

## **KORPORA ALS GRUNDLAGE FÜR DAS LEHREN UND LERNEN VON DEUTSCH ALS FREMDSPRACHE**

Carolina Flinz, Britta Hufeisen, Rita Luppi, Ruth M. Mell,  
Università degli Studi di Milano & Technische Universität Darmstadt

### **1. Korpora als Grundlage für das Lehren und Lernen von Deutsch als Fremdsprache**

Spätestens seit Lemnitzer / Zinsmeister (2015: 13) gilt die Definition von Korpus als einer Sammlung schriftlicher oder gesprochener Äußerungen in einer oder mehreren Sprachen in digitaler Form. Das wissenschaftliche Programm der Korpuslinguistik am Leibniz-Institut für Deutsche Sprache, welches das aktuell größte Korpus in deutscher Sprache vorhält, ist daher auch, „geleitet durch die explorative Analyse von sehr großen Sammlungen natürlichsprachlicher Daten neue Einsichten in die Strukturen, Gesetzmäßigkeiten, Eigenschaften und Funktionen von Sprache zu erlangen“. (<https://www.ids-mannheim.de/digspra/kl/>, 17.10.2021.) Seine Verantwortlichkeit liegt dabei darin, „den deutschen Schriftsprachgebrauch beständig und in angemessener Weise stichprobenartig im Deutschen Referenzkorpus zu dokumentieren“ (ebd.). Das Deutsche Referenzkorpus DeReKo bildet mit aktuell 50,6 Milliarden Wörtern (Stand 02.02.2021) die weltweit größte linguistisch motivierte Sammlung elektronischer Korpora und beinhaltet geschriebene Texte in deutscher Sprache aus der Gegenwart und der neueren Vergangenheit. Die Daten sind über die Korpusrecherchertools COSMAS II und KorAP kostenlos abfragbar und können als Basis für multiple linguistische Fragestellungen verwendet werden (<https://www.ids-mannheim.de/digspra/kl/projekte/korpora/>, 17.10.2021).

Der Einsatz von Korpora in den Wissenschaften birgt aber nicht nur fachbezogenes, sondern zudem auch ein großes pädagogisches Potential, das bisher kaum ausgeschöpft ist. So bietet die Analyse größerer – oder bei bei DeReKo sehr großer sprachlicher Datenmengen unter anderem neue Zugänge für die Sprachbeschreibung, für die Fremd- und Zweitspracherwerbsforschung, für die Entwicklung von Nachschlagewerken, Lehr- und Unterrichtsmaterialien sowie für die Unterrichtspraxis. Zwar werden schon lange Sammlungen authentischer Daten für die Fremdsprachendidaktik verwendet (vgl. Lüdeling / Walter 2009: 1), doch bietet die „digitale Form“ die Nutzung korpuslinguistischer Tools anstelle manueller Verfahren für die Analyse der sprachlichen Wirklichkeit und der Erarbeitung didaktischer Lehr- und Lernszenarien; denn Korpora bestehen zumeist nicht nur aus den digitalen Primärdaten, sondern sind eben auch durch Metadaten angereichert, welche Informationen über die lexikalischen Einheiten des Korpus enthalten, die unterschiedliche sprachliche Ebenen betreffen können. Je nachdem, auf welche Weise Korpusdaten und -methoden im Sprachunterricht verwendet werden (vgl. Römer 2006, 2008), bietet ihr Einsatz eine Reihe von unterschiedlichen Vorteilen. Systematische Sammlungen von authentischen Sprachdaten sowie der Gebrauch von bestimmten Verfahren für deren Analyse wie z.B. die Auswertung der Kontexte und das Zählen der Häufigkeit, werden schon seit langer Zeit in der Sprachdidaktik verwendet (vgl. Lüdeling / Walter 2010). Jedoch erst mit dem Aufkommen von Korpora im Sinne von Sammlungen von digitalisierten schriftlichen oder gesprochenen Äußerungen in einer oder mehreren Sprachen (vgl. Lemnitzer / Zinsmeister 2015: 13) und mit der Anwendung von korpuslinguistischen Methoden haben sich neue Möglichkeiten für die Fremdsprachendidaktik ergeben, vorerst im Bereich von Englisch als Zweit- und Fremdsprache (vgl. u.a. Mukherjee 2002; Corino 2014) und anschließend auch für das Deutsche (vgl. Fandrych / Tschirner 2007; Lüdeling / Walter 2009; Wallner 2013; Flinz 2021; Flinz / Hufeisen 2021).

Das reichhaltige thematische Portfolio und die Diversität der Arbeitsgebiete wie der Korpora wurde schon in der ersten Ausgabe unserer Zeitschrift deutlich, in der unterschiedliche Korpusarten zur Anwendung kamen: monolinguale und mehrsprachige Korpora, vergleichbare Korpora, mündliche, schriftliche sowie multimodale Korpora, allgemeine Korpora und Spezialkorpora. Das DWDS-Korpus, DeReKo, FOLK, das GeWiss-Korpus sowie eigenerstellte Korpora wurden zu unterschiedlichen Schwerpunkten für die DaF-Didaktik und im DaF-Unterricht benutzt: Kollokationen, Paronyme, unterschiedliche lexikalische Mittel, Sprachvergleich, der Ansatz des Mikro-Hörens sowie die Gestikanalyse für die DaF-Lehrerausbildung sind nur einige der dort behandelten Themen.

Unser Bestreben ist es, genau die große inhaltliche Breite des Themengebiets darzustellen und sie in den Beiträgen in unserer Zeitschrift abzubilden. So sollen neben wissenschaftstheoretischen und forschungsmethodischen Fragestellungen gerade auch anwendungsbezogene Beispiele rund um die Arbeit mit Korpora mit Bezug zu Deutsch als Fremdsprache vorgestellt werden. Zugleich wollen wir dabei immer wieder der Bezug zur Arbeit, Forschung und Lehre mit Korpora herstellen. Themenschwerpunkte sind daher auch Korpusaufbau und -verwendung für DaF, DaF-Didaktik und DaF-Erwerb, Bildungssprache/Wissenschaftssprache/Fachsprache oder korpusbasierte/korpusgestützte Fremd- und Zweitsprachenvermittlung.

## 2. Übersicht über die Beiträge

Die Beiträge der zweiten Ausgabe heben erneut hervor, wie vielgestaltig und themenreich die in KorDaF versammelten Beiträge sind: Fandrych / Meißner / Wallner beschreiben in ihrem Beitrag das Potential von Korpora der gesprochenen Sprache für die DaF/DaZ-Forschung und -Didaktik. Auch wenn die Entwicklung mündlicher Korpora in den letzten Jahren, sowohl was den Umfang, als auch was die Aufbereitung betrifft, große Fortschritte gemacht hat, wurde die direkte Nutzung der Korpora für die Entwicklung von Lehr- und Lernmaterialien, Curricula und für den Unterricht selbst bisher kaum verfolgt. Mit dem Projekt *ZuMult* (Zugänge zu multimodalen Korpora gesprochener Sprache) wurden hingegen für diese Zielgruppe spezielle Zugänge zu mündlichen Korpora geschaffen, durch die gezielt nach verschiedenen didaktisch relevanten Aspekten gesprochener Sprache in den Korpora gesucht werden kann. Dies wird für gesamte Korpora und damit für ein weites Spektrum an vielfältigen und nach individuellen Bedürfnissen selektierbaren Sprechereignissen ermöglicht. Im Beitrag werden drei Perspektiven eingenommen. Als erstes findet die Vorstellung der aktuellen Einsatzmöglichkeiten von Korpora gesprochener Sprache für den Unterricht statt: Im Fokus stehen insbesondere die im Projekt *ZuMult* entwickelten Werkzeuge *ZuMal* (Zugang zu Merkmalsauswahl von Gesprächen) und *ZuViel* (Zugang zu Visualisierungselementen für Transkripte). Anschließend wird erläutert, inwieweit Korpora als Grundlage für die Ermittlung von vermittlungsrelevanten Phänomenen in der gesprochenen Sprache, wie zum Beispiel Klitisierung, genutzt werden können. Schließlich wird auf die Möglichkeiten und Grenzen der Analyse (fortgeschrittener) gesprochener Lerner/-innensprache auf der Grundlage mündlicher Korpora eingegangen.

Ebel / Skoczec stellen in ihrem Aufsatz *Die Deutsche Aussprachedatenbank (DAD)* vor. Die Online-Datenbank, aktuelles Ergebnis der halleschen Kodifikationsarbeiten, bietet neben normphonetischen Angaben zum allgemeinen Wortbestand auch empirisch fundierte Eindeutschungsformen und -regeln etablierter und aktueller fremdsprachiger Lemmata an. Im Fokus sind u.a. der terminologische Apparat, der aktuelle Forschungsstand und die technischen Möglichkeiten der Datenbank. Es werden außerdem Anregungen für die sprachdidaktische Nutzung der Aussprachedatenbank gegeben. Dabei wird gezeigt, wie sie im Unterricht in Deutsch als Fremd-, Zweit- und Muttersprache sowohl als zuverlässige Referenzquelle für die normkonforme Aussprache als auch als schnell verfügbares und einfach durchsuchbares Nachschlagewerk genutzt werden kann.

Sabrina Bertollo beschäftigt sich in ihrem Beitrag mit dem *Atlas zur deutschen Alltagssprache* (AdA) und zeigt, wie er im DaF-Unterricht wirksam eingeführt werden kann, um die Bewusstheit über Sprachvariation auch in den Anfängerstufen zu fördern. Es wird erörtert, warum Sprachvariation einen Platz im DaF-Unterricht finden sollte und inwiefern der AdA sich dafür eignet. Zuletzt wird ein umfangreicher Unterrichtsentwurf ausführlich präsentiert. Der direkte Umgang mit dem Atlas durch die Lernenden, Arbeitsmethoden, technologische Aspekte, Lernziele sowie die unterschiedlichen Arbeitsphasen werden dabei betrachtet, sodass die Lerneinheit im DaF-Unterricht unmittelbar eingesetzt werden kann.

Eva Gredel fokussiert die kollaborative Textproduktion über digitale Tools, da diese in Zeiten von *Distance Teaching* noch mehr an Relevanz gewonnen hat. Auf Basis der Wikipedia-Korpora des Leibniz-Instituts für Deutsche Sprache (IDS), die Prozesse der Kollaboration in Form großer digital verfügbarer Datenmengen empirisch zugänglich machen, wird das Potential des Einsatzes von CMC-Korpora im DaF-/DaZ-Unterricht aufgezeigt. Drei didaktische Szenarien, die sprachliche Phänomene ganz unterschiedlicher Ebenen von Morphemen über syntagmatische Muster bis hin zu diskursiven Einheiten mithilfe der Wikipedia-Korpora präsentieren, werden vorgestellt und diskutiert. Im ersten Szenario werden morphologisches Wissen und Diskursfähigkeit korpuslinguistisch trainiert: Dabei geht es um nominale Determinativkomposita mit {Klima}, um Derivation und Kombination des Lexems *Nachhaltigkeit* und um Kontamination. Im zweiten Beispiel wird die Aushandlung sprachlicher (Un-)Höflichkeit korpuslinguistisch rekonstruiert; in einem dritten Fall werden Internet- bzw. plattformspezifische Vokabulare und syntagmatische Muster analysiert.

Susanne Kabatnik betrachtet in ihrem Beitrag Funktionsverbgefüge, da diese nicht nur im Deutschen frequent sind, sondern auch in typologisch unterschiedlichen Sprachen vorkommen. Nach einer Veranschaulichung der Behandlung von Funktionsverbgefügen in einer Auswahl an DaF-Lehrwerken behandelt werden, wird eine korpusbasierte quantitativ-qualitative Untersuchung des Funktionsverbgefüges (*eine Frage stellen*) vorgestellt. Dabei wird auch der Gebrauch des Verbgefüges im Kontext fokussiert. Anschließend werden Vorschläge für mögliche Herangehensweisen im Unterricht gegeben, die auf der quantitativen und qualitativen Korpusstudie basieren. Datengrundlage sind das Deutsche Referenzkorpus (DeReKo 2020) sowie das Wikipedia-Artikel-Korpus (2015) des Leibniz-Instituts für Deutsche Sprache. Abschließend wird diskutiert, wie die Untersuchungsergebnisse zur Verbesserung von Lern- und Lehrmaterialien beitragen können.

Fokus des Beitrags von Rainer Perkuhn sind zum einen Worthäufigkeitszählungen von unterschiedlichen großen Datensammlungen, die auch im Vergleich betrachtet werden. Verknüpft mit weiteren Datensammlungen ermöglichen diese eine differenzierte Bewertung für die didaktische Relevanz. Zum anderen bespricht der Autor Kollokations-/Kookkurrenzanalysen, die typische Formulierungsmuster ermitteln können. Nach einer getrennten Darstellung der beiden Herangehensweisen wird abschließend präsentiert, wie beide Perspektiven zusammengebracht werden können.

## Literatur und Ressourcen

Corino, Elisa (2014): Didattica delle lingue corpus-based. In: *EL.LE 3:2-Luglio*, 231-257. <http://doi.org/10.14277/2280-6792/99p> (17.10.2021).

Flinz, Carolina (2021): Korpora in DaF und DaZ: Theorie und Praxis. In: Flinz, Carolina / Hufeisen, Britta (Hrsg.): *Korpora in DaF und DaZ: Theorie und Praxis*. Themenausgabe Zeitschrift für Interkulturellen Fremdsprachenunterricht. Jahrgang 26 (April 2021). <https://ojs.tu-journals.ulb.tu-darmstadt.de/index.php/zif/article/view/1112/1108> (17.10.2021).

- Flinz, Carolina / Hufeisen, Britta (Hrsg.) (2021): *Korpora Deutsch als Fremdsprache 1*: 2021. <https://kordaf.tu-journals.ulb.tu-darmstadt.de/issue/10/info/> (17.10.2021).
- Fandrych, Christian / Tschirner, Erwin (2007): Korpuslinguistik und Deutsch als Fremdsprache. Ein Perspektivenwechsel. In: *Deutsch als Fremdsprache* 44: 4, 195-204.
- Leibniz-Institut für Deutsche Sprache (2021): Was ist Korpuslinguistik? <https://www.ids-mannheim.de/digspra/kl/> (17.10.2021).
- Leibniz-Institut für Deutsche Sprache (2021): Das Deutsche Referenzkorpus DeReKo. <https://www.ids-mannheim.de/digspra/kl/projekte/korpora/> (17.10.2021).
- Lemnitzer, Lothar / Zinsmeister, Heike (2015): *Korpuslinguistik. Eine Einführung*. Tübingen: Narr.
- Lüdeling, Anke / Walter, Maik (2009): Korpuslinguistik für Deutsch als Fremdsprache. Sprachvermittlung und Spracherwerbsforschung, 1-37. <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiter-innen/anke/pdf/LuedelingWalterDaF.pdf> (17.10.2021).
- Lüdeling, Anke / Walter, Maik (2010): Korpuslinguistik. In: Krumm, Hans Jürgen et al. (Hrsg.): *Handbuch Deutsch als Fremd- und Zweitsprache*. Berlin: Mouton de Gruyter, 315-322.
- Mukherjee, Joybrato (2002): *Korpuslinguistik und Englischunterricht. Eine Einführung*. Berlin u.a.: Peter Lang.
- Römer, Ute (2006): Pedagogical applications of corpora: some reflections on the current scope and a wish list for future developments. In: *Zeitschrift für Anglistik und Amerikanistik* 54, 121-134.
- Römer, Ute (2008): Corpora and Language teaching. In: Lüdeling, Anke / Merja Kytö (Hrsg.): *Corpus Linguistics. An International Handbook*. Vol. 1. Berlin / New York: Mouton de Gruyter, 112-131.
- Wallner, Franziska (2013): Korpora im DaF-Unterricht – Potentiale und Perspektiven am Beispiel des DWDS. In: *Revista Nebrija de Lingüística Aplicada* 13. <https://www.nebrija.com/revista-linguistica/korpora-im-daf-unterricht-potentiale-und-perspektiven-am-beispiel-des-dwds.html> (17.10.2021).