

Konstanze Marx / Georg Weidacher

Internetlinguistik

Ein Lehr- und Arbeitsbuch

2., aktualisierte und durchgesehene Auflage

narr STUDIENBÜCHER

narr\f
ranck
e\atte
mpto

narr STUDIENBÜCHER

Konstanze Marx / Georg Weidacher

Internetlinguistik

Ein Lehr- und Arbeitsbuch

narr\f
ranck
e\atte
mpto

2., aktualisierte und durchgesehene Auflage 2020
1. Auflage 2014

© 2020 · Narr Francke Attempto Verlag GmbH + Co. KG
Dischingerweg 5 · D-72070 Tübingen

Internet: www.narr.de
eMail: info@narr.de

ISSN 0941-8105

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

ISBN 978-3-8233-8117-4 (Print)
ISBN 978-3-8233-0236-0 (ePub)

Inhalt

- Zu diesem Buch
 - Vorwort zur zweiten Auflage
 - Vorwort zur ersten Auflage
- 1 Methoden der Internetlinguistik
 - 1.1 Das Internet als Datenpool
 - 1.2 Aber wem gehören die Daten?
 - 1.3 Das Zitat als eine annehmbare Lösung
 - 1.3.1 So zitiert man sprachliche Beispiele
 - 1.3.2 So zitiert man wissenschaftliche Publikationen
 - 1.4 Onlinedaten erheben – Durchaus eine Herausforderung
 - 1.4.1 WWW = Korpus?
 - 1.4.2 DIY: Eine Datensammlung selbst generieren
 - 1.4.3 Klick ins Feld: Einfach einmal nachfragen?
 - 1.4.4 Log-File-Analyse – Einfach mitschneiden?
 - 1.4.5 Digitale Ethnographie
 - 1.4.6 Auch Umwege führen zu digitalen Daten
 - 1.5 In der Bibliographie soll es dann so aussehen
 - 1.5.1 Die Online-Publikation im Literaturverzeichnis
 - 1.5.2 Quellenverzeichnis für Beispielbelege
- 2 Medientheorie des Internets
 - 2.1 Unsere Welt ist online
 - 2.2 Was ist ein Medium?
 - 2.2.1 Einer, der polarisiert: McLuhans Medienbegriff
 - 2.2.2 Wozu sind sie eigentlich da – die Medien?

- 2.2.3 Verbreitung! Kommunikation! Speicherung!
- 2.2.4 Von Kanälen, Sinnesmodalitäten und Codes: Elemente (technischer) Medien und Mediendefinition
- 2.2.5 Primäre, sekundäre und tertiäre Medien
- 2.2.6 Ist das Medium die Botschaft?
- 2.2.7 Von konventionalisierten kommunikativen Handlungen: Kommunikationsformen
- 2.2.8 So funktioniert es generell: Mediale Kommunikation – ein allgemeines Modell
- 2.3 Ist das Internet nun ein Medium?
 - 2.3.1 Das Internet als Netzwerk
 - 2.3.2 Web 2.0 und Social Media
- 2.4 Ja, das Internet ist ein Medium!
 - 2.4.1 Mehr noch: Das Internet ist ein Hybridmedium
 - 2.4.2 Von „designed spaces“ und Kommunikationsplattformen
 - 2.4.3 Das Netz: Alles ist möglich – oder doch nicht? Constraints und Affordances
 - 2.4.4 Von konventionalisierten kommunikativen Handlungen online
 - 2.4.5 So funktioniert es im Netz: Ein spezifisches Kommunikationsmodell
- 3 Sprache im Internet
 - 3.1 Weder Sondersprache noch Stil
 - 3.2 WhatsApp? Flashmob – approach, m8! Allerlei Fundstücke
 - 3.2.1 Abgeguckt: Indikatoren für Adaptionenprozesse
 - 3.2.2 Mitgespielt: Indikatoren für sprachliche Sensibilität und Reflexion
 - 3.2.3 Losgetextet: Indikatoren für Oraliteralität
 - 3.3 Netzbedingte Dynamik: Neue Bedeutung für Altbekanntes

- 3.3.1 Ein Freund, ein guter Freund?
- 3.3.2 Gefällt mir ist mehr als ein Ausdruck für Gefallen
- 4 Pragma-Internetlinguistik
 - 4.1 Ins Netz versetzt: Grundpfeiler der Pragmatik
 - 4.2 Sprachhandlungsphänomene
 - 4.2.1 Chat-Kommunikation aus pragmatischer Perspektive
 - 4.2.2 Performativität in MUDs
 - 4.2.3 Digitale Erweiterungen des Kommunikationsraums – (sprachlich) Freunde finden
 - 4.3 Spurensuche in der Postingflut
 - 4.3.1 Verschlüsseltes Gemeintes, verschlüsseltes Relevantes
 - 4.3.2 Gricesche Maximen als Dekodierungshilfe
 - 4.3.3 Emotionale Chiffren: E-Implikaturen
 - 4.4 Das Netz als Kontextkonstituente
 - 4.4.1 Nah und fern, privat öffentlich und alles parallel: Der mediale Kontext und die Äußerungssituation
 - 4.4.2 Ich kenne dich und ich kenne dich nicht: Zur Beziehung zwischen den Kommunikationsteilnehmer*innen
 - 4.4.3 Kognitionsinhärente Merkmale oder: Wie das Internet unser Denken prägt
- 5 Textlinguistik und das Internet
 - 5.1 Eine neue Textlinguistik für das Internet?
 - 5.2 Was ist ein Text überhaupt?
 - 5.3 Texte im Internet
 - 5.3.1 Verlinkt: Hypertextualität
 - 5.3.2 Vermischt: Multimodalität

- 5.3.3 Verflüssigt: Fluidity
- 5.3.4 Verteilt: Dialogizität
- 5.4 Kommunikationsformen und Textsorten im Internet
 - 5.4.1 Der „Tatort“ nicht mehr nur im TV – Websites zu traditionellen Medienangeboten
 - 5.4.2 BIG Fans – Fanseiten in Sozialen Netzwerken
 - 5.4.3 „Bin krank – Grippe“ – Entschuldigungs-E-Mails
 - 5.4.4 Das virtuelle Tagebuch: Verschiedene Blogs
- Bibliographie
- Register

Zu diesem Buch

Vorwort zur zweiten Auflage

Sechs Jahre nach Erscheinen des Arbeitsbuches Internetlinguistik war es höchste Zeit, Aktualisierungen vorzunehmen. Das haben wir mit Freude getan, denn die Entwicklungen in diesem Feld, dem inzwischen auch der Name „Digitale Linguistik“ zugewiesen wird, sind so spannend wie eindrucksvoll.

Es hat sich in drei große Bereiche ausdifferenziert: die Sprachverwendung in internetbasierten Kommunikationsumgebungen, die Interaktion zwischen Mensch und Maschine und die Generierung von Datenkorpora. Alle Bereiche werfen auch ethische Fragen auf. Als übergeordnete Fragestellungen lassen sich festhalten: Wie nimmt die Internettechnologie Einfluss auf unsere Sprache, auf unsere Kommunikation, auf unser Interaktionsverhalten und damit auch auf unsere Kultur? Mit welchen Methoden können wir diese Prozesse untersuchen? Wie können die Daten dazu erhoben, archiviert und für andere zugänglich gemacht werden? (Marx 2019a).

Dies im Blick haben wir inhaltliche Anpassungen vorgenommen, diese orientierten sich am von der Erstauflage vorgegebenen Rahmen. Die Entwicklungen im Forschungsfeld haben wir auch dort zu berücksichtigen versucht, wo wir Literaturhinweise auf weiterführende

Themen geben konnten. Durchaus bemerkenswert – das sei hier als Anekdote erwähnt – ist, dass wir in der Überzeugung, so ziemlich alle sprachlichen Belege ersetzen zu müssen, feststellten, dass die Beispiele zum Flughafen BER nicht an Aktualität verloren haben. Auch die Eigenschaften von Facebook haben in den letzten Jahren kaum noch Veränderungen erfahren. Das hat uns einerseits erstaunt und andererseits unseren Eindruck bestätigt, dass hier mehr und mehr ein Archiv für soziale Prozesse entsteht.

Die „Familie“ der Internetlinguistik-Bücher hat 2019 sowohl mit den bei Winter erschienenen Literaturhinweisen zur Linguistik als auch mit dem ebenfalls bei Narr erschienenen „Starter“ Zuwachs bekommen. Das wirkt auch auf die vorliegende 2. Auflage zurück. Im Grunde können wir das Bild der Geschwisterkinder wieder bemühen, die sich gegenseitig beeinflussen. Auch die „Großen“ übernehmen das, was sich bewährt, von den „Kleinen“. Das wird im vorliegenden Buch besonders im Kapitel 1 und im Kapitel 5 sichtbar. Wir danken bei dieser Gelegenheit unserem Lektor, Tillmann Bub, sehr herzlich für die Unterstützung bei der Umsetzung und die so wunderbare Zusammenarbeit. Fiona Makulik hat uns beim Erstellen der Druckvorlage geholfen, auch ihr danken wir herzlich.

A propos bewährt: Für hilfreiche Hinweise in den Rezensionen und die vielen anderen wertvollen Rückmeldungen, die uns in allen möglichen Formaten erreichten, danken wir unseren Kolleg*innen und Studierenden, unseren Leser*innen herzlich. Aus unserer Sicht hat das Buch die Tür zu einem Raum geöffnet, in dem seither viele inspirierende Begegnungen stattgefunden haben. Wir sind dafür sehr dankbar. Möge das weiter der Fall sein, möge der Forschungsgegenstand weiterhin viele beflügeln und die Vernetzung der Forschenden bewirken.

Greifswald und Graz, im Juni 2020
Georg Weidacher

Konstanze Marx,

Vorwort zur ersten Auflage

Eines ist wohl sicher: Ohne das Internet hätte dieses Buch nicht entstehen können. Das liegt zum einen daran, dass das Internet die Arbeit an diesem Buch praktisch erst ermöglicht hat. Schließlich sind Berlin und Graz geographisch nicht so günstig gelegen, dass regelmäßige Arbeitsbesprechungen in der einen oder der anderen Stadt eine Option gewesen wären. Das gemeinsame Projekt wuchs und gedieh also parallel in zwei verschiedenen Ländern in Büros etwa zwölf Zug- und zwei Flugstunden voneinander entfernt. Dabei ermöglichten uns die Kommunikationsfunktionen, die uns das WWW – insbesondere das Web 2.0 – bietet, sowohl den Zugriff auf unsere Daten als auch die Illusion, uns über das Manuskript verständigen zu können als säßen wir in einem Büro. Zum anderen liegt es daran, dass es ohne das Internet keine *Internetlinguistik* gäbe, sondern schlicht Linguistik. Dazu jedoch haben Sie schon eine Menge Lehrbücher gelesen.

Zur Entstehung

Das Wissen, das Sie aus diesen Büchern haben, wird Ihnen bei der Lektüre dieses Buches helfen, denn wir setzen linguistische Grundkenntnisse voraus. Dieses Buch ist also keine Einführung in die Linguistik, es richtet sich an Studierende, die auf sprachwissenschaftliches Basiswissen zurückgreifen können und Freude daran haben, dieses auf aktuelle sprachliche Phänomene im Internet anzuwenden. Dennoch werden wir nicht komplett auf Ausführungen zum theoretischen Hintergrundwissen verzichten, diese sind im Buch grau unterlegt und mit einem Häkchensymbol markiert (siehe auch die Symbolerklärungen am Ende dieser Einleitung).

Zu den
Leser*innen

Für Lehrende soll es als Inspiration für Seminarkonzeptionen dienen, mit den einzelnen Kapiteln gibt es eine Struktur für Unterrichtseinheiten vor. Zahlreiche

Übungsaufgaben dienen sowohl dem Selbststudium (weil sie im Folgetext diskutiert werden) oder können Anregung für das Seminalgespräch sein. Wir setzen voraus, dass die grundlegenden Funktionsweisen von Webanwendungen wie E-Mail, Sozialen-Netzwerkseiten oder Blogs bekannt sind.

Aber was ist das eigentlich, Internetlinguistik? Zum Gegenstar Was unterscheidet diese junge, moderne Disziplin von der Allgemeinen Linguistik, was von der Medienlinguistik? Womit beschäftigt sie sich und welche Forschungsfragen werden hier gestellt?

Beginnen wir doch an dieser Stelle gleich mit einer ersten kleinen Aufgabe: Geben Sie das Lexem *Internetlinguistik* einfach einmal in eine Internetsuchmaschine ein oder lassen Sie es uns anders formulieren: *Googeln* Sie es! In dieser Aufforderung steckt bereits ein Beispiel dafür, was im allgemeinen Verständnis mit *Internetlinguistik* assoziiert wird: Die Untersuchung einer sogenannten „Internetsprache“, *googeln* als eingedeutschtes Verb wäre entsprechend typisch dafür. In Kapitel 3 lüften wir allerdings ein offenes Geheimnis, wenn wir zeigen, dass es gar nicht die Internetsprache (oder auch *netspeak* oder *Internetslang*) gibt. Es gibt sie ebensowenig wie eine Mediensprache.

Nun, haben Sie gegoogelt? Wie sieht das Ergebnis aus? Sicherlich wurde Ihnen die Ankündigung zu diesem Buch angezeigt. Sie sind auch auf eine Tagung in Budapest gestoßen, in deren Rahmen wir einen gleichnamigen Workshop angeboten haben. Vielleicht finden Sie auch einen Hinweis auf ein Seminar, das in diesem Wintersemester an der TU Berlin stattfindet und von dessen Diskussionen dieses Buch hier profitiert. An dieser Stelle also schon der herzliche Dank an alle engagierten Seminarteilnehmer*innen.

Aber zurück zu Ihrer Suche: Eventuell sind Sie auf die „Gesellschaft für Interlinguistik e.V.“ gestoßen, ein Verein, der sich die Förderung von Plansprachen zur Aufgabe gemacht hat. Dieses Suchergebnis zeigt nur, dass selbst Google auf Nebenschauplätze ausweichen muss. Eine

untergeordnete Rolle für uns spielt übrigens, wie sich die Forschungsdisziplin Linguistik im Internet präsentiert (vgl. dazu bitte Cölfen/Cölfen/Schmitz 1997).

Stattdessen steht die Sprachverwendung im Internet im Mittelpunkt dieses Lehr- und Arbeitsbuches. *Internet* verwenden wir hier so wie in der einschlägigen Forschungsliteratur (siehe Runkehl/Schlobinski/Siever 1998: 3) üblich, sehr weit, und meinen damit alle durch das Internet verfügbaren Internetdienste (E-Mail, WWW u. a.). Die Termini *Web*, *WWW* oder *Netz* verwenden wir demzufolge ebenfalls synonym.

Als Aufgabe der Internetlinguistik betrachten wir die Untersuchung der Sprachverwendung in Abhängigkeit von der spezifischen Online-Umgebung. Welche wechselseitigen Einflüsse können wir feststellen? Wie gestaltet sich die Kommunikation über die verschiedenen Internetdienste? Was ist charakteristisch wofür? Wie müssen wir die bereits existierenden linguistischen Ansätze anpassen oder erweitern, um den neuen Kommunikationsraum theoretisch adäquat beschreiben zu können?

Wir befinden uns damit mitten in der Angewandten Linguistik und können folgerichtig nicht nur genuin sprachwissenschaftlich vorgehen, sondern greifen sinnvollerweise sowohl methodisch als auch theoretisch auf Nachbardisziplinen wie die Kommunikations- und Medienwissenschaft, die Psychologie und die Soziologie zurück. Damit werden die Bausteine, die in der modernen Linguistik verwendet werden, in einer anderen Umgebung neu zusammengesetzt und mit weiteren Bausteinen interdisziplinär ergänzt, um ein modernes und stabiles Gebäude zu errichten.

Die **Internetlinguistik** beschäftigt sich mit der Sprachverwendung im Internet und damit mit einem spezifischen kommunikativen Kontext, dessen Charakteristika in alle Analysen einfließen. Bei der Internetlinguistik handelt es sich um eine Schnittstellendisziplin, die – wie für die Angewandte Linguistik

typisch – neben linguistischen Zugängen, kommunikations- und medienwissenschaftliche Methoden kombiniert und durchaus auch sozio- und psychologische Fragestellungen motiviert.

Es ist unser Anliegen, eine Anleitung dafür vorzulegen, wie die isolierte Beobachtung sprachlicher Phänomene im Internet überwunden werden kann. Die Sprachverwendung im Internet ist nicht nur ein auf allen sprachwissenschaftlichen Beschreibungsebenen hochinteressanter Forschungsgegenstand, sondern auch über den Tellerrand der Linguistik hinaus.

Das haben viele andere vor uns bemerkt. Seit etwa 15 Jahren gibt es eine rege sprachwissenschaftliche Auseinandersetzung mit Themen wie Sprache und Kommunikation im Netz (Runkehl/Schlobinski/Siever 1998, Siever/Schlobinski/Runkehl 2005, Haase et al. 1997, Baron 2008a,b), Netzsprache (Crystal ²2006, 2011, Rosenbaum 1996), Mündlichkeit/Schriftlichkeit (Dürscheid 2003, ⁵2016, Storrer 2001a,b, Thaler 2003), Chat-Kommunikation (Beißwenger 2002, 2007, 2009, 2010a,b, 2013a,b, in press), syntaktischen Innovationen (Albert 2011, 2013), Bedeutungswandel (Dürscheid/Brommer 2013) oder den Charakteristika von Hypertexten (Storrer 2004, Jakobs/Lehnen 2005, Beißwenger/Storrer 2010, Jakobs 2011). Androutsopoulos (2003a,b) hat sich mit soziolinguistischen Aspekten auseinandergesetzt, Thimm (2000) hat eine Sammlung von Texten zur Sozialität im Netz veröffentlicht. Döring (²2003) definiert eine Sozialpsychologie des Internets als Forschungsgegenstand. Jakobs/Lehnen (2006, Jakobs 2013) befassen sich mit Fragen der Usability. Bei Dang-Anh/Einspänner/Thimm (2013), Siever (2001, 2006) oder Moraldo (2009) finden wir Untersuchungen zu spezifischen Kommunikationsformen wie Twitter. Es gibt Forschung zu Diskursen im Netz (Frass/Meier/Pentzold 2013, Thurlow/Mroczek 2011), zur Online-Liebe (Döring 2002, 2003, Ze'ev 2004, Marx 2012a,b) und sprachlichen Online-Gewalt (Fawzi 2009, ²2015, Gradinger 2010,

Katzer/Fetchenhauer/Belschak 2009, Marx 2012a, 2013a,b, Schwarz-Friesel 2013), zur digitalen (Un-)Ordnung (Runkehl 2013), zum Online-Verhalten der verschiedenen Generationen (Feufel/Stahl/Lee 2013, Ziefle/Jakobs 2013), zu Politikerauftritten im Netz (Weidacher 2007), zu kreativen Versuchen, Filter zu umgehen (Eliaz/Rozinger 2013), zu Identitätskonstruktionen (Gallery 2000, Götzke 2002, Döring 2000b, Thimm/Ehmer 2000), zur Theologie der Social Media (Ernst/Costanza 2012) oder zu ins Internet verlegten sozialen Ritualen, wie z. B. die Trauer (Jakobs/Ziefle 2010). Ein ganzes DFG-Forschungsnetzwerk, Empirikom, wurde gegründet, um empirische Methoden in der sprach- und kommunikationswissenschaftlichen Internetforschung zu diskutieren (Beißwenger 2012).

Wir könnten diese Aufzählung noch einige Seiten fortführen, die Publikationen zu allen nur vorstellbaren Facetten des Internets sind schier unüberblickbar. Das Internet erscheint wie eine Inspirationsquelle, die in absehbarer Zeit nicht versiegt. Wer von dieser Quelle trinken möchte, stellt schnell fest, dass das gar nicht so einfach ist, weil sie stark sprudelt und einen (Daten- oder auch Informations-)Strom speist, der Wellen schlägt, auf denen das Surfen durchaus eine Herausforderung darstellt. Wir haben uns mitreißen lassen von dieser Fülle an hochinteressanten Themen und Forschungsergebnissen. Um noch ein wenig im Bild zu bleiben: Es hat uns harte Ruderarbeit gekostet, schließlich am Ufer der Internetlinguistik anlegen zu können und zwar mit mehr Gepäck als zum Zeitpunkt unseres Aufbruchs. Einiges wurde zu Treibgut und das bedauern wir. Die Dinge aber, die wir retten konnten, weil sie oben schwammen oder wir schnell genug zugegriffen haben, haben wir sorgfältig aufbereitet. Dazu mussten wir sie manchmal auch miteinander verbinden, um sie wiederherstellen zu können. Ihren Eigentümer*innen danken wir. Es gab natürlich auch Dinge, die wir ersetzen mussten oder die komplett neu hinzukamen, weil sich während unseres Surfabenteuers die Notwendigkeit dazu ergab. Nun ist dieses Abenteuer vorerst

bestanden und Sie halten mit diesem Buch unsere kompakte Reisedokumentation in den Händen. Natürlich müssen wir uns auch fragen (lassen), warum wir die Dokumentation im Zeitalter des Internets überhaupt noch in eine Buchform gebracht haben. Ganz einfach, wir wollten anderen die Strapazen unserer Reise ersparen, sie sollen einen Aktiv-Urlaub antreten können – angeleitet von einem handlichen Reiseführer. Zudem, da geben wir uns konservativ, sind wir Anhänger*innen des guten alten Leseprozesses. Wir glauben, dass er Wunder bewirken kann, die auch positive Effekte auf alle prozeduralen Fähigkeiten unseres Gehirns haben. Im Umkehrschluss beobachteten wir im Selbstversuch und auch bei unseren Studierenden negative Auswirkungen auf die Gedächtnisleistungen oder die Konzentrationsfähigkeit beim staccatohaften Suchen nach Stichworten, wie es für die Onlinerezeption üblich ist, siehe dazu auch den sehr anschaulichen Erklärungsversuch von Nicholas Carr (2010a,b).

Jedes der folgenden Kapitel ist mit drei Leitfragen überschrieben, die die jeweilige Thematik des Kapitels aufgreifen. Am Ende eines jeden Kapitels finden Sie den „Speicherinhalt“, der einer Zusammenfassung der wesentlichen Punkte entspricht, und zusätzliche Übungsfragen. Aus didaktischen Gründen wird der Text innerhalb der Kapitel mehrfach von Aufgaben oder gerahmten Informationen unterbrochen. Eine Erläuterung der Symbole finden Sie unten stehend. Die einzelnen Kapitel dieses Buches bauen zwar aufeinander auf, bilden aber jedes für sich so geschlossene Einheiten, dass sie in selbst gewählter Reihenfolge gelesen werden können.

Zur Handhabung

 **Aufgabe**, für die die Lösung im Anschluss text gegeben wird. Für den optimalen Lerneffekt empfehlen wir dringend im Leseprozess innezuhalten und die Lösung für diese Aufgabe zunächst selbst herauszufinden. Fertigen Sie dazu Stichpunkte an.

 **Denkanregung.** Für diese Aufgaben geben wir keine Lösung vor, sie dienen als Anregung für die gründliche und auch kreative Reflexion und können sehr gut im Anschluss an die Rezeption eines Kapitels gelöst werden. Sie sollen zu einem tieferen Verständnis der diskutierten Themen beitragen.

 **Formulierungsvorschlag.** Im Kapitel 1 werden Vorschläge dazu unterbreitet, wie Textbausteine formuliert werden können, die für internetlinguistische studentische Arbeiten obligatorisch sind.

 **Zusatzinformation.** Diese Kästen enthalten Informationen, die uns im entsprechenden Zusammenhang als interessant erschienen und Inspiration zum Weiterdenken sein können.

 **Theoretisches Hintergrundwissen.** Die Informationen in den grauen Kästen beziehen sich auf linguistisches Vorwissen, das Ihnen in den Grundkursen Ihres Studiums vermittelt wurde. Es soll vorkommen, dass dieses im Laufe des Studiums wieder verdrängt wird, die grauen Boxen helfen Ihnen bei der Rekapitulation. In diesen grauen Kästen finden Sie auch Definitionen von anderen zentralen Begrifflichkeiten.

Zitat. Längere wörtliche Zitate sind ebenfalls gerahmt.

Weiterführende Literatur. Mit den Leseempfehlungen wollen wir Sie ermuntern,  Ihr theoretisches Wissen zu vertiefen und ein Gespür für Kontroversen in der Forschungsdiskussion zu bekommen. In dieser als Lehr- und Arbeitsbuch konzipierten Einführung können wir diese nämlich nicht aufgreifen.

Internetadressen. Das kleine Maus-Icon am Rand kennzeichnet Verweise  auf Internetadressen.

Wie in der Linguistik üblich und deshalb nicht noch einmal anhand von Symbolen erläutert, kennzeichnen einem Satz vorangestellte Fragezeichen eine zumeist semantische Markiertheit. Konzeptualisierungen sind mit Kapitälchen hervorgehoben, Großbuchstaben markieren Ergebnisse von Schlussfolgerungsprozessen.

Sie werden in diesem Buch eine Vielzahl von authentischen sprachlichen Beispielen finden, zur Kennzeichnung der Quelle im Fließtext haben wir auf Kurzzitate zurückgegriffen, d. h. dass auch die Internetadressen nur verkürzt angegeben wurden. Im Anhang dieses Buches finden Sie ein Quellenverzeichnis für die in der Arbeit verwendeten sprachlichen Belege, sofern sie nicht aus zugangsbeschränkten (privaten) Profilseiten stammen. Diese Quellenangaben sind jeweils der Beispielnummer zugeordnet, unter der sie im Fließtext erscheinen. Aus Datenschutzgründen sind bei Daten aus privaten Profilseiten alle Urheber*innen anonymisiert. Wenn innerhalb von sprachlichen Belegen Klarnamen (Vor- und Zunamen) auftauchen, die nicht veröffentlicht werden dürfen, wurden diese durch XYZ ersetzt.

Zur Illustration einiger weniger Annahmen wurden auch Beispiele konstruiert, diese sind dann daran zu erkennen, dass keine Quelle angegeben ist oder in der Quellenangabe darauf verwiesen wird, dass es sich um ein zu Veranschaulichungszwecken modifiziertes Beispiel handelt. Phänomene wie orthographische oder grammatische Fehler oder außergewöhnliche Schreibweisen treten so häufig auf, dass die Leserlichkeit gestört worden wäre, hätten wir sie jedes Mal mit einem [sic!] gekennzeichnet.

Wir danken Tillman Bub, Karin Burger, Bernd Villhauer und Celestina Filbrandt vom Narr-Verlag nicht nur für die vertrauensvolle, äußerst angenehme und vor allem konstruktive Zusammenarbeit, sondern auch für ihre Aufgeschlossenheit, ihre Geduld und Flexibilität. Unser Dank

Unsere Studierenden Toivo Glatz, Jonas Nölle, Isabella Knapp, Gerrit Kotzur, Julia Schirnhofer, Carina Stöckler und Alexander Zahrer gaben uns hilfreiche Rückmeldungen und stellten ihre Argusaugen zur Verfügung, ein großes Dankeschön dafür.

Unseren beiden wunderbaren Chefs, Monika Schwarz-Friesel und Paul R. Portmann-Tselikas, danken wir ebenso für die moralische und inhaltlich sehr wertvolle Unterstützung wie Thomas Wischnowski und Hildegard Weidacher-Gruber. Unseren Familien und Freunden können wir ohnehin gar nicht genug danken. Sie mussten über Monate zwei im kreativen Schreibprozess Befindliche ertragen. Obgleich sich noch Stoff für eine Fortsetzung auf unseren Schreibtischen stapelt, bleibt dieses Weihnachten der Rechner aus. Versprochen.

Berlin und Graz, im November 2013
Georg Weidacher

Konstanze Marx,

1 Methoden der Internetlinguistik

Lässt sich das WWW als Datenpool für sprachliche Belege nutzen?

Welche Datenerhebungsmethoden gibt es?

Wie sind Internetquellen zu zitieren und zu bibliographieren?

1.1 Das Internet als Datenpool

Das World Wide Web (WWW) ist ein für Sprachwissenschaftler*innen geradezu unerschöpflicher Datenpool, der sich zudem kontinuierlich neu generiert. Wissenschaftliche Fachtexte, journalistische Artikel, Kommentare von Leser*innen, virtuelle Plaudereien, Zeugnisse sozialer Kontaktpflege, Blogs oder Tweets sind gerade einmal den sprichwörtlichen Mausklick entfernt. Expert*innen und Lai*innen, Jugendliche und Erwachsene, politisch Motivierte, Spaßorientierte, Gelingweilte oder Wütende geben ihr Wissen und/oder ihre Überzeugungen preis, wollen sich unterhalten oder nutzen das quasi-öffentliche Forum, um ihren Gefühlen Ausdruck zu verleihen.

Bevor sich Web 1.0 und Web 2.0 etablierten, mussten spontan produzierte Sprachdaten mit viel Aufwand erhoben und mühsam transkribiert werden. Insbesondere dank zahlreicher Social-Media-Anwendungen (wie Facebook, YouTube, Twitter, Tumblr, Nutzer- und Diskussionsforen etc.) scheinen sich die Zugänglichkeitsbedingungen für Sprachdaten erheblich verbessert zu haben. Selbst E-Mail-Anbieter binden Kommentarfunktionen in ihre Nachrichtenportale ein. Das heißt, dass das Abfragen der E-Mails nicht nur damit verbunden werden kann, auf schnellem Wege neueste Informationen zu erhalten. Nutzer*innen können sich zudem gleich zu den im Nachrichtenportal präsentierten Inhalten äußern und liefern damit jede Menge sprachliches Material.

Unter **Scheinidentität** wird hier ein Profil verstanden, das selektive Informationen, wie etwa Interessen, Einstellung etc. enthält, in dem jedoch auf Klarnamen und ein Foto, auf dem der*die Profilurheber*in deutlich zu erkennen ist, verzichtet wird.

Auf diese Weise werden verschiedene Textsorten, wie Artikel, Nachrichtentexte oder Nutzerkommentare, deren Form und Inhalt für spezifische linguistische Fragestellungen relevant sein kann, relativ unproblematisch und mit wenig Zeitaufwand verfügbar. So scheint es unkompliziert, Zugang zu diesen sprachlichen Daten zu erhalten. Dennoch gestaltet sich die Korpusgenerierung schwierig. Das liegt auch daran, dass Nutzer nach wie vor anonym (in Foren oder Nutzerkommentaren) und/oder unter Scheinidentitäten (auf Twitter, Instagram und teilweise auch auf Facebook) agieren. Zwar trägt das sicherlich zum Abbau von Hemmungen bei und begünstigt, dass sprachliche Daten überhaupt entstehen und „veröffentlicht“ werden, es birgt aber auch methodische Nachteile: Erstens, ohne technische Hilfsmittel können die Urheber*innen der Texte kaum ermittelt werden, insbesondere in den Fällen, in denen sich die „Kommentator*innen“ Pseudonyme (vgl. dazu auch den Abschnitt zu Nicknames) geben. Zweitens kann die Situation, in der der Text entstanden ist, nicht kontrolliert werden, so dass keine Aussagen darüber getroffen werden können, ob es sich beispielsweise um eine spontan produzierte Äußerung handelt oder ob der*die Verfasser*in des Kommentars seinen*ihrer Text vorformuliert hat bevor er*sie ihn „veröffentlichte“. Eine dritte Schwierigkeit ergibt sich mit der Frage, inwieweit andere (und das schließt (Sprach-)Wissenschaftler*innen ein) überhaupt berechtigt sind, diese Daten für ihre Zwecke zu nutzen.

1.2 Aber wem gehören die Daten?

Im Umgang mit den Daten aus dem WWW drängt sich die Auffassung geradezu auf, dass die Daten, die hier veröffentlicht werden, jedem und jeder gehören. Sie sind zugänglich, ihre Urheber*innen haben sie „veröffentlicht“ und damit einem Publikum zur Verfügung gestellt. Keine anderen Merkmale aber treffen auf Informationen zu, die in Büchern veröffentlicht worden sind. Sie sind ebenfalls frei lesbar, ihre Vervielfältigung und Weiterverwendung ist jedoch gesetzlich geregelt. So entspricht es nicht der seriösen Forschungspraxis, Daten aus einem Buch schlicht zu übernehmen, ohne das zu kennzeichnen und auf den*die Autor*in zu verweisen. Die wenigsten wissen jedoch, dass auch für die Verwendung von Daten aus dem WWW eindeutige und verbindliche juristische Vorgaben zu beachten sind.



Das **Urheberrecht** räumt dem Schöpfer geistigen Eigentums ein Ausschließlichkeitsrecht an seinen eigenen kreativen Leistungen ein.

Es ist grundsätzlich davon auszugehen, dass „veröffentlichte“ Daten dem*der Urheber*in gehören und damit unter das Urheberrecht¹ fallen. In vielen Fällen sind also die Daten, die im Internet kursieren, gesetzlich bis 70 Jahre nach dem Tod des Urhebers/der Urheberin geschützt. Auch bei kleineren, einfachen Texten ist im Zweifel von einer sogenannten Schöpfungshöhe auszugehen, d. h. dass das „Werk“ – und darunter fallen dann auch internetbasierte (Kommunikations)beiträge – als geschützt gilt.

 Bei der **Schöpfungshöhe** handelt es sich um ein ungeschriebenes Tatbestandsmerkmal der „persönlich geistigen Schöpfung“ i.S.d. § 2 Abs. 1 UrhG. Geschützt sind Werke, weil es sich entweder um eine persönlich geistige Schöpfung handelt, ein Sprachwerk, ein Musikwerk oder eine schützenswerte Darstellung.

Wenn nun aber das Urheberrecht gilt, bedeutet das, dass wir die Daten, die sich uns wie auf einem „goldenen Tablett“ präsentieren, gar nicht oder nur mit Einverständniserklärung des Urhebers/der Urheberin für unsere Forschung verwenden dürfen?

Sollten wir tatsächlich eine Vervielfältigung anstreben, ist vorher das Einverständnis des Urhebers/der Urheberin einzuholen.

 Eine **Einverständniserklärung** sollte vor der Daten-Erhebung schriftlich fixiert werden und folgende Informationen beinhalten: Angaben über den*die Träger*in und Leiter*in des Forschungsvorhabens; Angaben zum Forschungszweck; Angaben zur Methode der Datenerhebung; Angaben zur weiteren Verwendung der Daten und der involvierten Personen; Zeitpunkt, zu dem die personenbezogenen Daten gelöscht werden; Erklärung über Freiwilligkeit und Möglichkeit zum Widerruf; explizite Einwilligungsformulierung; Ort, Datum und Unterschriften. (siehe Bankhardt 2010)

Aber: Wie können wir eine Einverständniserklärung von jemandem einholen, der*die sich a) nicht zu erkennen gibt, weil er*sie ein Pseudonym (Nickname) benutzt und der*die b) auf keinem (technologischen) Weg kontaktiert werden kann? Für viele ist es die fehlende Antwort auf diese Frage, die sie dazu verführt, sich der Daten einfach anzunehmen, sie ohne explizite Erlaubnis zu verwenden und weiterzuverarbeiten. Es entspricht dem z.B. bei Gatto (2014: 64) beschriebenen „praktischen Weg“: Sprachliche

Daten werden gesammelt, aber nicht verbreitet. Wenn die Daten nicht zugänglich gemacht werden (können), ist es allerdings schwierig, die Ergebnisse zu überprüfen.

Dieses Vorgehen ist vor allem deswegen so verbreitet, weil eine Kontrolle, was mit den Daten nach ihrer „Veröffentlichung“ im WWW geschieht, geradezu unmöglich ist. Zudem sind WWW-Daten weitaus leichter zugänglich als Daten, die in gedruckten Büchern veröffentlicht sind. Während hier der Erwerb oder zumindest das Aufsuchen einer Bibliothek (oder sogar eine Fernleihe, die mehrere Wochen in Anspruch nehmen kann) vonnöten ist, gibt es beim Zugang zu elektronischen Daten oftmals nicht einmal mehr die Barriere der Nutzerregistrierung.

In Form und Inhalt weichen WWW-Daten oftmals von klassischen, herkömmlichen Texten ab, von Texten beispielsweise, die als Romane oder wissenschaftliche Werke veröffentlicht werden. Autor*innen legitimieren sich hierbei im Idealfall durch eine besondere Kompetenz (Expert*innenwissen, Kreativität, Schreibstil etc.), die wiederum Rezipient*innen davon abhalten sollte, gedankliches Eigentum schlicht zu kopieren. Im WWW – besonders aber beim Surfen im Web 2.0 – entsteht möglicherweise gar kein Bewusstsein dafür, dass es sich hier um schützenswerte Werke handelt, selbst bei Beiträgen in Kommentarforen. Inhalt und Form scheinen oftmals keinen besonderen „Wert“ zu indizieren.

Heißt das nun, dass wir gesetzeswidrig handeln wenn wir sprachliche Daten für linguistische Analysen aus dem WWW kopieren um sie unter spezifischen Fragestellungen zu untersuchen?

Nein, denn es ist a) generell nicht ausgeschlossen, eine Einwilligungserklärung für Daten einzuholen. Das entspräche dem bei Gatto (2014) und Baroni et al. (2009) beschriebenen traditionellen, recht aufwändigen und eine hohe Frustrationstoleranz erfordernden Weg, weil

Antwortzeiträume lang werden können oder Antworten gar nicht eintreffen. Es gibt aber z. B. im Chat die Möglichkeit, bei Nutzer*innen direkt während des virtuellen Gesprächs anzufragen. Ebenso kann eine Erlaubnis stellvertretend bei Administrator*innen von Chat-Räumen oder Foren erfragt werden. Deren Kontaktdaten sind jeweils im Impressum der entsprechenden Seite aufgelistet. Dass über diesen „Umweg“ Kontakt zu den Nutzer*innen aufgebaut werden kann, ist nicht zu erwarten, da sich diejenigen, die Chat-Seiten etc. betreiben, dazu verpflichten, Daten von Nutzer*innen vertraulich zu behandeln. Da aber sprachliche Belege für wissenschaftliche Zwecke als Zitate gelten, ist es b) nicht gesetzeswidrig sie zu kopieren, wenn sie als Zitate ausgewiesen werden.

Aufgabe 1-1

Formulieren Sie ein Anschreiben an eine*n Administrator*in eines von Ihnen selbst gewählten Chat-Angebots (z. B. Knuddels, Jodel etc.) und bitten Sie darum, Daten von Nutzer*innen für Ihre wissenschaftliche Untersuchung verwenden zu dürfen.

1.3 Das Zitat als eine annehmbare Lösung

Wenn wir uns im Rahmen einer wissenschaftlichen Publikation mit sprachlichem Material inhaltlich auseinandersetzen, sind wir berechtigt, in unseren Veröffentlichungen darauf zu verweisen und es auch zu zitieren.

 Zulässig ist die Vervielfältigung, Verbreitung und öffentliche Wiedergabe eines veröffentlichten Werkes zum Zweck des Zitats, sofern die Nutzung in ihrem Umfang durch den besonderen Zweck gerechtfertigt ist. Zulässig ist dies insbesondere, wenn 1. einzelne Werke nach der Veröffentlichung in ein selbständiges wissenschaftliches Werk zur Erläuterung des Inhalts aufgenommen werden, 2. Stellen eines Werkes nach der Veröffentlichung in einem selbständigen Sprachwerk angeführt werden. (§ 51 UrhG)

Als Zitate kommen Inhalte von Webangeboten in Frage, die verschiedene Bereiche abdecken. Eine reine Zitatsammlung (ein Korpus, siehe Punkt 1.4.1) gilt allerdings nicht als Zitat.

1.3.1 So zitiert man sprachliche Beispiele

Die Quellenangabe für Beispielbelege enthält den Namen des Angebots (Forum, Chat, Soziales Netzwerk etc.), den Namen des*der Autor*in (alternativ ein Kürzel) und das Datum der Veröffentlichung. Optional ist die Uhrzeit der Veröffentlichung, deren Relevanz abhängig von der linguistischen Fragestellung ist.



Der **URL** (Uniform Resource Locator) wird umgangssprachlich auch als *Internet-* oder *Webadresse* bezeichnet.

Auf Grund ihrer Länge können URLs sperrig wirken und den Lesefluss beeinträchtigen. Es ist deshalb ratsam, nur den Namen des Angebots in das Kurzzitat aufzunehmen (1-1) und der wissenschaftlichen Arbeit ein Verzeichnis beizufügen, das die vollständigen URLs enthält, etwa nach dem Literaturverzeichnis (siehe dazu 1.5.2). Auf diese Vorgehensweise ist zu Beginn einer Arbeit (z. B. in der Einleitung) hinzuweisen.



Formulierungsvorschlag: In der vorliegenden Arbeit werden Internetquellen für sprachliche Belege in Kurzform angegeben, um den Lesefluss nicht zu beeinträchtigen. Die vollständigen URLs wurden auf einer Liste zusammengefasst, die sich im Anhang befindet.

Je nach Fragestellung kann es notwendig sein, sowohl sprachliche Belege aus gedruckten Quellen als auch aus digitalen Quellen anzugeben. Insbesondere wenn in diesem Zusammenhang Print- und Rundfunkmedien, die auch digital erscheinen und in Social-Media-Kanälen präsent sind, als Quellen herangezogen werden, sind die Kurzverweise auf digitale Quellen konsequent durch die Angabe der Top-Level-Domain, wie *.de* in (1-2) zu kennzeichnen. Natürlich sind auch die Namen der digitalen Ausgaben von Printmedien sprechend, z.B. Spiegel online oder Zeit online (siehe 1-3). Leser*innen können so ohne aufwendiges Suchen im Anhang bereits im Fließtext erschließen, ob der entsprechende Beleg gedruckten oder digitalen Quellen entnommen worden ist. Die Angabe der Uhrzeit der Veröffentlichung ist ein weiterer Hinweis dafür, dass es sich um eine digitale Quelle handelt (siehe dazu auch den Abschnitt zu Datum und Uhrzeit).

(1-1)