

Analyse zur Nutzung der GeWiss-Korpora

Stand Mai 2014

Daisy Lange & Franziska Wallner

Ein Jahr nach dem Erstrelease sollen mit diesem Arbeitspapier die derzeitigen Nutzer(innen) sowie ihre spezifischen Nutzungsinteressen analysiert und diskutiert werden. Die vorliegende Auswertung wurde auf Grundlage der im Rahmen der Nutzer(innen)registrierung für das GeWiss-Korpus erhobenen Informationen durchgeführt.¹ Die händische Autorisierung der Interessent(inn)en erfolgt nach einer Prüfung durch die Projektmitarbeiter(innen), sodass automatisch generierte, unseriöse sowie unvollständige Anfragen weitestgehend ausgeschlossen werden können.

Auf eine Analyse der Häufigkeit des Zugriffs auf die Daten durch die bis dato registrierten Nutzer(innen) muss an dieser Stelle verzichtet werden, da die zur Auswertung notwendigen Logfiles in der GeWiss-Datenbank (u.a. aus datenschutzrechtlichen Gründen) nicht gespeichert werden.

1. Zahl der Nutzer(innen)

Der Zugriff auf die GeWiss-Korpora ist seit März 2013 möglich. Seitdem ist eine kontinuierliche Nachfrage zu beobachten. Derzeit zählt GeWiss 235 Nutzer(innen).

2. Die Nutzer(innen)

Im Folgenden sollen die demographischen und institutionellen Angaben der registrierten Nutzer(innen) zusammengefasst werden. Berücksichtigt werden dabei sowohl die geographische Herkunft der Anwender(innen) als auch ihr institutioneller Hintergrund und beruflicher Status (Professionalisierungsgrad).

2.1. Herkunft nach Ländern

Zur Auszählung der Herkunft nach Ländern wurden bei der Analyse die entsprechenden Angaben zur Institutionszugehörigkeit herangezogen. Die reale geographische Herkunft der Nutzer(innen) wurde nicht gesondert erfasst und kann ggf. davon abweichen. 5 Registrierungen enthielten keine entsprechenden Angaben.

Ein Großteil der Nutzer(innen) (74 %) stammt aus dem deutschen Sprachraum, 169 davon aus Deutschland, 4 aus Österreich und 2 aus der Schweiz. Aus dem europäischen Ausland

¹ Der Zugang zu den GeWiss-Korpora erfordert eine einmalige und kostenfreie Anmeldung über das Portal <https://gewiss.uni-leipzig.de>, über welches auch die Korpora selbst abgerufen sowie ein Tool zur Konkordanzsuche genutzt werden kann. Unter Angabe ihres Namens, einer Adresse, einer Institution und eines Verwendungsgrundes erklären sich die Interessent(inn)en mit den Bedingungen der Nutzung sowie den Datenschutzrichtlinien einverstanden. Die Datenschutzrichtlinien sehen eine Nutzung ausschließlich zum Zweck der Wissenschaft, Forschung und Lehre vor und verbieten eine kommerzielle Verwendung der Daten.

sind 46 Nutzer(innen) registriert (9 Polen, 5 Großbritannien, 4 Bulgarien, 4 Dänemark, 3 Italien, 3 Russland, 2 Spanien, 2 Belgien, 2 Finnland, 2 Frankreich, 2 Tschechische Republik, 2 Türkei, 1 Litauen, 1 Norwegen, 1 Niederlande, 1 Schweden, 1 Ungarn). Die Zahl der außereuropäischen Nutzer(innen) beläuft sich auf 9 und verteilt sich kontinental wie folgt: 4 aus Asien (Libanon, Armenien, Pakistan und Südkorea), 3 aus Nordamerika (USA und Kanada) sowie 2 aus Südamerika (Kolumbien, Argentinien).

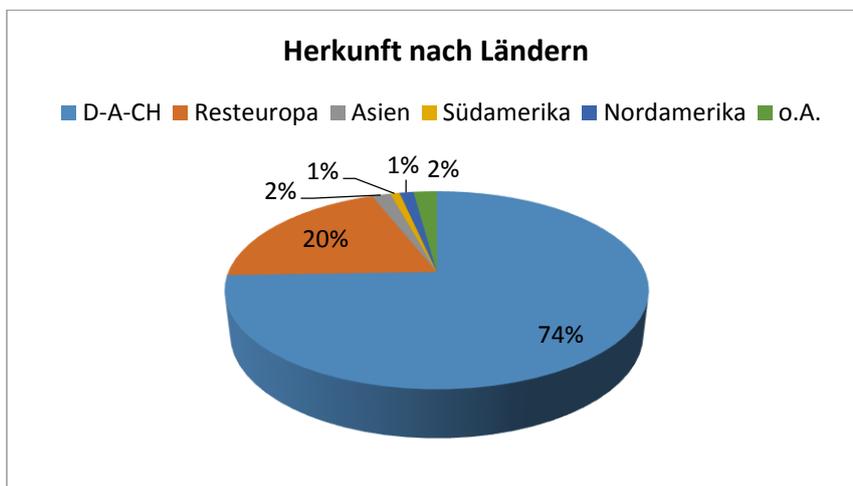


Abb. 2 Herkunft nach Ländern in Prozent

2.2. Institutionszugehörigkeit

204 (91%) der registrierten Nutzer(innen) sind an Hochschulen und Universitäten tätig bzw. als Student(innen) eingeschrieben. Daneben haben sich Mitglieder von Testinstituten (wie TestDaF und telc), Lehrbeauftragte verschiedener universitärer und außeruniversitärer Sprachlehrzentren sowie Angehörige unabhängiger Forschungsinstitute wie dem Institut für deutsche Sprache (IDS), dem Max-Planck-Institut für empirische Ästhetik und dem Zentrum für Allgemeine Sprachwissenschaft (ZAS) als Nutzer registriert.

2.3. Beruflicher Status, Professionalisierungsgrad

Der berufliche Status, d.h. in dem Fall der Grad an akademischer Professionalisierung, ist ein Kriterium, welches bei der Registrierung nicht notwendigerweise angegeben werden musste. Diese Informationen wurden im Rahmen der Auswertung der übrigen Registrierungsangaben zusammengetragen. Die Nutzer(innen) wurden hierfür in fünf übergreifende Kategorien eingeordnet²:

- a) Anwender(innen), die eine Professur innehaben (21);
- b) Privatdozent(inn)en (5);
- c) Anwender(innen) mit Dokortitel (50);

² Die jeweiligen Statusinformationen beziehen sich auf den Zeitpunkt der Registrierung. Zwischenzeitliche Änderungen können hier nicht berücksichtigt werden.

- d) Doktorand(inn)en (33);
- e) Post-Graduierte (34);
- f) Studierende (71).

Die Klasse der Post-Graduierten grenzt sich von den Doktorande(inn)en insofern ab, als für diese Nutzer(innen) nicht bekannt ist, ob Sie sich einem Promotionsprojekt widmen. Es handelt sich zum Großteil um Nutzer(innen), die entweder im wissenschaftlichen Betrieb (z.B. als Assistenten, Mitarbeiter oder wissenschaftliche Hilfskräfte) tätig oder aber in den erwähnten weiteren Institutionen in verschiedenen Funktionen beschäftigt sind. Darunter fallen aber auch die Interessent(inn)en, die sich zum Zwecke der Orientierung hinsichtlich einer Promotion für die Korpusnutzung registriert haben.

Für einen Teil der registrierten Nutzer(innen) (21) konnten derartige Informationen nicht anhand der Registrierungsangaben rekonstruiert werden.

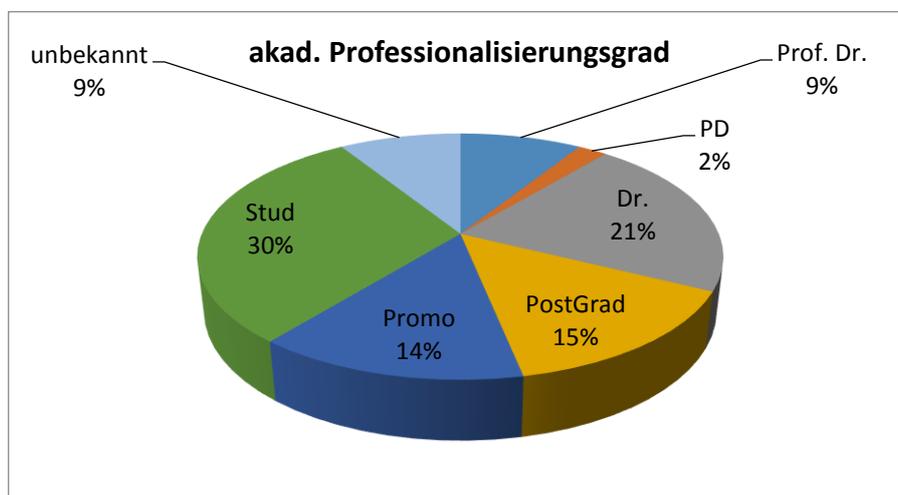


Abb. 3 beruflicher Status/akad. Professionalisierungsgrad in Prozent

3. Verwendungskontexte, Nutzungsinteressen

Bei der Registrierung ist eine Freitext-Eingabe des vorgesehenen Verwendungszweckes erforderlich. Die Ausführlichkeit, mit der dieser dargelegt wird, unterscheidet sich im Einzelfall sehr, sodass an dieser Stelle lediglich einige Schwerpunktbereiche aufgelistet werden können. Die folgenden Zahlenangaben korrespondieren dabei nicht zwangsläufig mit der Anzahl registrierter Personen, da für die individuelle Registrierung zum Teil mehr als ein Motiv der geplanten Nutzung angegeben wurde.

Studierende

Die 71 studentischen Nutzer(innen) geben an, im Rahmen ihres Studiums oder einer studentischen Tätigkeit (bspw. als Tutor(in), studentische Hilfskraft oder als Sprachmittler(in)) mit den Korpora zu arbeiten. 59% davon begründen ihre Registrierung mit der geplanten Datenverwendung in eigenen Qualifikationsarbeiten (Seminar- oder Modul-, BA-, MA- oder Examensarbeiten).

Graduierte

Unter den übrigen, bereits graduierten Nutzer(inne)n beabsichtigen 73% die Daten u.a. im Rahmen von Forschungsarbeiten und 57% diese u.a. zu didaktischen Zwecken zu nutzen, wobei 32% aller graduierten Nutzer zugleich beide Verwendungskontexte benennen.

Ein substantieller Anteil der Nutzer(innen), die die Daten in Lehrkontexten einsetzen, kann dem Bereich der Fremdsprachenvermittlung zugeordnet werden (74%). Innerhalb dieser Nutzergruppe ist das Fach Deutsch als Fremd- und Zweitsprache am häufigsten vertreten. Darüber hinaus wird das Korpus auch im Rahmen von Lehrveranstaltungen in den Fächern Germanistik und Soziologie sowie in korpuslinguistischen und sprachwissenschaftlichen Seminaren eingesetzt.

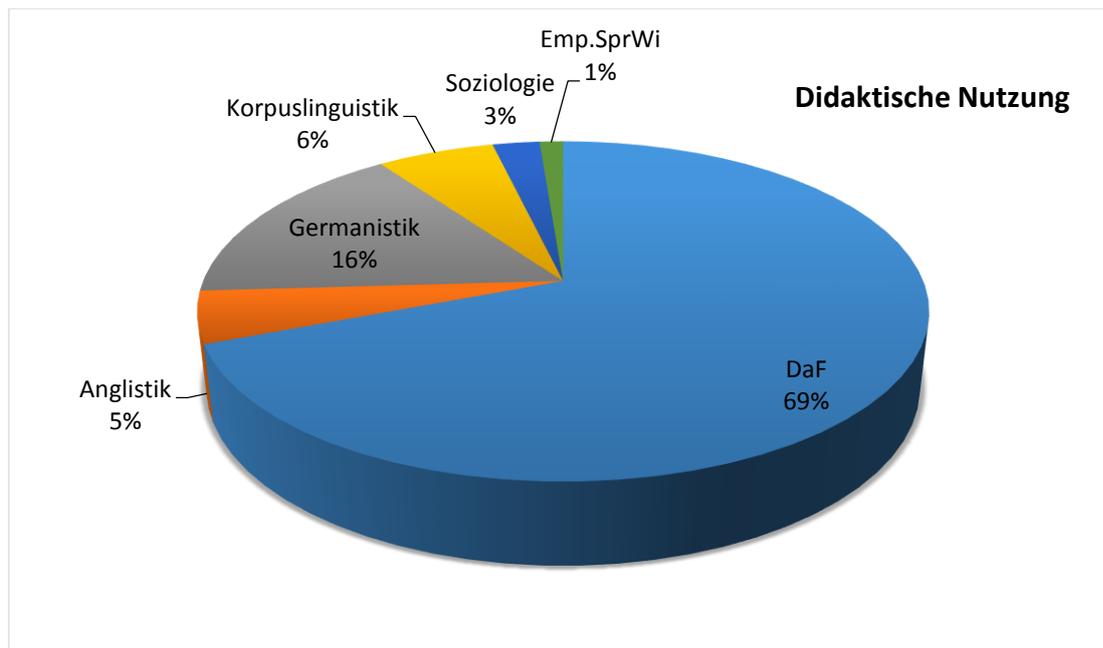


Abb.4 Nutzung für didaktische Zwecke

Bezüglich der Datennutzung für konkrete Forschungsvorhaben wurden nur in ca. der Hälfte der Fälle präzise Angaben gemacht. Aus diesen geht hervor, dass die Kerninteressen in den Teildisziplinen anzusiedeln sind, die sich mit Sprachvariation, darunter z.B. sozio- und pragmalinguistischen Inhalten, oder der kontrastiven Beschreibung verschiedener Diskursgestaltender Merkmale auseinandersetzen. Einen weiteren Schwerpunkt, den gewissermaßen die Art der abrufbaren Daten selbst nahelegt, bildet die Diskurs- und Gesprächsforschung.

Aus den weiteren Nennungen zeichnen sich vereinzelte Untersuchungsinteressen im Bereich der Grammatik, Alphabetisierung, Automatischen Sprachverarbeitung oder Phonetik/Phonologie ab. Auch Analysen vor dem Hintergrund methodischer und anwendungsbezogener Fragestellungen, allen voran auf dem Gebiet der Korpuslinguistik, werden gelegentlich als Motiv für die Registrierung genannt.

Die folgende Tabelle gibt einen groben Überblick über die Themenbereiche, innerhalb derer die Daten genutzt werden. Eine präzise Abgrenzung der Teildisziplinen ist nur teilweise möglich, sodass durchaus von Überschneidungsbereichen ausgegangen werden muss, die als solche nicht immer aus den Freitexteingaben rekonstruiert werden können.

Korpuslinguistik (Methodik, Design)	Clarín-d Erstellung und Aufbereitung von Korpora (Transkription, Annotation) Ressource i.H.a. didaktische Fragestellungen Grundwortschatzanalyse (akademischer Grundwortschatz)
Phonetik	Grundfrequenzanalysen Klangwechselforschung Reduktionen Rhythmusforschung
Kontrastive Linguistik	Vergleichend im Hinblick auf Genre <ul style="list-style-type: none"> - Erforschung institutioneller Kommunikation (akademischer vs. ausbildungsberuflicher Kommunikationsbereich) ➔ Vergleich zu Prüfungsgesprächen in der Altenpflegeausbildung ➔ Vergleich zu medizinischer Fachsprache, Arzt-Patienten-Kommunikation Vergleichend im Hinblick auf Sprach- und Kulturräume <ul style="list-style-type: none"> - Erforschung wissenschaftlicher Denkstile Vergleichend im Hinblick auf Sprachen und Sprachniveaus (als Sprechereigenschaften) <ul style="list-style-type: none"> - Beschreibung von Wissenschaftssprachen: Englisch, Portugiesisch, Italienisch, Polnisch - Metaphernanalyse Vergleichend im Hinblick auf Sprachgebrauch in fremdsprachigen Kontexten <ul style="list-style-type: none"> - Deutsch als L2 im britischen und bulgarischen akad. Kontext Vergleichend im Hinblick auf Disziplinen <ul style="list-style-type: none"> - Musterhafte Darstellungsformen in Literatur-, Sprachwissenschaft und DaF
Pragmatik	<ul style="list-style-type: none"> - Funktioneller Einsatz von mäßig-Adjektiven - Argumentationskompetenz - Humor in wissenschaftlichen Vorträgen
Grammatik	<ul style="list-style-type: none"> - Modalpartikeln - Verwendung des Genitivs
Diskurs- bzw. Gesprächsforschung	<ul style="list-style-type: none"> - Sprachökonomie - Konversationsanalyse, Dialogforschung - Gesprochene-Sprache-Forschung
Variationslinguistik im weitesten Sinn	<ul style="list-style-type: none"> - Skalierung von Bildungssprache (testwissenschaftlicher Hintergrund) - Fachsprache in internationalen Teams - Sprachwandel: Sprach- und Wortentwicklung - Wissenschafts- und Fachsprachforschung mit verschiedenen Schwerpunkten - Fachkulturforschung
Soziologie	<ul style="list-style-type: none"> - Alltägliche Praxis in der Wissenschaft

Für die Nutzer(innen)- und Nutzungsanalyse ist insgesamt festzuhalten, dass die Informationen, die im Rahmen dieses Papiers ausgewertet worden sind, lediglich einen ersten Einblick hinsichtlich der tatsächlichen Nutzer(innen)bedürfnisse gewähren. Dezidierte Nutzer(innen)analysen auf Grundlage detaillierterer sowie standardisierter Erhebungen und Auswertungen sind daher dringend erforderlich, um bspw. mit der Entwicklung entsprechender Werkzeuge und Portalfunktionen spezifischen Nutzungsinteressen entgegenkommen zu können.