

Zeitschrift für Interkulturellen Fremdsprachenunterricht

Sprachkorpora im Phonetikunterricht

Ulrike Gut

Studium der Anglistik und Psychologie in Mannheim und Cambridge (GB), Promotion in Mannheim und Habilitation in Freiburg. Assistenz an der Universität Bielefeld und Vertretungsprofessur an der Universität Osnabrück. Derzeit tätig als Juniorprofessorin an der Universität Freiburg. Forschungsschwerpunkte: Erst- und Zweitspracherwerb, Phonetik und Phonologie, Varietäten des Englischen, Korpuslinguistische Methoden.

E-Mail: ulrike.gut@anglistik.uni-freiburg.de

Erschienen online: 1. Mai 2007

© *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 2007

Abstract. Sprachkorpora, repräsentative Sammlungen authentischer Sprache angereichert mit zusätzlichen linguistischen Informationen, erfreuen sich einer zunehmenden Beliebtheit im Fremdsprachenunterricht. Während im fremdsprachlichen Wortschatz- und Grammatikunterricht Korpora bereits vermehrt als eine Möglichkeit des induktiven Lernens genutzt sowie als Werkzeuge zur Unterrichtsvorbereitung eingesetzt werden, ist diese Methode im Phonetikunterricht bisher noch weitgehend ungenutzt. Dieser Artikel beschreibt die Möglichkeiten des Einsatzes verschiedener Arten von Sprachkorpora im fremdsprachlichen Phonetikunterricht. Dabei werden sowohl die Vorteile der korpusbasierten Lehre für Lernende als auch der Einsatz von phonetisch annotierten Korpora für Lehrende zur Unterrichtsvorbereitung und der Entwicklung korpus-basierter Materialien vorgestellt.

1. Einleitung

Sprachkorpora sind Sammlungen gesprochener oder geschriebener Sprache in computerlesbarer Form, repräsentativ für eine Sprache oder Varietät (z.B. das Hochdeutsche) und angereichert mit zusätzlichen linguistischen Informationen, beispielsweise mit orthographischen oder phonetischen Transkriptionen, Transkriptionen der Wortart oder der syntaktischen Kategorien (z.B. McEnery & Wilson 2001). Da sie große Mengen authentischer Sprache enthalten, erfreuen sie sich nicht nur in der linguistischen Forschung sondern seit kurzer Zeit auch im Fremdsprachenunterricht einer zunehmenden Beliebtheit. Im korpusbasierten Unterricht werden die Lernenden im Gegensatz zum regelbasierten Unterricht ermuntert, sich mithilfe von Sprachkorpora muttersprachlichen Sprachgebrauch sowie sprachliche Regularitäten und Zusammenhänge auf den Gebieten des

Wortschatzes, der Grammatik und der Aussprache eigenständig zu erschließen. Derzeit werden in der internationalen Forschung zur Fremdsprachenpädagogik besonders die Vorzüge des induktiv-entdeckenden Lernens mithilfe von Sprachkorpora betont (u.a. Johns 1991, 1994; Leech 1997; Partington 1998; Flowerdew 2001; Ragan 2001; Tribble 2001; Todd 2001; Weskamp 2001; Mukherjee 2002; Bernardini 2004). Insbesondere sollen die Vorteile korpusbasierten Lernens darin liegen, dass die Aufmerksamkeit von Lernenden auf bestimmte sprachliche Strukturen gelenkt wird und sie sprachliche Zusammenhänge anhand authentischen Sprachmaterials selbst entdecken können. Dies soll zu einer erhöhten Sensibilität für die Strukturen und den Gebrauch von Fremdsprachen führen und die Lernerautonomie unterstützen. Beispielsweise kann neues Vokabular gelernt werden, indem Lernende ein Wort im Zusammenhang unterschiedlicher Äußerungen antreffen und sich dessen Bedeutung ohne Zugriff auf ein Wörterbuch erschließen. Ebenso können grammatikalische Regeln durch die Auseinandersetzung mit dem authentischen Sprachmaterial und ohne Rückgriff auf Grammatikbücher erworben werden. Es wird verschiedentlich postuliert, dass diese korpusbasierte Lernform der Arbeit von wissenschaftlichen Forschenden gleicht, da Lernende eine Fragestellung entwickeln und diese anhand realer Daten durch die Techniken des Beobachtens und der Interpretation beantworten müssen (u.a. Johns 1991, 1994; Leech 1997; Bernardini 2004). Dies soll den zusätzlichen Nutzen beinhalten, dass Lernende Fähigkeiten des kritischen Denkens erwerben, an computerbasierte linguistische Analysen herangeführt werden und durch selbstbestimmtes Arbeiten motivierter und erfolgreicher lernen.

Ein weiteres Anwendungsfeld von Korpora im Fremdsprachenunterricht liegt in ihrer unterstützenden Rolle für die Unterrichtsvorbereitung und Materialentwicklung. Hier sollen Sprachkorpora eine Ergänzung oder gar eine Alternative zu herkömmlichen Lehrbüchern und Übungsmaterialien darstellen, da die in den Korpora enthaltene Menge an muttersprachlichen Daten die tatsächliche Häufigkeit sprachlicher Strukturen und Regularitäten widerspiegelt, die nicht selten im Gegensatz zu in Schulbüchern aufgeführten grammatikalischen Regeln stehen. Beispielsweise wird in Schulbüchern zur englischen Grammatik nicht die häufigste Alternative von Verbtypen in *if*-Sätzen und den dazugehörigen Hauptsätzen gelehrt, wie Mindt (1997) durch eine korpusbasierte Studie feststellen konnte. Sprachkorpora sollen weiterhin dazu dienen, Lehrinhalte (neu) zu definieren: Lehrende können Korpora nutzen, um den muttersprachlichen Sprachgebrauch und die Häufigkeit verschiedenster grammatikalischer Strukturen in einer Sprache zu erheben und auf dieser Basis zu einer nach der Auftretenshäufigkeit erstellten Abfolge der Lehrinhalte für den Unterricht zu gelangen. Sprachkorpora können auch als Datenquelle für die Erstellung von authentischen Unterrichtsmaterialien und Übungen dienen, die – im Gegensatz zu manchen Beispielsätzen aus Lehrbüchern – den tatsächlichen Sprachgebrauch von Muttersprachlern und -sprachlerinnen widerspiegeln (z.B. Fox 1998; Osborne 2004). Ebenso eignen sich Korpora als Hilfestellung für die Korrektur von Lernerleistungen in schriftlichen und mündlichen Prüfungen, indem sie die Variationsvielfalt und -möglichkeiten muttersprachlicher Sprachverwendung aufzeigen (Mukherjee 2004).

Die überwiegende Anzahl bisheriger Erfahrungen mit korpusbasiertem Unterricht bezieht sich auf den Einsatz von Korpora für die Erweiterung des Wortschatzes oder den korpusbasierten Erwerb von pragmatischen oder morphosyntaktischen Eigenschaften von Sprache (u.a. Johns 1994; Fox 1998; Mukherjee 2002; Osborne 2004). Korpusbasiertes Lernen von Phonologie und Phonetik ist bisher sehr unterrepräsentiert (Gut 2005, 2006b). Es ist das Ziel dieses Artikels, Möglichkeiten des Einsatzes von Sprachkorpora im fremdsprachlichen Phonetikunterricht aufzuzeigen. Im folgenden Abschnitt werden einige für den deutschen Phonetikunterricht geeignete Korpora vorgestellt. Abschnitt 3 diskutiert verschiedene allgemeine Aspekte und Möglichkeiten des Lernens mit Korpora. Im Abschnitt 4 wird der Nutzen von Sprachkorpora für Lehrende in der Phonetik beschrieben. Abschnitt 5 präsentiert einige Beispiele für phonetische Übungseinheiten und Aufgaben, die im Unterricht Verwendung finden können. Abschließend werden die Grenzen des Einsatzes von Korpora im Phonetikunterricht diskutiert und ein Ausblick gegeben.

2. Geeignete Korpora für den deutschen Phonetikunterricht

Als Sprachkorpus wird eine Sammlung sprachlicher Daten bezeichnet, die sowohl durch ihren Umfang als auch durch die sorgfältige Auswahl der enthaltenen Sprache (z.B. bezüglich Sprechstil, Kommunikationssituation) sowie der enthaltenen Sprecher (z.B. bezüglich Geschlecht, Alter, sozialer Gruppenzugehörigkeit) als repräsentativ und als Standardreferenz für eine bestimmte Sprache oder Sprachvarietät gelten kann (McEnery & Wilson 2001). Ein repräsentatives Korpus für das gesprochene Hochdeutsch enthält beispielsweise mehrere hundert Aufnahmen von Monologen (Reden, Vorträgen), Dialogen (persönlicher und formeller Natur) und Gesprächen mit mehreren Teilnehmenden, mit männlichen und weiblichen Sprechern, die aus allen Altersgruppen stammen. Ein Korpus besteht aus den so genannten Primärdaten, die die Audio- oder Videoaufnahmen oder schriftlichen Texte umfassen, und zusätzlichen Annotationen. Annotationen bestehen aus linguistischen oder nicht-linguistischen Informationen. Linguistische Informationen sind beispielsweise orthographische oder phonetische Transkriptionen oder Klassifikationen von Wortarten oder einzelnen Lemmata. Nicht-linguistische Informationen bestehen zumeist aus Metadaten zur Korpuserhebung oder den aufgenommenen Sprechern wie beispielsweise deren Alter, Geschlecht, Muttersprachen oder Bildungsstand. Phonetisch annotierte Sprachkorpora sind üblicherweise zeitlich aligniert, d.h. jedes annotierte Element ist mit Zeitmarken versehen, die einen direkten Zugriff auf die entsprechenden Stellen der Audio- oder Videodaten erlauben. Abbildung 1 zeigt die Annotation der Äußerung „*Ich mag diese Sprache eigentlich*“. Unter dem Oszillogramm befinden sich zwei Annotationsebenen. Die obere beinhaltet eine orthographische Transkription der Äußerung und die Wortgrenzen jedes Wortes. In der unteren sind eine phonetische Transkription in SAMPA¹, der computerlesbaren Form von IPA (Wells et al. 1992), sowie die Grenzen jedes einzelnen Phonems enthalten.

Ulrike Gut, Sprachkorpora im Phonetikunterricht. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 12:2, 2007, 21 S.

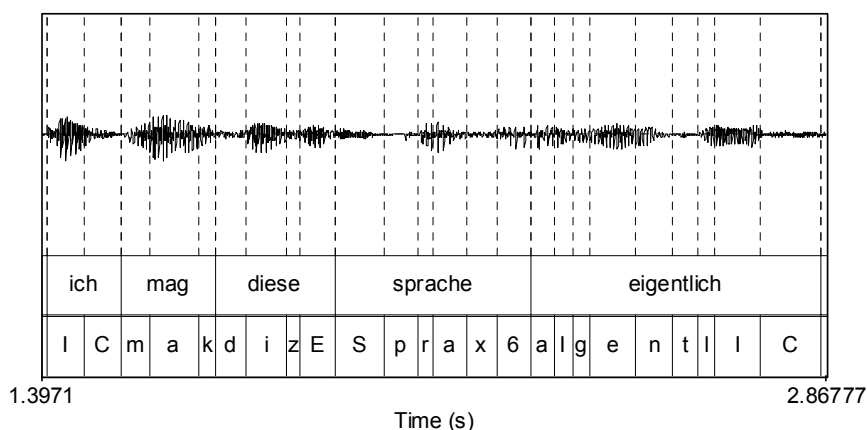


Abbildung 1. Annotation der Äußerung „Ich mag diese Sprache eigentlich“ mit Oszillogramm, orthographischer und phonetischer (SAMPA-)Transkription.

Mithilfe eines phonetischen Softwareprogramms wie *Praat* (<http://praat.org>) wird durch das Anklicken jedes Elements der entsprechende Teil des Audiosignals wiedergegeben. Durch elektronische Speicherung sind moderne Korpora computerlesbar und ermöglichen somit eine schnelle und (halb-)automatische Analyse der Daten. Die wissenschaftliche Nutzung von Sprachkorpora basiert auf dieser (halb-)automatischen Analyse der Annotationen. Mittels geeigneter Software oder Datenbankanfragen können die Inhalte von Korpora quantitativ-statistisch und qualitativ ausgewertet werden.

Es lassen sich verschiedene Arten von Korpora unterscheiden. In muttersprachlichen Korpora stammen die Primärdaten von Muttersprachlern der jeweiligen Sprache. Lernerkorpora hingegen enthalten Primärdaten, die von Lernern einer Sprache produziert wurden. Monolinguale Korpora bestehen aus Primärdaten nur einer Sprache oder Sprachvarietät. Multilinguale Korpora beinhalten Daten von mehr als einer Sprache oder Sprachvarietät. Die überwiegende Anzahl aller derzeit verfügbaren Korpora sind englische Korpora, jedoch werden zunehmend deutsche Korpora erstellt. Während die meisten Korpora orthographische Annotationen und Annotationen im Bereich der Syntax (Wortarten, syntaktische Konstituenten) und des Lexikons (Lemmata) aufweisen, sind phonetisch annotierte Korpora noch rar. Im Folgenden sollen einige deutsche Korpora, die phonetische Annotationen enthalten, vorgestellt werden: die muttersprachlichen Korpora Kiel-Korpus, die phonDAT-Korpora des BAS sowie das Lernerkorpus Leap.

2.1 Das Kiel-Korpus

Das Kiel-Korpus ist eine Sprachdatensammlung der deutschen Lese- und Spontansprache, welche seit 1990 an der Universität Kiel aufgezeichnet und annotiert

wird. Derzeit umfassen die CD-ROMs des Kiel-Korpus mehr als vier Stunden annotierter Lesesprache sowie knapp vier Stunden annotierter Spontansprache. Das spontansprachliche Korpus besteht aus Terminvereinbarungsdialogen 52 verschiedener Sprecher und umfasst orthographische und phonemische Transkriptionen sowie die Segmentierung der Primärdaten in Phoneme. Der Korpusteil mit gelesener Sprache beinhaltet 26 weibliche und 27 männliche Sprecher, die insgesamt 4932 Wörter vorlesen. Auch dieser Teil ist orthographisch und phonemisch transkribiert und umfasst eine Segmentierung einzelner Phoneme. Das gesamte Korpus ist aligniert; die Annotationen liegen in dem Format *xwaves+* vor, das in *Praat* eingelesen werden kann.

2.2 Das Bayrische Archiv für Sprachsignale

Das Bayrische Archiv für Sprachsignale (BAS) verfügt über eine Reihe von kostenpflichtigen muttersprachlichen deutschen Korpora, die teilweise spontane und teilweise vorgelesene Primärdaten enthalten. Das PhonDat2-Korpus beispielsweise wurde an den Universitäten Kiel, Bonn und München erstellt und besteht aus 3200 vorgelesenen Sätzen zum Thema Zugauskunft. Das Korpus ist repräsentativ für die verschiedenen regionalen Varianten des Deutschen in den drei genannten Gebieten und beinhaltet Sprache von insgesamt 16 Sprechern (sechs weiblich und 10 männlich) verschiedenen Alters, für die einige Metadaten zur Verfügung stehen. Das Korpus ist orthographisch, phonemisch und prosodisch annotiert. Jedes Wort ist orthographisch und jedes Phonem in SAMPA transkribiert. Zusätzliche Transkriptionen beziehen sich auf Betonung, Intonation und prosodische Phrasierung. Die Transkriptionen wurden automatisch mit dem Sprachsignal aligniert und können mit der entsprechenden Software im Unterricht genutzt werden.

2.3 Das LeaP-Korpus

Das LeaP-Korpus ist ein Lernerkorpus und wurde in den Jahren 2001 bis 2003 im Rahmen des LeaP (Learning Prosody in a Foreign Language) Projektes an der Universität Bielefeld erstellt. Es besteht aus einer Gesamtlänge von 12 Stunden und umfasst 359 annotierte Aufnahmen mit insgesamt 73941 Wörtern. Das Korpus ist bilingual und beinhaltet Sprache von 46 Lernenden des Englischen und 55 Lernenden des Deutschen. Zusätzlich wurden eine kleine Anzahl deutscher und englischer Muttersprachler aufgenommen. Die Korpusaufnahmen umfassen vier verschiedene Sprechstile: frei gesprochene Sprache in einem Interview (Dauer zwischen 10 and 30 Minuten), eine vorgelesene kurze Geschichte (Dauer ungefähr 2 Minuten), eine Nacherzählung der Geschichte (Dauer zwischen 2 und 10 Minuten) und eine vorgelesene Wortliste mit Nonsense-Wörtern (30 bis 32 Wörter). Die verschiedenen Lernergruppen, die im LeaP-Korpus repräsentiert sind, bestehen aus sehr fortgeschrittenen Lernenden (*near-natives*), die von Muttersprachlern praktisch nicht zu unterscheiden sind, aus Lernenden, die vor und nach einem Prosodiekurs aufgenommen wurden, und aus Lernenden, die vor und nach einem sechs- bis neunmonatigen Auslandsaufenthalt aufgenommen wurden. Ziel der Korpuserstellung war, eine repräsentati-

ve Lernergruppe zu erfassen, die sich möglichst gemischt bezüglich ihres Alters, Geschlechterverteilung, Muttersprachen, Lernniveau, Lerndauer, Alter bei Lernbeginn und nicht-linguistischen Faktoren wie Motivation, Musikalität und so weiter darstellt. In den Interviews wurde eine große Anzahl von Metadaten erhoben wie beispielsweise Daten

- zur Aufnahme (Datum, Ort, Interviewer und Sprache des Interviews)
- zum Sprachenlernenden (Alter, Geschlecht, Muttersprache, andere Sprachen, Alter beim ersten Kontakt mit der Zielsprache, Art des Kontakts [*formal* vs. *natürlich*], Dauer und Art von Auslandsaufenthalten, Dauer und Art von Unterricht in Prosodie [falls überhaupt], prosodisches Wissen)
- zur Motivation und Fähigkeiten der Lernenden (Gründe für den Spracherwerb, Integrationswilligkeit im Zielland, angegebene Wichtigkeit von Kompetenzen in der Aussprache im Vergleich zu anderen Teilgebieten von Sprache [Syntax, Wortschatz], musisches und schauspielerisches Interesse, Erfahrung und Fähigkeit)

Das LeaP-Korpus verfügt über acht verschiedene Annotationsebenen. Neben einer orthographischen und einer phonetischen Transkription wurden die Intonation, die vokalischen und konsonantischen Intervalle sowie die Grenzen aller Silben, Wörter, Intonationsphrasen und nicht-sprachlicher Ereignisse wie Geräusche und Gelächter annotiert. Wortarten und Lemmata-Annotationen wurden automatisch ergänzt. Abbildung 2 zeigt einen kleinen Ausschnitt aus dem LeaP-Korpus mit der orthographischen und der phonetischen Transkription.

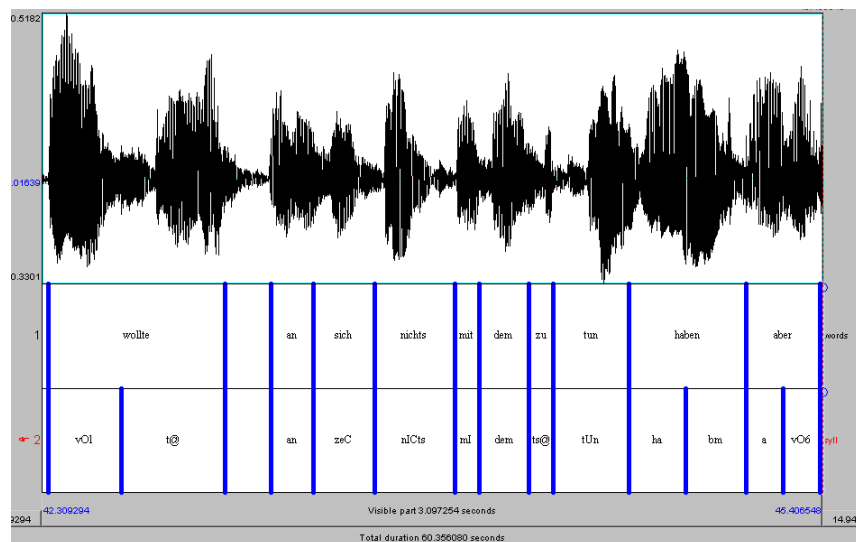


Abbildung 2. Annotation der Äußerung „wollte an sich nichts mit dem zu tun haben aber“ mit Oszillogramm, orthographischer und phonetischer Transkription.

Für die wissenschaftliche Nutzung ist das Korpus erhältlich im XML-basierten Format TASX (Milde & Gut 2002, 2004), zusammen mit einer Reihe automatischer Analysewerkzeuge. Korpusanfragen sind weiterhin möglich basierend auf einer SQL-Datenbank oder mit dem NITE Werkzeug *NXT Search* (Gut et al. 2004). Für den Einsatz des Korpus im Fremdsprachenunterricht sind die Annotationen zusätzlich erhältlich im *Praat* Format.

3. Korpusbasiertes Lernen

Eine zunehmende Anzahl von wissenschaftlichen Publikationen widmet sich dem Einsatz korpusbasierter Methoden im schulischen und universitären Fremdsprachenunterricht (u.a. Botley et al. 1996; Ghadessy et al. 2001; Kettemann & Marko 2002; Granger et al. 2002; Sinclair 2004; Connor & Upton 2004). Darin wird besonders betont, dass korpusbasierter Unterricht eine gute Möglichkeit des induktiven Lernens darstellt (z.B. Todd 2001). Sprachlernende können, wenn sie einer linguistischen Fragestellung mithilfe eines Korpus nachgehen, die Techniken der Datenbeobachtung und Interpretation erlernen und erhalten somit die Möglichkeit, wissenschaftliches Arbeiten wie die Erstellung einer Hypothese und deren Überprüfung anhand realer Sprachdaten kennen zu lernen (Johns 1991, 1994; Leech 1997). Diese „datengetriebene“ Arbeitsweise schult nicht nur ihre kritischen und methodischen Fähigkeiten, sondern fördert auch die Autonomie des Lernens. Beispielsweise können Lernende neue Wortbedeutungen – ohne die Bedeutung in einem Lexikon nachzuschlagen – erwerben, indem sie in einem Korpus das Wort in verschiedenen semantischen Kontexten antreffen. Dieser entdeckende Lernstil ermutigt Lernende in höchstem Maße, ihren eigenen Interessen nachzugehen und bietet ihnen die Möglichkeit, ihre Fähigkeiten weiter zu entwickeln (Bernardini 2004: 23).

Ebenso soll durch korpusbasierten Unterricht die Sprachbewusstheit gefördert werden. Diese wiederum kann motivationssteigernd wirken und somit den Lernerfolg erhöhen. In der Zweitspracherwerbsforschung wurde gezeigt, dass die Aufmerksamkeitslenkung auf eine bestimmte sprachliche Struktur Lernen fördert (z.B. VanPatten & Cadierno 1993; Ellis 1995; Robinson 1996). Durch die Beschäftigung mit echter Sprache im Gegensatz zu fiktiven Beispielen wird die Sensibilität für sprachliche Strukturen geschärft und *language awareness* kann sich entwickeln (Weskamp 2001).

Die neuesten Entwicklungen in der korpusbasierten Fremdsprachendidaktik unterstreichen die Vorteile des kombinierten Einsatzes von muttersprachlichen und Lernerkorpora im Unterricht. Während bei der Arbeit mit muttersprachlichen Daten in einem **top-down** Prozess Regeln des Sprachgebrauchs erlernt werden, bietet der zusätzliche Einsatz von Lernerkorpora, bestehend aus Sprache, die von Fremdsprachenlernern produziert wurde, eine **bottom-up** Perspektive, die es erlaubt, typische Lernerfehler zu erkennen und für eigene Fehler sensibilisiert zu werden (Granger & Tribble 1998; Meunier 2002; Nesselhauf 2004; Bernardini 2004; Osborne 2004). Beispielsweise ergibt ein Vergleich von englischen muttersprachlichen und fremdsprachlichen Daten, dass Lernende die

Wörter „*interesting*“ und „*important*“ zu häufig verwenden und Sätze in Texten zu häufig mit bestimmten Konnektoren (z.B. „*indeed*“, „*in fact*“, „*anyway*“) verbinden (Osborne 2004). Die Beschäftigung mit Lernerdaten ermöglicht es, typische Schwierigkeiten von Sprachlernern zu identifizieren und eine Sensibilität für eigene Fehler zu entwickeln oder zu steigern.

Eine Reihe von Publikationen betont die Vorzüge von Korpora im (zumeist englischen) Fremdsprachenunterricht für den Erwerb von Wortbedeutungen, Kollokationen, Phrasen, grammatischen Phänomenen wie dem Gebrauch von Zeiten, *count nouns* und *non-count nouns* und Konnektoren sowie Teilbereichen der Pragmatik (Partington 1998; Flowerdew 2001; Ragan 2001; Tribble 2001; Mukherjee 2002; Osborne 2004) und präsentiert konkrete Übungen und Unterrichtsmaterialien für diese sprachlichen Strukturen. Der Einsatz von Sprachkorpora im Phonetikunterricht und korpusbasierte phonetische Übungsmaterialien hingegen wurden bisher erst in wenigen Publikationen beschrieben (Gut 2005: 2006b). Im Folgenden sollen daher die Möglichkeiten und Grenzen des Einsatzes von Sprachkorpora im Phonetikunterricht dargestellt werden. Zunächst werden Bereiche aufgezeigt, in denen Lehrende von Sprachkorpora profitieren können. Daran schließt sich eine Beschreibung möglicher Unterrichtseinheiten und phonetischer Themen für einen korpusbasierten Unterricht an.

4. Sprachkorpora im Phonetikunterricht: Nutzung durch Lehrende

Für Phonetiklehrende hat der Einsatz phonetisch annotierter Korpora zwei herausragende Vorteile: zum einen können sie zuverlässige Informationen über tatsächliche Häufigkeiten phonetischer Phänomene sowohl in muttersprachlicher als auch in Lernerdsprache erhalten. Zum anderen ermöglicht ein Korpus ihnen, die akustischen Korrelate authentischer Sprache kennen zu lernen. Diese korpusbasierten Erkenntnisse können dazu führen, Prioritäten im fremdsprachlichen Phonetikunterricht neu zu setzen und zu anderen Maßstäben für die Rückmeldung und Bewertung der Aussprache der Lernenden zu gelangen. Einige dieser Möglichkeiten werden im Folgenden anhand konkreter Beispiele exemplarisch dargestellt.

Ein repräsentatives Korpus gibt Auskunft über die tatsächliche Variabilität der Aussprache einzelner Wörter von Muttersprachlern. Beispielsweise ergibt eine Suchanfrage an den muttersprachlichen Teil des LeaP-Korpus, dass das Wort *dem* in den Varianten [de:m], [dem], [dəm] und [m] auftritt, wobei die Variation der Aussprache besonders augenscheinlich in Abhängigkeit vom Sprechstil und somit der Sprechgeschwindigkeit steht. Auch bei einzelnen Lauten lassen sich unter Zuhilfenahme der spektrographischen Darstellungen erkennen, dass deren Realisation von einer Reihe von Faktoren bedingt wird: der Position des Lautes innerhalb einer Silbe, eines Wortes und einer Phrase, dem Betonungsgrad sowie der phonetischen Umgebung des Lautes. Ein /a/ am Anfang eines Wortes unterscheidet sich signifikant von einem /a/ wortmedial und in unbetonter Position, ein /a/ nach einem /n/ ist bedingt durch die Koartikulation, d.h. die artikulatorische Beeinflussung eines Lautes durch die Artikulation der vorangehenden und

nachfolgenden Laute, phonetisch sehr unterschiedlich von einem, das von einem /k/ gefolgt wird usw. Diese Erkenntnis, dass bei der Aussprache nicht einfach Phoneme aneinander gereiht werden, sondern dass sich einzelne Laute in verschiedenen Positionen bis zur Unkenntlichkeit voneinander unterscheiden und starken prosodischen und koartikulatorischen Effekten ausgesetzt sind, könnte Lehrende der Phonetik dazu ermuntern, den Fokus im Unterricht weniger auf die Aussprache isolierter Laute als vielmehr die Aussprache größerer prosodischer Einheiten (Silben, Wörter, Phrasen) zu legen.

Ebenfalls können Lehrende der Phonetik durch die Betrachtung eines phonetisch annotierten Sprachkorpus mit einer Sprachanalysesoftware wie *Praat* systematischen Einblick in die prosodischen Eigenschaften von Sprache erhalten. Prosodie – insbesondere Betonung und Intonation (oder Tonhöhenverlauf) – ist im Zweitsprachenunterricht noch immer sehr unterrepräsentiert und wird von Lehrenden gerne ausgespart. Die Nutzung eines Sprachkorpus und der Visualisierungsmöglichkeiten der Software zeigt, dass Prosodie zu den elementarsten Eigenschaften von Sprache gehört und in untrennbarer Interaktion mit der Aussprache der Sprachlaute steht. Der Tonhöhenverlauf einer Äußerung bestimmt nicht nur deren semantischen und pragmatischen Gehalt mit, sondern ist auch mit der Betonung einzelner Silben und der Länge einzelner Laute verknüpft. Bei der Betonung einzelner Silben im Deutschen spielen Vokallänge, Intensität einzelner Vokalformanten und Tonhöhe zusammen. Besonders im Vergleich mit fremdsprachlichen Äußerungen lassen sich so vermeintlich phonologische Probleme – wie die „fälsche“ Platzierung einer Betonung – als phonetische Probleme identifizieren: die Lernende betont die richtige Silbe, nutzt dazu jedoch zu wenig Intensität, so dass die Silbe unbetont „klingt“. Auch diese korpusbasierten Beobachtungen können dazu beitragen, den Fokus im Phonetikunterricht zu erweitern oder neu zu setzen.

Die korpusbasierte Analyse von Lerner Sprache ermöglicht einen systematischen Vergleich der segmentalen und prosodischen Eigenschaften verschiedener Äußerungen mit denen von Muttersprachlern. Die Analyse des LeaP-Korpus mittels der Datenbank oder der angebotenen Suchwerkzeuge ermöglicht die Beschreibung einer Reihe von unterschiedlichen phonetischen Phänomenen. Die Betrachtung der Silbifizierung ergibt beispielsweise, dass Deutschlernende nur zu 61,9% den Glottalverschlusslaut [ʔ] zwischen einen Wortstamm, der mit einem Vokal beginnt, und die Vorsilben *ver-* und *un-* wie in den Wörtern *verantworten* oder *unendlich* einfügen. In 38,1% der Fälle wird der dem Vokal vorangehende Konsonant als Onset der folgenden Silbe produziert, also [ve.rantvoɐ̯tən]. Dieses Phänomen kann auch wortübergreifend beobachtet werden, beispielsweise in der Realisierung von *fast ein* als [fas.tam] und *gibt es* als [gɪp.təs]. Ebenso kann die Vokalreduktion von Deutschlernenden anhand des LeaP-Korpus untersucht werden. Es zeigt sich, dass diese andere Vokale in unbetonten *-en* Silben, z.B. in *treffen*, realisieren als Muttersprachler: bei den Lernenden können neben Vokaltilgung, den Reduktionsvokalen [ə] und [ɐ], die auch in muttersprachlichem Deutsch produziert werden, Vollvokale wie [a] und [e] beobachtet werden.

Gleichzeitig gibt das LeaP-Korpus Auskunft über die tatsächlichen Auftretenshäufigkeiten einiger als typisch erachteter Lernerfehler. Die Realisierung von Konsonantenclustern, wie beispielsweise am Ende des Wortes *trumpfst*, stellt einen solchen Bereich dar. In experimentellen Studien wird berichtet, dass einige L2-Lernende diese Cluster vereinfachen, indem ein oder mehrere Konsonanten getilgt werden, indem sie einen Schwa einfügen und somit den Cluster aufbrechen, oder einen Schwa am Ende des Wortes hinzufügen (z.B. Hansen 2001; Lleó & Vogel 2004). Eine Analyse des LeaP-Korpus ergibt, dass es sich bei den letzten zwei Alternativen um seltene Sonderfälle handelt: von 4049 Konsonantenclustern, die aus zwei Konsonanten am Wortende bestehen, zeigt kein einziges einen eingefügten Schwa und nur 3,5% einen hinzugefügten. 32,45% aller Cluster sind teilweise getilgt, was der Häufigkeit der teilweisen Tilgungen in muttersprachlichem Deutsch (25,4%) nahe kommt. Deutliche Unterschiede in der Konsonantenclusterreduktion zwischen muttersprachlichen und fremdsprachlichen Sprechern liegen in der Realisation längerer Cluster (drei und vier Konsonanten), die signifikant häufiger von den Lernenden reduziert werden, und in der Realisation von Clustern innerhalb von Wörtern (z.B. in *selbstverständlich*), die ebenfalls häufiger von Lernenden reduziert werden.

Ein besonderes Augenmerk wird durch die Analyse typischer Aussprachefehler von Deutschlernenden mithilfe des LeaP-Korpus auf die verschiedenen Bereiche der Prosodie gerichtet. Das Korpus zeigt, dass es signifikante Unterschiede zwischen Muttersprachlern und Fremdsprachenlernern in den Gebieten der Phrasierung, der Intonation und des Tonumfangs gibt. Deutschlernende produzieren im Durchschnitt deutlich kürzere Intonationsphrasen als Muttersprachler, sowohl im Vorlesesprechstil als auch beim Nacherzählen. Viele dieser Intonationsphrasen enden bei Deutschlernern mit einem steigenden oder einem mittelhohen Ton, während in muttersprachlichem Deutsch Intonationsphrasen typischerweise mit einem fallenden Ton beendet werden. Wenn Deutschlernende einen fallenden Ton am Satzende produzieren, so hat dieser mit durchschnittlich 3,8 Semitönen einen deutlich weniger ausgeprägten Umfang als der von Muttersprachlern mit durchschnittlich 5,67 Semitönen. Auch der Tonumfang innerhalb von Phrasen ist deutlich kleiner bei Deutschlernenden, sowohl im Vorlesesprechstil als auch beim Nacherzählen.

Das LeaP-Korpus bietet weitere Möglichkeiten zur Entdeckung typischer Lernerfehler und für Vergleiche muttersprachlichen und fremdsprachlichen Deutschs. Die Vokalreduktion in unbetonten Silben wie des Vokals im Suffix –*en* oder –*em* in *laufen* oder *diesem* ist eines der phonetischen Gebiete, die als problematisch für Deutschlernende beschrieben werden (Kaltenbacher 1998; Gut 2003, 2006b). Eine Analyse des Korpus zeigt, dass im L2 Deutsch unbetonte Silben im Durchschnitt und in Relation zur Länge betonter Silben zu lang sind. Für Lehrende bedeutet dies, dass sie ihre Intuitionen bezüglich der Fehlerhäufigkeit einzelner phonetischer Phänomene an den im Korpus repräsentierten tatsächlichen Häufigkeiten messen können. Vielleicht führt dies in dem einen oder anderen Fall dazu, dass der Fokus des Phonetikunterrichts stärker auf typi-


sche Fehlergebiete wie die Vokalreduktion und die Intonation als auf Gebiete, die seltener zu Schwierigkeiten führen, gelegt wird.

Das LeaP-Korpus kann weiterhin genutzt werden, um den Einfluss der Muttersprache von Deutschlernenden auf deren Aussprache zu demonstrieren. Teilt man die im Korpus enthaltenen Sprecher in solche, deren Muttersprache Konsonantencluster erlaubt (z.B. Englisch) und solche, in deren Muttersprache es keine Konsonantencluster gibt (z.B. Chinesisch), zeigt es sich, dass die zweite Gruppe signifikant häufiger zusätzliche Schwas am Wortende – wie beispielsweise in [undə] für *und* – produziert. Sprachspezifische Unterschiede jedoch zwischen L1-Italienisch und L1-Chinesisch Sprechenden ergeben sich nicht. Ebenso unterscheiden sich L1-Englisch von L1-Italienisch und L1-Chinesisch Sprechern durch die Häufigkeit von Vokalreduktionen, die zwei letztgenannten Sprechergruppen untereinander jedoch nicht. Vergleicht man den Tonumfang von L1-Englisch-, Italienisch-, Chinesisch- und Russischsprechenden ergibt sich ein deutlicher Unterschied zwischen den L1-Englischsprechenden, die einen größeren Tonumfang aufweisen, und den anderen Gruppen, jedoch kein Unterschied innerhalb der zweiten Gruppe.

Diese Erkenntnisse der korpusbasierten Analyse von Lerner Sprache und des Vergleichs mit Muttersprache können dazu dienen, korpusbasierte Übungen für den Phonetikunterricht zu entwickeln. Einige Vorschläge dafür sind im nächsten Abschnitt angeführt.

5. Korpusbasierte Übungen für Lernende

Es bedarf nur weniger technischer Voraussetzungen, um Sprachkorpora im Phonetikunterricht einsetzen zu können. Zum einen muss ein Computerraum zur Verfügung stehen, dessen Computer mit Kopfhörern oder Lautsprechern ausgestattet sind, und auf denen die Software *Praat* (frei erhältlich unter <http://www.praat.org>) oder eine andere Sprachanalysesoftware (e.g. *wavesurfer*, *TASX-annotator*) installiert sind. Es dauert ungefähr 10 Minuten, um die Lernenden mit den Basisfunktionen von *Praat* vertraut zu machen. Die ersten Anweisungen können wie folgt aussehen:

- | |
|---|
| <ol style="list-style-type: none"> 1. <i>Praat</i> durch Klicken auf das icon  öffnen
Praat.exe 2. Eine Audiodatei und die passende (gleichnamige) TextGrid Datei einlesen (im <i>Praat</i> objects window auf <i>Read</i>, dann <i>Read from file</i> klicken). 3. Beide Dateien auswählen und auf <i>Edit</i> klicken. 4. Nun können Sie das Sprachsignal und die Transkriptionen sehen. 5. Klicken Sie auf ein Wort, um es auszuwählen, und auf den darunter liegenden grauen Balken, um es anzuhören. |
|---|

Nach dem vierten Schritt ist die in Abbildung 3 dargestellte Ansicht erreicht. Das Oszillogramm erscheint oberhalb des Spektrogramms zusammen mit anderen optionalen phonetischen Visualisierungen wie der Intonationskurve, den Vokalformanten oder der Intensität. Darunter sind die verschiedenen Ebenen der Annotation zu sehen. Die Anzahl und Anordnung der einzelnen Annotationsebenen und der unterschiedlichen Visualisierungen der phonetischen Eigenschaften der Sprachaufnahme können individuell eingestellt werden. Durch die Verwendung von einer phonetischen Analysesoftware wie *Praat* haben die Lernenden gleichzeitig Zugang zu den Korpusannotationen und den Originalaufnahmen. Durch das Anklicken jedes beliebigen Elements in den Annotationsebenen (z.B. das Wort *Löwe* in Abbildung 3) wird der entsprechende Teil der Aufnahme wiedergegeben. Jeder Lernende kann individuell kontrollieren, welche Teilstücke er oder sie sich anhören möchte und kann diese mehrfach abspielen oder zu anderen Elementen springen.

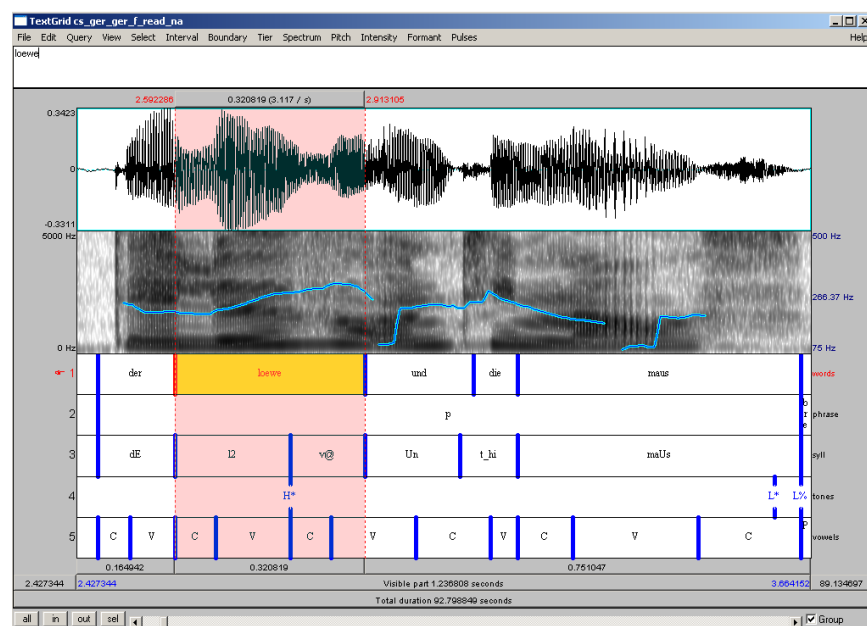


Abbildung 3. Von oben nach unten das Oszillogramm, das Spektrogramm mit Intonationskurve (blaue Linie) und fünf Annotationsebenen der Äußerung „Der Löwe und die Maus“. [Klangbeispiel 1](#)

Die zusätzlichen phonetischen Analysemöglichkeiten, die die Software bietet, können ebenso im Unterricht verwendet werden. Beispielsweise kann die Intonationskurve (dargestellt als die blaue Linie im Bereich des Spektrogramms in Abbildung 3) angezeigt werden und die Tonhöhe an jedem beliebigen Punkt der Intonationskurve gemessen werden. Auch können die Formanten aller Vokale auf diese Weise angezeigt und gemessen werden. Die exakte Dauer eines jeden Elements auf den einzelnen Annotationsebenen (beispielsweise jedes Wort, jede

Silbe oder jeder Vokal) wird auf der grauen Fläche unterhalb der Annotationsebenen angezeigt, sobald das Element ausgewählt wird. Abbildung 3 zeigt, dass das Wort *Löwe* genau 0.320819 Sekunden lang ist.

Im Folgenden werden einige korpusbasierte Übungen und Aktivitäten für den Phonetikunterricht für Deutschlernende vorgestellt, die sich besonders an fortgeschrittene Lernende wenden. Die Übungen sind unterteilt in die Gebiete: einzelne Sprachlaute, Intonation, Vokalreduktion und Betonung. Sie können sowohl im Präsenzunterricht als auch im Selbststudium eingesetzt werden. Ihnen gemein ist, dass sie das eigenständige Lernen fördern und es den Deutschlernern durch die relativ selbstbestimmte Beschäftigung mit phonetischen und phonologischen Phänomenen erlaubt, sich sprachliche Strukturen ohne Rückgriff auf ein Lehrbuch zu erschließen und darüber hinaus viele Phänomene, die nicht in Lehrbüchern erwähnt werden, zu entdecken.

5.1 Aktivitäten zu einzelnen Sprachlauten

5.1.1 Realisation von /ç/ und /x/

Um die muttersprachliche Aussprache und phonologische Verteilung einzelner Sprachlaute wie beispielsweise der Laute /ç/ und /x/ korpusbasiert zu üben, können Deutschlernende ein phonetisch annotiertes Korpus durchsuchen und sich die phonologische Regel oder regionalen Varianten selbst erarbeiten. Dazu wird ein Aufgabenblatt, auf dem alle zu untersuchenden Wörter und die entsprechende Instruktion aufgelistet sind, verteilt. Weiter fortgeschrittene Lernende können instruiert werden, sich die Spektrogramme der entsprechenden Laute zu betrachten und mit anderen, beispielsweise denen von /h/ zu vergleichen. Dies kann ihre perzeptiven Fähigkeiten für kleine phonetische Unterschiede schulen. Die Beschäftigung mit Spektrogrammen kann darüber hinaus zu der Erkenntnis führen, dass Sprache nicht aus einer Aneinanderreihung einzelner Phoneme besteht, sondern dass einzelne Sprachlaute sehr unterschiedliche Formen annehmen können bedingt durch ihre phonologische Umgebung. Aktivitäten zu einzelnen Sprachlauten können auch auf Lernerkorpora wie dem LeaP-Korpus basieren, um die Sensibilität für typische Lernerfehler zu erhöhen. Lernende werden beispielsweise instruiert, die Realisation dieser Phoneme durch die verschiedenen Sprecher im Korpus zu analysieren. Sie werden herausfinden, dass diese Phoneme manchmal durch andere Phoneme wie /h/ und /k/ ersetzt werden und ob diese Substitution durch die Position des Lautes innerhalb der Silbe und das phonetische Umfeld bedingt ist, z.B. wenn der Laut silbenfinal ist oder einem bestimmten Vokal folgt. Darüber hinaus kann die Frage untersucht werden, ob Lernende mit unterschiedlichen Muttersprachen unterschiedliche Substitutionen und Realisationen vornehmen.

5.1.2 Vokalqualität von /y:/ und /ø:/

Die Aussprache der gerundeten Vorderzungenvokale stellt für viele Deutschlernende eine Schwierigkeit dar, wie eine Untersuchung des LeaP-Korpus zeigt.

Auch für diese Laute können korpusbasierte Übungen erstellt werden, beispielsweise durch einen kombinierten bottom-up und top-down Ansatz. Mithilfe der muttersprachlichen und lernersprachlichen Aufnahmen der vorgelesenen und nacherzählten Geschichte im LeaP-Korpus können die Lernenden die Vokalqualität von /y:/ und /ø:/in beiden Sprechergruppen vergleichen. Diese Aktivität kann als rein auditive Analyse durchgeführt werden oder kann eine zusätzliche akustische Analyse der Formantenstruktur dieser Vokale beinhalten. Die ersten zwei Formanten, F1 und F2, die in *Praat* durch die roten gepunkteten Linien angezeigt werden, reflektieren die Zungenposition und Lippenstellung während der Artikulation und können somit Rückschlüsse auf artikulatorische Unterschiede zwischen Sprechern aufzeigen. Für diese Aufgabe werden die relevanten Vokale in den LeaP Aufnahmen, die auf der 'vowel' Annotationsebene annotiert sind, ausgewählt und die von *Praat* angezeigten Formantenhöhen in der Einheit Hertz gemessen (siehe Abbildung 4).

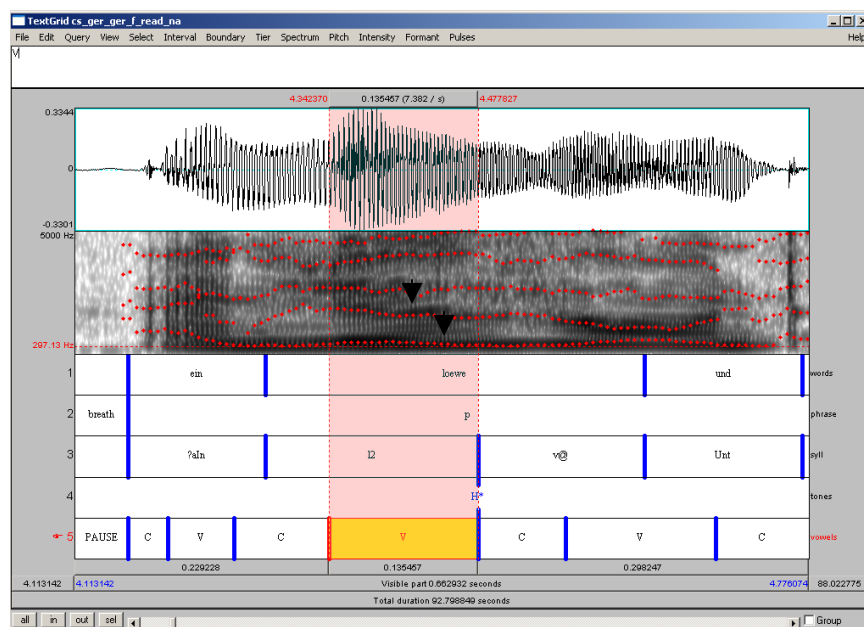


Abbildung 4. Messung des ersten und zweiten Formanten des Vokals im Wort Löwe.

5.2 Übungen zur Intonation

5.2.1 Phrasierung

Ebenfalls mit einer kombinierten bottom-up- und top-down-Übung können Deutschlernende die Unterschiede zwischen muttersprachlicher und lernersprachlicher Intonationsphrasierung im LeaP-Korpus entdecken. Dafür werden die vorgelesenen Geschichten ausgewählt. Die Annotationsebenen 'phrase' und 'words' enthalten Informationen darüber, welche Wörter vom Sprecher in Phra-

sen strukturiert werden. Die Lernenden können nun beispielsweise die durchschnittliche Länge einer Intonationsphrase eines Sprechers berechnen und werden entdecken, dass Muttersprachler im Vergleich zu Deutschlernenden längere Phrasen produzieren, Wörter in unterschiedliche Phrasen strukturieren und Phrasengrenzen an unterschiedlichen Positionen produzieren. Darüber hinaus lernen sie jedoch auch die Variabilität intonatorischer Phrasierung zwischen den einzelnen Muttersprachlern kennen, was als Grundlage für die Analyse des Zusammenhangs zwischen Phrasierung und Bedeutung dienen kann.

5.2.2 Tonumfang

Der Begriff Tonumfang bezeichnet das Ausmaß der Tonhöhenunterschiede zwischen den höchsten und tiefsten Tonhöhen-Punkten innerhalb einer Äußerung. Diese Punkte sind im Leap-Korpus auf der Annotationsebene 'hertz' transkribiert, wobei die exakte Tonhöhe in Hertz von *Praat* angezeigt wird, wenn der Cursor auf den entsprechenden Punkt im Spektrogramm bewegt wird (siehe Abbildung 5).

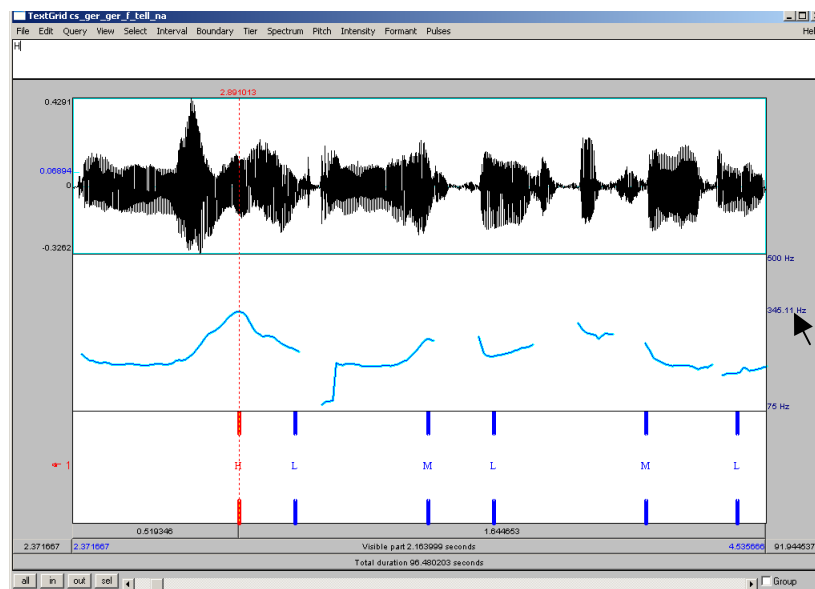


Abbildung 5. Tonhöhenumfang der Äußerung „Ein Löwe und eine Maus gingen spazieren.“ produziert von einer deutschen Muttersprachlerin. Die exakte Tonhöhe des H wird rechts angezeigt (siehe Pfeil). [Klangbeispiel 2](#)

Lernende können nun die Aufgabe erhalten, die Unterschiede zwischen muttersprachlichem und lernersprachlichem Gebrauch des Tonhöhenumfangs zu beschreiben. Durch einen Vergleich der identischen Sätze in den vorgelesenen Geschichten produziert von den deutschen Muttersprachlern und den Deutschlernenden zeigt es sich, dass der Tonhöhenumfang von Lernenden oft geringer

ist als der der Muttersprachler. Ein Beispiel dafür ist in den Abbildungen 5 und 6 gegeben.

In einer weiteren Übung können Deutschlernende die verschiedenen Diskursfunktionen des Tonhöhenumfangs analysieren. Dazu erhalten sie die Aufgabe, in der Geschichte oder den Nacherzählungen und Interviews des LeaP-Korpus Diskurselemente wie „neue Information“, „Elaboration neuer Information“ und „Hintergrundinformation“ zu identifizieren und den Tonhöhenumfang auf den entsprechenden Elementen zu messen. Auch diese korpusbasierte Aufgabe ermöglicht es Lernenden, eigenständig sprachliche Strukturen zu entdecken und die Beziehungsmuster zwischen Intonation und Pragmatik zu erkennen, da sie feststellen werden, dass neue Informationen üblicherweise mit einem größeren Tonhöhenumfang produziert werden als Hintergrundinformationen und Elaborationen. Als weitere Aufgabe bietet sich ein Vergleich des Tonhöhenumfangs in den verschiedenen Sprechstilen des LeaP-Korpus – gelesene Sprache, Nacherzählungen und frei gesprochene Sprache – an. Die Analyse der Muttersprachler zeigt, dass sich die Intonation vorgelesener Sprache deutlich von der anderer Sprechstile unterscheidet, was bei einigen Lernern jedoch noch nicht der Fall ist.

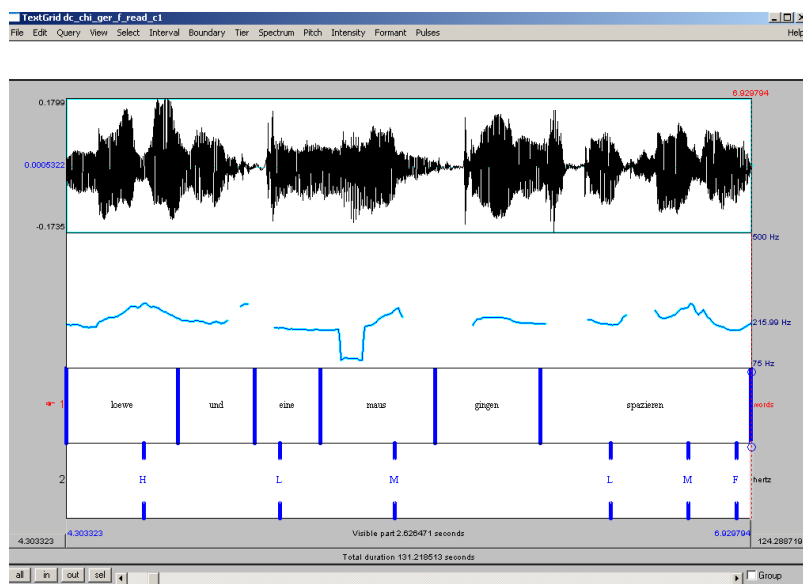


Abbildung 6. Tonhöhenumfang der Äußerung „Ein Löwe und eine Maus gingen spazieren.“ produziert von einer chinesischen Lernerin des Deutschen. [Klangbeispiel 3](#)

5.3 Vokalreduktion und Betonung

5.3.1 Betonte und unbetonte Vokale

Deutsche Wörter haben eine Silbe, die die Hauptbetonung trägt und möglicherweise weitere Silben, die eine Nebenbetonung tragen. Phonetisch unterscheiden sich betonte und unbetonte Silben durch die Länge und Qualität des Vokals sowie durch die Tonhöhe und Lautstärke des Vokals. Fortgeschrittene Lernende können die Aufgabe erhalten, in einem muttersprachlichen Korpus die Länge, mittlere Tonhöhe in Hertz und mittlere Lautstärke in dB einiger Vokale wie *a*, *o* und *e* in betonter (beispielsweise in *klaffen*, *Sonntag*, *Sender*) und unbetonter Position (wie in *Klavier*, *Bison*, *Rasen*) zu analysieren. Dabei werden sie herausfinden, dass Vokale in unbetonter Position kürzer sind als in betonter und dass sie in der Regel eine geringere mittlere Tonhöhe und Lautstärke aufweisen. Darüber hinaus kann ihnen die Aufgabe gestellt werden, zu erarbeiten, in welchen phonetischen Kontexten unbetonte Vokale ganz elidiert werden. Diese bewusste Auseinandersetzung mit Vokalreduktion und -elision hat das Ziel, durch die praktische multimediale (optische und akustische) Auseinandersetzung die Sprachbewusstheit auf diesem Gebiet zu fördern. Korpusbasierte Aufgaben haben jedoch auch den Vorteil, dass weitere, in der Aufgabe nicht fokussierte, sprachliche Strukturen erkannt werden können. So können Lernende eine Vielzahl weiterer phonetischer Besonderheiten betonter Vokale im Deutschen beobachten wie beispielsweise die Glottalisierung betonter Silben, die mit einem Vokal beginnen.

Eine weitere Aufgabe basiert auf dem Lernerkorpus *LeaP* und hat zum Ziel, Deutschlernende auf typische Lernerfehler auf dem Gebiet der Vokalreduktion aufmerksam zu machen. Ein Vergleich muttersprachlicher und fremdsprachlicher Daten zeigt den Lernenden, dass fehlende Vokalreduktion in unbetonten Silben einen der häufigsten Lernerfehler darstellt. Unbetonte Vokale in Lerner-sprache sind länger und haben eine unterschiedliche Qualität (Formantenstruktur) als die von Muttersprachlern. Weiterhin können Lernende Vokalreduktion von Lernern unterschiedlicher Muttersprachen vergleichen und die Gründe möglicher Unterschiede analysieren.

5.3.2 Betonung

Das Wahrnehmen und Produzieren von Betonung ist ein weiteres Gebiet der Phonetik, das sich mithilfe eines Korpus üben lässt. Besonders die Regeln der Satzbetonung, die pragmatische Funktionen hat, lassen sich anhand der Beschäftigung mit authentischem Sprachmaterial erschließen. Die muttersprachlichen Korpora mit vorgelesener Sprache können verwendet werden, um Lernende Regeln für die Position der Hauptsatzbetonung erarbeiten zu lassen. Ein Vergleich mit fremdsprachlichen Aufnahmen ermöglicht es, die größten Unterschiede zwischen den beiden Gruppen zu erarbeiten. Beispielsweise können in einer grundlegenden Übung die Anzahl betonter Wörter und die Position der Hauptbetonung innerhalb des gleichen Satzes produziert von einem Mutter-

sprachler und einem Deutschler analysiert werden. Es wird sich zeigen, dass viele Lernende mehr Wörter betonen als Muttersprachler und die Hauptbetonung auf andere Elemente legen. Eine daran anschließende Aufgabe beschäftigt sich dann mit der Beziehung zwischen Diskursfunktion eines Wortes und seiner Betonung. Dazu werden die einzelnen Wörter eines Satzes zunächst anhand ihrer Informationsstruktur (neu/alt) unterteilt und dann in Bezug zur Häufigkeit von Betonung gesetzt. Lernende werden herausfinden, dass in muttersprachlichem Deutsch neue Information üblicherweise betont wird und alte Information unbetont ist. In Lerner Sprache jedoch ist häufig sowohl neue als auch alte Information betont.

6. Zusammenfassung und Ausblick

Die pädagogische Nutzung von Sprachkorpora erfolgte bisher hauptsächlich in den Gebieten Vokabular, Pragmatik und Morphosyntax. Es war das Ziel dieses Beitrags zu zeigen, in welcher Form korpusbasierte Lehr- und Lernmethoden auch im Phonetikunterricht eingesetzt werden können. Einige phonetisch annotierte Korpora des Deutschen wurden vorgestellt und verschiedene Wege aufgezeigt, wie Korpora gewinnbringend im Phonetikunterricht eingesetzt werden können. Dies bezog sich sowohl auf den Einsatz der Korpora zur Unterrichtsvorbereitung und Themenauswahl als auch auf deren Verwendung als Werkzeug zum induktiven Lernen und der Entwicklung von Sprachbewusstheit für verschiedene phonetische Phänomene bei Lernenden.

Neben all den erwähnten Vorteilen des Einsatzes von Sprachkorpora im Phonetikunterricht sollen auch die Schwierigkeiten nicht unerwähnt bleiben: die Verwendung von Korpora zur Auswahl von Unterrichtsthemen und zur Unterrichtsvorbereitung von Lehrenden ist zeitaufwändig und erfordert eine gewisse Bereitschaft, sich in die technischen Erfordernisse einzuarbeiten. Um dies zu umgehen, kann jedoch auf die Fachliteratur der korpusbasierten Studien zum phonologischen Fremdsprachenerwerb (z.B. Gut 2006a) zurückgegriffen werden. Der Einsatz von Sprachkorpora als Methode im Unterricht erfordert ebenfalls eine Bereitschaft, sich den technischen Herausforderungen zu stellen oder geeignete technische Unterstützung anzufordern. Die Dokumentation der positiven Evaluationen korpusbasierten Phonetikunterrichts durch Lernende (z.B. Gut 2005, 2006b) sollte jedoch dafür genügend Anreiz sein.

Literatur

- Bernardini, Silvia (2004), Corpora in the classroom. In: John Sinclair (Hrsg.) (2004), *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins, 15-36.
- Botley, Simon, Glass, Julia, McEnery, Tony & Wilson, Andrew (1996), *Proceedings of Teaching and Language Corpora 1996*. Lancaster: UCREL technical papers volume 9.

- Connor, Ulla & Upton, Thomas (Hrsg.) (2004), *Applied Corpus Linguistics. A Multidimensional Perspective*. Amsterdam: Rodopi.
- Ellis, Nick (1995), *Implicit and Explicit Learning of Languages*. London: Academic Press.
- Flowerdew, Lynne (2001), The exploitation of small learner corpora in EPA materials. In: Mohsen Ghadessy, Alex Henry & Robert Roseberry (Hrsg.), *Small-corpus Studies and ELT*. Amsterdam: John Benjamins, 363-379.
- Fox, Gwyneth (1998), Using corpus data in the classroom. In: Brian Tomlinson (Hrsg.), *Materials development in language teaching*. Cambridge: Cambridge University Press, 25-43.
- Ghadessy, Mohsen; Henry, Alex & Roseberry, Robert (Hrsg.) (2001), *Small-corpus Studies and ELT*. Amsterdam: John Benjamins.
- Granger, Sylviane & Tribble, Christopher (1998), Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In: Sylviane Granger (Hrsg.), *Learner English on Computer*. London: Longman, 199-209.
- Granger, Sylviane; Hung, Joseph & Petch-Tyson, Stephanie (Hrsg.) (2002), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: Benjamins.
- Gut, Ulrike (2003), Non-native Speech Rhythm in German. *Proceedings of the ICPHS conference Barcelona*, 2437-2440.
- Gut, Ulrike (2005), Corpus-based Pronunciation Training. *Proceedings of the PTLC*, London.
- Gut, Ulrike (2006a), *Non-native Speech. A Corpus-based Analysis*. Habilitationsschrift, Universität Freiburg.
- Gut, Ulrike (2006b), Learner speech corpora in language teaching. In: Sabine Braun, Kurt Kohn & Joybrato Mukherjee (Hrsg.), *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*. Frankfurt: Lang, 69-86.
- Gut, Ulrike; Milde, Jan-Torsten; Voormann, Holger & Heid, Ulrich (2004), Querying Annotated Speech Corpora. *Proceedings of Speech Prosody 2004*, Nara, 569-572.
- Hansen, Jette (2001), Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics* 22, 338-365.
- Johns, Tim (1991), Should you be persuaded – two samples of data-driven learning materials. In: Tim Johns & Philip King (Hrsg.), *Classroom Concocting*. Birmingham: ELR Journal 4, 1-16.
- Johns, Tim (1994), From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. In: Terence Odlin (Hrsg.), *Perspecti-*

- ves on *Pedagogical Grammar*. Cambridge: Cambridge University Press, 293-313.
- Kaltenbacher, Erika (1998); Zum Sprachrhythmus des Deutschen und seinem Erwerb. In: Heide Wegener (Hrsg.), *Eine zweite Sprache lernen*. Tübingen: Narr, 21-38.
- Kettemann, Bernhard & Marko, Georg (Hrsg.) (2002), *Teaching and Learning by Doing Corpus Analysis*. Amsterdam: Rodopi.
- Leech, Geoffrey (1997), Teaching and language corpora: a convergence. In: Anne Wichmann, Steven Fligelstone, Tony McEnery & Gerry Knowles (Hrsg.), *Teaching and Language Corpora*. London: Longman, 1-23.
- Lleó, Conxita & Vogel, Irene (2004), Learning new segments and reducing domains in German L2 phonology: The role of the prosodic hierarchy. *International Journal of Bilingualism* 8, 79-104.
- McEnery, Tony & Wilson, Andrew (2001), *Corpus Linguistics*. Edinburgh: Edinburgh University Press (2. Aufl.).
- Meunier, Fanny (2002), The pedagogical value of native and learner corpora. In: Sylviane Granger, Joseph Hung & Stephanie Petch-Tyson (Hrsg.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: Benjamins, 119-141.
- Milde, Jan-Torsten & Gut, Ulrike (2002), The TASX environment: an XML-based toolset for time aligned speech corpora. *Proceedings of the Third International Conference on Language Resources and Evaluation, Las Palmas, 1922- 1927*.
- Milde, Jan-Torsten & Gut, Ulrike (2004); TASX - eine XML-basierte Umgebung für die Erstellung und Auswertung sprachlicher Korpora. In: Alexander Mehler & Henning Lobin (Hrsg.), *Automatische Textanalyse: Systeme und Methoden zur Annotation und Analyse natürlichsprachlicher Texte*. Wiesbaden: Verlag für Sozialwissenschaften, 249-264.
- Mindt, Dieter (1997), Corpora and the teaching of English in Germany. In: Anne Wichmann, Steven Fligelstone, Tony McEnery & Gerry Knowles (Hrsg.), *Teaching and Language Corpora*. London: Longman, 40-50.
- Mukherjee, Joybrato (2002), *Korpuslinguistik und Englischunterricht*. Frankfurt: Peter Lang.
- Mukherjee, Joybrato (2004), Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In: Ulla Connor & Thomas Upton (Hrsg.), *Applied Corpus Linguistics. A Multidimensional Perspective*. Amsterdam: Rodopi, 239-250.
- Nesselhauf, Nadja (2004), Learner corpora and their potential for language teaching. In: John Sinclair (Hrsg.), *How to Use Corpora in Language Teaching*. Amsterdam: Benjamins, 125-152.

- Osborne, John (2004), Top-down and Bottom-up Approaches to Corpora in Language Teaching. In: Ulla Connor & Thomas Upton (Hrsg.), *Applied Corpus Linguistics. A Multidimensional Perspective*. Amsterdam: Rodopi, 251-265.
- Partington, Alan (1998), *Patterns and Meanings. Using Corpora for English Language Research and Teaching*. Amsterdam: Benjamins.
- Ragan, Peter (2001), Classroom use of a systemic functional small learner corpus. In: Mohsen Ghadessy, Alex Henry & Robert Roseberry (Hrsg.), *Small-corpus Studies and ELT*. Amsterdam: Benjamins, 207-236.
- Robinson, Peter (1996), *Consciousness, Rules and Instructed Second Language Acquisition*. New York: Peter Lang.
- Sinclair, John (Hrsg.) (2004), *How to Use Corpora in Language Teaching*. Amsterdam: Benjamins.
- Todd, Richard (2001), Induction from self-selected concordances and self-correction, *System* 29: 1, 91-102.
- Tribble, Christopher (2001), Small corpora and teaching writing. In: Mohsen Ghadessy, Alex Henry & Robert Roseberry (Hrsg.), *Small-corpus Studies and ELT*. Amsterdam: Benjamins, 381-408.
- VanPatten, Bill & Cadierno, Teresa (1993), Explicit instruction and input processing. *Studies in Second Language Acquisition* 15: 2, 225-243.
- Wells, John; Barry, William; Grice, Martine; Fourcin, Adrian & Gibbon, Dafydd (1992), Standard computer-compatible transcription. *Technical Report No. SAM Stage Report Sen.3 SAM UCL-037*. University College London.
- Weskamp, Ralf (2001), *Fachdidaktik: Grundlagen und Konzepte, Anglistik-Amerikanistik*. Berlin: Cornelsen.

Anmerkungen

¹ Eine Liste, die IPA- und SAMPA-Symbole gegenüberstellt, ist erhältlich unter <http://www.ims.uni-stuttgart.de/lehre/teaching/2006-SS/Intonation-und-Informationsstruktur/symbole.pdf>.