

NOCH EINMAL: DIE GOOGLE-BUCHSUCHE

von Dominik Brückner

In Heft 3/2009 des SPRACHREPORTs erschien ein Beitrag mit dem Titel „Die Google Buchsuche als Hilfsmittel für die Lexikographie“¹, der jenseits der vieldiskutierten juristischen und kulturpolitischen Fragen den möglichen Nutzen der Google-Buchsuche für Lexikologen und Lexikographen in den Fokus rückte. In den drei Jahren, die seither vergangen sind, hat sich einerseits die Google-Buchsuche weiterentwickelt, andererseits hat die intensive Arbeit damit neue Erkenntnisse zutage gefördert.² Beides soll im Folgenden in Form einer Ergänzung des älteren Beitrags vorgestellt werden.

Dabei wird im Sinne einer Recherche mit Hilfe der Google-Buchsuche vorgegangen: Zunächst werden Sucheingabe und Ergebnisliste ins Blickfeld gerückt, vor allem unter den Aspekten der Ergebniszahlen und des Zugriffs auf einzelne Textstellen. In einem zweiten Abschnitt werden dann einzelne Probleme diskutiert, die im Zusammenhang mit in „vollständiger Ansicht“, in „eingeschränkter Vorschau“ und in „Auszügen“ vorhandener Bücher³ auftreten können. Zum Schluss wird, wie schon im ersten Beitrag zu diesem Thema, versucht, den Nutzen der Google-Buchsuche für Lexikographie und Lexikologie im Lichte dieser Erkenntnisse erneut kurz zu umreißen.

Sucheingabe und Ergebnislisten

Eingabe mehrerer Suchbegriffe

Bei jeder Google-Buchsuche muss bekanntlich berücksichtigt werden, dass hier eine Trunkierung, wie sie bei vergleichbaren Programmen heutzutage Standard ist, nicht vorgesehen ist. Daher muss z. B. jede Flexionsform oder Schreibvariante im Feld „mit **irgendeinem** der Wörter“ eigens eingegeben werden. Zu beachten ist dabei, dass die Eingabe auf 32 Formen beschränkt ist – was zunächst durchaus komfortabel erscheint, sich aber sehr schnell als zu limitiert erweisen kann, z. B. für denjenigen, der ein Verb auf *-ieren* in Texten des 18. Jahrhunderts, also mit der Schreibvariante *-iren* (und womöglich noch einigen anderen) suchen möchte. Dabei konnte beobachtet werden, dass

Google für jede Eingabe andere Ergebnislisten zusammenstellt, es also sein kann, dass man bei einer Eingabe von drei Formen Textstellen findet, die bei der Eingabe von zehn Formen nicht gefunden werden, selbst wenn die ursprünglichen drei Formen in diesen zehn enthalten sind.

Dies muss allerdings nicht immer so sein: Gesucht wurde am 12.2.2011 im Feld „mit **irgendeinem** der Wörter“ *Illuminierung Illuminirung* zwischen 1760 und 1780. Die Ergebnisliste umfasste sechs Einträge. In einem zweiten Schritt wurde im gleichen Feld und im gleichen Zeitraum nach *Jlluminierung Jlluminirung* gesucht, die Ergebnisliste umfasste erneut neun Einträge. In einem dritten Schritt wurden dann mit identischen Vorgaben *Jlluminierung Jlluminirung Illuminierung Illuminirung* gesucht, Google gab diesmal fünfzehn Ergebnisse aus. In diesen fünfzehn Ergebnissen waren genau die sechs Ergebnisse für *Illuminierung Illuminirung* sowie genau die neun Ergebnisse für *Jlluminierung Jlluminirung* enthalten. Dass unterschiedliche Lesungen (mit <I> bzw. <J>) auf verschiedene Scans des gleichen Buches (z. B. aus verschiedenen Bibliotheken) zurückzuführen sind, ist dabei eine naheliegende Vermutung, aber nicht zweifelsfrei zu klären. Aufgrund der Unsicherheit über die in den Ergebnislisten ausgegebenen Textstellen bedeutet das Ergebnis aber für den Benutzer, dass er in solchen Fällen immer sämtliche Varianten (in diesem Fall die <I>- und die <J>-Schreibung) suchen muss – am besten separat, wenn es auf Genauigkeit ankommt, was im Umkehrschluss bedeutet, dass die Möglichkeit, 32 Formen in das Suchfeld einzutragen, nicht allzu viel wert ist.

Wer sucht da eigentlich?

Diese Einschätzung muss umso mehr Gültigkeit beanspruchen, seit Google eigenwillig zusätzlich Formen mit ausgibt, die überhaupt nicht gefunden werden sollten: So fördert eine Suche nach *Trunkierungen* auch Ergebnisse für *Trunkierung, trunkiert, trunkierte und trunkierten* zutage, eine Suche nach *swer* erbringt Ergebnisse für *schwer*, und eine Suche nach *Imperatorin* erbringt Ergebnisse für *Imperator, in* – ohne

dass darum gebeten wurde und, gravierender, ohne dass der Benutzer die Möglichkeit hätte, diese Funktion abzustellen.⁴ Zwar kann man dann mit Hilfe der Tastenkombination Strg + F die Ergebnislisten nach *Trunkierungen*, *swer* oder *Imperatorin* durchsuchen, auch bietet Google eine Funktion „In den Ergebnissen suchen“ an, aber die erste Möglichkeit ist umständlich und die zweite reproduziert nur die Probleme, mit denen eine Google-Buchsuche ohnehin schon verbunden ist.

Diese Einstellung wurde mittlerweile zwar wieder zurückgenommen. Reste sind im System allerdings immer noch zu finden. Es ist nicht einzusehen, warum es nicht möglich sein sollte, beide Suchmethoden optional anzubieten.

Die zeitliche Eingrenzung und die Funktion „nach Datum sortiert“

Ein neues, auf den ersten Blick äußerst hilfreiches Tool ist die Möglichkeit, eine erzeugte Ergebnisliste chronologisch ordnen zu lassen. Die Funktion „nach Datum sortiert“ ermöglicht es, auf diese Weise schneller und gezielter Zugriff auf Textstellen aus bestimmten Zeiträumen zu nehmen.⁵

Mit diesem Tool ist allerdings eine recht ungewöhnliche Erscheinung verbunden: Aus unerfindlichen Gründen nämlich fügt die Google-Buchsuche im Zuge der chronologischen Sortierung der bereits generierten Ergebnisliste weitere Einträge hinzu, schätzungsweise etwa 20% bis 30%. Die Gründe hierfür sind zwar unklar, das Phänomen ist deshalb aber nicht weniger hilfreich: Es lohnt sich, die Ergebnisliste nach Datum sortieren zu lassen, auch wenn eine chronologische Sortierung nicht benötigt wird, da dies eventuell genau die Textstelle zutage fördert, nach der man ansonsten lange vergeblich gesucht hätte.

Man kann allerdings auch das umgekehrte Phänomen beobachten:

Eine Suche nach *Illuminatur Illuminaturen Illuminatur Illuminaturen* im Suchfeld „mit **irgendeinem** der Wörter“ mit der Vorgabe „zwischen 1880 und 1890“ ergab am 13. 11. 2010 eine Ergebnisliste, die nach Angaben der Google-Buchsuche „Ungefähr 689 Ergebnisse“ aufwies. Durch die chronologische Sortierung reduzierte die Google-Buchsuche diese 689 Ergebnisse auf null.

Eine genauere Überprüfung erbrachte das Resultat, dass die chronologische Sortierung offenbar nicht in allen Fällen funktioniert:

Der Sucheintrag *Imperialismus Imperialismen* im Feld „mit **irgendeinem** der Wörter“ mit der Vorgabe *Frantz* im Feld „**Autor** Bücher von folgendem Autor“ ergab am 1. 4. 2011 nach chronologischer Sortierung eine Ergebnisliste, die folgendermaßen sortiert war:

1873, 1882, 1902, 1921, 1924, 1925, 1967, 1859, 1862, 1970, 1859, 2005, 2009 [sic]

Die Liste ließ sich auch durch mehrmaliges Umsortieren nicht in eine chronologische Reihenfolge bringen. Bei der überschaubaren Menge von 13 Ergebnissen dieser Suche mag das Problem nicht gravierend sein, sieht man sich aber einer mehrere hundert Ergebnisse umfassenden Liste gegenüber, kann man die Konsistenz der Sortierung nicht ohne Weiteres prüfen. Verlässlich ist diese Funktion also nicht – und das umso weniger, wenn man Fälle wie die folgenden kennt:

Der Sucheintrag *Imperativismus Imperativismen* im Feld „mit **irgendeinem** der Wörter“ mit der Vorgabe „vor 1980“ erbrachte am 15. 11. 2010 eine Ergebnisliste, die 13 Einträge umfasste. Eine chronologische Sortierung der Liste mit Hilfe des Tools „nach Datum sortiert“ generierte weitere fünf, also insgesamt 18 Ergebnisse, was wenig überraschend war, da dieser Effekt zu diesem Zeitpunkt bereits bekannt war. Als frühestes Ergebnis wurde in dieser chronologisch sortierten Liste eine Textstelle aus dem Jahr 1970 angegeben.

Nun wurde die gleiche Suche erneut vorgenommen, allerdings mit dem Unterschied, dass jetzt nur Textstellen vor 1900 ausgegeben werden sollten. Aufgrund der Erfahrungen aus der ersten Suche war allerdings zu erwarten, dass diese Suche kein einziges Ergebnis generieren würde. Tatsächlich gab die Google-Buchsuche aber fünf neue Ergebnisse aus, der früheste Beleg stammte diesmal aus dem Jahr 1876.

Welche Textstellen zwischen 1900 und 1970 hatte das System unterdrückt? Und warum?

Um dies zu überprüfen, wurde nun in einem nächsten Schritt erneut im Feld „mit **irgendeinem** der Wörter“ nach *Imperativismus Imperativismen* gesucht, diesmal willkürlich zwischen 1890 und 1950. Diese Suche erbrachte (bereits chronologisch sortiert) 62 Ergebnisse, zwei Tage später (bereits chronologisch sortiert) 60 Ergebnisse und noch einmal eine $\frac{3}{4}$ Stunde später (bereits chronologisch sortiert) 59 Ergebnisse.⁶

In einem letzten Schritt wurde nun *Imperativismus Imperativismen* im gesamten Zeitraum zwischen 1891 und 1969 gesucht. Dies ergab zunächst 47, nach erfolgter chronologischer Sortierung 76 Ergebnisse. Das

Resultat: Google unterschlug bei der ersten Suche (vor 1980) insgesamt mindestens 81 Ergebnisse, einschließlich des „Erstbelegs“. (Suchen vor 1876 ergaben keine weiteren Ergebnisse – zumindest nicht an diesem Tag.)

Wie viele Ergebnisse gibt es wirklich?

Auch wenn die Google-Buchsuche am Beginn einer jeden Ergebnisliste eine ungefähre Ergebnisanzahl angibt („Ungefähr [...] Ergebnisse“), ist es aus mehreren, sehr verschiedenen Gründen unmöglich, sich auf diese Ergebnisanzahlen zu verlassen:

Das hat zunächst einige objektive Gründe – wenn Texte mehrfach vorhanden sind, ist es fraglich, was als ein Ergebnis zu gelten hat: Ist ein Buch in genau derselben Ausgabe mehrfach eingescannt worden, ist der Fall unproblematisch. Schwieriger wird es, wenn es sich um den gleichen Text in verschiedenen Ausgaben handelt oder um eine Textstelle, die in unterschiedlichen Versionen eines Texts vorkommt (etwa verschiedenen Fassungen eines Gedichts). Zitate werfen ähnliche Probleme auf. In all diesen Fällen zählt Google jedes Vorkommen des gesuchten Ausdrucks als ein eigenständiges Ergebnis, und es ist fraglich, wie das System mit solchen Problemen auch anders umgehen sollte. Eine genaue Trefferanzahl anzugeben, ist in solchen Fällen unmöglich. Es hängt vom Erkenntnisinteresse und der Fragestellung des Nutzers bzw. den sich daraus ergebenden methodischen Konsequenzen ab, was im Weiteren von ihm als eigenständiger Treffer gezählt werden kann und was nicht.

Ein weiteres objektives Problem sind Homographen, vor allem dann, wenn sie, wie in einem solchen einzelsprachübergreifenden System wie der Google-Buchsuche, Elemente mehrerer verschiedener Sprachen sind. Weltweit verbreiteten Entlehnungen, Internationalismen und zufällig in mehreren Sprachen identischen Wortgestalten begegnet man umso häufiger, je mehr Texte von Google eingescannt und über die Buchsuche verfügbar gemacht werden.

Dazu kommen Verlesungen der OCR, die ebenfalls kaum (und dann allenfalls durch eine technische Verbesserung) zu vermeiden sind. Abhängig von verschiedenen Faktoren wie der Schriftart, der Papierqualität usw. können sie bisweilen starken Einfluss auf die Ergebniszahlen haben.

Solche Faktoren stellen objektive Probleme dar, die zu lösen Google zumindest in einigen Fällen vor unüberwindliche Schwierigkeiten stellen dürfte. Andere dagegen sind eine Frage der Definition und sollten dem jeweiligen Benutzer anheim gestellt bleiben, wieder andere dürften zumindest nach und nach ver-

bessert werden können, die Lesefehler der OCR etwa oder die Einzelspracherkennung. Daneben gibt es aber eine Reihe von Problemen, die hausgemacht sind. Einige davon wurden auch bereits angesprochen, etwa die Tatsache, dass Google Textstellen bei einer zweiten Suche nicht wiederfindet, die Suchfunktion, die Ergebnisse mit ausgibt, die gar nicht gesucht worden waren, der Effekt, dass allein schon eine Umsortierung der Ergebnisliste eine Veränderung der Ergebnisanzahl mit sich bringt, sowie die Tatsache, dass Google bei zeitlichen Eingrenzungen im System vorhandene Textstellen nicht anzeigt.

Bisher noch nicht angesprochen wurde die Frage, was ein Ergebnis eigentlich genau ist: Google scheint in seinen Ergebnislisten ein Buch, in dem das Suchwort mindestens ein Mal bzw. mindestens eines der Suchwörter vorkommt, als ein Ergebnis zu zählen. Wenn sich ein Suchwort also 115 Mal in einem Buch findet, zählt dies ebenso als ein einziges Ergebnis wie wenn sich das Suchwort in einem anderen Buch nur ein einziges Mal findet.⁷

Ein weiteres, gravierenderes Problem besteht darin, dass die Google-Buchsuche die Ergebnislisten bei etwa 600 Einträgen zu „deckeln“ scheint. Das bedeutet, dass Google nur ca. 600 Ergebnisse anzeigt, selbst wenn deutlich mehr relevante Textstellen vorhanden sind. Bei den meisten Suchen fällt das gar nicht auf. Google gibt zwar an: „Ungefähr [...] Ergebnisse“, wenn man aber nicht alle Ergebnislisten durcharbeiten muss, kommt man über die 600 gar nicht hinaus. Wenn aber sehr umfangreiche Ergebnismengen mit Hilfe des neuen Tools „nach Datum sortiert“ chronologisch geordnet werden, bemerkt man, dass Textstellen aus ganzen Zeiträumen fehlen. Diese sind auch nicht zu erreichen, wenn man sich durch sämtliche Ergebnislisten hindurchklickt: Der Zugang zu den vermissten Ergebnissen ist nur durch eine neue Suche möglich, bei der der Suchzeitraum entsprechend eingeschränkt wird, nämlich so, dass die Suche nicht mehr als die etwa 600 Ergebnisse generiert. Dies kann insbesondere bei hochfrequenten Suchwörtern äußerst mühsam sein.

Aufgrund dieser Überlegungen muss man nicht in Bezug auf Häufigkeiten, sondern, gravierender, in Bezug auf das schiere Vorhandensein von Textstellen mit solchen Aussagen äußerst zurückhaltend sein, die sich allein auf Recherchen in der Google-Buchsuche stützen. Zwar ist es möglich, einige Probleme mittels verschiedener Hilfskonstruktionen zu umgehen, da Google aber nicht offenlegt, wie die Buchsuche funktioniert, können diese nicht in einer Weise eingerichtet werden, die es gestatten würde, verlässliche Ergebnisse zu generieren.

In „vollständiger Ansicht“ vorhandene Bücher

Ein weiteres Problem fällt dem Benutzer nur bei intensiver Nutzung der Google-Buchsuche ins Auge:

Bisweilen sind mehrere Bücher bzw. Bände als ein einziger Titel registriert. Dies kann zwei Gründe haben:

1. Zum einen werden von Bibliotheken aus unterschiedlichen Gründen physisch zunächst getrennt erschienene Bücher zusammengebunden. In solchen Fällen besteht meist eine unmittelbar einleuchtende Zusammengehörigkeit (etwa bei Zeitschriftenjahrgängen oder mehrbändigen Werken). Wenn diese dann von Google eingescannt werden, wird dies so ins System übernommen, d. h., der Scan wird als ein Buch in der Datenbank abgelegt.
2. Der zweite Grund kann weitaus problematischere Konsequenzen nach sich ziehen: Durch den Scanvorgang werden bisweilen auch digital Bücher zusammengesetzt, die in keinerlei Zusammenhang zueinander stehen. Das gravierendste Beispiel, das die Praxis zutage förderte, war eine theologische Abhandlung aus dem 17. Jahrhundert, die mit einem aktuellen Ratgeber zur Gartenpflege zusammengescannt worden war.

Bei der Übernahme von Daten aus dem zweiten Teil eines solchen Digitalisats kann es passieren, dass unbemerkt die Angaben einer falschen Titelseite übernommen werden, vor allem dann, wenn sich die zusammengebundenen oder -gescannten Bände im Layout ähneln. Solche Fälle sind zwar selten, kommen aber deutlich häufiger vor, als man erwarten möchte. Daher ist eine Überprüfung anhand der Seitenzahlen in jedem Fall anzuraten.

Handelt es sich tatsächlich um solche Fälle, muss das nicht unbedingt problematisch sein: Man sucht die Titelseite des Buches, in dem sich die benutzte Textstelle tatsächlich befindet, und zitiert danach. Existiert die entsprechende Titelseite nicht (das ist leider öfter der Fall), kann die Suche nach den korrekten bibliographischen Angaben allerdings schnell sehr problematisch werden.

Die Schwierigkeiten nehmen zu, wenn man Ergebnisse aus lediglich in „eingeschränkter Vorschau“ oder in „Auszügen“ vorhandenen Büchern übernehmen will oder muss. In solchen Fällen ist eine Überprüfung allein anhand der von Google bereitgestellten Mittel nicht möglich, da das Digitalisat nicht in Gänze durchsuchbar ist. Der Griff zum Buch ist unumgänglich.

Sollte man die Google-Buchsuche meiden?

Schon allein aufgrund ihrer schieren Größe ist die Google-Buchsuche mittlerweile ein schwer zu ignorierendes, zumindest jedoch äußerst verlockendes Hilfsmittel für den Lexikologen. Sie ermöglichte den Bearbeitern der zweiten Auflage des Deutschen Fremdwörterbuchs unter anderem das Auffinden des derzeit ältesten bekannten Belegs für „Hobby“ im Deutschen (in einem Brief Mendelssohns an Lessing aus dem Jahr 1763⁸, etwa 150 Jahre vor dem bislang vermuteten Entlehnungszeitraum) und die Belegung des weithin mit der Person und den Theorien Michel Foucaults in Verbindung gebrachten Wortes *Gouvernementalität* bereits im 19. Jahrhundert. Ebenso gelang der Nachweis des Gebrauchs von *Homöopathie* bereits im 18. Jahrhundert, deutlich vor der Einführung des von Samuel Hahnemann entwickelten Heilverfahrens, mit der die Bezeichnung bisher immer in Verbindung gebracht wurde.⁹

Darüber hinaus ermöglicht es oft erst eine Googlerecherche, zwischen einem (vielleicht von Wörterbuch zu Wörterbuch weitergetragenen) Okkasionalismus und einem tatsächlich gebräuchlichen Wort zu unterscheiden. Solche Fälle lagen etwa bei *Gagist*¹⁰, *helikopterisch* (DFWB Bd. VII, S. 188), *Holokauste* (ebd., S. 333ff.), *Humbug(g)er*, *humbug(g)en*, *humbugisch* (ebd., S. 476ff.) oder *hypermodern* vor, das im DFWB mit einem Beleg aus dem Jahr 1824 vertreten ist (ebd., S. 543 und 547), mittlerweile aber sogar noch einige Jahre früher nachgewiesen werden konnte. Dasselbe gilt für Zusammensetzungen. So verzeichnet das DFWB etwa neben den weithin bekannten *Hiobsbotschaft*, *Hiobsnachricht* und *Hiobspost* z. B. auch *Hiobszeitung*, *Hiobsmeldung* sowie *Hiobsbote* und *Hiobsbotschafter* (DFWB Bd. VII, S. 275ff.). Sicherlich mögen einige dieser Wörter marginal erscheinen, doch das wäre erst einmal zu beweisen. Noch vor wenigen Jahren wurden viele andere ebenfalls für marginal gehalten – während die genannten der Sprachregistration zum Teil nicht einmal bekannt waren.

Aber es sind nicht immer die spektakulären Entdeckungen, die ins Gewicht fallen. So sind erst mit Hilfe der Google-Buchsuche viele Ableitungen etwa von *Humor* (DFWB Bd. VII, S. 478-502) und *Humus* (DFWB Bd. VII, S. 502ff.) ausführlich belegbar. Und auch jenseits der lexikalischen und der morphologischen Ebene können neue Erkenntnisse zutage gefördert werden: So war es nicht nur möglich, die Bedeutung ‘Rangordnung der Engel’, spezifiziert zu ‘eine der drei mal drei Gliederungsebenen in der Rangordnung der Engel’ von *Hierarchie* ausführlich zu dokumentieren, die entsprechende Bedeutung konnte auch für einige

Ableitungen wie *hierarchisch* und *Hierarch* ('Engel') nachgewiesen werden (ebd., S. 246ff.).

Fazit

Wie sich bereits 2009 abzeichnete, kann die Google-Buchsuche zwar durchaus als ein Hilfsmittel der Lexikologie zum Einsatz kommen, allerdings nur bei entsprechender Vorsicht, d. h., wenn durch bestimmte Kautelen gesichert ist, dass die beschriebenen Probleme die Ergebnisse nicht verfälschen. Dass diese Kautelen keine Elemente des Systems selbst sind, sondern von außen, vom Nutzer in Form von Workarounds an es herangetragen werden müssen, ist wohl nicht nur aus Sicht des Linguisten als Mangel des Systems anzusehen.

In diesem Lichte betrachtet, aber auch im Lichte der ebenso unleugbaren Vorteile, die die schiere Textmenge mit sich bringt, die Google bietet, muss man eine Einschätzung dieses Tools von wissenschaftlicher Warte aus wohl wie folgt auf den Punkt bringen:

Die Google-Buchsuche ist kein Tool, das als elektronisches Korpus für lexikologisch-lexikografische Arbeiten eingesetzt werden kann, es ist vielmehr eines, das dem Lexikologen oder Lexikografen den Zugriff auf Texte und Textstellen erleichtert und in vielen Fällen überhaupt erst ermöglicht, die er anderweitig nicht oder nur schwerlich hätte auffinden können. Aus der Google-Buchsuche gewonnene Ergebnisse können keinesfalls als verlässlich erachtet werden und sollten daher mit google-externen Mitteln auf ihre Verlässlichkeit überprüft und zudem äußerst zurückhaltend und allenfalls ergänzend zur Anwendung kommen.

Anmerkungen

- ¹ Brückner, Dominik: Die Google Buchsuche als Hilfsmittel für die Lexikographie. In: SPRACHREPORT 3/2009, S. 26-31.
- ² Die folgenden Beobachtungen ergaben sich wieder aus der täglichen Arbeit am Deutschen Fremdwörterbuch (DFWB) und gehen auf ein internes Papier zurück, das von Herbert Schmidt, Katrin Ankenbrand und Dominik Brückner für den Zweck erarbeitet wurde, die Fähigkeiten der Mitarbeiter und Hilfskräfte des DFWB im Umgang mit der Google-Buchsuche einander anzugleichen, die Sensibilität für die Fehler und Fehlerquellen, die Google aufweist, zu erhöhen und letztlich den Umgang mit der Google-Buchsuche auf das nötigste zu reduzieren.
- ³ Im Folgenden wird versucht, den Ausdruck *Buch* soweit möglich in derselben Weise zu verwenden, wie Google das tut, um den Bezug zu dem, was Google ein „Buch“ nennt, zu wahren. Damit sei aber keinesfalls zum Aus-

druck gebracht, dass Google eine eindeutige, brauchbare oder auch nur auffindbare Definition von „Buch“ anbietet.

- ⁴ Wie hilfreich wäre daneben die Möglichkeit, satzzeichen-sensitiv zu suchen!
- ⁵ Nebenbei bemerkt ist vollkommen unklar, was die zweite Sortierungsmöglichkeit, die bei Google „Nach Relevanz sortiert“ heißt, leistet. Zwar hat es den Anschein, dass bei dieser Sortierungsweise deutsche Ergebnisse weiter „nach oben“ sortiert werden, auf welche Weise das geschieht, ist allerdings nicht deutlich.
- ⁶ Dass die Ergebniszahlen durch neu eingescannte Texte zunehmen, ist nachvollziehbar, dass sie aber abnehmen, muss dem Benutzer unverständlich bleiben.
- ⁷ Ähnliches gilt auch für die vollständige Ansicht der Bücher, hier in durchaus sinnvoller Weise: Erscheint ein Wort z. B. in den Kopfzeilen mehrerer Seiten, etwa als lebender Kolumnentitel, so wird dieses Vorkommen zwar gelb unterlegt, aber nicht durch eine Markierung am rechten Rand als Treffer markiert.
- ⁸ Wie einige der Frühbelege noch in der Form *Hobby-Horse*, s. Deutsches Fremdwörterbuch Bd. VII, *habilitieren* – *hysterisch*. Begonnen von Hans Schulz, fortgeführt von Otto Basler. 2. Auflage, völlig neu erarbeitet im Institut für Deutsche Sprache von Schmidt, Herbert/Brückner, Dominik/Nortmeyer, Isolde/Vietze, Oda (2010). Berlin/New York: de Gruyter, S. 328 (Im Folgenden zitiert als DFWB Bd. VII).
- ⁹ ebd., S. 347f. Diese Verknüpfung einer Wortgeschichte mit dem Aufkommen einer Theorie oder Methode, mit der das betreffende Wort in Zusammenhang steht, scheint ein verbreitetes Denkmuster zu sein. So schreibt z. B. Angelika C. Wagner in ihrem Beitrag „Gelassenheit und Handlungsfähigkeit. Über das Aufhören der Imperative als Essenz feministischer Therapie“ (in: Rommelspacher, Birgit (Hg.): *Weibliche Beziehungsmuster. Psychologie und Therapie von Frauen*. Frankfurt/New York 1987): „Das Verb „sich imperieren“ ist im deutschen Sprachraum bislang ungebräuchlich und bietet uns deshalb die Chance, es als Ausdruck für einen präzise zu umschreibenden Bewußtseinsprozeß zu gebrauchen, den Prozeß, in dem das Bewußtsein sich selber imperativisch vorschreibt, daß „X sein muß“, bzw. „Y nicht sein darf““ (S. 159). Die Rede ist dabei von einem Verb, das seit 1487 belegt ist. Vergleichbar ist auch der bekannte Fall des deutschen Wortes *Fernseher*, das bereits im 18. Jahrhundert vielfach belegt ist (in den Bedeutungen ‚Prophet‘ und ‚Fernrohr‘), vgl. z. B. auch Grimm, Jacob/Grimm, Wilhelm: *Deutsches Wörterbuch*. Bd. III, Sp. 1539.
- ¹⁰ *Deutsches Fremdwörterbuch*. Bd. VI, *Gag* – *Gynäkologie*. Begonnen von Hans Schulz, fortgeführt von Otto Basler. 2. Auflage, völlig neu erarbeitet im Institut für Deutsche Sprache von Strauß, Gerhard/Schmidt, Herbert/Brückner, Dominik/Nortmeyer, Isolde/Vietze, Oda unter Mitarbeit von Kämper, Heidrun (2006). Berlin/New York: de Gruyter, S. 4f.

Der Autor ist wissenschaftlicher Mitarbeiter am Institut für Deutsche Sprache in Mannheim.