdrucksform weniger komplex. Wenn ein Patient beispielsweise darüber spricht, dass er traurig ist, wieder-

PRO

»Vor allem die ›kleinen‹ Wörter wie ›ich‹, ›wir‹, ›ein‹ oder ›alle‹ lassen Rückschlüsse auf die Persönlichkeit zu«

James W. Pennebaker, University of Texas in Austin

KONTRA

»Es sind gerade die seltenen Sprachmerkmale, die am meisten offenbaren«

Brigitte Boothe, Universität Zürich

holt er das Wort ›traurig‹ in der Regel nicht sehr oft, sondern beschreibt seinen Zustand anders. Deshalb sind bereits scheinbar kleine Gewichtungen relevant«, erklärt der Forscher.

Die eigentliche Schwachstelle von Pennebakers Methode liegt in der Interpretation der Daten. Was bedeutet eine Häufung von Pronomen? Ist jemand, der viele Verneinungen benutzt (»ich tue das nicht ungerne«), tatsächlich vorsichtiger oder besonnener, als jemand der Zusammenhänge direkt formuliert (»das ist mir lieber«)? Pennebaker und Wolf sind sich einig: Methoden wie das LIWC seien nur erste Annäherungen – die Deutung der Befunde stecke noch in den Kinderschuhen.

Methodenvergleich

Vergleiche mit etablierten Messmethoden, die die Persönlichkeit eines Probanden durch Befragung erkunden, sollen die Auslegungen zusätzlich absichern. Markus Wolf hat die Aussagekraft des deutschen LIWC außerdem mit anderen elektronischen Verfahren verglichen – unter

anderem mit dem »Dresdner Angstwörterbuch« des Psychologen Hendrik Berth von der Technischen Universität Dresden. Es orientiert sich stark an den Gottschalk-Gleser-Skalen. »Die Übereinstimmung war überraschend gut«, sagt Wolf. Auch die Einschätzungen menschlicher Gutachter weiche nicht nennenswert davon ab. »Und das«, ergänzt der Psychologe, »obwohl das LIWC weder Doppeldeutigkeiten noch Verneinungen berücksichtigt.«

Tatsächlich stoßen sich viele Kritiker daran, dass Pennebakers Software keinen Unterschied dazwischen macht, ob jemand »ich bin traurig« oder »ich bin nicht traurig« sagt – beide Aussagen enthalten ein Wort der Kategorie »negative Emotion«. Die Befürworter des Verfahrens gehen jedoch davon aus, dass sich solche Fälle statistisch »herausmitteln«, wenn die Textsammlung groß genug ist.

»Es sind aber gerade die kleinen Hinweise, die seltenen Ereignisse, die am meisten offenbaren«, sagt dagegen Brigitte Boothe, die an der Universität Zürich die Sprache von Patienten in

»Ich« und der Tod

Verraten sich Selbstmordgedanken bereits in der Sprache eines Gedichts? James W. Pennebaker und seine Kollegin Shannon Stirmann von der University of Pennsylvania in Philadelphia untersuchten rund 300 Werke von neun suizidalen und neun nicht suizidalen amerikanischen Lyrikern – mit eindeutigen Ergebnis: Jene, die sich später das Leben genommen hatten, verwendeten fast doppelt so oft die Wörter »ich«, »meiner«, »mir« und »mich«. Vergleichsweise wenig nutzten sie dagegen »Wir«-Pronomen. Auch Begriffe für kommunikative Tätigkeiten gebrauchten sie in ihrem literarischen Schaffen seltener.

Die Wortwahl lasse auf eine übertriebene Selbstbezogenheit und gesellschaftliche Isolation schließen, meint Pennebaker, was zu Theorien über die Auslöser suizidalen Verhaltens passe. Überraschend war aus Sicht der Forscher, dass beide Gruppen in etwa gleich oft von negativen Gefühlen schrieben.

Obamaish versus McCainese: Sprache im US-Wahlkampf

Die elektronische Analyse von Interviews, Debatten und Reden von Barack Obama und John McCain während der Präsidentschaftskandidatur ergab klare Unterschiede in der Ausdrucksweise. Wörter folgender Kategorien traten jeweils vermehrt auf:

McCain: emotional, direkt, zukunftsorientiert	Obama: reserviert und überlegt
häufige Verwendung von: 1. Person Singular (I, me, my) Futur (will, must) Konjunktionen (and, but, also) verbindender Wortsinn (all, both) negative Affekte (sad) positive Emotionen (enjoy, happy)	häufige Verwendung von: Artikel (a, the) Verneinungen (never, no) ausschließender Wortsinn (without, except) Angst und Wut (worry, fight) Vorsicht (probably, careful) Wahrnehmung (look, listen) Bewegung (run, carry) Richtung (around, eastern) Zeit (age, beginning) Tod (casualty, grievance)

QUELLE: WORDWATCHERS.WORDPRESS.COM

Wie LIWC Texte untersucht

als ich an die uni gekommen bin, ging es los mit diesem gefühlsmischmasch: aufregung, angst, neugier und so. vor allem war ich so froh, dass ich endlich aus der schule rauskam. ich hab die schule gehasst! im ersten moment war das irgendwie für mich wie eine flucht in ein neues leben. aber ich hatte natürlich noch gar keine ahnung, was das für ein leben sein sollte. da war so viel neues und unbekanntes. und ehrlich gesagt kam ich anfangs überhaupt nicht mit meinen kommilitonen klar. da hab ich mir gedacht, geht das jetzt genauso wieder weiter oder wie!?

Die Farben markieren beispielhaft eine Auswahl von Kategorien, nach denen LIWC das Protokoll einer Therapiesitzung aufgliedert (Doppeldeutigkeiten haben wir nicht berücksichtigt):

1. Person Singular Präpositionen Artikel Konjunktionen Emotionswörter

LITERATURTIPPS

Pennebaker, J.W., Chung, C.:

Computerized Text Analysis of Al-Qaeda Transcripts. In: K. Krippendorff & M. Bock (Hg.): A Content Analysis Reader. Thousand Oaks, Sage 2007.

Pennebaker, J.W., Stirman, S.:

Word Use in the Poetry of Suicidal and Nonsuicidal Poets. In: Psychosomatic Medicine 63, S. 517–522, 2001.

Wolf, M. et al.: Computerergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count. In: Diagnostica 54, S. 85–98, 2008.

WEBLINK

www.liwc.net

James W. Pennebaker und Kollegen widmeten ihrem Sprachanalyseprogramm LIWC (Linguistic Inquiry and Word Count) eine Homepage, auf der man kostenlos eigene Texte analysieren lassen kann.

der Psychotherapie erforscht. Eine computer-gestützte Auswertung der Daten sieht sie daher kritisch. »Wenn in einem Text gehäuft ›wir‹ auftaucht – heißt das, der Patient ist sozial gut eingebunden? Oder ist er schüchtern und sucht in der Gruppe nach ›Geleitschutz?‹ Genau solche Fragen seien wichtig, wolle der Therapeut einen Zugang zu seinem Patienten finden.

Analyse nach dem Google-Prinzip

»Wir haben versucht, unsere Einteilungen so gut wie möglich mit den Einschätzungen menschlicher Gutachter abzugleichen. Aber auch diese sind nicht unfehlbar«, hält Pennebaker seinen Kritikern entgegen. »Letztlich ist es wie bei den Wirtschaftswissenschaftlern«, findet der Sprachforscher: »Wir stellen immer nur begründete Vermutungen über eine vielschichtige Materie an. Für eine standardisierte Interpretation ist es noch zu früh.«

Einen Ausweg könnten ausgefeiltere Methoden bieten. So entwickelten Computerlinguisten in den vergangenen Jahren Programme, die selbstständig relevante Informationen über Stil und Inhalt sammeln. Sie funktionieren nach dem »Google-Prinzip«: Anstatt dem Computer ein Kategoriensystem vorzugeben, erlernen zum Beispiel so genannte Faktorenanalyse-Verfahren im Alleingang, welche Gruppierungen eine Rolle spielen. In unüberschaubaren Textmassen fahnden sie nach gehäuft auftretenden Wörtern und Wortgruppen. So lässt sich eine etwaige Voreingenommenheit der Forscher weitgehend ausschalten.

Zusammen mit seinem Kollegen Sherlock Campbell von der Yale University in Connecti-

cut berechnete Pennebaker 2003 per Latent Semantic Analysis ein Maß für die Ähnlichkeit der Tagebucheinträge von Traumapatienten aus drei verschiedenen Studien der letzten Jahre. Pennebaker und Cambell fragten sich, welche Textmerkmale sich bei jenen Patienten verändert hatten, die auf dem Weg der Besserung waren und nur noch selten ihren Arzt konsultierten. Wieder zeigte sich, dass inhaltliche Komponenten keine Rolle spielten: Der Faktor, der am deutlichsten mit dem Genesungsfortschritt in Zusammenhang stand, war – die Verwendung von Pronomen. Probanden, deren Aufschriebe von Mal zu Mal Veränderungen in der Erzählperspektive zeigten, ließen sich in der Folgezeit am seltensten ärztlich behandeln. Möglicherweise bewältigt ein Patient traumatische Erlebnisse schneller, wenn er seine Situation aus der eigenen und aus der Sicht anderer beschreibt.

Der Erfolg solcher Studien spricht dafür, alles dem Computer zu überlassen, doch Pennebaker mahnt zur Vorsicht: Fehlen inhaltliche Vorgaben wie bestimmte linguistische Kategorien, berechnen die neuen »Google-Programme« unverständliche Ergebnisse: »Für uns ist beispielsweise ›1. Person Singular‹ ein gut nachvollziehbares Konstrukt: ›mich‹, ›meine‹, ›mir‹ gehören in dieselbe Reihe. Für den Computer fällt aber auch ›wir‹ in diese Kategorie, obwohl es die Mehrzahl anzeigt.«

Daher hält auch Markus Wolf altbewährte Analyseprogramme wie LIWC derzeit für die besten Verfahren, um reproduzierbare Effekte zu erzielen – diese positive Beurteilung bleibt unter Experten jedoch nach wie vor umstritten: Ein Ansatz, der den jeweiligen Kontext eines Wort-, Satz- oder Textzusammenhangs unbeachtet lässt, erscheint vielen Sprachforschern suspekt.

Versöhnlich stimmt die Sichtweise von Roderrick Hart vom College of Communication der University of Texas in Austin: Wenn zwei Leute eine Stadt kennen lernen wollen, besteigt der eine vielleicht einen Hubschrauber – der andere geht zu Fuß. Bei der Sprachanalyse sei es ähnlich: Wer mit Programmen wie LIWC viele Texte »überfliegt«, mag Dinge verpassen, die beim persönlichen Lesen auffallen – im Gegenzug aber erhält er einen unschätzbaren Ausblick auf das gesamte Panorama. ~

Jan Dönges ist Linguist und Wissenschaftsjournalist in Heidelberg.

 www.gehirn-und-geist.de/audio