

## ELEKTRONISCHE KORPORA ALS RESSOURCE FÜR DEN FREMDSPRACHENUNTERRICHT

Sabine Braun & Angela Chambers

*Universität Tübingen / University of Limerick*

In diesem Beitrag geht es um Möglichkeiten der Nutzung von Korpora im Sekundarschulbereich. Nach einem Überblick über einschlägige Korpusressourcen, Analyseverfahren und *Tools* werden in knappen Zügen die Grundlagen der Korpusnutzung im Sprachlernkontext skizziert und anschließend verschiedene Möglichkeiten für die Nutzung von Korpora gesprochener und geschriebener Sprache illustriert.

### 1. Einleitung

Im Laufe der letzten Jahrzehnte wurde es möglich, umfangreiche Sammlungen authentischer Texte (Korpora) in elektronischer Form anzulegen und linguistisch auszuwerten. Dadurch treten lexikalische und grammatische Sprachverwendungsmuster zutage, die durch die Betrachtung einzelner Texte oder durch die Befragung muttersprachlicher Intuitionen nicht oder nur ansatzweise zugänglich sind.

Als Datenquelle für die linguistische Forschung werden elektronische Korpora bereits seit den 1960er Jahren genutzt. Ihr Einsatz für Sprachlehrzwecke ist dagegen eine relativ junge Entwicklung. Zunächst schärfte die Auswertung von Korpora den Blick für den realen Sprachgebrauch. So zeigte die Analyse gesprochener Korpora, dass die in Lehrbüchern enthaltenen Materialien die Lernenden häufig nicht angemessen auf relevante Szenarien mündlicher Kommunikation vorbereiteten. Seit Ende der 1980er Jahre ging es dann – angeregt durch die Publikationen von Johns (1986) und Tribble & Jones (1990) – zunehmend um die Frage, wie Korpora direkt von Lehrenden und Lernenden genutzt werden können. Einführende Darstellungen zur Nutzung von Korpora im Sprachlernkontext finden sich in Bowker & Pearson (2002), Hunston (2002) sowie Mukherjee (2002). Mukherjee wendet sich insbesondere an den Sekundarschulbereich. Ferner bietet die ICT4LT-Webseite, die einen Überblick zur Nutzung von IKT im Fremdsprachenunterricht gibt und ebenfalls sehr praktisch ausgerichtet ist, zwei Kapitel zum korpusbasierten Sprachenlernen.

### 2. Grundlagen

#### 2.1 Korpusressourcen

Die meisten elektronischen Korpora wurden bislang für das Englische entwickelt. Das erste ent-

stand 1960 an der Brown University und gilt bis heute als gutes Beispiel für ein ausgewogenes Korpus. Es enthält 500 jeweils 2000 Wörter umfassende Textausschnitte aus 15 Textgattungen (*Genres*). Ab etwa 1980 wurden immer größere Korpora einschließlich der Megakorpora *Bank of English* und *British National Corpus* entwickelt, die heute als Referenzkorpora für das Englische gelten. Daneben gibt es zahlreiche kleinere fach- oder genrespezifische Korpora (z.B. *BNC Baby*). Zur Untersuchung gesprochener Sprache wurden in jüngerer Vergangenheit Korpora entwickelt, in denen Transkripte sowie Audio- oder Videodaten verfügbar sind (z.B. *MICASE*).

Korpora können annotiert werden, um ihre Erschließung effizienter zu gestalten. Die häufigste Art der Annotation ist die Wortarten-Annotation. Diese kann heute mit entsprechenden Computerprogrammen (engl. *tagger*) automatisch vorgenommen werden. Die Genauigkeit liegt je nach Programm und Textgattung bei 90-98 %.

Zur Illustration möglicher Anwendungsbereiche von Korpora im Sprachlernkontext werden in diesem Beitrag ein englisches Korpus gesprochener Sprache und ein französisches Korpus geschriebener Sprache verwendet. Das *ELISA English Language Interview Corpus* umfasst 60.000 Wörter und enthält Transkripte sowie Videos narrativer Interviews, in denen Muttersprachler des Englischen über ihren Lebensweg berichten. Das *Corpus Chambers-Rostand de français journalistique (CRFJ)* umfasst eine Million Wörter und beinhaltet Presstexte aus drei französischen Zeitungen der Jahrgänge 2002 und 2003. Diese beiden Korpora, die kostenlos und frei zugänglich sind, wurden speziell als Ressource für den Sprachlernkontext entwickelt. Sie sind inhaltlich homogener als die großen, mehrere Genres abdeckenden Korpora und dadurch für Lernende leichter erschließbar.

Eine gute Alternative zu den genannten Korpora stellen webbasierte Textarchive wie das *Project Gutenberg* und das *Oxford Text Archive* dar, aus denen digitale Texte und Kleinkorpora (z.B. literarische Werke) in vielen Sprachen heruntergeladen werden können.

#### 2.2 Analyseverfahren und *Tools*

Zur Recherche in Korpora aus Textarchiven oder auf CD-ROM benötigt man sogenannte Konkordanzprogramme. Online zugängliche Korpora haben i.d.R. eigene, webbasierte Suchmasken.

Ein guter Ausgangspunkt ist die Häufigkeit der in einem Korpus verwendeten Wortformen (z.B. *got* vs. *gotten*) oder Lexeme (z.B. GET als Be-

deutungseinheit mit den Wortformen *get, gets, getting, got, gotten*). Ein wichtiges quantitatives Analyseverfahren ist deshalb die Erstellung von Wortlisten, die wahlweise alphabetisch oder nach Vorkommenshäufigkeit sortiert werden können. Wortlisten können gefiltert werden, um z.B. alle Funktionswörter (oder auch Zahlen, Eigennamen u.a.) auszuschließen. Hierzu benötigt man eine Stopliste, die die auszuschließenden Wörter enthält. Außerdem können Wortlisten lemmatisiert werden, um alle Formen eines Lexems unter ihrer Grundform (ihrem Lemma) zusammenzufassen. Hierfür benötigt man eine Lemmaliste, die die Zuordnung enthält (z.B. *get, gets, getting, got, gotten* → *get*). Stop- und Lemmalisten sind im Internet erhältlich (z.B. <http://devoted.to/corpora>). Häufigkeitsbasierte Wortlisten bilden auch die Basis für Schlüsselwort-Analysen, bei denen – durch Vergleich mit einem Referenzkorpus – das für ein Korpus charakteristische Wortschatzprofil ermittelt werden kann (z.B. für ein literarisches Werk).

Ein weiteres wichtiges Analyseverfahren ist die Erstellung einer Konkordanz, d.h. die Sammlung aller Fundstellen einer Suchform in einem Korpus. Die Suchform kann ein Wort, ein Wortteil oder eine Phrase sein. Die Fundstellen werden üblicherweise zusammen mit dem umliegenden sprachlichen Kontext zitiert. Eine verbreitete Darstellungsform ist das KWIC-Format (*keyword in context*). Die gängigen Konkordanzprogramme ermöglichen dem Benutzer, die Spannweite des umliegenden Kontextes selbst festzulegen (und sie auch auf ganze Sätze oder Absätze auszudehnen) und die Konkordanz nach verschiedenen Kriterien zu sortieren.

1.	the new hospital. A	<b>couple</b>	of councillors thought
2.	flew up here a	<b>couple</b>	of days ago, and
3.	me down for a	<b>couple</b>	of days.' You adjust
4.	doing that for a	<b>couple</b>	of decades. In Melbourne
5.	or I have a	<b>couple</b>	of horses to ride
6.	a rest for a	<b>couple</b>	of hours, maybe a
7.	aboriginal guide for a	<b>couple</b>	of hours. They go
8.	years ago, my first	<b>couple</b>	of jobs were minimum
9.	horse for just that	<b>couple</b>	of minutes, all right?
10.	course, you have a	<b>couple</b>	of seasons in the
11.	nursery first, for a	<b>couple</b>	of sessions, so that
12.	we also had a	<b>couple</b>	of staff from here
13.	that we had a	<b>couple</b>	of years ago, the
14.	work there for a	<b>couple</b>	of years in an
15.	a volunteer for a	<b>couple</b>	of years. And then

Abb. 1: KWIC-Konkordanz für *couple* aus dem ELISA-Korpus, sortiert nach dem ersten und zweiten Wort rechts vom Suchwort (1R und 2R).

Konkordanzen eignen sich als Ausgangspunkt für qualitative Analysen von lexiko-grammatischen Verwendungsmustern, Kollokationen und Bedeu-

tungsnuancen der Suchform. Abb. 1 zeigt z.B., dass *couple* im ELISA-Korpus typischerweise in der Phrase (*for*) *a couple of* mit der Bedeutung „ein paar“ vorkommt, häufig eine Zeitdauer quantifiziert (*days, decades, hours, minutes, seasons, years*) und als Subjekt mit der 3. Person Plural des Verbs gebraucht wird (vgl. Zeile 8).

Zusätzlich unterstützen einige Programme und Online-Korpora die quantitative Analyse von Kollokationen, d.h. sie ermitteln die Stärke der kollokativen Bindung zwischen einem Suchwort und seinen Nachbarn (engl. *collocates*). Die Liste der signifikantesten Nachbarn im Umfeld von *couple* in der *Bank of English* bestätigt das von der ELISA-Konkordanz vermittelte Bild.

	Nachbar	Häufigkeit des Nachbarn im Korpus	gemeinsame Häufigkeit
1.	a	973 489	5 552
2.	of	1 100 578	5367
3.	years	45 988	724
4.	ago	13 590	401
5.	weeks	9 026	386
6.	days	18 379	362
7.	months	13 169	248
8.	times	16 389	240
9.	hours	9809	205
10.	last	45548	274
11.	for	376522	832
12.	married	4 311	135
13.	minutes	9 865	131
14.	there	149133	391
15.	next	23 175	143

Abb. 2: Nachbarn von *couple* in der *Bank of English*

Bei den hier vorgestellten Korpusanalyseverfahren steht das „vertikale Lesen“ im Vordergrund. Mit Blick auf den Sprachlernkontext ist jedoch hervorzuheben, dass vor allem beim Arbeiten mit kleineren Korpora auch eine Kombination dieser Verfahren mit der diskursbasierten Erschließung der Texte in einem Korpus sinnvoll ist, um Diskursmerkmale zu untersuchen, die sich nicht in Wortlisten oder Konkordanzen manifestieren.

### 3. Anwendung

#### 3.1 Theoretisch-methodische Grundlagen

Im Rahmen eines „Kommunikativen Ansatzes“ gilt es weithin als akzeptiert, dass Lernende Zugang zu authentischen Texten benötigen und nicht (ausschließlich) mit konstruierten Beispielen arbeiten, wie sie in vielen Grammatiken und Lehrbüchern verwendet wurden. Korpora bieten sich als Datenquelle an. Widdowson (2000: 7) weist allerdings darauf hin, dass die Beispiele in Konkordanzen von ihrem ursprünglichen Kontext isoliert sind und dass es deshalb bei der Nutzung für Sprachlernzwecke einer *pedagogic mediation* (2003: 102-103), also einer didaktisch ge-

leiteten Lösung bedarf, um den Lernenden die kontextuelle Einbettung der Korpusdaten zu erleichtern. Die beiden Korpora, die in diesem Beitrag zur Illustration herangezogen werden, enthalten z.B. Texte aus jeweils nur einer Gattung (Lebensgeschichten im ELISA-Korpus sowie Pressetexte im CRFJ). Beim Arbeiten mit Materialien (Konkordanzen, Wortlisten) aus diesen Korpora haben Lernende entsprechend mehr Anhaltspunkte für die Interpretation der Daten, als dies bei einem großen Korpus, das viele Textgattungen abdeckt, möglich wäre. Die beiden Korpora können somit als *pedagogically relevant corpora* (Braun 2005: 49) gelten.

Ein zweites wichtiges Argument für den Einsatz von Korpora bzw. Konkordanzen im Fremdsprachenunterricht leitet sich daraus ab, dass in ihnen charakteristische Sprachverwendungsmuster, insbesondere häufig wiederkehrende (idiomatische) Wendungen und Fügungen (wie z.B. *a couple of* im ELISA-Korpus) gut sichtbar werden. In der Psycholinguistik, Korpuslinguistik und der Zweitspracherwerbsforschung herrscht Einigkeit über die Bedeutung solcher Bausteine sowohl im Erstsprache- als auch im Zweitspracherwerb. Sinclair (1991: 109) unterscheidet zwischen dem *open-choice-principle*, nach dem ein Sprecher eine Äußerung aus einzelnen ihm jeweils zur Verfügung stehenden Segmenten (Substantiven, Verben, Präpositionen usw.) produziert, und dem *idiom principle*, nach dem ein Sprecher bei der Äußerungsproduktion auf vorgefertigte Bausteine zurückgreift. Sinclair zufolge kommt dem *idiom principle* größere Bedeutung zu; der Fremdsprachenunterricht steht dagegen häufig noch im Zeichen des *open-choice*-Prinzips.

Drittens können Konkordanzen durch die großen in Korpora enthaltenen Datenmengen eine Vielzahl von Beispielen zur jeweils untersuchten Fragestellung liefern. Cobb (1997: 302-303) sieht hierin eine Bereicherung des kommunikativen Ansatzes: der Lernerfolg wird durch die Mehrfachillustration gefördert. Cobb (1997) und Gaskell & Cobb (2004) betonen außerdem die Effizienz von Konkordanzen: Konkordanzlisten ermöglichen eine Analyse sprachlicher Ausdrucksmittel in verschiedenen (Mikro-)Kontexten, für die im Lernalltag meist nicht genügend Zeit bleibt.

Zusätzlich zu den genannten Vorteilen steht die Nutzung von Korpora bzw. Konkordanzen durch die Lernenden auch im Einklang mit aktuellen Erkenntnissen der Sprachlernforschung. So wird in konstruktivistischen und kognitiven Ansätzen die Bedeutung von Lernaktivitäten hervorgehoben, in denen die Lernenden eine aktive Rolle

einnehmen. Ferner ergibt sich aus der Forderung nach Lernerautonomie die Notwendigkeit zur Stärkung der Eigenverantwortlichkeit der Lernenden für den Lernprozess.

Für den Einsatz von Korpora im Sprachlernkontext ergeben sich hieraus grundsätzlich zwei Möglichkeiten: Bei einer *indirekten* Nutzung erstellt die Lehrperson mithilfe eines Korpus Arbeitsblätter wie zum Beispiel die in diesem Artikel dargestellten Konkordanzen und Wortlisten, die sich für explorierendes Lernen eignen (vgl. auch Kennedy 2003). Die *direkte* Nutzung eines Korpus durch die Lernenden erfordert Übung im Umgang mit einem Konkordanzprogramm; diese Art der Korpusnutzung war bereits Gegenstand einiger Untersuchungen im Hochschulkontext, die im Allgemeinen positive Ergebnisse lieferten (Chambers 2005; Cobb 1997; Gaskell & Cobb 2003).

### 3.2 Nutzung von Korpora gesprochener Sprache

Beim Arbeiten mit gesprochenen Korpora sind für den Schulkontext Korpora interessant, die nicht nur Transkripte, sondern auch Audio- oder Videosequenzen enthalten und in denen – neben der Erstellung von Wortlisten und Konkordanzen – auch der Zugriff auf einzelne Textpassagen leicht möglich ist. Dies macht die Korpora z.B. für den Einsatz im Rahmen von Hörverstehensaktivitäten attraktiv.

In ELISA sind vergleichbare Passagen aus verschiedenen Interviews sowie die zugehörigen Videosequenzen (z.B. zu „Werdegang“, „Arbeitsalltag“, „Zukunftsvorstellungen“) über einen Themenindex zugänglich.

Durch die Möglichkeit, ähnliche Teile aus verschiedenen Korpusdaten heranzuziehen, ergibt sich eine reichere Materialbasis als beim Arbeiten mit Einzeltexten (Gavioli & Aston 2001). Hiervon kann je nach Kompetenzniveaustufe in unterschiedlicher Dosierung Gebrauch gemacht werden. Im ELISA-Korpus sind außerdem Sprecher verschiedener Varietäten des Englischen vertreten, so dass sich Hörverstehensaufgaben auf die Anforderungen der Lehr-/Bildungspläne (Verstehen verschiedener nationaler und regionaler Aussprachevarianten) abstimmen lassen.

Wichtig ist die gründliche Vorauswahl geeigneter Passagen durch die Lehrperson. Dabei bietet sich die Erstellung von Wortlisten zu den betreffenden Passagen an. Eine gefilterte Wortliste ohne Funktionswörter gibt einen raschen Überblick über den Kernwortschatz der Passage, der ggf. zur Vorentlastung einer Verstehensaufgabe mit den Schülern erarbeitet werden muss.

Des weiteren können Korpora wie ELISA zur Analyse von Merkmalen der gesprochenen Sprache genutzt werden. Hier hat ein Korpus – gegenüber einem Einzeltext – den Vorteil, dass charakteristische Merkmale klarer hervortreten. Allerdings stellt das vergleichende Arbeiten mit verschiedenen Texten eine Herausforderung dar, an die die Schüler schrittweise herangeführt werden müssen. So kann in einem ersten Schritt zunächst mit einem Text(teil) gearbeitet werden, der den Schülern bereits durch vorausgegangene Hörverstehensaktivitäten vertraut ist. In einem zweiten Schritt können gemeinsam aufgestellte Hypothesen anhand weiterer Passagen (aus anderen Korpus-texten) überprüft werden.

Die textbasierte Arbeit kann hierbei auf vielfältige Weise durch wortlisten- und konkordanzbasierte Analysen ergänzt werden. Im Gegensatz zu Wortlisten für die Vorentlastung von Verstehensaufgaben sind hier ungefilterte Listen sinnvoll. Sie lassen sofort einige für die gesprochene Sprache charakteristische Funktionswörter deutlich werden. Abb. 3 zeigt die Wortliste (Rang und Häufigkeit) des Interviews "A tour guide from Ayers Rock" aus dem ELISA-Korpus.

Rang	Wort	Häu	Rang	Wort	Häu
1	AND	69	16	ABOUT	17
2	THE	65	17	IT	16
3	TO	56	18	FOR	13
4	A	43	19	JUST	13
5	OF	37	20	ABORIGINAL	11
6	YOU	28	21	HAVE	11
7	I	26	22	THAT	11
8	PEOPL	24	23	UP	11
	E				
9	IS	23	24	WE	11
10	IN	22	25	BE	10
11	IT'S	22	26	LOTS	10
12	OUT	22	27	WHICH	10
13	SO	21	28	DO	9
14	THEY	21	29	GET	9
15	HERE	19	30	GOOD	9

Abb. 3: Wortliste des Interviews „A tour guide from Ayers Rock“

Die Wortliste kann bei der Auswahl geeigneter Interviewpassagen helfen, anhand derer sich z.B. der Gebrauch von *so* (Rang 13), *just* (Rang 19) oder der Quantifikatoren *lots of* (Rang 26), *a lot of*, *a bit of* gut illustrieren lässt. Eine solche Passage ist in Abb. 4 dargestellt.

**So, just** to give you **a bit** more information about me ... Before I came out here and before I became a tour guide, I grew up in a country town. And growing up in country towns, you play football and end up working on farms. **So I** started out as a stock man, big

hats, big boots, tight jeans, working with cattle and sheep. Spent **a fair bit of** time just doing that. Did **a bit of** modelling for amateur stuff to promote wool, trying to sell wool around the world. The area I'm from grows some of the best fine wool in the world. **So** that was good fun. A friend of mine who owns a farm, who's also doing **a bit of** bus driving, said 'Hey, you'd like this? It's good fun, you get to meet **lots of** people, **lots of** girls.' He said 'Why don't you give it a go?' **So**, I used to drive Sydney to Cairns. It's about three thousand kilometres and **a lot of** fun. But I've always wanted to visit the Northern Territory - they call it frontier country out here - **so I just** thought I'd come on out about nine months ago, being single it's easy. You **just** pack up and come here. **So** that's fantastic.

Abb. 4: Ausschnitt aus dem Interview "A tour guide from Ayers Rock"

Die Passage eignet sich außerdem zur Diskussion weiterer Merkmale gesprochener Sprache, die sich mit korpusbasierten Verfahren weniger gut erschließen lassen (z.B. die elliptischen Strukturen *Spent a fair bit of time..., Did a bit of modelling...*).

Ein anderes Beispiel für den Nutzen dieser Verfahren ist der Gebrauch von *that* (Rang 22 der Wortliste). Einzelne Interviewabschnitte illustrieren die Funktionen von *that* nicht ausreichend. Die Konkordanz in Abb. 5 zeigt dagegen auf einen Blick, dass *that* in dem Interview als Relativpronomen, Demonstrativadjektiv und Demonstrativpronomen, nicht aber als Konjunktion verwendet wird. Dieser Befund kann der Verwendung in der Schriftsprache gegenübergestellt werden und den Ausgangspunkt für eine grundsätzlichere Diskussion der Unterschiede zwischen Mündlichkeit und Schriftlichkeit bilden.

1.	is the resort town	<b>that</b>	we live in. This
2.	the original people	<b>that</b>	were here, which
3.	They can see things	<b>that</b>	we can't see. They
4.	I'd be good at	<b>that</b>	job. So, they offered
5.	here gave themselves	<b>that</b>	name of 'we people'
6.	bucks and moved into	<b>that</b>	. Because of that I
7.	into that. Because of	<b>that</b>	I got to meet
8.	of time just doing	<b>that</b>	. Did a bit of
9.	in the world. So	<b>that</b>	was good fun. A
10.	in the morning. And	<b>that</b>	can kill people.
11.	got half way through	<b>that</b>	. So I do sunrise.

Abb. 5: Verwendung von *that* im dem Interview „A tour guide from Ayers Rock“

### 3.3 Nutzung schriftsprachlicher Korpora

Zeitungskorpora wie das CRFJ können als besonders ergiebige Quelle für die Erstellung von (explorativen) Lernmaterialien auf der Basis authentischer Ressourcen gelten: Einerseits sind Beispiele aus Zeitungstexten aufgrund des hohen Bekanntheitsgrades dieser Textgattung für Schü-

ler gut zu erschließen. Andererseits finden sich in Zeitungstexten trotz ihrer hauptsächlich schriftsprachlichen Ausrichtung auch Elemente gesprochener Sprache (Zitate), sodass diese Textgattung viele Facetten der Alltagssprache repräsentiert.

In den meisten Arbeiten zur didaktischen Nutzung von Korpora geht es um den Einsatz konkordanzbasierter Materialien mit fortgeschrittenen Lernenden. Mit Blick auf den Sekundarschulbereich ist jedoch interessant, dass sich mithilfe von Konkordanzen auch ganz einfache Sprachverwendungsmuster besser und vielfältiger illustrieren lassen als in traditionellen Lehrwerken üblich. So kann bereits die kleine Konkordanz in Abb. 6 wichtige Aspekte der Komparativbildung im Französischen veranschaulichen. Die mit *que* bzw. *de* gebildeten Konstruktionen sind im Deutschen jeweils mit *als* wiederzugeben.

Les sites chimiques ont souvent **plus d'**un siècle  
Ils sont **plus** anciens **que** nos sites pétroliers.  
Depuis la chute du mur de Berlin, il y a **plus d'**une décennie,  
L'actualité est toujours **plus** forte **qu'**une fiction  
il a décidé d'y investir **plus de** 900 millions d'euros  
Les arbres vivent **plus** longtemps **que** les hommes  
Chaque année Chamonix accueille **plus de** trois millions de touristes  
il y a des choses **plus** graves **que** le rugby dans la vie  
**plus de** 2 millions de personnes  
les policiers étaient probablement **plus** nombreux **que** les spectateurs

Abb. 6: Komparativkonstruktion mit *de* und *que* im Französischen aus dem CRFJ. Die Konkordanzen aus dem CRFJ wurden mit WordSmith Tools 4.0 erstellt.

Wichtig ist eine gute Planung bei der Erstellung vorgefertigter Konkordanzen (für Arbeitsblätter) durch die Lehrperson. Insbesondere bei weniger fortgeschrittenen Schülern ist auf Verständlichkeit und klare Fokussierung der Konkordanzdaten zu achten. Die Konkordanz in Abb. 6 kann in einem kleinen Korpus durch die Suche nach *plus* und anschließendes Löschen unerwünschter bzw. unpassender Konkordanzzeilen erstellt werden. Zur Einschränkung der Suche (in größeren Korpora) erlauben gängige Konkordanzprogramme auch die Definition komplexer Suchanfragen (engl. *queries*). In diesem Fall wurde nach Verbindungen von „*plus* + 0 oder 1 Wort + *que* oder *de*“ gesucht. Die hier entstandene, gemischte Konkordanz eignet sich als Basis für die Erstellung von Lückenübungen, in denen *de* und *que* ausgelassen werden.

Diese grundlegenden Techniken sind vielseitig anwendbar. So lassen sich auch komplexere grammatische Konstruktionen extrahieren, die für fortgeschrittene Lernende interessant sind. Die Konkordanz in Abb. 7 zeigt beispielsweise die Ver-

wendung des *ne explétif* in Komparativkonstruktionen.

L'audience baissera ? Moins qu'on **ne** le craint  
l'économie est bien plus délabrée qu'on **ne** le pensait  
Le Président entre donc en campagne plus tôt qu'il  
**ne** le prévoyait  
Il est moins lié à l'extrême droite qu'il **ne** le fût  
Mer et Fillon plombent plutôt Raffarin qu'ils **ne** le  
mettent en valeur.  
l'Europe, elle, s'est révélée beaucoup moins  
vertueuse qu'elle **ne** le prétendait

Abb. 7: Verwendung des *ne explétif* im CRFJ

Diese Beispiele verdeutlichen den Nutzen von Korpora als Materialquelle bzw. zur Bereicherung eines kommunikativen Ansatzes. Sie zeigen, wie Korpora von der Lehrperson genutzt werden können, um Schülern zur Vertiefung eines Themas eine sinnvolle Auswahl und Menge authentischer Beispiele an die Hand zu geben. Allerdings ist bei der Bearbeitung von Konkordanzen, insbesondere beim Löschen von Zeilen, darauf zu achten, dass kein verzerrtes Bild entsteht. Konkordanzen liefern keineswegs immer ein homogenes Bild. Ein Beispiel hierfür ist die Konkordanz in Abb. 8, die belegt, dass *la moitié* und *la majorité* sowohl mit der Singularform als auch mit der Pluralform des Verbs verwendet werden.

Plus de la moitié de la population vit au-dessous du  
seuil de pauvreté  
Près de la moitié d'entre eux vivent avec un seul de  
leurs parents  
la moitié d'entre elles vit uniquement des prestations  
sociales.  
Plus de la moitié des Palestiniens vivent au-dessous  
du seuil de pauvreté  
Jacques Chirac [...] a dit : " Je sais que la grande  
majorité des Corses veulent rester français  
j'ai rappelé au Ministre qu'en 1993, une certaine  
majorité avait voté  
une majorité des Américains doutent de la politique  
de George W. Bush  
une écrasante majorité (93 %) craint une nouvelle  
catastrophe sur nos côtes.  
La majorité des criminologues est unanime  
La majorité des détenus ont entre 30 et 40 ans

Abb. 8: *la majorité* und *la moitié* mit Singular- und Pluralform des Verbs

Durch die Einbeziehung sprachlicher Variationen in den Unterricht kann insbesondere bei fortgeschritteneren Lernenden das Verständnis für Sprache als ein komplexes und sich stetig wandelndes System gefördert werden. Hierdurch wird wiederum die Herausbildung eines deskriptiven Ansatzes in der sprachlichen Analyse unter-

stützt. Nach entsprechender Vorbereitung auf die Arbeit mit authentischen Daten können Schüler auch selbst in einem Korpus wie dem CRFJ recherchieren, um – in Einzel- oder Gruppenarbeit – ganz gezielt bestimmte Sprachgebrauchsmuster zu untersuchen.

Eine weitere Möglichkeit der direkten Nutzung eines Korpus durch die Schüler ergibt sich im Rahmen der eigenen Textproduktion. Hier kann ein Korpus als Nachschlagewerk benutzt werden. Ein Zeitungskorpus wie das CRFJ ist hier besonders hilfreich, da es Texte zu aktuellen Themen enthält, die häufig auch Gegenstand schriftlicher Schülerarbeiten sind. Die Suche nach der Form *Verts* im CRFJ ergibt zum Beispiel 70 Fundstellen, von denen sich 57 auf die politische Partei beziehen. In diesen Suchergebnissen finden Schüler, die über die Grünen schreiben, eine Fülle einschlägiger zielsprachlicher Wendungen. Eine Auswahl ist in Abb. 9 dargestellt.

Quant aux Verts, mouvement anti-establishment par essence  
La croisade des Verts contre le Village des marques  
il a été élu en 1999 sur une liste des Verts  
les Verts ont maintenant sept députés au Parlement,  
au lieu d'un seul  
les Verts redoutent aussi « le ballet des autobus et des voitures

Abb. 9: Ausgewählte Vorkommen von *Verts* im CRFJ

Schließlich lässt sich die eigenständige Recherche von Schülern in einem Korpus auch bei der Überarbeitung eines zuvor von der Lehrperson korrigierten Textes nutzen. Eine entsprechende Untersuchung von Chambers & O'Sullivan (2004) zeigte, dass Lernende ein Korpus in diesem Zusammenhang als nützliches Hilfsmittel erachteten und es u.a. zur Korrektur von Fehlern bei präpositionalen Anschlüssen und Interferenzfehlern verwendeten. Die Grundlage bildete ein kleines Korpus thematisch einschlägiger Texte. Als abschließendes Beispiel soll eine Konkordanz aus einem kleinen, etwa 50.000 Wörter umfassenden Korpus französischer Wissenschaftstexte (wissenschaftliche Artikel und Auszüge aus studentischen Abschlussarbeiten französischer Muttersprachler) dienen, das an der University of Limerick aufgebaut wurde. Das *Corpus Chambers-Rostand d'articles de recherche* befindet sich derzeit im Aufbau und ist ab September 2006 im Oxford Text Archive erhältlich. Die Suche nach *comme nous* in diesem Korpus ergab die in Abb. 10 dargestellten Wendungen.

Comme nous l'avons vu  
comme nous l'avons vu au chapitre 1  
comme nous l'avons vu plus haut (2 occurrences)  
comme nous l'avons indiqué plus haut  
comme nous l'avons mentionné plus haut  
comme nous l'avons déjà signalé  
Comme nous l'avons déjà souligné  
Comme nous l'avons déjà dit  
comme nous le verrons plus loin

Abb. 10: *Comme nous* in einem 50.000-Wörter-Korpus französischer Wissenschaftstexte

Im Gegensatz zu den 15 Vorkommen von *comme nous* im CRFJ bieten die Beispiele aus dem viel spezifischeren Wissenschaftskorpus viele nützliche Wendungen für das Verfassen eines wissenschaftlichen Textes in der Fremdsprache. Ebenso fördert beispielsweise die Suche nach *nous* in diesem Korpus eine Reihe von Verben zutage, die in wissenschaftlichen Arbeiten charakteristischerweise zur Beschreibung wissenschaftlicher Tätigkeiten benutzt werden (z.B. *nous analyserons*, *nous aborderons*, *nous nous pencherons sur*). Solche Beispiele machen deutlich, dass für den Sprachlernkontext eine Reihe von Korpora – allgemeinere und spezifischere – erforderlich sind, um verschiedene Bedürfnisse und Lernziele effizient zu unterstützen.

#### 4. Ausblick

Die Integrationsmöglichkeiten von Korpora reichen vom gelegentlichen, ergänzenden Rückgriff auf ein Korpus im Rahmen der Unterrichtsvorbereitung, etwa zur Anreicherung einer Lehrbucheinheit mit zusätzlichem Material, bis hin zur regelmäßigen Nutzung von Korpora durch die Schüler, z.B. als zusätzliches Nachschlagewerk für die eigene Textproduktion. Bei der praktischen Umsetzung sind jedoch noch einige Grenzen zu überwinden. Ganz entscheidend ist zunächst, dass die Lehrperson über einschlägige Korpuskenntnisse verfügt; hier sind Hochschulausbildung und Lehrerfortbildung entsprechend gefordert. Außerdem bestehen noch erhebliche Defizite hinsichtlich der Verfügbarkeit geeigneter Korpusressourcen. So sind einfach und kostenlos zugängliche sowie sprachdidaktisch relevante Korpora noch spärlich bzw. nicht immer leicht zu finden. Angesichts der bereits existierenden frei benutzbaren, webbasierten Konkordanzprogramme wären webbasierte Referenzmaterialien denkbar, die Lehrenden und Lernenden bedarfsgerechten Zugang zu Korpussoftware und Korpora bieten. Zur Absicherung künftiger Entwicklungen sowie als Quelle für praktische Anregungen sind schließlich weitere empirische

Studien zum Einsatz von Korpora im Sekundar-  
schulbereich erforderlich.

#### LITERATURVERZEICHNIS

- Bowker, L. & Pearson, J. (2002). *Working with specialized language: a practical guide to using corpora*. London, Routledge.
- Braun, S. (2005). From pedagogically relevant corpora to authentic language learning contents. *ReCALL* 17 (1), 47-64.
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology* 9 (2), 111-125.
- Chambers, A. & O'Sullivan, I. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL* 16 (1), 158-172.
- Cobb, T. (1997). Is there any measurable learning from hands on concordancing? *System* 25 (3), 301-315.
- Davies G. (ed.) *Information and communications technology for language teachers (ICT4LT)*. Slough, Thames Valley University [Online]. Available from: <http://www.ict4lt.org> (14.12.2005).
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System* 32 (3), 301-319.
- Gavioli, L. & Aston, G. (2001). Enriching reality: language corpora in language pedagogy. *ELT Journal* 55 (3), 238-246.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, Cambridge University Press.
- Johns, T. (1986). Micro-concord: a language learner's research tool. *System* 14 (2), 151-162.
- Kennedy, G. (2003). *Structure and meaning in English*. Harlow, Longman/Pearson.
- Mukherjee, J. (2002). *Korpuslinguistik und Englischunterricht: Eine Einführung*. Frankfurt a.M., Peter Lang.
- Tribble, C. & Jones, G. (1990). *Concordances in the classroom: a resource book for teachers*. Harlow, Longman.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, Oxford University Press.
- Widdowson, H.G. (2000). On the limitations of linguistics applied. *Applied Linguistics* 21 (1), 3-25.
- Widdowson, H.G. (2003). *Defining issues in English language teaching*. Oxford, Oxford University Press.

#### KORPORA

- Braun, S. (ed.) (2004). *ELISA English Language Interview Corpus as a Second-Language Application*. <http://www.uni-tuebingen.de/elisa> (14.12.2005).
- Chambers, A. and Rostand, S. (eds.) (2005). *Le Corpus Chambers-Rostand de français jour-*

*nalistique*. <http://www.ota.ahds.ac.uk/texts/2491.html> (14.12.2005).

- Bank of English. Collins Wordbanks Online, Harper Collins. <http://www.collins.co.uk/books.Aspx?group=153>. Kostenlose Online-Recherche in der Demo-Version: <http://www.collins.co.uk/Corpus/CorpusSearch.aspx> (14.12.2005).
- British National Corpus (BNC). Oxford University Computing Services. <http://www.natcorp.ox.ac.uk/> kostenlose Online-Recherchen: <http://sara.natcorp.ox.ac.uk/lookup.html> oder <http://view.byu.edu> (14.12.2005).
- Burnard, L. (ed.) (2004). *BNC Baby* [CD-Rom]. Oxford, Research and Technology Service, Oxford University. <http://www.natcorp.ox.ac.uk/babyinfo.html> (14.12.2005).
- Simpson, R. et al. (eds.) (2002). *MICASE: The Michigan Corpus of Academic Spoken English*. University of Michigan. <http://www.lsa.umich.edu/eli/micase> (14.12.2005).

#### RESSOURCEN

- ICT4LT. Information and Communications Technology for Language Teachers. <http://www.ict4lt.org/en/index.htm> (14.12.2005).
- Oxford Text Archive. <http://ota.ahds.ac.uk/> (14.12.2005).
- Project Gutenberg. <http://www.gutenberg.org/> (14.12.2005).
- Tim Johns' Website. [http://www.ecml.at/projects/voll/our\\_resources/graz\\_2002/ddrivenlrning/whatisddl/resources/tim\\_dll\\_learning\\_page.htm](http://www.ecml.at/projects/voll/our_resources/graz_2002/ddrivenlrning/whatisddl/resources/tim_dll_learning_page.htm) (14.12.2005).

#### KONKORDANZPROGRAMME

- Antconc: <http://www.antlab.sci.waseda.ac.jp> (kostenlos) (14.12.2005).
- ConcApp: <http://www.edict.com.hk/PUB/concapp> (kostenlos) (14.12.2005).
- Wordsmith Tools: <http://www.lexically.net/wordsmith/index.html> (14.12.2005).
- MonoConc Pro: <http://www.athel.com/> (14.12.2005).