Helmut Feldweg & Ralf Kibiger & Christine Thielen

Zum Sprachgebrauch in deutschen Newsgruppen*

Mit der weltweiten Vernetzung von Rechnersystemen hat sich ein völlig neues Kommunikationsmittel für die unterschiedlichsten Sprachgemeinschaften eröffnet, das sich immer regeren Gebrauchs erfreut. In sogenannten NetNews, die einer überdimensionalen (Bildschirm–)Zeitung gleichen, können Teilnehmer zugleich Leser und Autor von Nachrichten sein. Zur Strukturierung der zahlreichen Nachrichten hat sich eine hierarchische Organisation in sogenannten **Newsgruppen** durchgesetzt, die bestimmte Themenbereiche abdecken, wie *Freizeit, Humor, Spiel, Sport, Soziales, Kultur, Politik* und *Computer*.

An der Universität Tübingen werden seit Juni 1992 Beiträge aus deutschen Newsgruppen gesammelt, um damit ein Korpus der deutschen Gegenwartssprache zu erstellen, das eine bisher wenig beachtete Textsorte enthält. Dieses **Newskorpus** eignet sich besonders gut für computerlinguistische Untersuchungen. Zum einen ermöglicht seine maschinenlesbare Form einen direkten Zugriff auf das Sprachmaterial, zum anderen bietet diese Textsorte einen linguistisch interessanten Untersuchungsgegenstand. Es wird ein inhaltlich breites Spektrum abgedeckt, und anhand der Texte lassen sich Sprachvarietäten beobachten, die vor allem unter dem formalen Aspekt der gesprochenen und geschriebenen Sprache zu erfassen sind (vgl. [Coseriu 1973]). In erster Linie handelt es sich hier um Schriftsprache, die aber aufgrund der Kommunikationssituation beeinflußt wird durch Merkmale gesprochener Sprache.

Kommunikationsbedingungen gesprochener Sprache (vgl. [Koch/Österreicher 1990]) wie Dialogizität, Spontaneität und Emotionalität liegen in vielen Newsgruppen vor. In manchen Newsgruppen beeinflußt sogar eine gewisse Privatheit und Vertrautheit der Kommunikationsteilnehmer die Äußerungen, besonders in Gruppen, an denen sich nur wenige Teilnehmer aktiv beteiligen und die deshalb schon sehr vertraut miteinander sind, auch wenn sie sich wahrscheinlich niemals persönlich kennengelernt haben. Beiträge aus Newsgruppen entstehen unter einer ähnlichen Kommunikationssituation wie der des Briefeschreibens und folgen in ihrem Aufbau der Briefstruktur. Davon weicht jedoch das Medium ab, das dabei verwendet wird, nämlich ein elektronisches. Der Zugang zum Internet ist an Universitäten für Studenten meist kostenfrei und die Postzustellung erfolgt wesentlich schneller als auf dem üblichen Postweg. Dies führt dazu, daß selbst Nachrichten, die lediglich aus einem Satz bestehen, verschickt werden und so die Kommunikationssituation der eines direkten Dialoges sehr nahe kommt. Dieses elektronische Medium bietet ganz andere Möglichkeiten bzw. schränkt auch einige Möglichkeiten ein (z.B. darstellbare Zeichen), was sich direkt auf den Sprachgebrauch auswirkt (Kapitel 5).

*erschienen in: Osnabrücker Beiträge zur Sprachtheorie (OBST), 50 (1995), S. 143–154.

Im folgenden geben wir einen kurzen Überblick über das in Tübingen erstellte Newskorpus (Kapitel 1–2), bevor wir auf unsere Untersuchungen zum Sprachgebrauch in Newsgruppen im einzelnen eingehen (Kapitel 3–5).

1 Das Tübinger Newskorpus

Das Korpus ist im Rahmen des am Seminar für Sprachwissenschaft der Universität Tübingen durchgeführten Forschungsprojektes *Korpusgestützte Entwicklung lexikalischer Wissensbasen (ELWIS)* entstanden (vgl. [Storrer u.a. 1993]). In diesem Projekt wurden deutschsprachige Textkorpora in einem Umfang benötigt, wie sie damals (1992) nicht allgemein zur Verfügung standen. Zur Behebung dieses Datenmangels wurden kurzerhand nächtlich alle über den News–Server am Zentrum für Datenverarbeitung der Universität Tübingen verbreiteten deutschsprachigen Newsgruppen archiviert. Nicht berücksichtigt werden bei der Archivierung alle diejenigen Gruppen, die aufgrund ihrer Bezeichnung auf fast ausschließlich nicht–natürlichsprachlichen Inhalt schließen lassen: das sind in erster Linie Gruppen mit Programmkode (Binär– oder Quellkode) und grafischen Daten. Die Archivierung begann im Juni 1992 und wird weiterhin fortgesetzt. Der Datenbestand hat inzwischen längst die GigaByte–Grenze überschritten.

Als empirische Grundlage für die in diesem Beitrag durchgeführte erste Untersuchung zum speziellen Sprachgebrauch in Newsgruppen greifen wir aus diesem Korpus den kompletten Jahrgang 1993 mit allen archivierten Newsgruppen heraus. In diesem Jahr standen uns über das Rechenzentrum die vollständigen Hierarchien der Newsgruppen *cl.**, *cnet.**, *de.**, *fido.**, *maus.**, *stgt.** und *zer.** zur Verfügung. Es handelt sich dabei um insgesamt 647 verschiedene Newsgruppen mit zusammen über 433000 News–Artikeln. Der sperrige Forschungsgegenstand belegt 733 Megabyte auf der Festplatte. Und die Datenmenge nimmt ständig zu: allein von Januar bis Dezember des Jahres 1993 hat sich das monatliche Datenaufkommen verdreifacht. Trotz des bereits erwähnten Ausschlusses offenkundig nicht–natürlichsprachiger Daten findet sich in diesem Korpus weiterhin ein erheblicher Anteil nicht–sprachlichen Materials. Um diese Störsignale herauszufiltern, wurden die jedem News–Artikel vorangestellten News–Header bei der sprachlichen Analyse nicht berücksichtigt, und vom verbleibenden Rumpf der News–Artikel wurden nur solche Zeilen verarbeitet, die sich ausschließlich aus den Zeichen des Alphabets und den Satzzeichen zusammensetzten. Mit diesem Verfahren konnten eine große Menge störenden Materials, wie Listen, Programmkode und vor allem Zeichnungen, wie sie sich häufig in den Signaturen der News–Artikel finden, eliminiert werden.

Bevor wir uns dem in den News verwendeten Wortschatz zuwenden, sei hier noch etwas über die Schreiber und Leser der News–Artikel gesagt. Soziologische Erhebungen zu diesem Bereich sind uns nicht bekannt, doch darf wohl davon ausgegangen werden, daß es sich überwiegend um junge und relativ gut ausgebildete Teilnehmer handelt. Der Anteil von Studenten, die heute vielerorts über einen einfachen und unentgeltlichen Zugang zum Internet verfügen, ist in diesen Gruppen sehr hoch. Allerdings nimmt auch im deutschsprachigen Raum der Anteil der kommerziellen Nutzer und der über Modem und Datex–J vernetzten Heimcomputernutzer stetig zu. Eine Analyse der Absenderangaben der News–Schreiber des Korpus liefert 43000 verschiedene Adressen, jeder Schreiber ist demnach mit durchschnittlich 10 Artikeln im Korpus vertreten. Klassifiziert man die Schreiber nicht nach Individuen, sondern nach den Rechnern bzw. lokalen Netzen, von denen die Artikel versandt wurden, dann lassen sich 13000 Herkunftsorte lokalisieren. Darüber, wer welche News liest bzw. abruft, läßt sich allenfalls spekulieren. Statistiken darüber werden nicht geführt, dem steht — vom in Deutschland geltenden Datenschutzrecht abgesehen — schon die anarchische Organisationsform der News–Verteilung entgegen. Es darf aber sicherlich davon ausgegangen werden, daß die Gemeinde der Leser die Anzahl der Schreiber um ein vielfaches übertrifft.

2 Wortschatz

Das 93er Newskorpus hat einen Umfang von 43 Millionen Wörtern. Insgesamt treten im Korpus 540000 verschiedene 'Wortformen' auf, wobei diese Notation von 'Wortformen' auch eine nicht zu vernachlässigende Anzahl von ad–hoc–Bildungen, Eigennamen, Abkürzungen, Schreibfehlern und fremdsprachigen Wörtern umfaßt. Zum Vergleich: die Datenbank des Centre for Lexical Information berücksichtigt 360000 Wortformen des Deutschen ([Burnage 1990]).

Bei der Suche nach für das Genre typischem Wortschatz bietet sich der Vergleich von Wortranglisten verschiedener Texttypen an. Dabei werden Häufigkeitslisten jedes Texttyps nach absteigender Häufigkeit sortiert und jedem Wort dieser geordneten Liste ein Rang zugewiesen: das häufigste Wort erhält den Rang 1, das zweithäufigste 2 usw. Schließlich werden die verschiedenen Ranglisten auf signifikante Abweichungen der Positionen einzelner Wortfomen untersucht (s. Anhang B).

Wir haben uns für einen Vergleich mit Texten der *Frankfurter Rundschau* entschieden, die uns für den Berichtszeitraum von Juli 1992 bis März 1993 ebenfalls in maschinenlesbarer Form zur Verfügung steht. Dieses Zeitungskorpus umfaßt 35 Millionen Wörter mit 680000 verschiedenen Wortformen.

Beim Vergleich der Ranglisten von News– und Zeitungskorpus zeigten sich die folgenden Auffälligkeiten:

*Out*wörter in den News, aber *In*wörter im Zeitungskorpus: Wörter wie *Kita, Ballett, Sondersitzung, Gastspiel, Weinfest, gratulierte, fuerchtete, erwirken, ererbt, Kalkar, Aspiranten, erbauten, zuenftig, einleitete, autowaschen* sind zwar hochfrequent im Zeitungskorpus, in den News jedoch praktisch nicht vertreten. Bei vielen dieser Formen handelt es sich um typische Begriffe der Lokalberichterstattung, einer Textsorte, die in einer Tageszeitung natürlich weit überdurchschnittlich vertreten ist. Anderes ist auf stilistische Eigenheiten der Zeitungsschreiber zurückzuführen, denn die Artikel einer Tageszeitung werden von einer vergleichsweise kleinen Gruppe von Schreibern verfaßt. Eher zufällig ist die hohe Frequenz von *Kalkar* in der Frankfurter Rundschau: der Schnelle Brüter selbst war 1993 kein Thema mehr (sonst wäre er auch in den News anzutreffen gewesen), doch hatte man in der Hanauer Brennelementefabrik noch mit für Kalkar vorgesehenen Brennelementen zu tun. Kein Zufall dürfte es allerdings sein, daß es in den News weitaus weniger *zünftig* zugeht.

Aufschlußreicher für den Sprachgebrauch in den NetNews ist eine Liste von *In*wörtern in den News, die in der Frankfurter Rundschau weit weniger vertreten sind: *ich, man, du, mir daß, mal, mich, also, schrieb, wohl, jemand, dir, meine, bin, muß, einfach, vielleicht, bitte, eigentlich, leider, mein, hast, wirklich, weiss, hab, warum, problem, paar, dich, meiner, natuerlich, eben, meinte.* Die vielen Pronomina der ersten und zweiten Person sind ebenso ein deutliches Zeichen für die dialogorientierte Struktur des Newskorpus wie die vielen Verbformen des Sagens und Meinens. Letztere sind auch für die hohe Frequenz von *daß* mitverantwortlich. Der starke Einfluß der gesprochenen Sprache zeigt sich auch in der sehr häufigen Verwendung Vagheit ausdrückender Formen wie *ziemlich, glaube, finde, irgendwie, irgendwo, denke, irgendwelche* usw. Trotz des dialogorientierten Charakters sind sich die Schreiber offensichtlich durchaus dessen bewußt, daß sie sich auch an ein großes Publikum wenden bzw. von einem solchen beobachtet werden. Eindeutige, falsifizierbare Aussagen werden häufig vermieden, zumal der Umgangston in manchen Newsgruppen harsch ist.

3 Zur allgemeinen Struktur von Artikeln in Newsgruppen

Obwohl Newsgruppen sehr stark dialogorientiert sind, handelt es sich dabei nicht um direkte Dialoge, bei denen die Teilnehmer gleichzeitig anwesend sind. Die Dialoge sind zeitversetzt, und daher ähnelt der grobe Aufbau eines Artikels dem Aufbau eines Briefes in traditioneller Form. Zu Beginn des Artikels werden sowohl persönliche Anreden bzw. Grußformeln (“Hi Inke!”, “Hallo Peter”) als auch Anreden oder Grußformeln allgemeiner Art (“Hi!”, “Hallo Ihr alle!”) benutzt. Am Ende eines Artikels stehen üblicherweise Abschiedsformeln und meistens der Autor des Artikels (“tschuess, Alex”). Abschieds– und Grußformeln sind häufig Fremdsprachen entnommen (“ciao”, “byebye”, “hi”), dabei der gesprochenen Sprache angepaßt (“tschau”). Das gleiche gilt für dialektale (“moin”, “tach”), und umgangssprachliche (“hallo”, “tschuess”, “tschoe”) Gruß–/Abschiedsformeln. Aber auch standardsprachliche Formeln lassen sich finden (“gruss”), oft auch in abgekürzter Form (“mfg” = mit freundlichen Grüßen). Abschieds– und Grußformeln sind die Wortformen, die im Vergleich zum Wortschatz der FR (s. Anhang B) sehr häufig vorkommen, was auf den verwendeten Briefstil zurückzuführen ist. Wird direkt auf einen bestimmten Beitrag bezug genommen, dann wird zu Beginn des Artikels (oft automatisch) die genaue Referenz angegeben, der dann Auszüge aus dem zitierten Artikel folgen, die meist durch spitze Klammer ‘>’ markiert sind (s. Anhang A).

Bei neuen Benutzern folgt nach der Begrüßung auch häufig der Hinweis, daß sie neue Benutzer sind. Dies geschieht offensichtlich mit der Absicht, andere Benutzer nachsichtig zu stimmen, falls die neuen Benutzer unbeabsichtigterweise die sog. “Netiquette” (Network Etiquette) verletzen, ein Verhaltenskodex für Usenet–Benutzer (vgl. [Salzenberg 1994], [Vielmetti 1994]).

4 Gesprochene Sprache in Newsgruppen

Die Dialoge wirken relativ spontan: jemand schreibt zu einem bestimmten Thema, was ihm gerade so einfällt, und ein anderer, der gerade diese Nachricht liest und dem vielleicht auch etwas zu diesem Thema einfällt, antwortet. Es wird versucht, die Situation eines nahen (distanzlosen) Dialoges nachzuempfinden, auch wenn das rein technisch nur eingeschränkt möglich ist. Dennoch wirkt sich dies unmittelbar auf den Sprachgebrauch aus, die gesprochene Sprache beeinflußt die geschriebene Sprache. Die häufigsten Auffälligkeiten sind:

- Verschmelzungen von zwei Wortformen zu einer (mit Markierung des Schwas –e im Auslaut):

`gibt es, war es, geht es...`	→ `gibts, wars, gehts, ...`
`wenn es`	→ `wenns`
`kannst du, hast du, willst du...`	→ `kannste, haste, willste, ...`
`kriegst einen, hast einen`	→ `kriegst'n, hast'n`
`nicht einmal`	→ `nichtmal`

- Verkürzungsformen:

`eine, einen, einem ...`	→ `ne, nen, nem, ...`
`glaube, frage, habe...`	→ `glaub, frag, hab, ...`

- Interjektionen:

  `hmpf, he, nja, na, aeh, hmm, nu, eh, ach, oh, ...`

- Dialektausdrücke:

  `moin, tach, bischen, denne, pladde, dat, wat, ...`

- Verben des Sagens (Schreibens), Meinens und Denkens:

  `meinst, meintest, findest, glaubst, schreibst, sprichst, finde, ...`

Auffällig ist auch eine häufige Verwendung der Indefinitpronomina bzw. Adverbien, die mit *irgend–* gebildet werden. Wir vermuten dahinter die Absicht, sich lieber etwas vage auszudrücken, weil die Benutzer nie genau wissen, wer gerade die Nachricht liest, außer dem, der vielleicht direkt angesprochen wird, weil man sich auf einen früheren Beitrag bezieht. Es scheint uns auf jeden Fall eine Verwendung zu sein, die eher dem umgangssprachlichen (gesprochenen) Sprachgebrauch angehört. In wissenschaftlichen Abhandlungen wäre eine Verwendung von *irgendwelch–*, wie im folgenden Beispiel, kaum denkbar:

```
Die neckischen Anspielungen auf irgendwelche Saetze von irgendwelchen
Leuten gibt es kaum noch [...]

                    [Newsgruppe:  zer.z-netz.miteinander.kontakte]
```

Für Dialoge ist weiterhin kennzeichnend der hohe Anteil von Fragen, Aufforderungen und Bitten ("bitte", "danke"). Dies kann auch für Newsgruppen festgestellt werden.

5 Newsgruppen–spezifische Sprachelemente

Spezifisch für die Sprache in Newsgruppen ist die ausgiebige Benutzung von Zitaten. Dies wird durch eigens dafür entwickelte Software ("newsreader") unterstützt, indem der zu beantwortende Artikel automatisch kopiert und durch spitze Klammern eingerückt wird. Außerdem wird das Zitat durch eine genaue Angabe identifiziert. Aufgrund dieser Erleichterung der Beantwortung von bereits Gesagtem/Geschriebenem werden Dialoge häufig in folgender Form dargestellt, indirekte Rede kann somit weitgehend vermieden werden:

```
Am 01.01.93 meinte {ANYBODY@ANYWHERE}
zu dem Thema 'Re^2:Reicht es jetzt?' :
>>  Hey Joerg !
Hai Juergen !   Frohes neues auch an all die anderen hier im Brett !!
>>  Ja, dann wollen wir mal alle in die Haende spucken und
>>  frisch ans Werk gehen, nich. Auch Udo, denn ich da
>>  wohl etwas scharf angegangen hab am letzten Tag des
>>  Jahres, vielleicht hatver ja einen Korn oder Aeppelwoi
>>  drauf getrunken .
Nun, also das muss ich aber leider auch sagen, war wohl ein bisschen
SEHR scharf. Aber, was soll's ... trinkt euch einen zusammen und das
Problem ist geritzt.
>>  Ist denn schon jemand hier, oder seid Ihr alle noch
>>  fleissig am Feiern ??
Feiern ?? Haette ich besser mal gemacht ... aber meinereiner musste ja
unbedingt in den Schnee fahren ueber die Feiertage.
War echt aetzend, kalt und ... kalt und noch viel kaelter ...
nee, da lob ich mir unser nass--kaltes Sauerland *grins* Muss aber
auch sagen, ich habe einen grossen Bogen um dieses beruehmt
beruechtigte Landsberg gemacht ... hinterher steht der Martin (apropo
... lebt der etwa noch ??? ) mit seinem Golf hinter einer Ecke und
lauert nur auf mich ... nene, das war mir dann doch zu gefaehrlich
hehe ...
                Radkaeppchen und der boese boese Golf    hihi ...
>>  Also aufhoeren und schnell an die Klimperratur und
>>  Toene geschwungen, mit Rahmspinat, ohne Rahmspinat und
```

```
>>  moeglichst vielen Pizza's !!
Womit soll ich den aufhoeren ? Mache doch garnix ... (Unschuld
heucheln ... dupdidu ...).
                Aber, den Spinat lassen wir dann aber mal weg, oki ??
>>  Gruss
>>  Juergen

                Ciao ... Joerg
```

Es sieht so aus, als ob der Dialog von beiden Seiten gleichzeitig geführt wird, die Kommentare von Jörg sind jedoch zeitversetzt zu einem Beitrag von Jürgen, der automatisch in diesen Artikel hereinkopiert wurde und aus dem Jörg nur einige Äußerungen ausgewählt hat, auf die er dann näher eingeht. Der Informationsaustausch ist innerhalb eines Artikels einseitig, denn Jörg reagiert auf bereits gemachte Äußerungen von Jürgen, aber Jürgen kann nicht unmittelbar auf Jörgs Kommentare reagieren, dazu hat er erst wieder in einem neuen Artikel die Gelegenheit. Dies ermöglicht einen höheren Grad an Reflektiertheit, die Äußerungen sind demnach nicht ganz so spontan wie in gesprochenen face–to–face–Dialogen. In den Artikeln, die wir stichprobenartig gelesen haben, zeigt sich dann auch, daß in Newsgruppen eher selten abgebrochene oder unvollständige Sätze vorkommen, was wir jedoch bisher nicht statistisch nachweisen können, da die Texte in keiner Weise syntaktisch annotiert sind (z.B. mit Wortarten) und uns kein Korpus der gesprochenen Sprache als Vergleich zur Verfügung steht.

Wie bereits erwähnt, wird in Newsgruppen versucht, einen direkten Dialog nachzustellen. Da der visuelle Kanal einer face–to–face–Kommunikation bei Newsgruppen jedoch fehlt, wird von manchen Benutzern versucht, zumindest einen Teil der Mimik durch im ASCII–Code darstellbare Symbole, sogenannte "smilies", auszudrücken. Hierbei gibt es Symbole, welche die äußere Erscheinung des Schreibers wiedergeben, z.B.:

- daß dieser eine Brille oder einen Bart trägt:

  `8-), :-{`

- Symbole, welche Gefühlszustände o.ä. wiedergeben sollen:

  Freude, Lachen:

  `:-)`

  Mißbilligung, Trauer:

  `:-(`

  Wut:

  `>:-O`

  Lächerlichkeit:

  `<:-)`

Sehr häufig finden auch in deutschsprachigen Newsgruppen englische Abkürzungen Verwendung:

```
IMO        → in my opinion (meiner Meinung nach)
BTW        → by the way (nebenbei gesagt)
PM         → personal mail (persönliche elektronische Post)
```

Die häufige Verwendung von Abkürzungen hat einen ihrer Ursprünge in den technischen Gegebenheiten des Mediums: viele Benutzer, die nicht über einen unentgeltlichen Zugang zum Internet verfügen, waren (und sind) über relativ langsame Schnittstellen und Telefonleitungen mit dem Netz verbunden. Die Benutzung der englischen Sprache kann wohl dadurch erklärt werden, daß die Entwicklung und der Aufbau von Computernetzwerken sowie die Nutzung durch eine breite Öffentlichkeit zuerst in den USA stattfand. Die ersten deutschen Benutzer haben wahrscheinlich an internationalen, d.h. englischsprachigen Newsgruppen teilgenommen und bestimmte Ausdrücke dann nach dem Aufbau deutscher Newsgruppen weiterbenutzt. Ähnlich oft finden sich dann auch andere englischsprachige Ausdrücke aus dem Bereich der 'Newsgruppenkultur' z.B. *snail mail* oder *smail* für *Briefpost*, *flame* für *polemischer Kommentar*, *posten* für *senden* usw. Englische Verben werden dabei häufig nach deutscher Flexionsmorphologie flektiert (z.B. *poste, gepostet*), ein Phänomen, das sich vor allem in der Computersprache beobachten läßt.

Die Verwendung elektronischer Medien wie der NetNews bringt für den deutschen Sprecher/Schreiber einen Nachteil mit sich: Die Verwendung von Umlauten und Eszett (ß) wird von mancher Software nicht unterstützt, weil sie in erster Linie für die englische Sprache entwickelt wurde. Sie werden in den Beiträgen, so gut es geht, vermieden und durch ihre Entsprechungen (a/o/u + e, ss) ersetzt, wobei sich immer mehr eine Darstellung des *ß* durch *sz* durchsetzt (z.B. *dasz, musz*).

Zusammenfassend läßt sich sagen, daß der Sprachgebrauch in Newsgruppen sehr stark von der Kommunikationssituation des Dialoges beeinflußt wird und viele Merkmale gesprochener Sprache aufweist. In ihrer Struktur ähneln die Artikel der traditionellen Briefform, es werden jedoch sprachliche Mittel verwendet, die durch das elektronische Medium unterstützt werden. Ein Vergleich mit einem entsprechenden Korpus von Briefen der deutschen Umgangssprache könnte weiteren Aufschlußüber den unterschiedlichen Sprachgebrauch in News und Briefen geben. Der im Anhang wiedergegebene News–Artikel illustriert, daß der Sprachgebrauch in den News selbst Gegenstand von Diskussionen in Newsgruppen ist.

Literatur

[Burnage 1990] Burnage, Gavin: *CELEX. A Guide for Users.* Nijmegen: CELEX – Centre for Lexical Information, 1990.

[Coseriu 1973] Coseriu, Eugenio: *Einführung in die strukturelleBetrachtung des Wortschatzes.* Tübingen: Narr, 1973.

[Koch/Österreicher 1990] Koch, Peter/Österreicher, Wulf: *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch.* Tübingen: Niemeyer, 1990. (Romanistische Arbeitshefte; 31)

[Salzenberg 1994] Salzenberg, Chip: *What is Usenet?* Regelmäßig in den Newsgruppen news.announce.newusers, news.admin.misc und news.answers versandter Beitrag. Letzte Änderung vom 5. Juli 1994 von Mark Moraes.

[Sander–Beuermann 1993] Sander–Beuermann, Wolfgang: *Informationskarusell. Usenet: Lese– und Schreibgewohnheiten der Benutzer.*. In: iX Multiuser Multitasking Magazin, Oktober 1993, S. 96–109.

[Storrer u.a. 1993] Storrer, Angelika/Feldweg, Helmut/Hinrichs, Erhard: *Korpusunterstützte Entwicklung lexikalischer Wissensbasen.* Sprache und Datenverarbeitung 17 (2) 1993, S. 59–72.

[Vielmetti 1994] Vielmetti, Edward: *What is Usenet? A second opinion.* Regelmäßig in den Newsgruppen news.announce.newusers, news.admin.misc und news.answers versandter Beitrag. Letzte Änderung vom 16. August 1994 von Mark Moraes.

Anhang

A Beispiel–Artikel einer News–Gruppe (soc.culture.german)

```
Article: 12312 of soc.culture.german
Path: newsserv.zdv.uni-tuebingen.de!news.belwue.de!news.dfn.de!zeus.rbi...
From: schrod@iti.informatik.th-darmstadt.de (Joachim Schrod)
Newsgroups: soc.culture.german
Subject: Re: Are Germans more blunt than Americans
Date: 23 Mar 1994 15:15:50 GMT
Organization: TH Darmstadt, FG Systemprogrammierung
Lines: 61
Distribution: world
Message-ID: <2mpmf6$1om@rs18.hrz.th-darmstadt.de>
References: <2mn829$uhh@rs18.hrz.th-darmstadt.de>
<1994Mar23.073358.13088@lmpsbbs.comm.mot.com>
NNTP-Posting-Host: spice.iti.informatik.th-darmstadt.de

In article <1994Mar23.073358.13088@lmpsbbs.comm.mot.com>,
michaelr@ecs.comm.mot.com (Michael Roetzer) writes:
> In article <2mn829$uhh@rs18.hrz.th-darmstadt.de>
schrod@iti.informatik.th-darmstadt.de (Joachim Schrod) writes:
> >
> >Hmm, wenn ich mir so die Ausarbeitungen des letzten Seminars ansehe,
> >dass ich veranstaltet habe; mit den Deutsch-Kenntnissen unserer
> ?????                 ???? ????             ?
> >Studenten -- die ja immerhin Abitur haben -- ist es auch nicht
>            ??                              ??
> >besonders weit her. (Wobei ich explizit deutsche Studenten meine,
> >nicht ihre auslaendischen Kommilitonen.)
> >
> >Eine Arbeit war dabei, da war ich froh, als ich nach zwei Seiten den
>                      ?????????????????????????????????????
> >ersten vollstaendigen Satz vorgefunden habe. Von korrekter
> >Kommasetzung traeume ich nur... Bzw. von einer Einschraenkung der
>                               ?????????
> >Anglizismen auf ein vernuenftiges Mass (d.h., auf Fachtermini)...
>                                                     ???????????  ??
>
> Eigentor?

Denke ich nicht... das ist aber evtl. eine interessante Frage: Wie
drueckt man sich in den News aus?
```

News ist fuer mich ein Medium, in dem ich mich umgangssprachlich
ausdruecke. Ausserdem ueberarbeite ich meine Texte dort selten --
schliesslich ueberarbeite ich ja auch keine Beitraege, die ich an
einer Kneipentheke von mir gegeben habe. [<-- Beachte bitte die
Benutzung des Perfekts, das ist hessische Umgangssprache. :-)] Das
ist in meinen Augen auch sehr sinnvoll, da dadurch der
Kommunikationscharakter von News gegenueber dem Mitteilungscharakter
betont wird. In einem Vortrag druecke ich mich wiederum anders aus.

Falls ich mich sehr missverstaendlich -- oder 'mal wieder zu flapsig
fuer einige empfindliche Gemueterchen -- ausgedrueckt habe, kann ich
das in einem Followup klarstellen. (Was ich z.B. in einem anderen
Artikel in diesem Thread auch hoffentlich gemacht habe.) [<-- Beachte
bitte die Benutzung von 'Thread', wo ich in einer
nicht-jargon-behafteten Umgebung eher Diskussionsstrang benutzen
wuerde. Ich moechte mich strikt von der Deutschtuemelei distanzieren,
die unbedingt _alles_ ins Deutsche uebersetzen will. IMO muss die
Auswahl der Begriffe auf das Medium und den Leser/Zuhoerer-Kreis
angepasst werden, ich kann also in den News sehr wohl einen
restringierten Code benutzen. Btw, als Gedankenstriche benutzt man
wirklich Halbgeviertstriche und normalen Wortzwischenraum, cf.
Tschichold.]

Eine Seminarausarbeitung ist aber etwas ganz anderes. Zumindest ich
habe an eine wissenschaftlich gepraegte Arbeit andere Ansprueche als
an einen locker-flockig runtergehackten News-Artikel. Sowohl
inhaltlich als auch sprachlich... (Falls Du ein Beispiel haben
willst, kannst Du Dir ja meine Antwort auf den Artikel von Uwe
Nimscheck ansehen.)

Ach ja, einige Deiner Fragezeichen sind definitiv falsch plaziert.
Dafuer hast Du andere Stellen uebersehen... ;-)

--
Joachim
Darmstadt, Germany
<schrod@iti.informatik.th-darmstadt.de>

B Wortformen mit hoher Abweichung der Ränge in den beiden Korpora (News, FR)

100 Wortformen, ohne Computerfachausdrücke und englische Wortformen, geordnet nach absteigendem Grad der Abweichung; Groß–/Kleinschreibung wurde nicht berücksichtigt.

No	Wortform	Rang/News (a)	Rang/FR (b)	Abweichung (|a-b|)/a
1	muss	108	96521	892
2	dass	41	28220	687
3	hallo	77	23828	308
4	ciao	334	98468	293
5	moin	381	-	261
6	solltest	549	-	181
7	hi	216	38237	176
8	mfg	622	-	159
9	meinst	645	-	154
10	tschuess	669	-	148
11	tschau	731	-	135
12	gibts	432	51050	117
13	dasz	906	-	109
14	schonmal	934	-	106
15	ne	244	23041	93
16	nich	923	77326	82
17	schriebst	1210	-	81
18	deinem	576	43736	74
19	deiner	375	28187	74
20	hmpf	1339	-	73
21	glaub	1356	-	72
22	gruss	174	12399	70
23	hast	136	9582	69
24	brauchst	925	64367	68
25	bekommst	1492	-	66
26	meintest	1503	-	65
27	nen	471	31295	65
28	bischen	1527	-	64
29	frag	1536	-	64
30	kannst	204	13352	64
31	deine	199	12653	62
32	sowas	261	16343	61
33	hab	149	9228	60
34	nem	1305	80420	60
35	grosse	627	37994	59
36	irgendjemand	1199	72491	59
37	poste	1665	-	59
38	findest	1605	93105	57

No	Wortform	Rang/News (a)	Rang/FR (b)	Abweichung (|a-b|)/a
39	glaubst	1693	97571	56
40	schreibst	1769	-	55
41	gepostet	1786	-	54
42	obige	1806	-	54
43	folgendes	251	13523	52
44	koenntest	1868	-	52
45	wars	1479	79328	52
46	tach	1898	-	51
47	wenns	1818	94591	51
48	benutze	741	37868	50
49	haettest	1512	77980	50
50	du	21	1046	48
51	dir	94	4339	45
52	obigen	2052	96244	45
53	verkaufe	1076	50423	45
54	letztens	2197	-	44
55	soetwas	2182	-	44
56	muessten	1938	84484	42
57	immernoch	2335	-	41
58	danke	387	16211	40
59	deines	2243	93526	40
60	wirst	705	29139	40
61	hattest	2488	-	39
62	kenn	2178	88117	39
63	willst	547	22403	39
64	tschoe	2610	-	37
65	wuerdest	2315	90190	37
66	dein	281	10531	36
67	gehts	1695	62243	35
68	habs	2712	-	35
69	sagst	2836	-	34
70	deinen	530	18039	33
71	denne	2896	-	33
72	gabs	2956	-	32
73	waers	2996	-	32
74	wirds	2997	-	32
75	garnicht	1290	41446	31
76	kittel	989	31728	31
77	nimmst	3059	-	31
78	haemmernde	3220	-	30
79	he	642	20310	30
80	kennst	1962	61987	30
81	nichtmal	3138	-	30
82	schnipp	3166	-	30

No	Wortform	Rang/News (a)	Rang/FR (b)	Abweichung (|a-b|)/a
83	waehre	3207	-	30
84	weiss	147	4698	30
85	zeile	757	23492	30
86	nimm	1509	46744	29
87	scheinst	3318	-	29
88	bescheuert	3384	-	28
89	darfst	2209	64295	28
90	kannste	2334	70004	28
91	pladde	3383	-	28
92	serbian	3363	-	28
93	thanx	3366	-	28
94	haste	2535	72592	27
95	machst	1400	39357	27
96	musz	3522	-	27
97	tipperte	3478	-	27
98	waaaahnsinn	3549	-	27
99	boese	3075	83334	26
100	dich	166	4605	26