

Christoph Beeh
Szeged

Schnittstelle Korpora und Fremdsprachenunterricht

Zur Arbeit mit Lernerkorpora des Deutschen als Fremdsprache

DOI: 10.14232/fest.bassola.23

ABSTRACT

Der vorliegende Aufsatz zeigt überblicksartig auf, welches vielfältige Potenzial Korpora für den Einsatz im Fremdsprachenunterricht bieten und wo ihr Einsatz auf Schwierigkeiten stößt. Der Fokus liegt auf der vergleichenden Betrachtung des Einsatzes von Daten aus Belegsammlungen versus größer angelegten öffentlich zugänglichen Korpora. Zentrale Hypothesen sind, dass die Arbeit mit eigenen Belegsammlungen das naheliegendste bedeutungsvolle Bindeglied zwischen Autonomieförderung und datenbasiertem Lernen im Fremdsprachenunterricht ist und dass ein Fokus auf Transfer- und Interferenzphänomene vielversprechendes Potenzial bietet.

1. Typen von Korpora

Der allgemeinen Definition von Lemnitzer / Zinsmeister (2015: 39) folgend sind Korpora „Textsammlungen mit kompletten Texten oder zumindest mit sehr großen Textausschnitten“. Sie werden typischerweise digital bearbeitet und verfügbar gemacht, möglichst kontrolliert und systematisch erhoben¹, und sind darüber hinaus „oft, aber nicht immer repräsentativ für den Gegenstand, auf den sie sich beziehen, durch Metadaten erschlossen, linguistisch annotiert.“ Je nach ihren Eigenschaften können unterschiedliche Korpora abweichend voneinander klassifiziert werden (vgl. für eine mögliche Typologie ebd.: 137). Für den vorliegenden Aufsatz ist von folgenden Korputypen die Rede:

¹ Zur Schwierigkeit von Systematik bzw. Kontrolle der vielfältigen Faktoren, die mit der Erstellung von Korpora einhergehen vgl. Granger (2008: 263f.).

- (1) **Referenzkorpora** sind Korpora, die „die Eigenschaften des dadurch repräsentierten Gegenstandes möglichst gut abdecken“ (ebd.: 141). Für den Untersuchungsgegenstand ‚deutsche Sprache‘ bspw. zählt zu diesen das Deutsche Referenzkorpus, das „mit 42 Milliarden Wörtern (Stand 13.11.2018) die weltweit größte linguistisch motivierte Sammlung elektronischer Korpora mit geschriebenen deutschsprachigen Texten aus der Gegenwart und der neueren Vergangenheit“ (DeReKo o. J.) ist.
- (2) **Lernerkorpora** sind als Spezialkorpora zu betrachten (vgl. Storrer 2011: 227) und umfassen all jene Korpora, die danach klassifiziert werden, dass die darin enthaltenen Texte nichtmuttersprachlich sind. Zur weiteren Spezifizierung können „externe Kriterien (mehr als eine L1, gesprochene oder geschriebene Texte, Erhebungssituation, Lernstand) geauso wie interne Kriterien (Texte, die bestimme sprachliche Strukturen aufweisen)“ (Lüdeling et al. 2008: 68) herangezogen werden. Granger (2008: 261) fügt hinzu, dass die gewählten Aufgabenstellungen so offen wie möglich gehalten sein sollten, um die Lernerinnen und Lerner dazu zu veranlassen, ihren eigenen Wortlaut zu wählen („to choose their own wording“), statt auf vorformulierte Chunks zurückzugreifen.
- (3) Als **sog. lokale Lernerkorpora** werden Datensammlungen verstanden, die Lehrkräfte vor Ort von Texten ihrer eigenen Gruppen ad hoc erstellen und die nicht öffentlich zugänglich sind. Da sie von den Kernkriterien eines linguistischen Korpus abweichen, sind sie im engeren Sinne nicht als Korpus zu betrachten: Sie werden nicht unbedingt systematisch und kontrolliert erhoben, sind darum nicht oder nur beschränkt repräsentativ für ihren Bezugsgegenstand, sind von kürzerer Lebensdauer, weisen in der Regel keine Metadaten auf und werden je nach Zielsetzung unterschiedlich (ggf. uneinheitlich) und eher nach didaktischen Kriterien annotiert. Hat sich im angelsächsischen Fachdiskurs der Begriff ‚local learner corpora‘ etabliert (vgl. Seidlhofer 2002; Mukherjee / Rohrbach 2006), ist im deutschsprachigen i. d. R. von Belegsammlungen die Rede. Im Folgenden wird die Bezeichnung aus dem Englischen übertragen, um zu zeigen, welch vielfältig einsetzbare Textsammlungen sie sein können. Die durch den Wortlaut erreichte Einreichung in die Kategorie eines Korpus, wenn auch unter Anführungszeichen, weist gleichermaßen aus, wie anspruchsvoll und voraussetzungsreich diese Arbeit ist, sodass damit

der Appell einhergeht, den sachkundigen Umgang mit Korpora unter Lehrkräften wie auch unter Lernerinnen und Lernern systematisch zu fördern.

2. Korpora und Fremdsprachenunterricht

Je nach didaktischer Zielsetzung können Korpora vor bzw. nach dem Fremdsprachenunterricht (FSU) oder im FSU selbst zum Einsatz kommen. Jedoch sind nicht alle Korpora barrierefrei recherchierbar.² Als öffentlich zugängliche Lernerkorpora können im Sinne der obigen Definitionen all jene Korpora gelten, die sowohl nichtmuttersprachliche Texte sammeln und annotieren, als auch, meist kostenfrei, Dritten zugänglich sind. Kommen die Lernerinnen und Lerner auch nicht direkt in Berührung mit den Korpora, so können sie doch der Aus- und Fortbildung des Lehrerpersonals dienen. Mindestens als „Referenzquelle“ werden muttersprachliche Korpora oder vergleichbare Lernerkorpora zu Orientierungsgrößen, die die „Allmacht“ intuitiver Urteile ablösen (Lüdeling / Walter 2009: 6) und das Bewertungsverhalten empirisch absichern. Zwar lag der Schwerpunkt bislang auf lexikogrammatischen Phänomenen, jedoch zeichnet sich eine Ausweitung der Anwendungsbereiche etwa auf Pragmatik³ und Phonetik⁴ ab. Im Folgenden muss auf die Diskussion des Korpus-Einsatzes vor und nach dem FSU aus Platzgründen verzichtet werden. Schwerpunktmaßig sollen darum die Einsatzmöglichkeiten im FSU angerissen (2.1) sowie auf ihre Grenzen hin untersucht werden (2.2). Auf dieser Grundlage folgt die Vorstellung einer Musteraufgabe zur Arbeit mit sog. lokalen Lernerkorpora (2.3).

² Das an der Humboldt-Universität zu Berlin befindliche fehlerannotierte Lernerkorpus (Falcko) ist im Hinblick darauf angelegt worden, durch frei zugängliche Korpusdaten auch in Fremdspracherwerbsforschung und Sprachvermittlung eine Lücke zu schließen (vgl. Siemen / Lüdeling / Müller 2006: 1). Eine Übersicht deutschsprachiger Korpora findet sich in Lemnitzer / Zinsmeister (2015). Eine groß angelegte Übersicht weltweiter Korpora mit Vermerk zur Zugänglichkeit findet sich auf der Homepage der Katholischen Universität Löwen / Belgien (vgl. Centre for English Corpus Linguistics 2019).

³ Pate für diesen Trend steht die 2017 zum ersten Mal erschienene Zeitschrift *Corpus Pragmatics*, vgl. <https://link.springer.com/journal/41701> (eingesehen am 14.03.2019).

⁴ Vgl. Trouvain / Zimmerer (2016) für phonetische Lernerkorpora bzw. Kána (2014: 117-124) für die phonetische Auswertung anderer Korpora, die sich für einen DaF-Einsatz didaktisieren lässt.

2.1 Einsatzmöglichkeiten im FSU

Die Arbeit mit Korpora hat als Methode unter dem Begriff *data driven learning* (DDL) Eingang in den FSU gefunden. In der Literatur ist dieser Ansatz vielfältig reflektiert und kritisiert worden, jedoch gibt es keine einheitliche Einordnung.⁵ Die Entwicklung des DDL hat, an der Nomenklatur unschwer zu erkennen, im angelsächsischen Raum gegenüber anderen Sprachen einen Vorsprung. Im Deutschen hat das Konzept als „datengesteuertes Lernen“ (Lüdeling / Walter 2009: 6) Eingang in die Fachliteratur gefunden.

Als Pionier des DDL gilt der Brite Tim Johns, dessen Definition (vgl. Mukherjee 2009: 166) von der Arbeit mit computer-generierten Konkordanzen ausgeht, mit denen die Lernerinnen und Lerner Regularitäten der Gestaltung („patterning“) der Zielsprache entdecken sollten. Die Aufgabe der Lehrkraft besteht hierbei darin, einen entsprechenden Kontext zu schaffen, in dem die Lernerinnen und Lerner sich entfalten können, um auch metakognitive Strategien auszubilden (vgl. Tognini-Bonelli 2001: 44). Neben allgemein induktiven und deduktiven Ansätzen (vgl. Gilquin / Granger 2010: 1) sind mehrere methodische Ansätze zur Nutzung von Korpora entwickelt worden, von denen aber die induktiven überwiegen. Allem voran stellt die Arbeit mit Konkordanzprogrammen die Hauptkomponente des DDL dar (vgl. Gilquin / Granger 2010: 5). Die Lernerinnen und Lerner sollen dabei idealerweise von einzelnen sprachlichen Instanzierungen auf generelle Muster schließen, die das jeweilige Lernziel darstellen. Gabrielatos (2005: 10–17) unterscheidet hierbei zwischen zwei Versionen des DDL und liefert zahlreiche Beispiele mit: Eine harte Version findet am Computerbildschirm statt und eine weiche Version wäre, wenn die Lehrkraft die Arbeit am Computer im Vorfeld erledigt und zur Bearbeitung im FSU lediglich einen Ausdruck mitbringt. Diesen zwei Spielarten entsprechend bietet es sich zur Vermittlung des sachkundigen Umgangs mit Korpora einerseits an, Korpusarbeit als vorrangiges Lernziel zu betrachten, andererseits

⁵ Gabrielatos‘ (2005) Differenzierung zwischen text- und korpusbasiertem DDL ist so zu verstehen, dass im FSU Computer mit entsprechender Software zum Einsatz kommen können, aber nicht müssen. O’Keefe et al. (2007: 24-30) ordnen die Arbeit mit Korpora klar dem *computer assisted language learning* unter, während Huang (2011: 92f.) DDL wiederum als Teil eines *corpus-aided discovery learning* beschreibt.

Korpursarbeit als Methode zum Erreichen anderer Ziele einzusetzen. Für den ersten Fall eignet sich eine von Lüdeling / Walter (2009: 12) vorgeschlagene fünfschrittige Progression zur Arbeit mit Häufigkeitslisten:

- (1) Nachvollzug von einfachen Korpusanalysen [...]
- (2) Interpretation von kleineren Listen, die beispielsweise ausgewählte Strukturen vergleichen [...]
- (3) Fokussierung auf sprachliche Phänomene und relevante Kontexteigenschaften in größeren Listen [...]
- (4) Erstellung eigener Listen auf der Basis eines sprachlichen Unterrichtsgegenstandes [...]
- (5) Extraktion eigener Häufigkeitslisten in selbst zusammengestellten Korpora. (Lüdeling / Walter 2009: 12)

Diese fünf Einzelschritte können und sollten zu konkreten Didaktisierungszwecken wieder in Unterschritte gegliedert werden, da das Erreichen des Lernziels, der autonome Umgang mit Korpora, sicher mehrere Unterrichtseinheiten umfasst.

Mit einem „DDL-basierten Miniprojekt“ liefert Mukherjee (2009: 167f.) eine beispielhafte Aufgabe kleineren Umfangs⁶, die aber auch auf das Aktivwerden der Lernerinnen und Lerner ausgerichtet ist:

WORKSHEET

handsome good attractive pretty tall successful

1. In pairs, choose three of these adjectives and have a look at their concordances.
 - Which words do they most frequently occur with?
 - Do they occur as part of a series of adjectives?
2. Discuss your findings with other pairs, then with the rest of the class. (Mukherjee 2009: 167f.)

⁶ Bei gründlicher Vorbereitung und Begleitung der Lehrkraft kann diese Aufgabe auch bewerkstelligt werden, ohne dass der autonome Umgang mit Korpus-Werkzeugen oder das sachkundige Erstellen von Frequenzlisten im Mittelpunkt stehen.

Neben der Erweiterung des Wortschatzes versprechen sich Gilquin / Granger (2010) vom DDL den Erwerb oder zumindest die Schärfung von generellen kognitiven Fertigkeiten, deren Relevanz über den FSU hinausgeht. Im Zentrum steht dabei eine Bewusstmachung zu erlernender Strukturen bzw. auch der eigenen Fehlerstrukturen (vgl. Gabrielatos 2005: 18; Brdar-Szabó 2010: 526ff.). Indem Lernerinnen und Lerner ihre eigenen Texte mit L1-Texten oder fehlerannotierten L2-Lernertexten abgleichen und interimssprachliche Merkmale („misuse“, „overuse“, „underuse“) analysieren, entwickeln sie Strategien zur Selbstkorrektur. Diese „korrektive Funktion“ (Gilquin / Granger 2010: 1) des DDL führt zur Verbesserung der eigenen produktiven wie kreativen Schreibfertigkeiten (vgl. O’Keeffe et al. 2007: 192) und möglicherweise auch, im Falle bereits behandelter Strukturen, zum Aufbrechen von fossilisierten Fehlern (vgl. Nesselhauf 2004: 140).⁷ Dass Fehleranalyse und -behebung eigens zu übende Fertigkeiten sind, geht auch deutlich aus Rauzs‘ (2019, in diesem Band) Untersuchung zu Übersetzungsbewertung und Fehlerkorrektur hervor.

In diesem Zusammenhang wird aber noch eine weitere Stärke des datengesteuerten FSU augenfällig. Schon Nesselhauf (2005: 269) ruft in Erinnerung, dass DDL gerade am Grenzverlauf zwischen Syntax und Lexik, der „kollokativen Grenze“ („collocational border“: Gilquin / Granger 2010: 6; vgl. Lüdeling / Walter 2009: 10f.), am effektivsten sei. Kollokationen und ihre Rolle im Bereich DaF sind gut untersucht (vgl. für das Deutsche als fremde Wissenschaftssprache Kispál 2014), stellen aber eine mögliche Fehlerquelle im FSU dar, was darin begründet liegen mag, dass sie „teilweise flexible Kombinationen“ sind, die „„unwahrscheinlich“ häufig zusammen (auch nah beieinander) auftreten“ (Lüdeling / Walter 2009: 11). Nesselhauf (2005: 252) betont dabei, dass am Anfang jeglicher Phraseodidaktik – und damit teilen DDL und Phraseodidaktik den gleichen Ausgangspunkt – die Bewusstmachung sprachlicher Strukturen stehe. Da L2-Lerndauer und Aufenthalte im Land der Zielsprache nur wenig positiven Einfluss auf das Lernen von Kollokationen haben, komme es im FSU

⁷ Walsh (2010: 340f.) weist auf die im FSU oft vernachlässigte Entwicklung der Fertigkeiten Korrekturlesen und Fehlerkorrektur hin und zitiert eine Studie von Tankó (2004), die durch den Einsatz korpusbasierter Methoden eine Verbesserung ungarischer Englischlernerinnen und -lerner im Gebrauch von adverbialen Konnektoren aufzeigt.

vor allem auf Prinzipien wie Wiederholung und Systematizität an (ebd.: 264f.).⁸ Konkrete Aufgaben sollten dabei so aussehen, dass in Form oder Bedeutung ähnliche Kollokationen innerhalb der L2 – aber auch sprachvergleichend zur jeweiligen L1 und relevanten L3 – kontrastiert werden, und um simple Zuordnungsaufgaben zu vermeiden, schlägt Nesselhauf (ebd.: 266ff.) vor, komplexe Raster mit multiplen Lösungsmöglichkeiten anzulegen, in denen weniger nach Nomen als nach Verben gefragt werde:

| | a laugh | a smoke | an experience | a trip |
|------|---------|---------|---------------|--------|
| take | | | | |
| make | | | | |
| have | | | | |
| do | | | | |

Tab.1: Einübung von Verben in Kollokationen nach Nesselhauf (2005: 267)

Aufgabentypen, die auf dem Korrigieren fehlerhafter Strukturen beruhen, seien nur dann angebracht, wenn die in Frage gestellten Abweichungen relevant für die Gruppe sind, d.h. tatsächlich in der Lernersprache auftauchen (ebd.: 270f.; vgl. Walsh 2010: 333).⁹ Es liegt auf der Hand, dass (sog. lokale) Lerner-korpora hierfür besonders gewinnbringend eingesetzt werden können, beson-ders in heterogenen Gruppen.

Bei der Gestaltung der Aufgaben, die entweder durch die Lehrkraft („teacher-led“) oder die Lernerinnen und Lerner selbst („learner-led“) angeleitet sein können (Gilquin / Granger 2010: 5), bietet der Grad an Lernerautonomie eine zusätzliche Möglichkeit zur Unterscheidung. Hierbei wird deutlich, dass der Umgang mit Korpora nicht nur entdeckendes (vgl. auch Tognini-Bonelli 2001: 44), sondern dadurch auch autonomes Lernen fördern soll (Mukherjee

⁸ Auch neuere Erkenntnisse in der Fremdsprachendidaktik legen diese Praxis nahe. Im Grund-lagenwerk „Sprachenlernen und Kognition“ zu einer kognitiven Sprachendidaktik sieht Sabine De Knop (2017: 114) Kollokationen und ähnliche Einheiten als „rekurrente, bedeutungsvolle Struk-turen einer Sprache“ an, die als solche im Ganzen zu lernen sind.

⁹ Über Form und Bedeutung hinaus sollte ggf. auch auf weitere Dimensionen der Kollokatio-nen hingewiesen werden, bspw. auf pragmatische und prosodische Eigenschaften.

2009: 166; vgl. auch: Nesselhauf 2004: 140) – zumal, und auf diesen nicht unerheblichen Faktor verweist Tognini-Bonelli (2001: 45), „in Zeiten hoher Lernerzahlen“. Dass beim Lernziel Autonomie in der Praxis des FSU allerdings Defizite in der Zusammenbringung fachdidaktischer und allgemein-erzieherischer Diskurse bestehen, konstatiert Schmenk (2018):

Nur weil [...] eine Lernende ohne fremde Hilfe eine Fremdsprache lernt, Sprachlernstrategien anwendet oder fremdsprachliche Aufgaben bewältigt, ist das Etikett ‚Autonomie‘ für diese Form der Selbststeuerung irreführend. Mehr noch: Das Etikett blendet [...] die fremdbestimmten Elemente aus, die zur Ausübung von Selbststeuerung entwickelt werden mussten (Schmenk 2018: 19).

Erstrebenswert sei es darum, curriculare Vorgaben zu persönlichen Zielen zu machen. Die Arbeit mit sog. lokalen Lernerkorpora dürfte in diesem Sinne das naheliegendste Bindeglied zwischen Autonomieförderung und DDL sein, da das Korpus nicht zu einer anonymen Masse verschwimmt (vgl. Braun 2006: 26, zit. n. Gilquin / Granger 2010: 2), sondern die eigenen Texte vorliegen. Um nämlich vorgeschriebene Kursziele zu persönlichen Zielen zu machen, müssen sie für die Lernerinnen und Lerner bedeutungsvoll sein. Aus dem eigenen Entwurf (für bspw. einen Brief, eine Übersetzung etc.) einen guten Text zu produzieren wird sicher als umso bedeutungsvollere Aktivität bewertet, je mehr der Text seiner Funktion gerecht wird¹⁰. Die Arbeit mit jeglichem authentischen Material hat dabei zwar gegenüber der Arbeit mit fingierten Texten potenziell den Vorteil der Relevanz, wenn sie die Lernerinnen und Lerner nicht zur Überforderung führt. Doch auch die Authentizität der Texte garantiert weder Lernerfolge noch stellt sie die Motivation der Lernerinnen und Lerner sicher. Seidlhofer (2002: 220) resümiert: „there are no global solutions to motivational problems.“ Da es eine Vielzahl an Variablen für den Einsatz von Korpora im FSU gibt, muss den Lernerinnen und Lernern klar sein, was das Lernziel sein soll, auf welches das Korpus nach entsprechenden Belegen durchsucht wird. Dies erfordert, dass als global geltende Prinzipien des FSU in Bezug auf lokale Gegebenheiten hin über-

¹⁰ Dies ist dann der Fall, wenn etwa der Brief verschickt werden oder die Übersetzung jemandem weiterhelfen soll (vgl. Tomlinson 2018).

prüft werden sollten, da sie andernfalls bedeutungslos werden. Diesem Leitsatz gemäß hat Seidlhofer (ebd.) die Arbeit mit Korpora um einen Ansatz namens *learning-driven data* erweitert, der vom Postulat eines archetypischen Lerners absieht und davon ausgeht, dass alle Pädagogik „has to be local, designed for specific learners and settings“ (ebd.). Korpora dienen insbesondere dort, wo die Zielsprache nicht gesprochen wird, als Quelle verdichteten L2-Kontakts („condensed language exposure“: Gabrielatos 2005: 10), die sich einerseits auf Regelbildung hin untersuchen lässt, andererseits, dank der vertikal wie horizontal lesbaren Konkordanzachsen, in idealer Weise die Zusammenghörigkeit formaler und funktionaler Parameter sichtbar macht (vgl. Tognini-Bonelli 2001: 41). Wird die Interimssprache der Lernerinnen und Lerner als eigengesetzliches System ernstgenommen, so können insbesondere sog. lokale Lernerkorpora dazu eingesetzt werden, nicht nur theorieverhaftet über mögliche Fehlerquellen irgendwelcher Lernerinnen und Lerner zu reden, „but to work **with them**, on a corpus to which **they themselves** have contributed“ (Seidlhofer 2002: 1; Hervorhebung C.B.).

Eine letzte wichtige Facette in der Arbeit mit öffentlich zugänglichen und / oder sog. lokalen Lernerkorpora im FSU ist das direkte Geben und Annehmen von Feedback. Korpusbasierte Aufgaben und Übungen zu Fehleranalyse und -korrektur sind bereits eine Form von Feedback. Stärke und Schwäche zugleich kann bei sog. lokalen Lernerkorpora sein, dass die Texte die eigenen oder die von Kursnachbarinnen und -nachbarn sind. Während die einen motiviert sind zu helfen und selbst Hilfe anzunehmen, werden die anderen sich vor Kritik scheuen. Wird dieses Format nicht gründlich erklärt und geübt, bleiben nicht nur die Fehler undiskutiert, sondern auch die Lernatmosphäre kann sich verschlechtern. Abhilfe schafft in diesem Fall der Einsatz von muttersprachlichen oder groß angelegten öffentlich zugänglichen Lernerkorpora, bei denen die Schwelle der persönlichen Bekanntschaft nicht gegeben ist. Dabei ist formatives Feedback¹¹ (vgl. Strum et al. 2018: 80f.) mit detaillierten, metasprachlichen Erklärungen effektiver als globale Urteile (z. B. falsch / richtig oder Notenvergabe) und sollte „vorwärts gerichtet“ sein, also im Sinne der Förderung der Ausbildung von Lernstrategien stets weitere Lernziele anregen und die Schritte dorthin aufzuzeigen helfen (ebd.: 80).

¹¹ tw. auch intelligentes bzw. explizites Feedback, s. Heift / Chapelle 2012: 563.

2.2 Grenzen und Einschränkungen

Für den Einsatz von größeren L1- oder Lerner korpora vor, nach oder während des FSU stellt der sachkundige Umgang eine offensichtliche, aber ernst zunehmende Hürde dar. Als *corpus literacy*¹² bezeichnet Mukherjee (2009: 173f.) „die Fähigkeit, zur Klärung sprachlicher Fragen und zur Lösung linguistischer Probleme selbstständig und kompetent geeignete Korpora und relevante korpuslinguistische Programme nutzen zu können.“ Das heißt, dass ‚corpus literacy‘, wie andere Fertigkeiten auch, gelernt und geübt werden muss. Storrer (2011) liefert eine überblicksartige Einführung einer „sachkundigen Auswertung von Korpusdaten“ (ebd.: 238) für „Sprachinteressierte und ‚Spracharbeiter‘ in Verlagen und Bildungsinstitutionen“ (ebd.: 217). Diese Sachkundigkeit auch unter Lernerinnen und Lernern zu fördern ist daher sicher ein umso sinnvoller Unterfangen, je mehr selbständige Arbeits- und Lerntechniken mitvermittelt werden sollen, zumal die jüngeren Generationen zu den „digital natives“ bzw. der „net generation“ zählen und oft bereits große Affinität sowie intuitives Geschick zur Benutzung von Computersoftware mitbringen.¹³ Nesselhauf (2004: 140f.) weist zudem darauf hin, dass die Erstellung sog. lokaler Lerner korpora zeitliche Ressourcen beansprucht, die nicht allen zur Verfügung stehen. Trotz der unschlagbaren Nähe zur Lernmotivation ist auch zu bedenken, dass selbst anonymisierte Textsamples innerhalb der Gruppe erkannt werden und so zu unangenehmen, ja lernhemmenden Situation führen können.

Die Konfrontation mit fehlerhaftem Sprachgebrauch, räumt Nesselhauf (2004: 139) ferner ein, „is only useful to the extent that focused negative evidence supports language learning.“ Was die Arbeit an fossilisierten Fehlern angeht, so dürfte der FSU hierbei aufgrund mehrerer Faktoren auf Schwierigkeiten treffen. Fossilisierungen wurden zunächst durch Selinker (1972: 215f.)

¹² Terminologisch soll im Folgenden Mukherjees englischsprachiger Begriff ‚corpus literacy‘ im Deutschen mit Storrers Begriff der korpusbezogenen Sachkundigkeit wiedergegeben werden.

¹³ Insofern es expliziter Bestandteil des Curriculums wird, dieser Komponente somit Zeit und technische Ressourcen im Lehrplan eingeräumt werden müssen, wird der sachkundige Umgang mit Korpora („corpus literacy“) gewiss nur an ausgewählten Orten zum Lernziel. Ob Schulen und freie Bildungsträger darunter sind, ist fragwürdig. Zu erwarten wäre dies eher an sprachwissenschaftlichen Einrichtungen an Universitäten und in Übersetzer-Ausbildungen.

im Rahmen der fachdiskurslichen Etablierung des Begriffs Interimssprache als Merkmal derselben definiert. Dabei handelt es sich um

Sprachformen, die der Norm der Zielsprache nicht entsprechen, jedoch stabile Elemente der Interlanguage bleiben. Sie sind eine Art des ‚Stillstands‘, durch den bestimmte Eigenschaften der Interlanguage (hier: Fehler) trotz intensiven Unterrichts oder jahrelangen Aufenthalts in der Zielsprachigen Umgebung erhalten bleiben [...]. (Bawej 2009: 25f.)

Eine entsprechende Analyse durch die Lehrkraft, welche Abweichungen tatsächlich Fossilisierungen darstellen, bedarf eines engen Betreuungsverhältnisses mit einer Stundenanzahl, die gewährleistet, dass die fraglichen Strukturen im Sinne Nesselhaufs systematisch und repetitiv eingeübt werden. Doch selbst dann sind Bewusstmachung und Einübung kein Garant dafür, dass festgefahrenen Strukturen wirklich gelockert würden. Han (2012: 485) weist vielmehr darauf hin, dass „[i]nstruction can sometimes promote or induce fossilization.“ Vor diesem Hintergrund überrascht es nicht, dass das Metzler Lexikon für Fremdsprachendidaktik (Surkamp 2017: 87) Fossilisierungen sogar in toto als für die Fremdsprachendidaktik „irrelevant“ ausweist, da das Auftreten derselben auch an kognitiven Barrieren liegen könne, die die Motivation der Lerner betreffen und somit nicht in den kontrollierbaren Wirkungskreis des FSU fallen. Es scheint vielversprechender, das Augenmerk statt auf Fossilisierungen auf **Transferphänomene** (vgl. Gürbeyaz 2012: 30ff.) und **Interferenzfehler** (vgl. Weinreich 1953: 1; Oebel 2004: 214; Bawej 2009: 25) – und damit auf das mehr oder weniger „latent vorhandene kontrastive Sprachbewusstsein“ (Oebel 2004: 218) der Lernerinnen und Lerner und seine gewinnbringende Aktivierung im FSU – zu richten. Freilich kann es sich bei Interferenzfehlern auch um fossilisierte Fehler handeln. An dieser Stelle scheint jedoch ein Perspektivenwechsel eingetreten zu sein, der den Blick von individuellen (fossilisierten) Fehlern auf möglichst empirisch abgesicherte typische Abweichungen von Lernerinnen und Lernern bestimmter Muttersprachen umschwenkt. Interferenzen in der Sprache betreffen potenziell alle Lernerinnen und Lerner, sind nicht auf individuelle Faktoren, sondern auf die L1-Struktur zurückzuführen und haben daher Relevanz für den FSU, insbesondere für die Entwicklung von Curricula und Lehrmaterialien. Beispielhaft ist eine Studie von Gürbeyaz (2012: 202), in

der Fehlerkorpora spanisch- und türkischsprachiger DaF-Lernender angelegt wurden, die „eine erhebliche Zahl von Interferenzfehlern“ enthalten. Dabei

fällt bei der Auswertung der Fehleranteile in einzelnen Kategorien auf, dass türkischsprachige Lerner in einigen Kategorien trotz ihres im Durchschnitt höheren Kenntnisstands mehr Interferenzfehler produzierten, und zwar in den Bereichen, in denen die ES [= Erstsprache, C.B.] und FS [= Fremdsprache, C.B.] sich stark voneinander unterscheiden. (ebd.)

Neben Interferenzfehlern aus der L1, die „besonders im Anfangsstadium“ (ebd.: 203) auftreten, seien auch Fehler aus L3-Sprachen sowie intralinguale Fehler zu verzeichnen, woran insgesamt sichtbar wird, „dass Transfer ein Selektionsprozess sei, dass also Lerner transferfähige Strukturen in ihrer ES von nicht transferfähigen unterscheiden und die ersteren auf die FS übertragen“ (ebd.: 202).

Letztlich sind Schwierigkeiten insbesondere beim Umgang mit Variation und Varietäten zu erwarten. Hier muss zum einen eine weitreichende Kompetenz der Lehrkräfte vorausgesetzt werden, um Normabweichungen als Varianten, Varietäten, Code-Switching o. Ä. zu erkennen, zum andern verwischen diese die für zu benotende Leistungsanforderungen an Schulen oder Hochschulen benötigte Klarheit von Falsch oder Richtig. Auch sollte bei der Entwicklung der rezeptiven und produktiven Fertigkeiten je zwischen gesprochener und geschriebener Sprache differenziert werden (vgl. Fandrych / Thurmail 2018: 16). Wildfeuer-Eller / Wildfeuer (2018: 15) zeigen auf, dass einschlägige DaF-Lehrwerke die muttersprachliche Vielfalt des deutschsprachigen Gebiets bereits berücksichtigen und plädieren für einen Modus Vivendi, „der einen sinnvollen Einsatz von Variation im Fremdsprachenunterricht zum Ziel hat, damit dieser nicht fundamental an der Sprachwirklichkeit vorbeigeht.“ O’Keeffe et al. (2007: 190) gehen noch einen Schritt weiter und weisen darauf hin, dass Kreativität und Sprachspiele nicht „the preserve of the native speaker“ seien. Auch das Rezipieren kreativer Korpusbelege könne die L2-Produktion fördern (ebd.: 192), bedarf aber wiederum gut geschulter Lehrkräfte, die diese Belege auch spontan verstehen, einordnen und erklären können.

2.3 Muster zur Arbeit mit sog. lokalen Lernerkorpora

Lüdeling / Walter (2009: 9–13) und Sripicharn (2010: 381) legen nahe, Korpusdaten vor allem quantitativ auszuwerten, um anhand von Frequenzlisten (sowohl aus L1- als auch aus L2-Korpora) Phänomene von under- oder overuse miteinander zu vergleichen. Bei sog. lokalen Lernerkorpora sind eine entsprechende Menge an Daten und Erhebungsbedingungen für quantitative Auswertungszwecke zwar nicht gegeben, trotzdem werden sich den begleitenden Lehrpersonen, die das Sprachverhalten ihrer Lernerinnen und Lerner oft bereits gut kennen, Tendenzen zeigen. Diese Tendenzen können für mehr oder weniger komplexe Aufgaben- und Übungstypen im eigenen FSU nutzbar gemacht werden, um Bewusstmachung, autonomes und entdeckendes Lernen, Fehlerkorrektur und den sachkundigen Umgang mit Korpora („corpus literacy“) in unterschiedlichen Sozialformen zu fördern. In Anlehnung an Walsh (2010: 342) haben Übungen zu fehlerhaften Texten folgenden Musteraufbau:

- (1) Identifizieren und markieren Sie Fehler.
- (2) Begründen und kategorisieren Sie die Fehler.
- (3) Korrigieren Sie die Fehler.
- (4) Überprüfen Sie Ihre Korrektur bzw. Textstellen, bei denen Sie unsicher sind, anhand eines Korpus.

In unterschiedlichen Sozialformen kann dann ein Katalog an (lexikogrammatisch, pragmatisch, stilistisch) auffälligen oder falschen Chunks bzw. Sätzen, die die Gruppe selbst produziert hat, unter der Aufsicht der Lehrkraft bearbeitet werden. Einzelne Schritte können ausfallen oder übersprungen werden, indem bspw. die Fehler schon vorab markiert werden. Das Kategorisieren der Fehler kann fein- oder grobkörniger ausfallen, je nachdem, wie viel metasprachliche Kenntnisse vorhanden sind, oder ganz entfallen. Zur Überprüfung der Korrekturen wären neben den üblichen Hilfsmitteln (Grammatiken und Wörterbüchern) in erster Linie L1-Korpora zu Rate zu ziehen, etwa nach Anleitung in Storrer (2011). Jedoch ist auch an dieser Stelle das kontrastive Arbeiten mit L2-Korpora möglich: Sripicharn (2010: 373) verweist darauf, dass Lernerinnen und Lerner „can exploit a learner corpus [...] to compare and contrast their own writings with texts retrieved from native speaker corpora, which can help

them correct language mistakes in their writing tasks.“ Hierzu sind allerdings ein relativ hoher Grad an Fehlerbewusstheit und sichere Kenntnis der metasprachlichen Kategorien zu empfehlen. Des Weiteren ist zu erwarten, dass die Funde in Schritt 4, je nach Kontext – etwa bei Regionalismen, Archaismen, Mehrfachkonnotation etc. –, weitere Fragen aufwerfen.

Nesselhauf (2004) liefert eine weitere Beispielprozedur dafür, wie entdeckendes Lernen wirkungsvoll bei under- und overuse-Phänomenen eingesetzt werden kann, indem die Relevanz für die eigene Textproduktion sichtbar gemacht wird: Hat die Lehrkraft für ein behandeltes Unterrichts-Phänomen geeignete Fehler identifiziert, präsentiert sie zunächst konkordanzartig Belege aus einem Lerner- und einem Referenzkorpus. Die Lernerinnen und Lerner sollen daraufhin Unterschiede zwischen diesen beiden herausarbeiten. Sind die Unterschiede wenig offensichtlich, sollten zielführende Fragen vorbereitet werden. Als mögliche Variante schlägt Nesselhauf vor, zuerst Treffer aus dem L1-Korpus zu präsentieren, danach aus dem L2- oder dem sog. lokalen Korpus, um einen Überraschungseffekt anzuregen:

This variant seems particularly useful if the native speaker use is well-known to the learners but the learners' frequent deviance is not (for example in areas of underuse). If the learner data, **ideally from the group itself**, is presented shortly after the native speaker data, the surprise about the fact that a pattern that seems entirely obvious is often not used or misused in the language of learners could make the insight longer-lasting. (Nesselhauf 2004: 144; Hervorhebung C.B.)

Abschließend, um zu verhindern, dass Lernerinnen und Lerner den fehlerhaften Sprachgebrauch erinnern, sollten auf die Konfrontation stets Übungen wie Lückentexte o. Ä. zur Konsolidierung der Zielstruktur folgen.

3. Fazit

Es lässt sich festhalten, dass aufgrund der Nähe zu den Lernerinnen und Lernern die Arbeit mit sog. lokalen Lerner korpora das naheliegendste bedeutungsvolle Bindeglied zwischen Autonomieförderung und datenbasiertem Lernen

im Fremdsprachenunterricht gelten kann, insbesondere wenn formatives Feedback und Fehlerkorrektur als Aufgabentyp wichtige Faktoren sind. Zur Förderung des sachkundigen Umgangs mit Korpora (‘corpus literacy’) für Lehrende und Lernende können sie als intermediärer Schritt – nicht isoliert, sondern stets in sinnvoll angelegter Ergänzung – zum Umgang mit Referenz- und Lerner korpora angesehen werden. Im Einzelnen bedürfen diese Annahmen jedoch empirischer Absicherung. Dazu kommt, dass das Erforschen von Transfer- und Interferenzphänomenen vielversprechendes Potenzial bietet, das über den zu engen Fokus des Blicks auf Fossilisierungen hinausreicht. Für die FSU-Praxis muss beachtet werden, dass sich die Anwendung des DDL darüber hinaus auf o.g. Bereiche beschränken sollte, denn für allgemeinere grammatische Bereiche wie Tempus oder Aspekt eigne sich dieser Ansatz nicht (vgl. Nesselhauf 2004: 142). Haben Lernerinnen und Lerner kein fundiertes Grundwissen in linguistischer Analyse, können schnell falsche Konsequenzen gezogen werden. Und auch das Bild zu lernender Regelfälle kann durch authentische Belege eher getrübt als klarer werden (vgl. Tognini-Bonelli 2001: 44), insbesondere wenn es sich dabei um Sprachvariations- und Sprachwandelphänomene handelt, aber auch dann, wenn es sich um allgemein eher unerwünschte Elemente (‘noise’) handelt wie Schimpfwörter oder lyrischen, antiquierten oder anderweitig spezifischen Sprachgebrauch. Zwar kann auf diese Weise ein im FSU oft unterentwickeltes „sprachliches Bewusstsein für das Verhältnis von Norm und Varianz“ und die Unterschiede zwischen Varietäten gefördert werden“ (Lüdeling / Walter 2009: 8). Doch setzt dies eine entsprechend vorbereitete und reflektierte Auswahl durch die Lehrenden voraus. Letztlich stellt genau diese Fähigkeit zum ‚sachkundigen Lesen‘ eines Korpus (‘corpus literacy’) selbst für Lehrende eine recht hohe Hürde dar, etwa wenn es um die Unterscheidung zwischen den Urteilen „nicht im Korpus repräsentiert“ und „gibt es nicht“ geht (vgl. Nesselhauf 2004: 144; Gilquin / Granger 2010: 8). Hier ist zum einen die Gefahr zur Übergeneralisierung entdeckter Einzelfälle gegeben (vgl. Sripicharn 2010: 382), zum anderen dürfte die Formulierung einer sprachlich akzeptablen Alternative schwerfallen, wenn nicht gerade ein fehlerannotiertes Korpus vorliegt, das bereits muttersprachlich überprüfte Zielhypothesen an die Hand gibt.

4. Literatur

- Bawej, Izabela (2009): Sprachliche Fehler und ihre Ursachen im Lichte dreier Spracherwerbsmodelle. In: *Linguistica Bidgostiana VI*, 19–28.
- Braun, Sabine (2006): „ELISA: a pedagogically enriched corpus for language learning purposes“. In: dies. / Kohn, Kurt / Mukherjee, Joybrato (Hrsg.) (2006): *Corpus Technology and Language Pedagogy*. Frankfurt am Main: Peter Lang, S. 25–47.
- Brdar-Szabó, Rita (2010): Nutzen und Grenzen der kontrastiven Analyse für Deutsch als Fremd- und Zweitsprache. In: Krumm, Hans-Jürgen et al. (Hrsg.): *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch*. Berlin / New York: de Gruyter (HSK 35.2), 518–531.
- Centre for English Corpus Linguistics (2019): Learner Corpora around the World. Louvain-la-Neuve: Université catholique de Louvain. <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html> (eingesehen am 09.03.2019).
- De Knop, Sabine (2017): Chunking und Dechunking. In: Roche, Jörg / Suñer, Ferran (Hrsg.): *Sprachenlernen und Kognition. Grundlagen einer kognitiven Sprachdidaktik*. Tübingen: Narr Francke Attempto, 112–118.
- DeReKo (o. J.): Ausbau und Pflege der Korpora geschriebener Gegenwartssprache: Das Deutsche Referenzkorpus – DeReKo. <http://www1.ids-mannheim.de/kl/projekte/korpora> (eingesehen am 22.03.2019).
- Fandrych, Christian / Thurmail, Maria (2018): Grammatik im Fach Deutsch als Fremd- und Zweitsprache. Grundlagen und Vermittlung. Berlin: Erich Schmidt.
- Gabrielatos, Costas (2005): Corpora and language teaching: just a fling or wed-ding bells? *TESL-EJ* 8 / 4, 1–35. <http://www.tesl-ej.org/ej32/a1.html> (eingesehen am 09.03.2019).
- Gass, Susan / Mackey, Alison (Hrsg.) (2012): *The Routledge Handbook of Second Language Acquisition*. London u.a.: Routledge.
- Gilquin, Gaëtanelle / Granger, Sylviane (2010): How can data-driven learning be used in language teaching? In: O’Keeffe / McCarthy (Hrsg.), 359–370.
- Granger, Sylviane (2008): Learner corpora. In: Lüdeling / Kytö (Hrsg.), 259–275.
- Gülbeyaz, Esin Işıl (2012): Spracherwerb und Fehleranalyse. Eine korpusanalytische Studie. Frankfurt am Main u.a.: Peter Lang.

- Han, ZhaoHong (2012): Fossilization – A classic concern of SLA research. In: Gass / Mackey (Hrsg.): 476–490.
- Heift, Trude / Chapelle, Carol (2012): Language learning through technology. In: Gass / Mackey (Hrsg.): 555–569.
- Huang, Li-Shih (2011): Language learners as language researchers: the acquisition of English grammar through a corpus-aided discovery learning approach mediated by intra- and interpersonal dialogues. In: Newman, John et al. (Hrsg.): Corpus-based Studies in Language Use, Language Learning, and Language Documentation. Amsterdam / New York: Brill, 91–122.
- Káňa, Tamáš (2014): Sprachkorpora in Unterricht und Forschung DaF / DaZ. Brno: Masarykova univerzita.
- Kispál, Tamás (2014): Wissenschaftssprachliche Kollokationen in Seminararbeiten ausländischer Germanistikstudierender. In: Ferraresi, Gisella / Liebner, Sarah (Hrsg.): SprachBrückenBauen. 40. Jahrestagung des Fachverbandes Deutsch als Fremd- und Zweitsprache an der Universität Bamberg 2013. Göttingen: Universitätsverlag (Materialien Deutsch als Fremdsprache 92), 235–250.
- Lemnitzer, Lothar / Zinsmeister, Heike (2015): Korpuslinguistik: eine Einführung. 3., überarb. und erw. Aufl. Tübingen: Narr Francke Attempto.
- Lüdeling, Anke / Doolittle, Seanna / Hirschmann, Hagen / Schmidt, Karin / Walter, Maik (2008): Das Lernerkorpus Falko. In: Deutsch als Fremdsprache 2, 67–73.
- Lüdeling, Anke / Kytö, Merja (Hrsg.) (2008): Corpus Linguistics. An international handbook. Volume 1. Berlin / New York: de Gruyter (HSK 29).
- Lüdeling, Anke / Walter, Maik (2009): Korpuslinguistik für Deutsch als Fremdsprache. Sprachvermittlung und Spracherwerbsforschung: <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiterinnen/anke/pdf/LuedelingWalterDaF.pdf> (eingesehen am 22.03.2019).
- Muhkerjee, Joybrato / Rohrbach, Jan-Marc (2006): Rethinking Applied Corpus Linguistics from a Language-pedagogical Perspective: New Departures in Learner Corpus Research. In: Kettemann, Bernhard / Marko, Georg (Hrsg.): Planing, Gluing and Painting Corpora: Inside the Applied Corpus Linguist's Workshop. Frankfurt am Main: Peter Lang, 205–232.
- Mukherjee, Joybrato (2009): Anglistische Korpuslinguistik. Eine Einführung. Berlin: Erich Schmidt.

- Nesselhauf, Nadja (2004): Learner Corpora and their potential for language teaching. In: Sinclair (Hrsg.), 125–152.
- Nesselhauf, Nadja (2005): Collocations in a Learner Corpus. Amsterdam / Philadelphia: Benjamins.
- O'Keeffe, Anne / McCarthy, Michael / Carter, Ronald (2007): From Corpus to Classroom. Language Use and Language Teaching. Cambridge: Cambridge University Press.
- O'Keeffe, Anne / McCarthy, Michael (Hrsg.) (2010): The Routledge Handbook of Corpus Linguistics. London u.a.: Routledge
- Oebel, Guido (2004): L1- bzw. L2(Englisch)-Interferenzfehler japanischer DaF-Lerner Fehleranalyse von ‚Dornröschen‘-Nacherzählungen. In: Deutsch als Fremdsprache 4, 213–219.
- Rauzs, Orsolya (2019, in diesem Band): Übersetzungsbewertung und Fehlerkorrektur aus der Perspektive von Studierenden.
- Schmenk, Barbara (2018): Aporien des Lern- und Bildungsziels Autonomie im Kontext des institutionalisierten Fremdsprachenunterrichts. In: Fremdsprachen Lehren und Lernen 47/1, 10–24.
- Seidlhofer, Barbara (2002): Pedagogy and local learner corpora. Working with learning-driven data. In: Granger et al. (Hrsg.): Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching. Amsterdam / Philadelphia: Benjamins, 213–234.
- Selinker, Larry (1972): Interlanguage. In: International Review of Applied Linguistics 10/2, 209–231.
- Siemen, Peter / Lüdeling, Anke / Müller, Frank Henrik (2006): Falko – ein fehlerannotiertes Lernerkorpus des Deutschen. In: Proceedings of Konvens 2006, Konstanz.
- Sinclair, John (Hrsg.) (2004): How to use corpora in Language Teaching. Amsterdam / Philadelphia: Benjamins.
- Sripicharn, Passapong (2010): How can we prepare learners for using language corpora? In: O'Keeffe / McCarthy (Hrsg.), 371–384.
- Storrer, Angelika (2011): Korpusgestützte Sprachanalyse in Lexikographie und Phraseologie. In: Knapp, Karlfried et al. (Hrsg.): Angewandte Linguistik: ein Lehrbuch. 3., vollst. überarb. und erw. Aufl. Tübingen / Basel: A. Francke, 216–239.

- Strum, Afra / Lindauer, Nadja / Sommer, Tim (2018): Rückmelden. Aufgaben- und lernzielbezogenes Feedback. In: Der Deutschunterricht 3, 2–10.
- Surkamp, Karola (Hrsg.) (2017): Metzler Lexikon Fremdsprachendidaktik. Ansätze – Methoden – Grundbegriffe. 2., aktualisierte und erw. Aufl. Stuttgart: Metzler.
- Tankó, Gyula (2004): The use of adverbial connectors in Hungarian university students' argumentative essays. In: Sinclair (Hrsg.), 157–181.
- Tognini-Bonelli, Elena (2001): Corpus Linguistics at Work. Amsterdam / Philadelphia: Benjamins.
- Tomlinson, Brian (2018): What should meaning-focused mean? In: Bouckaert, Marina et al. (Hrsg.): Meaning-Focused Materials for Language Learning. Newcastle upon Tyne: Cambridge Scholars Publishing, 5–18.
- Trouvain, Jürgen / Zimmerer, Frank (2016): Phonetische Lernerkorpora und ihr Nutzen im DaF-Bereich – eine Fallstudie. In: Deutsch als Fremdsprache 53 / 4, 204–212.
- Walsh, Steve (2010): What features of spoken and written corpora can be exploited in creating language teaching materials and syllabuses? In: O'Keeffe / McCarthy (Hrsg.), 333–344.
- Wildfeuer-Eller, Nicole / Wildfeuer, Alfred (2018): Denken in Kontinua – Überlegungen zum Umgang mit Norm und Variation im DaF- und DaZ-Unterricht. In: Zielsprache Deutsch 45/1, 3–30.
- Weinreich, Uriel (1953): Languages in Contact. Findings and Problems. Den Haag u.a.: de Gruyter.