

- Temporalpartikeln (z.B. *erst, noch, schon*);
- Antwort- oder Satzpartikeln (z.B. *ja, doch, eben*);
- Vergleichspartikeln (z.B. *wie, als*);
- Interjektionspartikeln (z.B. *au, oh je*);
- Negationspartikeln (z.B. *kein, nicht*);
- Infinitivpartikel *zu*.

Wie kaum eine andere Wortart beziehen die Partikeln ihre Bedeutung durch ihren Kontext. Die Partikeln tragen nichts zur propositionalen Bedeutung einer Äußerung bei, wie man an dem folgenden Beispiel sieht:

- (17) Was macht Peter jetzt (*eigentlich*)?

Auch ohne die Modalpartikel *eigentlich* ist der Satz als Frage über Peeters momentane, z.B. berufliche, Aktivitäten verständlich. Die Partikel erfüllt hier die Funktion, die Frage als Eröffnung eines neuen Themas zu markieren, z.B. in einem Dialog wie dem folgenden:

- (18) A: Stell dir vor, da steht: 100 km Stau! – B: Na, da wirds wieder gekracht haben. – A: Sag mal, wie hoch ist man eigentlich versichert, wenns mal so richtig kracht?<sup>96</sup>

Sie hat hier also gesprächssteuernde Funktion und sichert außerdem die Kohärenz, die bei einem abrupten Themenwechsel sonst gefährdet wäre. Modalpartikeln können außerdem dazu verwendet werden, um die Haltung des Sprechers zum Beispiel zum Wissen oder den Haltungen der Gesprächspartner zu signalisieren:

- (19) Und sie bewegt sich *doch*.  
 (20) Das kann schon mal etwas wackeln. Sie bewegt sich *ja*.

In Beispiel (19) signalisiert der Sprecher seine Annahme, dass die Gesprächspartner seine Behauptung (bisher) nicht teilen, in Beispiel (20) hingegen wird signalisiert, dass das Behauptete auch den Gesprächspartnern bekannt ist. Die Redundanz der Äußerung wird dadurch abgemildert.

Die Analyse und Beschreibung von Partikeln ist eine besondere Herausforderung für die theoretische Linguistik, die Lexikographie, die maschinelle Sprachverarbeitung und für die Sprachlehre<sup>97</sup>.

<sup>96</sup> Dieses leicht modifizierte Beispiel entstammt Thurmair (1989), S. 176.

<sup>97</sup> Wir gehen im Abschnitt zur Sprachlehre u.a. auf eine Arbeit zur didaktischen Vermittlung des Gebrauchs von Modalpartikeln ein.

Viele Partikeln stellen die theoretische Linguistik und Lexikographie vor schwierigere Aufgaben: Da sie nichts zur Proposition einer Äußerung, in der sie auftauchen, beitragen, müssen die kontextreferentiellen Funktionen dieser sprachlichen Einheiten bestimmt und beschrieben werden. Dies muss in so allgemeiner Weise geschehen, dass möglichst alle Verwendungsweisen bzw. Verwendungssituationen mit dieser Beschreibung abgedeckt werden. Die Gefahr einer solchen generischen Beschreibung ist, dass sie zu allgemein und damit wertlos wird. Will man andererseits das Spezifische der Verwendungskontexte aller Modalpartikeln erfassen, läuft man Gefahr, das Gemeinsame aller Verwendungsinstanzen in den Details zu verlieren.

Das Problem der Analyse von Partikeln stellt sich in verschärftem Maße bei der maschinellen Analyse natürlicher Sprache. Die spezifische „Bedeutung“ bzw. ihre pragmatische Funktion kann nur erfasst werden, wenn Wissen über den Kontext der jeweiligen Äußerung vorhanden ist. Selbst wenn dieses Kontextwissen nicht in ein Computerlexikon gehört, so doch zumindest eine lexikalische Beschreibung, die Angaben zu den möglichen Vorkommenskontexten umfasst.

Viele lexikalische Elemente der Partikelklasse gehören mehreren Unterklassen an und einige darüber hinaus auch anderen Wortklassen<sup>98</sup>. Die Verwendung dieser lexikalischen Elemente muss einerseits von einer ander abgegrenzt, andererseits miteinander in Beziehung gesetzt werden. Einige Typen von Partikeln, z.B. die Modalpartikeln, treten zudem in zahlreichen Kombinationen auf. Maria Thurmair<sup>99</sup> listet weit über 100 Kombinationen auf, von denen viele aber nur eingeschränkt akzeptabel seien.

All diese Aspekte von Partikeln machen diese zu einem besonders guten Gegenstand für korpuslinguistische Untersuchungen.

Umfangreiche linguistische und lexikologische Studien zu den Partikeln von Wolfgang Weydt erschienen 1979 und Anfang der 1980er Jahre<sup>100</sup>. Gerhard Helbig<sup>101</sup> widmet den Partikeln ein eigenes Wörterbuch. Darüber hinaus ist die Frage der angemessenen Übersetzung der Abtönungspartikeln in eine andere Sprache ein wichtiges Problem. König, Stark und Requardt füllen eine Lücke mit ihrem deutsch-englischen Spezialwörterbuch zum Wortschatzbereich der Adverbien und Partikeln<sup>102</sup>.

<sup>98</sup> So kann z.B. *doch* als Antwortpartikel und als Modalpartikel entsprechend der Klassifikation von Helbig und als Konjunktion verwendet werden.

<sup>99</sup> Vgl. Thurmair (1989).

<sup>100</sup> Vgl. Weydt (1979), Weydt (1983) u.a.

<sup>101</sup> Helbig (1994).

<sup>102</sup> Vgl. König et al. (1990).

Mittlerweile sind zahlreiche Monographien und detaillierte Arbeiten auch zu einzelnen Partikeln oder zu Partikelgruppen erschienen<sup>103</sup>. Von besonderem Interesse sind dabei die Bedeutungs- oder Funktionskontraste nah verwandter Partikel<sup>104</sup>. Ulrike Nederstigt kritisiert aber zu Recht, dass linguistische Arbeiten zu den Partikeln allzu oft auf erfundene oder konstruierte Beispiele aufbauen. Diese mögen als Testmaterial zur Ermittlung von Akzeptabilitätsurteilen oder zur Ermittlung von Kontrasten geeignet sein, erscheinen aber in vielen Fällen unnatürlich und können nicht das wiedergeben, was in authentischen Gesprächen geschieht<sup>105</sup>. Die Situation hat sich in den letzten Jahren gebessert, was auch durch die bessere Verfügbarkeit von Korpora geschriebener, vor allem aber auch gesprochener Sprache bedingt ist.

Thurmair, die in einer Monographie die Kombinierbarkeit von Modalpartikeln untersucht<sup>106</sup>, erwähnt einen Vorschlag von Collins aus dem Jahre 1938:

It would be an alluring task to pick out in German a certain number of simple particles, combine them in pairs or triplets or even larger groups, and try to discover which groups are the most commonly used, which have the most characteristic functions, and which cannot be combined with which, or at least not in a particular order. (Zit. nach Thurmair (1989), S. 203).

Tatsächlich ist durch die Existenz sehr großer Korpora und der entsprechenden Werkzeuge zu ihrer Analyse nun die Möglichkeit gegeben, zumindest eine der von Collins gestellten Fragen zu beantworten, nämlich die nach der Vorkommenshäufigkeit, Reihenfolge und Bindungsstärke einzelner Partikelkombinationen.

Collins schneidet außerdem wichtige Fragen an, mit deren Beantwortung erst begonnen wurde:

- Sind die Restriktionen, denen die Kombinierbarkeit von Abtönungspartikeln unterliegt, systematisch zu beschreiben? Dies betrifft sowohl die Möglichkeit des Kovorkommens zweier Abtönungspartikeln als auch die Reihenfolge ihres Auftretens. Die allgemeinste Beschränkung des Kovorkommens ist dadurch gegeben, dass zwei Partikeln, deren Modus inkompatibel ist, nicht zusammen auftreten.<sup>107</sup>
- Wenn zwei Abtönungspartikeln miteinander in einem Satz vorkommen können, ist ihre Abfolge durch Prinzipien beschreibbar, die sich

<sup>103</sup> Einen guten und aktuellen Überblick gibt das Literaturverzeichnis in Möllering (2004).

<sup>104</sup> Z.B. die der Gradpartikeln *auch* und *noch*, die im Mittelpunkt der Monographie von Ulrike Nederstigt 2003 stehen.

<sup>105</sup> Vgl. Nederstigt (2003), S. 12.

<sup>106</sup> Vgl. Thurmair (1989).

<sup>107</sup> Vgl. Helbig (1994), S. 76 und Thurmair (1989), S. 204f.

aus ihren Merkmalen ergeben? Helbig bildet zwar Distributionsklassen für eine Reihe der Partikeln, um deren Reihenfolgebeziehung bei der Kettenbildung zu erfassen, die Belege für seine Hypothesen sind allerdings wenig überzeugende Eigenkonstruktionen.<sup>108</sup>

Thurmair verwendet eine Menge von semantischen Merkmalen, nach denen die einzelnen Partikeln im ersten Teil ihrer Arbeit klassifiziert werden. Mit diesen Mitteln sollen Selektionsrestriktionen und Kombinationspräferenzen beschrieben werden. Diese semantischen Merkmale gehen in die Partikelkombinationen ein.

Die Einzelbedeutungen der Partikeln werden nach Auffassung der Autorin zur Gesamtbedeutung der Partikelkombinationen addiert:

Es soll hier davon ausgegangen werden, daß eine Kombination der Partikel A mit der Partikel B eine Addition ihrer Bedeutung und damit ihrer Merkmale bedeutet; d.h. also, daß sich die Kombinationen in ihre Einzelteile zerlegen lassen. (Thurmair (1989), S. 205)

Thurmair führt dieses Programm in ihrer Monographie dadurch aus, dass sie die Bedeutung und Funktion der einzelnen Modalpartikeln<sup>109</sup> und im Anschluss daran die akzeptablen, bedingt akzeptablen und inakzeptablen Kombinationen beschreibt<sup>110</sup>. Sie beschließt ihre Arbeit mit einer Synopse der Partikelkombinationen (S. 278), mit einer Übersicht über die Distribution der Kombinationen über verschiedene Satztypen (S. 282, Tabelle 13) und einer Übersicht über Stellungenregeln für einzelne Modalpartikeln (S. 285-289). Ob das Postulat der additiven Bedeutung von Partikelkombinationen durchzuhalten ist, bleibt unklar. Hier setzt die Kritik von Lemnitzer<sup>111</sup> an, der davon ausgeht, dass Partikelkombinationen als komplexe Lexeme nicht transparent und analysierbar sind. Lemnitzer 1997 präsentiert als Fallstudie die Kombinationen mit der Modalpartikel *denn*, Lemnitzer 2001 untersucht, ähnlich wie Thurmair, systematisch alle Kombinationen, konzentriert sich aber bei den Einzeldarstellungen auf die häufigsten. Er trifft allerdings keine Aussage zur psychologischen Plausibilität seiner Vermutung, dass es sich hier um Mehrwortlexeme handelt. Im Zentrum seiner Arbeit steht vor allem die Analyse und (computer-)lexikographische Erfassung dieser komplexen sprachlichen Einheiten.

Die Aneignung des korrekten Gebrauchs von Partikeln im Zuge des Erstspracherwerbs verfolgt Ulrike Nederstigt am Beispiel der Grad-

<sup>108</sup> Vgl. Helbig (1994), S. 75f.

<sup>109</sup> Abschnitt 2, S. 94-202.

<sup>110</sup> Abschnitt 3, S. 203-284.

<sup>111</sup> Vgl. vor allem Lemnitzer (2001).

partikeln *auch* und *noch*<sup>112</sup>. Der Erwerb des komplexen sprachlichen Wissens, das für die korrekte Verwendung der Partikeln notwendig ist, demonstriert die Autorin am Gebrauch dieser Partikeln durch erwachsene Sprecher. Verschiedene linguistische Versuche, die phonologischen, syntaktischen, semantischen und diskursiven Aspekte des Gebrauchs dieser Partikeln zu beschreiben und zu erklären, stellt die Autorin auf den Prüfstand. Sie verwendet hierfür mehrere Korpora gesprochener Sprache, von denen einige den Sprachgebrauch Erwachsener, andere den Sprachgebrauch eines Mädchens in der Phase zwischen dem zweiten und vierten Lebensjahr wiedergeben<sup>113</sup>. Die Verwendung von Korpora, vor allem solcher der gesprochenen Sprache, erlaubt ihr, gemessen an den bisherigen linguistischen Arbeiten zum Thema, einen neuen Blick auf die von ihr beschriebenen Partikeln<sup>114</sup>. Die wichtigsten Erkenntnisse sind:

- Die Gradpartikel *noch* hat mehr Bedeutungen als gemeinhin angenommen, es werden insgesamt neun Bedeutungen unterschieden:
  - additiv (*Nimm dir noch einen Nachtsch, bitte!*);
  - additiv, weiteres Element einer Menge (*Davor habe ich noch einen Termin.*);
  - additiv, vor einer Wende (*Damals haben wir noch ohne fließend Wasser gewohnt.*);
  - temporal (*Der Platz ist noch frei.*);
  - temporal, perfektiv (*Wir können uns gern noch in diesem Monat treffen.*);
  - temporal, mit abnehmender Menge von Objekten (*und jetzt brauche ich noch eine Siebenerleiste.*);
  - mit *einmal*, repetitiv (*Dann sind wir noch einmal Karussell gefahren.*);
  - mit *einmal*, restititiv (*Knöpf die Jacke am besten noch einmal auf.*);
  - mit *einmal*, additiv (*Wir sollten uns dann nochmal einen Tag Zeit nehmen, wenn das heute nicht klappt.*)<sup>115</sup>.

<sup>112</sup>Vgl. Nederstigt (2003). Die Autorin spricht von „focus particles“, dies entspricht aber der von Helbig definierten Klasse der Gradpartikeln.

<sup>113</sup>Zu den verwendeten Korpora vgl. Kapitel 3.2, S. 80–83 (Korpora der Erwachsenensprache) und Kapitel 7.2, S. 211 (Korpus der Kindersprache) in Nederstigt (2003).

<sup>114</sup>Kapitel 6 ist überschrieben mit „A fresh look at focus particles“.

<sup>115</sup>Vgl. S. 100–106. Die hier gewählten Beispiele sind erfunden, Nederstigt präsentiert authentische, aber auch etwas komplexere Beispiele.

Der Kontrast zwischen diesen Bedeutungen korrespondiert mit Unterschieden in den phonologischen und syntaktischen Merkmalen der Partikel bzw. mit Unterschieden in den Verwendungskontexten<sup>116</sup>.

- Bei der Gradpartikel *auch* unterscheidet die Autorin zwischen einer betonten Variante und einer unbetonten Variante. Die betonte Variante weist Merkmale einer Antwortpartikel auf. Ihr Gebrauch signalisiert, dass einer vorhergehenden positiven (oder negativen) Antwort eine weitere positive (oder negative) Antwort hinzugefügt wird, wie in dem folgenden Beispiel:

(21) Mitte der Woche habe ich AUCH nicht so gerne...

Der Kontext dieser Äußerung ist die telefonische Suche nach einem Termin für ein Treffen. Der Sprecher hat bereits vorher einige Terminvorschläge negativ beschrieben. Das Beispiel wird hier als eine weitere Ablehnung interpretiert<sup>117</sup>.

Die unbetonte Variante weist die typischen Merkmale einer Gradpartikel auf, besonders eine größere Variabilität in der Wortstellung, wie die beiden folgenden Beispiele zeigen<sup>118</sup>.

(22) Wir können sonst auch Freitag oder Samstag nehmen.

(23) Auch Freitag oder Samstag können wir sonst nehmen.

- Wenn man die betonte und unbetonte Variante der Fokuspartikel *auch* als zwei verschiedene lexikalische Elemente betrachtet, dann wird eine homogene Beschreibung der Funktionen der unbetonten Variante und der Fokuspartikel *noch* möglich. Es bleiben semantische Unterschiede zwischen beiden Partikeln, die aber ihrer Substituierung in eine Klasse nicht entgegenstehen.

Die Untersuchung von Spracherwerbsdaten in Hinblick auf die beiden Partikeln – und der Partikel *auch* in beiden Betonungsvarianten – zeigt, dass die Unterscheidung der beiden Varianten von *auch* kognitiv plausibel ist: Die betonte Variante von *auch* wird früher erworben als die unbetonte Variante, und letztere wiederum in etwa zur gleichen Zeit wie die Partikel *noch*<sup>119</sup>. Nederstigt stützt ihre Erkenntnisse auf Langzeitaufzeichnungen der Sprachentwicklung eines Kindes namens *Caroline*. Caroline beginnt mit einem Jahr und neun Monaten, *AUCH* zu

<sup>116</sup>Vgl. Kapitel 4.2.2, S. 167–171.

<sup>117</sup>Vgl. S. 190f. Die Großschreibung der Partikel in unserer Wiedergabe des Beispiels signalisiert, dass sie betont ist.

<sup>118</sup>In Anlehnung an Nederstigt, weitere Beispiele in ihrer Arbeit auf S. 200

<sup>119</sup>Vgl. Abschnitt 9.4.1, vor allem Abbildung 9.7 auf Seite 340.

verwenden, und mit gut zwei Jahren, ungefähr zur gleichen Zeit, *auch* und *noch*.

Durch die gründliche qualitative Analyse von Korpusbelegen, die aus mehreren Korpora gesprochener Sprache stammen, gelingt es der Autorin, hinsichtlich einer der beiden untersuchten Partikeln (*auch*) eine Unterscheidung zwischen zwei Varianten zu treffen. Der Schnitt, den sie macht, erlaubt es, eine Variante dieser Partikel konsistent in das System der Fokuspartikeln, wie sie sie nennt, zu integrieren.

## 8 Besondere Textsorten

Das linguistische und auch korpuslinguistische Interesse gilt auch immer besonderen Textsorten. In den letzten Jahren sind Arbeiten zum Sprachgebrauch in den neuen Medien populär geworden. Eine häufig untersuchte Frage ist die, ob die Sprachverwendung in verschiedenen Anwendungen des Internet – E-Mail, Foren, Chatträume, Gästebücher<sup>120</sup> – als schriftlich oder als mündlich zu charakterisieren ist. Ergebnis dieser Diskussion ist es, dass mittlerweile, einer Idee von Koch und Oesterreicher folgend<sup>121</sup>, zwischen medialer und konzeptueller Schriftlichkeit bzw. Mündlichkeit unterschieden wird. Die Merkmale konzeptueller Mündlichkeit werden von Lothar Lemnitzer und Karin Naumann anhand von Chatprotokollen herausgearbeitet<sup>122</sup>. Chatprotokolle spielen auch in einigen anderen Arbeiten eine wichtige Rolle als empirische Basis.

Schriftliche Dokumente als Ergebnis computervermittelter Kommunikation sind relativ einfach zu beschaffen, man braucht z.B. nur die Diskussion in einem der vielen Chatträume mitszuschneiden. Es besteht aber die bedenkliche Tendenz, dass sich jeder sein eigenes Korpus baut. Nichts liegt näher, als die eigene E-Mail-Korrespondenz als Datenbasis zu verwenden<sup>123</sup>. Dieser Trend zur Beliebtheit der Datenauswahl erschwert auf längere Sicht die Replizierbarkeit und Generalisierbarkeit der auf Grund dieser Daten gewonnenen Erkenntnisse. Das *Dortmunder Chatkorpus*<sup>124</sup> verspricht diesem Missstand zumindest für diese Kom-

<sup>120</sup> Der „Kommunikationsform E-Mail“ wurde mittlerweile ein ganzer Sammelband gewidmet, vgl. Ziegler (2002), ebenso der Kommunikationsform Chat, vgl. Beißwenger (2001). Ein weiterer Sammelband widmet sich dem „Sprachwandel durch Computer“, vgl. Weingarten (1997).

<sup>121</sup> Vgl. Koch und Oesterreicher (1994).

<sup>122</sup> Vgl. Lemnitzer und Naumann (2001), Abschnitt 4.

<sup>123</sup> Dies lässt sich etwa an Beiträgen zum oben genannten Sammelband zur „Kommunikationsform E-Mail“ beobachten.

<sup>124</sup> Vgl. Kapitel 5.

munikationsform abzuhefen, wenn es gelingt dieses Korpus als Referenzkorpus zu etablieren.

Auf anderen Gebieten ist es weitaus schwieriger, an ausreichende Datenmengen zu kommen. Ein gutes Beispiel hierfür sind die SMS genannten Kurznachrichten, die über mobile Telefone gesendet werden. Der ungeheuer hohen Zahl an SMS, die in Deutschland jährlich versendet werden, steht die Schwierigkeit entgegen, solche Botschaften systematisch zu erfassen – das Medium ist nahezu so flüchtig wie die gesprochene Sprache. Dennoch gibt es erste Arbeiten zu dieser Kommunikationsform. Als Beispiele sollen hier nur Schwitalla (2002) und Doering (2002) genannt werden. Diese Arbeiten können allerdings zum jetzigen Zeitpunkt nicht mehr als eine erste Erkundung sein, und die Korpora, auf denen sie basieren, sind alles andere als ausreichend groß und stabil.

Harald Burger<sup>125</sup> untersucht die Kommunikation auf den Seiten eines öffentlich zugänglichen Gästebuchs einer Fernseh-Talkshow begleitenden Website<sup>126</sup>. Sein Interesse gilt dem Wechselspiel von öffentlicher und privater Kommunikation, ihrer Inszenierung und Wahrnehmung. Einen ähnlichen Zugang wählt Thomas Niehr<sup>127</sup>, der die politische Diskussion auf der Homepage eines Bundestagsabgeordneten auswertet. Gästebücher und Homepages sind als Ort der Kommunikation sicher zu ungewöhnlich, um daraus eine besondere Kommunikationsform zu destillieren. Es ist dennoch interessant und wichtig, dieses Medium als eine Form der Interaktion im World Wide Web, bzw. als eine Spielart der computervermittelten Kommunikation, zu beobachten.

Die neuen Medien sind sicher der auffälligste Ort für die Entstehung neuer bzw. untersuchenswürdiger Textsorten oder Kommunikationsformen. Es gibt aber eine Reihe interessanter korpuslinguistischer Arbeiten zu sprachlichen oder stilistischen Aspekten in traditionellen Textsorten oder Kommunikationsformen. Als Beispiele seien hier lediglich erwähnt: Beiträge von Richard Glahn zu Geschäftsberichten als Form der Unternehmenskommunikation<sup>128</sup> sowie zur Sprache deutscher und amerikanischer Talkshows<sup>129</sup>; die Arbeit von Peter Schlobinski und Florian Fiene zu Fußballfanfanzines<sup>130</sup>; eine Analyse des *Dissensus*, einer besonderen Form der Kommunikation unter Jugendlichen, von Arnulf Deppermann und

<sup>125</sup> Vgl. Burger (2002).

<sup>126</sup> Zum verwendeten Korpus s. S. 167.

<sup>127</sup> Vgl. Niehr (2003).

<sup>128</sup> Vgl. Glahn (2003).

<sup>129</sup> Vgl. Glahn (2004).

<sup>130</sup> Vgl. Schlobinski und Fiene (2000). Fanzines sind Magazine von Fangruppen.

Axel Schmidt<sup>131</sup>; die Arbeiten von Christa Dorn zur Textsorte Erpresserbrief, die sich auf ein Korpus dieser Gattung beim Bundeskriminalamt stützt<sup>132</sup>.

## 9 Sprachlehre

Im fünften Kapitel haben wir die Dichotomie von Korpora in der Sprachlehre schon erwähnt: Sie umfassen sowohl muttersprachliche Korpora, die als Datenressource im Unterricht eingesetzt werden können, als auch Korpora, die den Zweitspracherwerb dokumentieren, also Sprache von Nichtmuttersprachlern enthalten.

Joybrato Mukherjee<sup>133</sup> beschreibt in seiner Einführung in die Korpuslinguistik ausführlich, wie Korpora für den Englischunterricht eingesetzt werden können, sowohl in der Unterrichtsvorbereitung als auch im Unterricht selbst. Sie dienen als Quelle für natürliche Beispiele und geben dem Sprachlerner frühzeitig Kontakt zur natürlichen Sprachverwendung.

Diese Verwendungsweise bietet sich insbesondere auch für die Erstellung von Lehrbüchern an. Dieter Mindt<sup>134</sup> analysiert Lehrbücher für den Englischunterricht, die an deutschen Schulen eingesetzt werden, und stellt fest, dass sie teilweise irreführend dahingehend sind, dass weniger häufig verwendete Formen früher eingeführt werden als die eigentlich gängigen. Dadurch entsteht beim Lernen ein falsches Gewicht. Als Negativbeispiel stellt er das Englische *going to*-Futur vor, das in mehreren Standardlehrbüchern früher eingeführt wird als das viel häufiger verwendete *will*-Futur. Er argumentiert, dass Lehrwerke, die auf der Basis von korpusbestimmten quantitativen Untersuchungen von Wortschatz und Verwendungsweisen erstellt werden, solche Verzerrungen nicht enthalten<sup>135</sup>. Guy Aston<sup>136</sup> nennt diese Verwendung von Korpora in der Lehre, bei der der Sprachlerner keinen direkten Zugang zu den Korpora bekommt, den *Hidden-Kulissen-Ansatz* (*Behind the Scenes Approach*). Er kontrastiert ihn mit dem *Auf der Bühne-Ansatz* (*On Stage Approach*), bei dem der Lerner direkt mit dem Korpus arbeitet. Hier

<sup>131</sup> Vgl. Deppermann und Schmidt (2001).

<sup>132</sup> Vgl. Dorn (2003).

<sup>133</sup> Vgl. Mukherjee (2002).

<sup>134</sup> Vgl. Mindt (1996).

<sup>135</sup> Bereits seit 1980 werden im Rahmen des *COBUILD*-Projekts – eine Kooperation zwischen einem Verlag und der Universität Birmingham – in einem korpusbasierten Ansatz Materialien und Referenzwerke für den Englischunterricht für Nichtmuttersprachler erstellt (Sinclair, 1987).

<sup>136</sup> Vgl. Aston (2000), auch <http://sslmit.unibo.it/~guy/barc.htm>.

kann *Data-Driven Learning* zum Einsatz kommen, d.h. Lerner leiten von den Daten Generalisierungen ab, die sie dann auf die Analyse neuer Daten anwenden. Die Analyse von Sprache wird so direkt mit ihrer natürlichen Verwendung gekoppelt. Technische Voraussetzungen dafür sind ein Korpus, ein Konkordanzwerkzeug und Werkzeuge zur eigenen Datenextraktion. Konkrete Anwendungsszenarien sind das Nachschlagen von Wortverwendungen im Satzkontext für die Textproduktion und -rezeption, das systematische Untersuchen bestimmter Sprachverwendungen oder Grammatikkonstruktionen und das ‚genüssliche Schmökern‘ (*serendipitous exploration*). Sogar eine Art enzyklopädischer Verwendung ist möglich, da man durch das Korpus Informationen zu bestimmten Orten oder Personen erhalten kann, sowie über die Kultur der Sprachgemeinschaft, wenn z.B. nach Stereotypen und Vorurteilen geforscht wird<sup>137</sup>. Aston geht auch auf begleitende Effekte des Korpus-einsatzes im Klassenzimmer ein, z.B. den kommunikativen Aspekt bei gemeinsamer Körperarbeit (Korpusanfrage, Finden von Mustern, Interpretation usw.). Als Zielgruppe für diese Art von Korpus-einsatz im Unterricht empfiehlt er fortgeschrittene (erwachsene) Lerner und Lehrer, da es z.B. schwieriger ist, Konkordanzzeilen zu interpretieren, als ein Lernerwörterbuch zu lesen<sup>138</sup>.

Im dänischen *Visual Interactive Syntax Learning*-Projekt (kurz: VISL)<sup>139</sup> kommen linguistisch annotierte Korpora direkt zum Einsatz, wenn auch nicht ganz offen ‚auf der Bühne‘, wie Aston es beschreibt hat. Auf den Projektseiten im Internet kann man online verschiedene Grammatikübungen in mehr als 25 Sprachen ausführen<sup>140</sup>. Die Übungen basieren zum Teil auf manuell vorannotierten Sätzen, zum Teil auf großen, automatisch geparsten Korpora. Dem Lerner kann dadurch eine enorme Vielfalt an authentischem Übungsmaterial angeboten werden. Neben den Syntaxübungen enthält die Seite auch eine Reihe von Sprachspielen, die sehr ansprechend aufgebaut sind. Es gibt z.B. ein kleines Fellknäuel, den *Grammar Man*, den man durch ein Labyrinth von Wortarten leiten muss, ohne einem Gespenst zu begegnen. Der richtige Weg wird jeweils durch einen Beispielsatz vorgegeben, den man aber zuerst analysieren muss. Im Hintergrund des System läuft ein kategorialgrammatischer Parser<sup>141</sup>, der den Sätzen eine Dependenzanalyse zuweist.

<sup>137</sup> Stubbs (1996).

<sup>138</sup> Ein Beispiel für eine Konkordanz finden Sie in Abschnitt 3 des vierten Kapitels.

<sup>139</sup> <http://visl.sdu.dk/visl/de>.

<sup>140</sup> Vgl. Bick (2005).

<sup>141</sup> Vgl. Karlsson (1990).

Ein Beispiel für den zweiten Typ von Korpusersatz beim Sprachenlernen ist das Berliner Projekt *FehlerAnnotiertes LernerKORpus des Deutschen als Fremdsprache* (FALKO). Im vierten Kapitel sind wir kurz auf die Annotation des Lernerkorpus eingegangen. Im Umfeld von FALKO entstehen zur Zeit mehrere Arbeiten zum Zweitspracherwerb und der Didaktik von Deutsch als Fremdsprache, Maik Walter<sup>142</sup> zum Beispiel untersucht in seinem Dissertationsprojekt Satzkonnektoren wie *da*, *weil* oder *obwohl*, deren Verwendung gemeinhin als Indikator für die Niveaueinstufung von Lerner genutzt wird. Die Frage, ob Konnektoren, tatsächlich gute Indikatoren sind, versucht Walter korpusbasiert und im Vergleich mit Daten von Muttersprachlern zu klären. Die bisherige Korpusauswertung zeigt systematische Abweichungen in der Wortstellung und der Konnektorenwahl. In diesem Bereich werden wir in Zukunft interessante Forschungsbeiträge erwarten können.

Wir stellen im Folgenden zwei Arbeiten vor, in denen das Potenzial von Korpusanalyse und didaktischer Aufbereitung von Belegen für den Sprachunterricht demonstriert wird. Es handelt sich also in beiden Fällen um Korpusarbeit *Hinter den Kulissen*.

Die erste Arbeit bezieht sich auf Modalpartikeln, die zweite auf Präpositionen. Für den Sprachlerner stellen Modalpartikeln eine besondere Herausforderung dar. Sie sind weder allein dem Lexikon noch der Grammatik zuzurechnen, ihre Funktion kann deshalb nicht einfach durch Verwendung der entsprechenden Referenzwerke erschlossen werden. Zweitens ist das komplexe Wechselspiel zwischen Partikelfunktion, Kontext und Kontext nicht leicht zu verstehen. Gerade dieses Wechselspiel kann nur anhand von authentischen Beispielen vermittelt und verstanden werden<sup>143</sup>. Möllering begegnet diesen Problemen mit einem fremdsprachendidaktischen Programm, das auf die Verwendung authentischer Beispiele setzt. Als Materialgrundlage dienen ihr vor allem Korpora gesprochener Sprache<sup>144</sup>, da Modalpartikeln vor allem im gesprochenen Deutsch verwendet werden. Sie ermittelt die Vorkommenshäufigkeit aller Modalpartikeln in diesen Korpora und widmet die weiteren Ausführungen den häufigsten Partikeln: *eben*, *nur*, *denn*, *schon*, *doch*, *mal*, *aber auch*, *ja*. Für jede Partikel erarbeitet sie Arbeitsblätter auf der Basis von authentischen Belegen. Diese Arbeitsblätter sollen den Lernern helfen: a) die Verwendung der einzelnen lexikalischen Einheiten als

<sup>142</sup> Vgl. Walter (in Vorb.).

<sup>143</sup> Vgl. hierzu Möllering (2004), Kapitel 1.

<sup>144</sup> Zu den verwendeten Korpora s. S. 101-104. Auf S. 249 diskutiert Möllering einige Schwächen des von ihr verwendeten Korpus. Es sei erstens relativ klein und zweitens sei die überaus hohe Frequenz von *ja* dessen häufiger Verwendung als Gesprächspartikel in Telefondialogen geschuldet.

Modalpartikeln von den anderen Verwendungen dieser Einheiten, z.B. als Konjunktion oder als Gradpartikel, zu unterscheiden und b) Funktion und Bedeutung der Modalpartikeln zu verstehen<sup>145</sup>. In Kapitel 5 dieser Arbeit werden die partikelbezogenen Lehrmaterialien vorgestellt und diskutiert. Die Materialien wurden in der Praxis erprobt, die Einstellung der Schüler zum Lernen an authentischem Sprachmaterial wurde evaluiert. Möllering sieht sich mit ihrer Arbeit in einem Trend der Sprachlehre, die im Sprachlehrer eher einen Vermittler als einen Wissensproduzenten sieht und Sprachlernen als aktive Auseinandersetzung der Lernenden mit authentischen Äußerungen der Zielsprache<sup>146</sup>.

Randall Jones<sup>147</sup> ist an Präpositionen aus der Perspektive der Sprachlehre interessiert. Ziel seiner Studie ist es, die Beschreibungen und Lernhilfen in Lehrbüchern und Lernergrammatiken, die Präpositionen betreffen, mit den Ergebnissen der Analyse eines Korpus gesprochener Sprache zu vergleichen. Für seine Untersuchungen verwendet er ein an der Brigham Young University erstelltes Korpus des gesprochenen Deutsch<sup>148</sup>. Er betrachtet die neun am häufigsten im Korpus vorkommenden Präpositionen: *hinter*, *neben*, *zwischen*, *unter*, *vor*, *über*, *an*, *auf* und *in*<sup>149</sup> und stellt fest, dass eine solche Korpusanalyse andere Informationen zu Tage fördert, als sie in Sprachlehrwerken vermittelt werden. Im Detail:

- Die prototypische Unterscheidung von Ort und Richtung hilft bei der Bestimmung des Kasus, den die Präposition regiert, wenig, weil bei fast allen Präpositionen die wenigsten Vorkommen sich diesem Schema zuordnen ließen. Die meisten Vorkommen hatten keine klare lokale oder direktionale Bedeutung. Viele Präpositionen sind Teil von Präpositionalergänzungen von Verben oder Teil von idiomatischen Wendungen. In diesen Fällen ist der regierte Kasus aber nicht regelhaft erschließbar;
- die Verwendung des Akkusativs und die Verwendung des Dativs sind bei keiner Präposition ausgewogen. Bei *hinter* dominierte der Dativ mit über 80 Prozent, bei *über* der Akkusativ mit über 99 Prozent. Diese quantitativen Tendenzen zu kennen kann für Lerner wichtig sein;

<sup>145</sup> Der Autorin geht es ausdrücklich nicht darum, die aktive Verwendung der Partikeln einzutüben, sondern nur darum, das Verstehen zu erleichtern, vgl. S. 244.

<sup>146</sup> Möllering (2004), S. 250.

<sup>147</sup> Vgl. Jones (2000).

<sup>148</sup> Vgl. S. 118.

<sup>149</sup> Vgl. Tabelle 1 auf S. 120.

- die Präpositionen selbst kommen unterschiedlich oft vor – am seltensten *hinter* und am häufigsten *in*<sup>150</sup>. Diese Erkenntnis mag vor allem für Muttersprachler banal sein, sie wird aber für den Lerner durch die Gleichbehandlung der Präpositionen in vielen Lehrbüchern verdeckt. Jones schlägt hier ein Vorgehen vom Häufigeren zum Seltenen vor.

Als Fazit schlägt Jones den verstärkten Einbezug von Korpora, gesprochener und geschriebener Sprache, für die Sprachlehre oder doch zumindest für die Erstellung von Lehrwerken vor, da sie das Verständnis der komplexen Maschinerie des Deutschen erleichtern<sup>151</sup>.

Diese Arbeiten leisten einen wertvollen Beitrag zu Forschungen, die den Lernprozess nicht aus der Sicht der kognitiven Leistungen der Lernenden, sondern aus der Sicht der Besonderheiten des authentischen Sprachgebrauchs betrachten. Es bleibt zu hoffen, dass diese Erkenntnisse bei den Verlagen, die Lehrmaterialien für Deutsch als Fremdsprache erstellen, auch ankommen.

## 10 Fazit

Wie wir eingangs erwähnt haben, können die in diesem Kapitel dargestellten Arbeiten zum Teil als gute, zum Teil als schlechte Beispiele korpuslinguistischer Forschung aufgefasst werden. Wir wollen die methodischen Tendenzen, die in diesen Arbeiten deutlich werden, hier zusammenfassen und daraus Empfehlungen für eine gutes methodisches Arbeiten ableiten.

Zunächst fällt auf, dass viele Arbeiten sich auf kleinere Korpora stützen, die sich überwiegend im Besitz der Autoren befinden bzw. für diese zum Zweck der Untersuchung erstellt wurden. Es ist auch oft nicht klar, ob die Korpora digital vorliegen und maschinell ausgewertet wurden. Es ist im Prinzip nichts gegen die intellektuelle Auswertung eines (kleinen) Gesamtkorpus einzuwenden. Diese Methode erschwert aber die Überprüfung oder Reproduktion der Ergebnisse. Das einzige Bewertungskriterium ist in diesem Fall die Plausibilität der Ergebnisse. Die linguistische Forschung wird auch weiterhin auf Spezialkorpora angewiesen sein, die ad hoc zum Zwecke einer bestimmten Untersuchung zusammengestellt werden. Es sollte aber gefordert werden, dass diese Spezialkorpora a) digital erfasst und b) begleitend zur Publikation der Öffentlichkeit zur Verfügung gestellt werden, soweit keine urheberrechtlichen oder personenrechtlichen Gründe dagegen sprechen. Die

<sup>150</sup> Vgl. Tabelle 2, S. 141.

<sup>151</sup> S. 142.

Publikation der Daten kann entweder über die Homepage des Forschers oder über eine zentrale Sammel- und Dokumentationsstelle für Korpora geschehen. Eine solche Stelle existiert allerdings noch nicht. Auch die Beschreibung dieser Korpora mit Metadaten ist wünschenswert.

Ein ähnliches Problem ergibt sich, wenn nicht zu wenig, sondern zu viel Daten zur Verfügung stehen. Dies ist bei Forschungen zur computervermittelten Kommunikation der Fall. Hier besteht die Tendenz, Daten in wenig kontrollierter und opportunistischer Weise zu sammeln. Auch dies erschwert letztendlich die Generalisierbarkeit der gewonnenen Erkenntnisse. Dem könnte durch den Aufbau textsorten- oder medien-spezifischer Referenzkorpora abgeholfen werden. Dies ist freilich nicht die Aufgabe einzelner Wissenschaftler, sondern muss institutionell geregelt werden. Einzelne Forscher können und sollten zu einem solchen Referenzkorpus beitragen.

Es gibt nach wie vor nicht *das* Referenzkorpus des Deutschen, wie es etwa das *British National Corpus* für das Britische Englisch war und ist. Die meisten Forscher verwenden die Korpora des Instituts für deutsche Sprache. Dies bedeutet auf der anderen Seite eine gewisse Verantwortung für dieses Institut, geeignete Daten und Korpora zur Verfügung zu stellen. Die Dienstleistungen des Instituts für deutsche Sprache und deren Dokumentation lassen aus unserer Sicht zu wünschen übrig. Als einen Beitrag zur Verbesserung der Situation stellen wir eine Bedienungsanleitung für das Recherchesystem COSMAS auf die dieses Buch begleitende Webseite<sup>152</sup>. Wir hoffen außerdem, mit der Darstellung der Korpuslandschaft des Deutschen in Kapitel 5 dazu beitragen zu können, dass die erwähnten Korpora stärker genutzt werden.

Es gibt kaum einen (korpus-)linguistischen Bereich oder Fragenkomplex, dem sich mehrere Arbeiten widmen. Am ehesten ist dies bisher im Bereich der Modalpartikeln geschehen. Gerade die in den vorhergehenden Kapiteln beschriebenen Probleme mit Korpusdaten als Grundlage linguistischer Erkenntnisse sollten zur Reproduktion bzw. Kontrolle einmal erzielter Ergebnisse ermuntern. Von den hier beschriebenen Arbeiten leistet dies nur Lemmitzer 2001 mit Bezug auf die Arbeit von Thurmair 1989. Verstehen Sie als Leser dieses Buches dies auch als Aufforderung, die hier beschriebenen Arbeiten und daraus gewonnenen Erkenntnisse selbst zu überprüfen.

Wenige Arbeiten machen bisher von den Möglichkeiten Gebrauch, die die linguistische Annotation von Korpora für die weitergehende Forschung schafft. Beispielhaft seien hier die Arbeiten von Nöcker 2003 sowie von Zinsmeister und Heid 2003 erwähnt. Nicht jede Untersuchung

<sup>152</sup> <http://www.lemmitzer.de/lochar/KoLi>.

profitiert von der Existenz einer linguistischen Annotation, zumal diese nicht immer hundertprozentig korrekt ist und oft auch nicht tief genug geht. Wir empfehlen aber, die Verwendung der in den vorhergehenden Kapiteln beschriebenen annotierten Korpora bzw. Baumbanken für die Analyse zu erwägen. Man kann sich u.U. auf Vorkommen von Wörtern in einer bestimmten Wortart oder in einer bestimmten syntaktischen Position beschränken und so die Genauigkeit der Datenextraktion und -analyse verbessern.

Nicht in allen Arbeiten wird das Verhältnis von quantitativer und qualitativer Analyse reflektiert. Ein Musterbeispiel ist hier wieder die Arbeit von Niederstigt 2003, die für alle analysierten Wörter eine die kompletten Korpusdaten umfassende quantitative Analyse vornimmt, für die darauf folgende qualitative Analyse aber für jedes Wort eine gleich große Anzahl von Belegen auswählt. Letzteres erlaubt ihr, die Analyse der beschriebenen Wörter vergleichbar zu machen. Die saubere Trennung beider Aspekte sollte bereits Gegenstand des Forschungsdesigns sein und vor der Auswahl der Korpora und weiteren Analysemit-teln stehen.

Letztendlich müssen auch die grundsätzlichen Fragen beantwortet werden, die wir in den vorhergehenden Kapiteln aufgeworfen haben: Ist ein Korpus überhaupt geeignet zur Beantwortung der Forschungsfragen? Gibt es Alternativen oder Ergänzungen? In welchem Verhältnis stehen die ausgewählten Korpusdaten zum beschriebenen Gegenstand, sind Generalisierungen über die Korpusdaten hinaus möglich? Diese grundsätzlichen Fragen werden in den hier beschriebenen Arbeiten keinesfalls ausgeblendet, sie könnten u.E. aber stärker reflektiert werden.

## 11 Weiterführende Literatur

Es gibt leider bisher für das Deutsche kein Buch, das die in diesem Kapitel dargestellten korpuslinguistischen Ansätze und Themengebiete in geschlossener Form präsentieren würde. Zur Zeit ist ein Band *Korpuslinguistik* in der Reihe *Handbücher zur Sprache und Kommunikationwissenschaft* in Arbeit, der vermutlich im Laufe des Jahres 2007 erscheinen wird. Die Themen dieses Bandes sind allgemeiner Natur und nicht speziell auf das Deutsche bzw. die germanistische Linguistik zugeschnitten. Ebenfalls im Jahre 2006 wird eine *Einführung in die Korpuslinguistik* für Germanisten von Carmen Scherer erscheinen. Wir wissen leider wenig über den geplanten Inhalt des Buches, denken aber, dass sich die Lektüre des Buches ergänzend zu diesem Buch lohnen wird. Ansonsten ist ein regelmäßiger Blick in die germa-

nistischen Fachzeitschriften zu empfehlen. Ergiebige Quellen sind die Zeitschriften *Deutsche Sprache*, *Zeitschrift für germanistische Linguistik*, *Muttersprache*, *Osnabrücker Beiträge zur Sprachtheorie*. Man sollte außerdem die Beiträge in der englischsprachigen Zeitschrift *Corpus Linguistics* zur Kenntnis nehmen, wenn man up-to-date bleiben möchte. Die Studienbibliographie zur Korpuslinguistik von Lenz (2000) ist ebenfalls eine gute Quelle, der Berichtszeitraum endet allerdings mit dem Jahr 1999. Zur Verwendung von Korpora, die die Sprache von Zweitsprachlern dokumentieren, verweisen wir auf die ausführliche Online-Bibliographie von Sylviane Granger (<http://cecl.fltr.ucl.ac.be/learnercorpusbibliography.html>).

## 12 Aufgaben

1. Sie wollen untersuchen, wie oft verschiedene orthographische Varianten eines Wortes verwendet werden, oder, anders formuliert, welche Variante eines Wortes überwiegt. Sie wählen das Web als Korpus und wollen eine Suchmaschine verwenden, um anhand der gelieferten Treffer zu jeder Variante eine ungefähre Abschätzung der Verwendungshäufigkeit vorzunehmen. Arbeiten Sie mit den folgenden Beispielen: a) *Buddyliste / Buddy-Liste / Buddy Liste*, b) *Musikdownload, Musik-Download, Musik Download*, oder wählen Sie ein eigenes Beispiel. Testen Sie die Suchmaschinen [google \(www.google.de\)](http://www.google.de), [msn \(http://search.msn.de\)](http://search.msn.de) und [alltheweb \(http://www.alltheweb.com\)](http://www.alltheweb.com). Welche Ergebnisse bringt die jeweilige Trefferliste? Prüfen Sie einige Treffer, auch solche, die weiter hinten in der Liste stehen. Sind die Treffer korrekt? Sind Sie, nach Durchsicht der Ergebnisse, der Meinung, dass eine oder mehrere der Suchmaschinen sich für solche linguistischen Untersuchungen eignen?
2. Für eine Untersuchung zu Anglizismen im Deutschen möchten Sie aus einem Korpus möglichst viele Anglizismen extrahieren. Welche Möglichkeiten sehen Sie, Anglizismen von nativen deutschen Wörtern zu unterscheiden, ohne jedes einzelne Wort zu überprüfen?
3. Bearbeiten Sie die Vorsilbe *zwischen* als Vorsilbe zu Verben wie z.B. *zwischenfinanzieren*. Suchen Sie Belege aus einem Korpus oder aus dem Web. Verfassen Sie einen Wörterbuchartikel für dieses Präfix. Erarbeiten Sie eine Übung für den Sprachunterricht.
4. Betrachten Sie die E-Mail in Ihrem Postfach als eine Art Korpus. Diskutieren Sie, wenn möglich in einer Gruppe, nach welchen Textsorten Sie diese Mail sortieren könnten. Untersuchen Sie auch die



Header Ihrer Mail. Welche Informationen aus dem Header lassen sich für eine Klassifikation der Nachrichten in Textsorten nutzen?

5. Auf unserer begleitenden Webseite haben wir Listen von möglichen Kollokanten für einige Schlüsselwörter bereitgestellt. Die Liste der Kollokanten wurde mit statistischen Mitteln aus einem sehr großen Korpus extrahiert. Wählen Sie aus diesen Listen alle Wortpaare aus Schlüsselwort und Kollokant aus, die Sie für die Aufnahme in ein Wörterbuch für würdig halten. Markieren Sie das Stichwort, unter dem Sie die Kollokation einordnen würden. Vergleichen Sie die Ergebnisse mit Ihren Kollegen und ermitteln Sie, wie hoch die Übereinstimmung ist. Vergleichen Sie Ihre Ergebnisse auch mit den Kollokationen in einem ein- oder zweisprachigen Wörterbuch. Welche Kollokationen zum Stichwort fallen Ihnen ein und welche Kollokationen finden Sie im Wörterbuch, die in der Liste nicht enthalten sind?
6. Eine eigene, etwas systematischere, aber keinesfalls erschöpfende Untersuchung des *Wortwarte-Korpus* förderte die folgenden Wörter mit BinnenGroßSchreibung zutage:  
*eBoy, eBook eGovernment, eLearning, GamerInnen, geWAPnet, Li-nuxTag, MUDDer, WinNT.*  
 Klassifizieren Sie diese Einheiten nach den Motiven, die zu diesen Bildungen führten. Fallen Ihnen weitere Beispiele ein? Nehmen Sie Stellung zu der Frage, ob die Rechtschreibnorm solche Formen zulassen sollte.

## Man muss sich mit einem Korpus anfreunden – Erfahrungen von Linguisten mit Korpora

Im Folgenden präsentieren wir die Antworten einiger Linguisten auf vier Fragen, die wir im Dezember 2005 stellten. Die Fragen beziehen sich auf die Erfahrungen, die die Befragten bei ihrer linguistischen Forschung mit Korpora gemacht haben, und auf Empfehlungen, die sie Neulingen der Korpuslinguistik geben. Einige der Befragten haben per Mail geantwortet, einige haben sich zu einem mündlichen Interview bereit erklärt. Dies ist einer der Gründe, warum die Antworten unterschiedlich lang ausgefallen sind. Wir haben in den meisten Fällen die Antworten gekürzt, da sie sonst den Rahmen dieses Kapitels gesprengt hätten. Die vollständigen Texte bzw. Transkripte der Interviews können Sie auf der dieses Buch begleitenden Webseite (<http://www.lemmitzer.de/lothar/KoLi>) nachlesen.

Wir haben versucht, Interviewpartner zu gewinnen, die Experten auf mindestens einem der mit der Korpuslinguistik verbundenen Gebiete sind, nämlich theoretische Linguistik, Lexikographie, Computerlinguistik und Sprachlehre. Dass die Biographien der meisten Interviewten in irgendeiner Weise mit Tübingen verbunden ist, hängt sicher auch mit der Biographie und dem momentanen Arbeitskontext der Autoren dieses Buches zusammen.

Wir stellen zunächst die befragten Experten vor und fassen im Anschluss daran alle Antworten zu jeweils einer Frage in einem Abschnitt zusammen. Wir haben die Beiträge dabei alphabetisch geordnet.

### 1 Porträts

**Julia Berman** Studium der Germanistischen Linguistik, Romanistik und Deutsch als Fremdsprache an der Universität München, anschließend Wissenschaftliche Angestellte am Institut für Maschinelle Sprachverarbeitung und Mitglied im Graduiertenkolleg *Linguistische Grundlagen der Sprachverarbeitung* in Stuttgart. Promotion im Bereich der Syntax des Deutschen im Jahr 2000. Zur Zeit wis-