



Protokoll zum Vortrag von Joybrato Mukherjee zum Thema „English Learner Corpora: Linguistic Analysis and Language-Pedagogical Implications “ (10.07.2008)

1. Vortrag

Der Gastredner der Sitzung vom 10.07 ist Prof. Dr. Joybrato Mukherjee von der Justus Liebig Universität Giessen. In seinem Vortrag geht Mukherjee auf die Bedeutung und Möglichkeiten von Englisch-Korpora allgemein und auf die speziellen Anwendungsgebiete und -Möglichkeiten von English-learner corpora im genaueren ein.

Bei einem Korpus handelt es sich um eine Sammlung von Texten, welche als repräsentativ für die jeweilige Sprache, den Dialekt oder die Untergruppe einer Sprache gelten. Diese Sammlung sprachlicher Daten wird dann zur linguistischen Analyse der jeweiligen Sprache genutzt. Obwohl es Sprachkorpora schon seit wesentlich längerer Zeit gibt, ist es erst mit der Erfindung von Computern möglich geworden, komplexe Strukturen des Sprachgebrauchs mithilfe von Computer Korpora zu identifizieren und zu analysieren. Im Gegensatz zu anderen Methoden ermöglichen es Computer Korpora, bestimmte sprachliche Strukturen und Wendungen auf verlässliche Art und Weise zu zählen, zu vergleichen, zu sortieren und zu annotieren.

Die Korpuslinguistik unterscheidet sich in ihrem Ansatz damit deutlich von dem der Generativen Grammatik. Während diese Schule der Linguistik vor allem von der Intuition des Sprechers einer Sprache ausgeht, ist die Hauptannahme in Korpuslinguistik, dass unsere Intuition als Muttersprachler nicht immer verlässlich ist. Um festzustellen, wie der tatsächliche Sprachgebrauch ist, verlässt sich die Korpuslinguistik daher nicht auf Intuition sondern auf konkrete zählbare Beobachtungen.

Anhand verschiedener Beispiele (Gebrauch von 'would' in if-clauses, am häufigsten verwendeter Kontext des Worts 'order') demonstriert Herr Mukherjee, dass Häufigkeiten des Sprachgebrauchs durch bloßes Sprachgefühl nicht immer richtig einzuschätzen sind. Des weiteren werden Wörter häufig in größeren Sinnzusammenhängen verwendet. Daher ist es essentiell, dass auch in der linguistischen Analyse der Fokus nicht auf einzelnen Wörtern, sondern auf größeren Sinnabschnitten liegt.

Mittlerweile gibt es sehr viele verschiedene Korpora, die sich vor allem in ihrer Größe und in der in ihnen enthaltenen Sprache/Dialekt/Varietät unterscheiden. Der British National Corpus enthält 100 Millionen Wörter schriftlicher und mündlicher Sprache und ist damit einer der größten Korpora. Andere, teilweise frei verfügbare sogenannte „Mega-Korpora“ sind COBUILD, ICE und MICASE.

Der praktische Nutzen der Korpuslinguistik zeigt sich unter anderem darin, dass heutzutage auch Wörterbücher auf den Informationen aus großen Korpora basieren. Relevant ist dies vor allem im Hinblick auf Informationen bezüglich der Häufigkeit eines Wortes und der Reihenfolge verschiedener Bedeutungen sowie im Bezug auf gängige Wortverbindungen und Beispielsätze. Für den Englischunterricht kann Korpuslinguistik in zweierlei Hinsicht wichtig sein: zum einen kann sie die Inhalte des Unterrichts bestimmen, wie beispielsweise durch die Verbesserung von Lehrbüchern und die Reihenfolge des Lernens von irregulären Verben. Zum anderen kann sie methodisch genutzt werden, indem Korpusmaterialien selbst für den Unterricht verwendet werden.

In der zweiten Hälfte seiner Präsentation geht Herr Mukherjee genauer sogenannte Learner-Corpora ein, eine spezielle Art von Korpora welche gesammelte Daten von Lernern einer Sprache enthalten. Hierbei unterscheidet man zwischen Reference Learner Corpora, welche die Daten von unbekanntem Lernern enthalten, und Local Learner Corpora, die Sprachdaten der eigenen Lernergruppe enthalten. Des Weiteren lässt sich zwischen Lernerkorpora der gesprochenen und der geschriebenen Sprache unterscheiden. Der International Corpus of Learner English (ICLE) enthält 2,5 Millionen Wörter aus Essays von Englischstudenten verschiedener Muttersprachen im dritten oder vierten Lehrjahr. Vergleicht man bestimmte Daten dieses Korpus mit einem vergleichbaren Muttersprachler-Korpus, so lassen sich beispielsweise sehr gut zu häufig oder aber zu selten verwendete Begriffe und Formulierungen aufzeigen. In ähnlicher Weise zeigt ein Vergleich eines Learner Corpus of spoken English (LINDSEI) mit einem Muttersprachlerkorpus, dass der Gebrauch von Füllwörtern entscheidend ist für die gefühlte Sprachfertigkeit eines Lernenden. Wie Mukherjee betont, gelten all diese Schlüsse jedoch nur solange man der von Muttersprachlern verwendeten Sprach- und Ausdrucksweise eine Normfunktion beimisst.

Lokale Korpora wie beispielsweise der Giessen-Göttingen Local Learner Corpus of English (GLLC) sind zwar bedeutend kleiner als Referenzkorpora, ermöglichen es jedoch, Veränderungen und Fortschritte einer bestimmten Gruppe von Lernenden über längere Zeiträume zu analysieren und den Sprachgebrauch dieser Gruppe zu beschreiben. Durch die Annotation von Fehlern ist es möglich, typische und häufige Fehler der Lernenden genau zu bestimmen und auch hier mögliche Verbesserungen zu messen. Diese Korpora bieten daher die Möglichkeit, den Englischunterricht auf den tatsächlichen Sprachgebrauch einer Gruppe anzupassen. Wie Herr Mukherjee betont, wäre auf diesem Gebiet eine verstärkte Kooperation von LuL und Linguisten auf lokaler Ebene wünschenswert.

2. Diskussion

Viele der in der Präsentation vorgestellten Aspekte und Potentiale von Korpuslinguistik werden in der darauffolgenden Diskussion wieder aufgegriffen. Teilweise handelt es sich dabei um praktische Fragen, welche die Anwendungsmöglichkeiten von Korpora für LuL und Lernende betreffen, wie die Frage nach einem englischen Kernwortschatz auf Basis von Korpora oder nach geeigneten Programmen zum Aufbau eigener lokaler Korpora. Wie Herr Mukherjee erklärt gibt es bereits solche Programme, die es beispielsweise ermöglichen, die Facharbeiten verschiedener Schüler/Klassen in eine Art lokalen Lernerkorpus umzuwandeln.

Eine weitere Frage bezieht sich auf die Kooperation von Wissenschaft und Schule im Hinblick auf lokale Korpora. Hier betont der Referent, dass diese aufgrund des enormen Arbeitsaufwands noch stark von persönlichen Beziehungen und einem abzusehenden Mehrwert für die involvierten Lehrer abhängig ist. Seiner Meinung nach ist es unrealistisch, dass jeder Englischlehrer seinen eigenen lokalen Korpus aufbaut und diesen für den Englischunterricht der jeweiligen Klasse nutzt. Wesentlich wichtiger sei es, dass angehende Lehrer schon während des Studiums mit den Grundlagen von Korpora in Kontakt kommen und so später leichter die bereits vorhandenen digitalen Angebote nutzen können.

Ein weiterer Diskussionspunkt bezieht sich auf den Gebrauch von Füllwörtern zur Verbesserung der Redeflüssigkeit. Im Sinne der Korpuslinguistik müssten diese verstärkt auch von Lernenden in den Sprachgebrauch integriert werden, da sie die gefühlte Redeflüssigkeit und Sprachkompetenz stark beeinflussen und im Vergleich zu Muttersprachlern deutlich „zu wenig“ benutzt werden. Hier wenden jedoch einige Studierende ein, dass sie den gezielten Gebrauch von Füllwörtern im Englischunterricht nicht als sinnvoll betrachten, da die Schüler zunächst lernen sollten, sich klar und präzise auszudrücken und man den Gebrauch von Füllwörtern vor allem durch einen Auslandsaufenthalt sehr schnell lernen könne. An dieser Debatte werden zwei Probleme deutlich: zum einen scheint es sich um ein Generationsproblem zu handeln, gerade viele ältere Englischlehrer kritisieren den Gebrauch von Füllwörtern im Unterricht und mahnen die Schüler an, diese zu vermeiden. Dies lässt sich vielleicht teilweise auf einen geringeren direkten Kontakt zu englischen Muttersprachlern während der eigenen Ausbildung zurückführen. Zum anderen wird daran jedoch auch deutlich, dass über die Zielnorm des Englischunterrichts sehr unterschiedliche Auffassungen existieren. Die Frage, welche Art von Englisch im Klassenraum gelehrt und verwendet werden soll, ist bisher nicht einheitlich beantwortet worden und der Druck steigt sowohl in Hinblick auf die gesprochene Sprache als auch auf das schriftliche Englisch. Letztlich hängt es bisher noch stark von der Grundauffassung der LuL ab, welche Schwerpunkte gesetzt werden und ob Korpora als ein Mittel zur Unterrichtsgestaltung und -planung genutzt werden oder nicht.