

HEFT
2.20
15
MÜNCHEN

Spiegelungen

ZEITSCHRIFT FÜR
DEUTSCHE KULTUR UND GESCHICHTE
SÜDOSTEUROPAS



Konzepte des Kollektiven

SPIEGELUNGEN-PREIS FÜR MINIMALPROSA 2020 / JOSEF BALAZS /
KATHARINA EISMANN / KRISTIANE KONDRAT / HORST SAMSON /
HELLMUT SEILER

Spiegelungen

ZEITSCHRIFT FÜR
DEUTSCHE KULTUR UND GESCHICHTE
SÜDOSTEUROPAS

HEFT 2.2020
JAHRGANG 15

Konzepte des Kollektiven

Herausgegeben von Florian Kühner-Wielach
unter Mitwirkung von Enikő Dác, Angela Ilić und Tobias Weger
im Auftrag des Instituts für deutsche Kultur und Geschichte
Südosteuropas an der Ludwig-Maximilians-Universität München

Redaktion: Georg Aesch (Literatur), Enikő Dác (Ressortleitung Literatur),
Ralf Grabuschnig, Klaus Hübner, Angela Ilić (Ressortleitung Kultur),
Florian Kühner-Wielach (verantwortlicher Redakteur), Doris Roth (Rezensionen),
Joachim Schneider, Tobias Weger (Ressortleitung Wissenschaft)

Wissenschaftlicher Beirat: Dr. Lilia Antipow, Dr. habil. Mathias Beer, Prof. Dr. Andrei
Corbea-Hoişie, Prof. Dr. Elisabeth Knipf-Komlósi, Dr. Oxana Matychuk, PD Dr.
Aurelia Merlan, Prof. Dr. Maren Röger, Dr. Harald Roth, Prof. Dr. Irena Samide



**Institut für deutsche Kultur
und Geschichte Südosteuropas**
an der LMU München

VERLAG
FRIEDRICH PUSTET

Die Zeitschrift *Spiegelungen* setzt die
Süddeutschen Vierteljahresblätter (1952 – 2005) fort.

Gefördert von:  Die Beauftragte der Bundesregierung
für Kultur und Medien

Abbildung vordere Umschlagseite: Ausschnitt aus der Ausschlagkarte von Arnold Hillen
Ziegfeld im Band von Richard Bahr: Deutsches Schicksal im Südosten. Hamburg 1936

Die wissenschaftlichen Beiträge von Cornelia Eisler, Gerhard Seewann, Maren Röger,
Hans-Christian Petersen, Isabel Röskau-Rydel und Tobias Weger wurden im Rahmen
eines doppelblinden Peer-Review-Verfahrens nach internationalen Standards begutachtet.

Die *Spiegelungen* sind in ERIH PLUS (European Reference Index for the Humanities
and Social Sciences) indiziert.

IMPRESSUM

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;
detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

© 2020 Institut für deutsche Kultur und Geschichte Südosteuropas an der
Ludwig-Maximilians-Universität München und Verlag Friedrich Pustet, Regensburg

Die Zeitschrift und alle in ihr enthaltenen Beiträge sind urheberrechtlich geschützt. Mit Ausnahme der
gesetzlich zugelassenen Fälle ist eine Verwertung ohne Einwilligung des Verlages strafbar.

Die namentlich gezeichneten Beiträge geben die Meinung der Autorinnen und Autoren wieder.

Redaktion »Spiegelungen«
Institut für deutsche Kultur und Geschichte Südosteuropas
an der Ludwig-Maximilians-Universität München
Halskestraße 15, D-81379 München
Telefon: +49 (0) 89 780609-0, Fax: +49 (0) 89 780609-22
Zusendungen an die Redaktion werden erbeten an: redaktion@ikgs.de
Redaktionshinweise auf www.ikgs.de
Für unverlangt eingegangene Manuskripte wird keine Haftung übernommen.

ISSN 1862-4995
ISBN 978-3-7917-3169-8

Diese Publikation ist auch als E-Book erhältlich
eISBN 978-3-7917-7319-3 (pdf)

Bestellung, Vertrieb und Abonnentenverwaltung:
Verlag Friedrich Pustet, Regensburg
Gutenbergstraße 8, D-93051 Regensburg
Tel. +49 (0) 941 92022-0, Fax +49 (0) 941 92022-330
verlag@pustet.de
www.verlag-pustet.de

Preis des Einzelheftes: € 17,- zuzüglich Porto- und Versandkosten; E-Book (pdf) € 12,99
Jahresbezug: € 28,- zuzüglich Porto- und Versandkosten; E-Journal (pdf) € 21,-

Bankverbindung: Verlag Friedrich Pustet, Sparkasse Regensburg,
IBAN DE37 7505 0000 0000 0002 08, BIC BYLADEM1RBG

Kündigung des Jahresabonnements nur schriftlich bis 1.10. zum Ende des jeweiligen Kalenderjahres.

Umschlaggestaltung & Layout: Martin Veicht, Regensburg
Satz: Vollnhals Fotosatz, Neustadt a. d. Donau
Druck und Bindung: Friedrich Pustet, Regensburg
Printed in Germany 2020

INHALT

Editorial	7
------------------	-------	---

WISSENSCHAFT

Konzepte des Kollektiven

TOBIAS WEGER:	Einleitung	9
CORNELIA EISLER:	»Making sense of data«. Digitale Sprachdaten und die Begriffsgeschichte(n) der »Deutschen im Ausland«	11
GERHARD SEEWANN:	Deutsche, (Donau-)Schwaben, Ungarndeutsche	27
MAREN RÖGER:	Kollektivbegriff mit deutlichen Grenzen. Die Buchenlanddeutschen	35
HANS-CHRISTIAN PETERSEN:	Wechselnde Loyalitäten, völkische Traditionen und kritische Selbstreflektion. Zur Begriffsgeschichte der Bessarabiendeutschen	47
ISABEL RÖSKAU-RYDEL:	Galiziendeutsche	59
TOBIAS WEGER:	Karpatendeutsche. Zur Migration eines ethnischen Kollektivbegriffs	69

Aufsätze

ALFRED WILDFEUER:	<i>Was woafse dos woafse</i> – Oberdeutsche Siedlungen in der ukrainischen Oblast Transkarpatien und im rumänischen Banater Bergland	79
-------------------	--	-------	----

Institutionen

JOHANNA HOLZER:	Thematisches Netzwerk Mehrsprachigkeit. Ein Gespräch mit Claudia M. Riehl	101
-----------------	--	-------	-----

Projektwerkstatt

VIKTÓRIA NAGY, SZIMONETTA WALDHAUSER:	Wörterbuch der ungarndeutschen Mundarten (WUM) – Werkstattbericht	105
BEÁTA MÁRKUS:	»Sie waren nie in ihrem Leben Deutsche«. Lokale Archivreisen zur Deportation der Ungarndeutschen in die Sowjetunion 1944/1945. Ein atypisches Fallbeispiel	109

Rezensionen 119

Ulf Brunnbauer, Klaus Buchenau: Geschichte Südosteuropas (*Aleksandar Jakir*) / Bruno Dobrić: Novine i časopisi na njemačkom jeziku u Istri (1871–1918). Pula, Opatija, Brijuni [Zeitungen und Zeitschriften in deutscher Sprache in Istrien (1871–1918). Pola, Abbazia, Brioni] (*Rudolf Preinerstorfer*) / Alexandru Dragomir: Chronos. Notizbücher über Zeit (*Mădălina Diaconu*) / Frank Henschel: »Das Fluidum der Stadt ...«. Urbane Lebenswelten in Kassa/Košice/Kaschau zwischen Sprachenvielfalt und Magyarisierung 1867–1918 (*Tobias Weger*) / Brigitte Klosterberg, István Monok (Hgg.): Die Hungarica Sammlung der Franckeschen Stiftungen zu Halle. Alte Drucke: 1495–1800. Bearbeitet von Attila Verók (*Christian Rother*) / Ingrid Schiel: Frei – Politisch – Sozial: Der Deutsch-Sächsische Frauenbund für Siebenbürgen 1921–1939 (*Ingeborg Szöllösi*)

Berichte 137

Internationale Tagung: »Zwischen Kronen und Nationen. Die zentral-europäischen Priesterkollegien in Rom vom Risorgimento bis zum Zweiten Weltkrieg« (*Theresa Gillinger*)

LITERATUR

Spiegelungen-Preis für Minimalprosa 2020	147
<i>Kristiane Kondrat</i> : Der ultimative Beweis	166
<i>Josef Balazs</i> : Gedichte	170
<i>Katbarina Eismann</i> : Gedichte	175
<i>Horst Samson</i> : Gedichte	178
<i>Hellmut Seiler</i> : Flüchtigkeit	186

FEUILLETON

Aspekte

CARMEN-FRANCESCA BANCIU:	Theaterzeiten. Ein Bericht über meine Erfahrung als Schriftstellerin, Co-Autorin und Performerin von <i>Ein Land voller Helden</i>	189
RENATA SAKOHOESS:	»Mit Kopf und Herz ein Eremit geblieben«. Die Pressburger Autorin Therese Schröer (1804–1885) und ihre Briefe an Karl von Holtei	196
RENATE LUNZER:	Melancholie über Österreichs einstigem Kriegshafen Pula/Pola. Seearsenal und Schiffsbauwerft Uljanik sind bankrott	199

Fokus: Rijeka – Europäische Kulturhauptstadt 2020–2021

IRVIN LUKEŽIĆ:	Arthur Steinacker, Bankier und Fabrikant aus Rijeka	206
----------------	---	-----

Personalia

ANDREA KOLBUS:	Hans Dama zum Geburtstag	209
ENIKŐ DÁCZ:	Ein kritischer Denker und Vermittler. Zum Tod von Prof. Dr. Andreas Oplatka	213

Besprechungen

Jürgen van Buer, Josef Balazs (Hgg.): Der befestigte Glaube. Kirchenburgen in Siebenbürgen (<i>Lilia Antipow</i>) / Ilse Hehn: Roms Flair in flagranti (<i>Franz Heinz</i>) / Miljenko Jergović: Die unerhörte Geschichte meiner Familie (<i>Anna Hodel</i>) / Natalka Sniadanko: Frau Müller hat nicht die Absicht, mehr zu bezahlen (<i>Alexander Kratochvil</i>) / Friedemann Spicker (Hg.) unter Mitarbeit von Angelika Spicker-Wendt: Beziehungsweisen. Elazar Benyoëtz: Ein Porträt aus Briefen (<i>Klaus Hübner</i>) / Aleš Šteger: Logbuch der Gegenwart (<i>Klaus Hübner</i>) / Iris Wolff: Die Unschärfe der Welt (<i>Klaus Hübner</i>)	215
---	-----

FORUM

Aus dem IKGS	233
--------------------	-----

Neuerscheinungen des IKGS	237
------------------------------------	-----

Editorial

Mit dem Schwerpunktthema „Konzepte des Kollektiven“ widmen wir uns in dieser Ausgabe exemplarisch Fragen der Benennung von Gruppen in Südosteuropa. Selbstverständlich geht es um viel mehr als um die bloße Nomenklatur. Die kollektive Zuschreibung, sei sie nun endogen oder exogen, akademisch induziert oder im Alltagsdiskurs etabliert, ist untrennbar mit der Wahrnehmung einer Gruppe verbunden – und oft eine Art Handlungsanweisung: Werdet Euch ähnlich, fühlt Euch zusammengehörig! Unter der redaktionellen Leitung von Tobias Weger haben sich sechs Autorinnen und Autoren auf die Spuren dieses Phänomens begeben und inspirieren mit ihren verschiedenen Perspektiven zu einer vertieften Auseinandersetzung.

Im Feld der Sprachwissenschaft bewegen sich die weiteren Beiträge: Alfred Wildfeuer (Augsburg) untersucht Alltagsdiskurse aus einer philologischen Perspektive und wirft für uns einen vergleichenden Blick auf Transkarpatien und das Banat. Im Interview mit Johanna Holzer stellt unsere Vorstandsvorsitzende Prof. Claudia M. Riehl die Internationale Forschungsstelle für Mehrsprachigkeit an der LMU München vor, deren Leiterin sie ist. Im Werkstattbericht zum Wörterbuch der ungarndeutschen Mundarten (WUM) von Viktória Nagy und Szimonetta Waldhauser rückt einmal mehr die Dialektforschung in den Mittelpunkt. Beáta Márkus stellt in ihrem Arbeitsbericht lokale Archivquellen zur Deportation der Ungarndeutschen in die Sowjetunion 1944/1945 vor.

Im Literaturteil dieser Ausgabe werden neben Beiträgen unserer etablierten Autorinnen und Autoren erstmals die mit dem Spiegelungen-Preis für Minimalprosa 2020 ausgezeichneten Texte von Natalie Buchholz, Mariana Codruț und Halyna Jazenko in ihren deutschsprachigen Fassungen veröffentlicht. Darüber hinaus können Sie die deutschsprachigen Versionen weiterer, von den Juries sehr weit vorne gereihter Texte nachlesen.

Im Feuilleton spiegelt sich wie immer das breite Spektrum unserer Themen wider: Es geht mit Carmen-Francesca Banciu um eine persönliche Perspektive auf zeitgenössisches Theater, Renata SakoHoess stellt uns die Pressburger Autorin Therese Schröer vor, und mit Renate Lunzer werfen wir einen melancholischen Blick auf die ehemalige k. u. k. Hafenstadt Pula. Nicht zuletzt fügt Irvin Lukežić mit einem Porträt Arthur Steinackers, Bankier und Fabrikant aus Rijeka, unserem Kulturhauptstadt-Mosaik einen weiteren Baustein hinzu.

Rijeka behält angesichts der schwierigen Umstände den Status der Europäischen Kulturhauptstadt auch im Jahr 2021. Darum werden wir unter der redaktionellen

EDITORIAL

Leitung von Angela Ilić unseren Rijeka-Fokus weiterführen – die Beiträge sind, wie auch die Besprechungen und weitere ausgewählte Texte, auf www.spiegelungen.net frei zugänglich. Wir sind froh, dass wir gemeinsam mit dem Verlag Friedrich Pustet diesen „hybriden“ Weg gehen können und auf diese Weise eine noch breitere Leserschaft erreichen.

Die *Spiegelungen* erfahren, nachdem wir vor einiger Zeit unser Doppelblindes Peer-Review-Verfahren eingeführt haben, zwei weitere Aufwertungen: Einerseits verfügt unsere Zeitschrift ab dieser Ausgabe über einen Redaktionsbeirat, der der Redaktion in konzeptionellen Fragen beratend zur Seite steht. Zum anderen sind die *Spiegelungen* seit Kurzem offiziell in ERIH PLUS (European Reference Index for the Humanities and Social Sciences) indiziert. Auch aus der Redaktion gibt es Neuigkeiten: Ab der nächsten Ausgabe wird Tobias Weger die Aufgabe des verantwortlichen Redakteurs übernehmen. Florian Kühner-Wielach bleibt Herausgeber und arbeitet weiter in der Redaktion mit.

Eine aufschlussreiche Lektüre wünscht

Ihre

Spiegelungen-Redaktion

Konzepte des Kollektiven

Benennungen deutscher Gruppen in Südosteuropa, deren Genese, Funktion und Bedeutung

Wer sich mit deutscher Kultur und Geschichte in Südosteuropa befasst, stößt unweigerlich auf Kollektivbegriffe, mit denen eine lokalisierbare Gruppenzuschreibung sprachlich mitkonstituiert worden ist. Bei genauerer Betrachtung fällt auf, dass – abgesehen von den bereits ins Mittelalter zurückreichenden Begriffsbildungen um Zipser und Siebenbürger Sachsen – viele Gruppenbenennungen, die etwa auch in den Namen von Landsmannschaften, Kultureinrichtungen, Presseorganen und anderen Institutionen und Medien auftauchen, Produkte des 19. und des frühen 20. Jahrhunderts sind, ja zum Teil sogar erst im Laufe des 20. Jahrhunderts etabliert worden sind. Mit der ethnischen Zuschreibung »Deutsche« wurden die Namen von Regionen beziehungsweise Verwaltungsgliederungen verbunden – etwa im Falle der »Bukowinadeutschen«, »Galiziendeutschen«, »Bessarabiendeutschen« oder »Dobrudschadeutschen«. In anderen Fällen wählte man für das Kompositum geografische Komponenten wie Gebirgszüge – man denke etwa an die »Alpendeutschen«, die »Sudetendeutschen« oder die »Karpatendeutschen«. Das Verständnis moderner Nationalstaaten und die Idee des Minderheitenschutzes führte zu Meta-Kollektiven wie »Ungarndeutsche«, »Rumäniendeutsche«, »Slowakeideutsche« oder »Jugoslawiendeutsche«, die zum Teil historisch, konfessionell und sozial sehr heterogene Teilgruppen subsumierten. Um die Mitte des 20. Jahrhunderts versuchte man sogar, durch den Sammelbegriff »Südostdeutsche« noch eine weitere Metaebene einzuführen. Wieder andere Bezeichnungen gehen auf Außenperspektiven und Raffungen zurück – so wird von »Donauschwaben« gesprochen, obwohl deren Vorfahren nicht ausschließlich aus dem schwäbischen Sprachgebiet nach Südosteuropa ausgewandert sind.

Konzepte des Kollektiven haben ihre Geschichte, und sie erfüllen unterschiedliche Funktionen: Als deskriptive Begriffe können sie dem Bedürfnis entspringen, Angehörige einer bestimmten Ethnie, die in einer Region lebten, unter einer Bezeichnung zusammenzufassen und damit eine allgemeine Verständigungsgrundlage zu liefern. Nicht selten entspringen sie dem seit dem 19. Jahrhundert im deutschen Diskurs verbreiteten Bestreben, die Deutschen als eine Gesamtheit unterschiedlicher Stammesgemeinschaften zu imaginieren. Konzepte des Kollektiven können aber auch einen normativen Anspruch verfolgen, indem sie terminologisch einen Verge-meinschaftungsprozess festschreiben und bestimmte Menschen als nicht der Gruppe

zugehörig erklären. Das mussten im 18. Jahrhundert bereits die von den Habsburgern vor allem aus Kärnten und dem Salzkammergut nach Siebenbürgen verbannten Protestanten, die »Landler«, erfahren. Mit den seit Jahrhunderten in der Region lebenden Siebenbürger Sachsen bildeten sie als Lutheraner zwar eine Konfessionsgemeinschaft, doch wurden sie nicht in deren nach ständischen Prinzipien verfasstes Kollektiv aufgenommen. Mit der zunehmenden Bedeutung des »Rassegedankens« im Zuge nationalistischer und alldeutscher Bestrebungen in den letzten beiden Jahrzehnten des 19. Jahrhunderts erhielten Konzepte des Kollektiven ebenfalls einen distinguierenden, ausgrenzenden Charakter. Ausgeschlossen wurden häufig zwei- oder mehrsprachige Gemeinschaften ohne eindeutiges Identitätsbekenntnis, vor allem aber die Juden, auch dann, wenn sie sich selbst als Träger deutscher Kultur und Geschichte empfanden. Kollektivbegriffe, die um 1900 in Gebrauch kamen, beschränkten sich demnach auf christliche Deutsche. Die Zwangsmigrationen des 20. Jahrhunderts verstärkten oftmals die kollektiven Kohäsionsprozesse und den zugehörigen Sprachgebrauch, indem sie emotional fundierte »Schicksalsgemeinschaften« hervorbrachten.

Die Aufsätze des Themenschwerpunkts befassen sich exemplarisch mit Konzepten des Kollektiven in Bezug auf die Deutschen in Südosteuropa und spüren deren Genese, Funktion und Bedeutung in unterschiedlichen Zusammenhängen nach. Mit einem innovativen Ansatz geht Cornelia Eisler (Oldenburg) der Frage nach, wie digitale Sprachdaten für die Ermittlung von Begriffsgeschichten der »Deutschen im Ausland« nutzbar gemacht werden können, und zeigt dabei Möglichkeiten, aber auch Grenzen dieses Vorgehens auf. Gerhard Seewann (München) präsentiert in der Perspektive einer »longue durée« die gewandelten Bezeichnungen für die deutschen Siedler in Ungarn. Mit dem Terminus »Bukowinadeutsche« befasst sich Maren Röger (Augsburg), während Hans-Christian Petersen (Oldenburg) der Begrifflichkeit »Bessarabiendeutsche« nachspürt. Die Bezeichnung »Galiziendeutsche« untersucht Isabel Röskau-Rydel (Krakau/Kraków); Tobias Weger (München) behandelt die kollektive Benennung »Karpatendeutsche«, die im Laufe des 20. Jahrhunderts eine räumliche Verengung erfahren hat.

Mit dieser Zusammenstellung strebt das IKGS keine enzyklopädische Vollständigkeit aller Begrifflichkeiten an, sondern möchte anhand ausgewählter Beispiele das Bewusstsein für die Zeitgebundenheit und die Konstruktion solcher Konzepte schärfen.

Tobias Weger

»Making sense of data«¹ Digitale Sprachdaten und die Begriffs- geschichte(n) der ›Deutschen im Ausland‹

Von Cornelia Eisler, Bundesinstitut für Kultur und Geschichte der Deutschen
im östlichen Europa (BKGE), Oldenburg

Die Bitte des Herausgeberteams für diesen Beitrag zum Thema »Kollektive Begriffe« bestand darin, im Sinne einer Top-Down-Perspektive zu untersuchen, inwieweit die Implementierung bestimmter Begrifflichkeiten im Deutschen Reich oder auch Österreich Einfluss auf deutsche Minderheiten in der Zwischenkriegszeit hatte und möglicherweise Normierungen erfolgten. Die Fragestellung klang spannend und herausfordernd; doch wie kann die Vorgehensweise aussehen, wenn plötzlich Archive und Bibliotheken pandemiebedingt schließen und zentrale Quellen nicht mehr direkt verfügbar sind? Die wissenschaftlichen Arbeitsweisen sind längst hybrid: Findbücher durchblättern und in Datenbanken recherchieren, analoge Dokumente durchsehen und digitale Quellen hinzuziehen. Nun standen vorerst nur noch die digitalen Angebote zur Verfügung, und diese Situation bot umso mehr den Anlass, eben diese sehr viel genauer auf ihre Verwendbarkeit hin zu prüfen. Daher entspricht der Beitrag insofern der ursprünglichen Bitte, als er sich den Kollektivbegriffen widmet, aber stärker methodisch angelegt das Potenzial digitaler Ressourcen für die Untersuchung ihres sprachlichen Gebrauchs erörtert.

Im Zentrum des Beitrages stehen das lexikalische Informationssystem samt der Referenz- und Zeitungskorpora des Digitalen Wörterbuchs der deutschen Sprache (DWDS)² und der Google Books Ngram Viewer³ – verbunden mit der Frage, inwieweit sich automatisierte Verfahren und Angebote aus der Korpuslinguistik⁴ für

1 Fred Gibbs, Trevor Owens: The Hermeneutics of Data and Historical Writing. In: Jack Dougherty, Kristen Nawrotzki (Hgg.): Writing History in the Digital Age (Digital Humanities/digitalculturebooks). Ann Arbor 2013, S. 159–170, hier: S. 160.

2 Berlin-Brandenburgische Akademie der Wissenschaften (Hg.): DWDS – Digitales Wörterbuch der deutschen Sprache. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart, <<https://www.dwds.de/>>, 14.2.2020.

3 Google Books Ngram Viewer, <<https://books.google.com/ngrams/info>>, 24.4.2020.

4 Die Korpuslinguistik als empirische Methode beinhaltet heute die elektronische Speicherung großer Textmengen im sprachlichen Kontext, die gezielt durchsuchbar sind und als Grundlage für statistisch basierte

(kultur-)historische beziehungsweise begriffsgeschichtliche Fragestellungen nutzen lassen. Der Anglist und Sprachwissenschaftler Christian Mair hofft, dass unter dem »interdisziplinären Dach der Digital Humanities die digitale Aufbereitung von Sprachdaten in Zukunft auf eine Weise erfolgt, die Anschlussmöglichkeiten in alle relevanten fachlichen Richtungen«⁵ bietet. In seinem Beitrag *Erfolgsgeschichte Korpuslinguistik?* merkt er zudem an, dass die »Suchabfrage in digitalen Textdaten [...] heute kein Spezialwissen einer kleinen Avantgarde von Experten mehr [ist], sondern essenzielle Kulturtechnik, ohne die niemand bestehen kann, der sich als Teil der modernen Welt versteht.«⁶ Historiker*innen wiederum sind angehalten, so die Vorstellung von Fred Gibbs und Trevor Owens, den Suchprozess und auch neue Suchmodi, etwa im Sinne der *Serendipity*, stärker in ihrem hermeneutischen Prozess zu berücksichtigen.⁷ Womöglich ließen sich somit die methodischen Nachteile bezüglich der Textauswahl für die Begriffsgeschichten von Reinhart Koselleck ausgleichen.⁸

Einen leichten Einstieg für Nicht-Linguist*innen erlauben das DWDS und in noch stärkerem Maße Google Books Ngram Viewer aufgrund ihrer freien Zugänglichkeit sowie der Quantifizierung und automatischen Visualisierung der Suchergebnisse. Die Suchinterfaces und Einstellungen sind auf ein überschaubares Maß begrenzt. Es gibt die Möglichkeit, die Rohdaten anzeigen zu lassen oder eine Glättung (bei Google mit Abstufungen) vorzunehmen. Die Suche kann *case-sensitive* erfolgen, also Groß- und Kleinschreibung berücksichtigen. Zudem ist die Einstellung des Zeitbereiches nach Jahren möglich.

Mittels umfangreicher, im Volltext durchsuchbarer Textkorpora im DWDS und im Google Books Ngram Viewer scheint das Anlegen eines spezifischen Korpus, in diesem Fall für die Thematik »Auslandsdeutschtum«, nicht notwendig zu sein, um die Frage nach der Genese und Entwicklung des Konzeptes in spezifischen Kontexten und mit diversen Begrifflichkeiten sowie Bedeutungszuschreibungen beantworten zu können. Die bestehenden Korpora wurden daher im Sinne der Forschungsfragen genutzt und gleichzeitig auf ihre Anwendungsmöglichkeiten hin geprüft. Die systematischen Suchanfragen führten indes von Beginn an zu widersprüchlichen Resultaten. Die vermeintlich schnelle Abfrage entwickelte sich zu einer längeren Recherche, ohne dass – um es vorwegzunehmen – ein überzeugendes Ergebnis vorliegt.

Analysen der Linguistik dienen. Vgl. Ulrike Haß-Zumkehr: Das Wort in der Korpuslinguistik. Chancen und Probleme empirischer Lexikologie. In: Vilmos Ágel u. a. (Hgg.): Das Wort. Seine strukturelle und kulturelle Dimension. Festschrift für Oskar Reichmann zum 65. Geburtstag. Tübingen 2002, S. 45–70, hier: S. 46; Unter einem Korpus werden Texte verstanden, die spezifisch für linguistische Fragestellungen oder sprachhistorische Analysen zusammengestellt werden. Korpora bieten »eine Sammlung von Textdaten, also Sprache im Kontext, die dem Zweck der linguistischen Auswertung dient und eine quantitative Auswertung von (qualitativen) sprachlichen Merkmalen zulässt.« Hagen Hirschmann: Korpuslinguistik. Eine Einführung. Stuttgart 2019, S. 2, S. 16.

5 Christian Mair: Erfolgsgeschichte Korpuslinguistik? In: Marc Kupietz, Thomas Schmidt (Hgg.): Korpuslinguistik. Berlin, Boston 2018, S. 5–25, hier: S. 11.

6 Ebenda, S. 8f.

7 Vgl. Gibbs, Owens: The Hermeneutics, S. 162, S. 165. – Zur zufälligen, glücklichen Entdeckung, die während einer Suche gar nicht intendiert war und zum »Trendwort Serendipity« vgl. Rolf Lindner: Serendipity und andere Merkwürdigkeiten. In: vokus 22 (2012) H. 1, S. 5–11.

8 Zur Begrenztheit der »Begriffsgeschichten« und der Möglichkeiten zur Erweiterung von Textbelegen vgl. Kathrin Kollmeier: Begriffsgeschichte und Historische Semantik, Version: 2.0. In: Docupedia-Zeitgeschichte, <http://docupedia.de/zg/kollmeier_begriffsgeschichte_v2_de_2012>, 6.7.2014; vgl. grundsätzlich: Reinhart Koselleck: Begriffsgeschichten. Studien zur Semantik und Pragmatik der politischen und sozialen Sprache (Mit zwei Beiträgen von Ulrike Spree und Willibald Steinmetz sowie einem Nachwort zu Einleitungsfragmenten Reinhart Kosellecks von Carsten Dutt). Frankfurt am Main 2006.

›AUSLANDSDEUTSCHE‹

Dass die Thematik 1845 mit einem nicht normierten Vokabular aufgegriffen wurde, legten Rainer Münz und Rainer Ohliger vor einigen Jahren in ihrem Beitrag zum Erinnerungsort »Auslandsdeutsche« dar. Sie verwiesen auf die *Verhandlungen der Germanisten zu Frankfurt am Main* im September 1846, in denen zwar nicht wortwörtlich, aber doch explizit das Thema verhandelt wurde.⁹ Des Weiteren erwähnen sie die kurz zuvor erschienene Publikation *Die Verbreitung des deutschen Volkes über die Erde* von Wilhelm Stricker (1816–1891). Der Arzt und Bibliothekar, der sich mit Fragen der Auswanderung und der räumlichen Dimension eines deutschen Sprachgebietes beschäftigt hatte, bedauerte darin, dass die »deutsche Bevölkerung auswärtige[n] Staaten [... als] Dünger«¹⁰ diene, und offenbarte seine Abneigung gegen alles »Fremde« sowie sein Streben nach der Reinheit der Sprache. Seit 1847 gab er zur Verbreitung seiner Ansichten die Zeitschrift *Germania, Archiv zur Kenntniß des deutschen Elements in allen Ländern der Erde* heraus.¹¹ Das Thema »Deutsche im Ausland«, mannigfaltig umschrieben, findet sich demnach spätestens seit der Mitte des 19. Jahrhunderts in deutschsprachigen Publikationen.¹²

Ein viel früheres Erscheinungsdatum lässt auf den ersten Blick allerdings die Suche nach »Ausland(s)deutsche(n)« im Google Books Ngram Viewer vermuten. Bereits 1835 soll – den Belegen zufolge – vom »Schutzbund für Grenz- und Auslandsdeutsche« die Rede gewesen und 1847 die Bezeichnung »Auslanddeutsche« in der *Allgemeinen Zeitung Augsburg/München* erwähnt worden sein.¹³ Eine Prüfung des ersten Beispiels zeigt auf, dass der Suchtreffer für »Auslandsdeutsche« zu einer Publikation von Friedrich von Warnstedt, *Ueber Alterthums-Gegenstände, auf welche die Gesellschaft für die Sammlung und Erhaltung vaterländischer Alterthümer die Aufmerksamkeit ihrer Mitglie-*

9 Vgl. Rainer Münz, Rainer Ohliger: Auslandsdeutsche. In: Etienne François, Schulze Hagen (Hgg.): Deutsche Erinnerungsorte. Band 1. München 2001, S. 370–388, hier: S. 370f. In den »Verhandlungen« schlug die Historische Sektion vor, die Deutschen, die nicht in deutschen Bundesstaaten lebten, einzubeziehen und Einfluss auf ihre Literatur im »deutschen Sinne« zu nehmen: »Abgesehen von dieser Masse von Deutschen in Amerika, sind es die Deutschen im Elsaß, Lothringen, den Niederlanden, die Deutschen jenseits des Niemen, die Deutschen in Böhmen, in Ungarn und Siebenbürgen, welche ein Recht auf die Forterhaltung ihrer deutschen Nationalität und angeborenen Muttersprache haben.« Verhandlungen der Germanisten zu Frankfurt am Main am 24., 25. und 26. September 1846. Frankfurt am Main 1847, S. 106f.

10 Wilhelm Stricker: Die Verbreitung des deutschen Volkes über die Erde. Ein Versuch. Leipzig 1845, S. V.

11 Die Zeitschrift erschien bis 1850 und ging dann über in *Der Deutsche Auswanderer*. Vgl. Germania. Archiv zur Kenntnis d. dt. Elements in allen Ländern d. Erde (Urheberrechtsfreie Bände digitalisiert oder im Digitalisierungsprozess 1.1847–3.1850), <urn:nbn:de:vbv:12-bsb10015779-4>, 21.5.2020; Zu den biografischen Angaben vgl. Rudolf Jung: Lemma: »Stricker, Wilhelm«, in: Allgemeine Deutsche Biographie (1893), <http://www.deutsche-biographie.de/pnd101441592.html>, 16.4.2015; Siehe auch Richard Mai: Auslandsdeutsche Quellenkunde. 1924–1933. Berlin 1936, S. IV.

12 Neben Münz, Ohliger: Auslandsdeutsche, S. 370, gibt auch Bradley D. Naranch an, dass die Begriffe »Auslandsdeutsche« und die »Deutschen im Ausland« seit den 1850er-Jahren allmählich als konkurrierende Begriffe erschienen. Bradley D. Naranch: Inventing the Auslandsdeutsche. Emigration, Colonial Fantasy, and German National Identity, 1848–71. In: Eric Ames, Marcia Klotz, Lora Wildenthal (Hgg.): Germany's Colonial Past. Lincoln, London 2005, S. 21–40, hier: S. 26; Die Bezeichnung »Deutsche im Ausland« lässt sich bislang in der deutschsprachigen Presse der Regionen selbst ab 1875 bzw. 1883 nachweisen, wie etwa im *Brünner Tagesboten* (Jg. 25, Nr. 177, 6.8.1875) oder in der *Kaschauer Zeitung* (Werbeanzeige, Jg. 12, Nr. 146, 22.12.1883). Vgl. Volltextsuche der Bibliothek des Digitalen Forums Mittel- und Osteuropa (DIFMOE). Digitale Bibliothek, <https://www.difmoe.eu/d/>, 11.5.2020.

13 Google Books Ngram Viewer und Links auf Google Books: »Auslandsdeutsche« und »Auslanddeutsche« 1800–2012 in German 2012, <https://books.google.com/ngrams/graph?content=Auslandsdeutsche%2BAuslanddeutsche&year_start=1800&year_end=2008&corpus=20&smoothing=0&share=&direct_url=t1%3B%2C%28Auslandsdeutsche%20%2B%20Auslanddeutsche%29%3B%2C%20#t1%3B%2C(Auslandsdeutsche%20%2B%20Auslanddeutsche)%3B%2C%20>, 24.5.2020.

die zum großen Teil urheberrechtlich geschützt sind.¹⁸ Der Zugang erfolgt über eine Art Online-Portal mit Suchwerkzeugen des DWDS, die über die vorhandenen, eingebundenen Korpora hinweg eingesetzt werden können. Sie bleiben allerdings auch auf die im Portal angebotenen beschränkt.¹⁹ Obgleich sie nicht als repräsentative Datenmengen gelten können, werden die Texte der Korpora kontrolliert zusammengestellt, um bestimmte Textsorten, Zeiträume sowie Themen und somit den Sprachgebrauch abzudecken.²⁰ Die Beratung und Auswahl erfolgt durch Wissenschaftler und Wissenschaftlerinnen der Arbeitsgruppe DWDS und Mitglieder der Berlin-Brandenburgischen Akademie der Wissenschaften.²¹ Nicht nur aus europäisch-ethnologischer Sicht wäre folglich zu fragen, wie die Text-Auswahl aufgrund der Beteiligten einzuordnen ist und welche Auswirkungen dieser Ansatz schließlich auf die Suchergebnisse haben könnte. So ist anzunehmen, dass unter diesen Umständen der Bereich der Expert*innen- oder Eliteforschung eine zentrale Rolle spielen dürfte. Im Vergleich dazu wirkt die Erstellung der Texte in Google Books und somit für den Google Books Ngram Viewer auf den ersten Blick demokratischer. Doch zugleich wirft das von einem US-amerikanischen Konzern durchgeführte Projekt bezüglich seiner Finanzierung, der Sammlung privater und öffentlicher Daten sowie des Urheberrechts Fragen auf.²² Für weiterführende kulturanthropologische Untersuchungen ist hier durchaus Potenzial gegeben, doch vorrangig sollen in diesem Beitrag die Möglichkeiten zur Ergänzung der Begriffsgeschichte der ›Deutschen im Ausland‹ mit Hilfe digitaler Sprachdaten diskutiert werden. Trotz der angedeuteten Vorbehalte bildet der Google Books Ngram Viewer einen Bestandteil der Betrachtungen, da er zu den leicht zugänglichen Suchwerkzeugen gehört, die online und frei zur Verfügung stehen, und bereits Eingang in wissenschaftliche Studien und Recherchen gefunden hat.²³

Es handelt sich beim Google Books Ngram Viewer um eine Suchmaschine *über* und zugleich um ein Visualisierungswerkzeug *für* Datensätze, die aus retrodigitalisierten, mit Annotationen versehenen Texten zusammengestellt werden. Die Suchergebnisse werden automatisiert in einem Verlaufsdiagramm dargestellt, dessen x-Achse als Zeitstrahl dient und dessen y-Achse die relative Menge an Treffern im Vergleich

18 Siehe Berlin-Brandenburgische Akademie der Wissenschaften: Digitales Wörterbuch der deutschen Sprache (DWDS), Referenzkorpora, <<https://www.dwds.de/d/k-referenz>>, 19.4.2020; Siehe auch Stefan Hartmann: Deutsche Sprachgeschichte. Grundzüge und Methoden. Tübingen 2018, S. 50.

19 Hirschmann zufolge gehört das DWDS zu den Suchwerkzeugen, die als »zentral distribuiert, korpuspezifisch, annotationsspezifisch und weniger anfragesprachenmächtig einzustufen« sind, und zwar im Gegensatz zu korpus-unspezifischen Werkzeugen, wie etwa die open-source, browserbasierte und über verschiedene Betriebssysteme greifende Such- und Visualisierungsarchitektur ANNIS oder die freie Software zur Konkordanz- und Textanalyse AntConc. Hirschmann: Korpuslinguistik, S. 105.

20 Ein Referenzkorpus muss »die folgenden Ansprüche erfüllen: Es sollte kontrolliert zusammengestellt sein, so dass wesentliche Parameter (wie Textsorte, Alter der Texte, Themen usw.) über die Gesamtdaten gesehen ausgewogen bzw. homogenisiert sind oder anhand von Metadaten berücksichtigt werden können und je nach Auswertungsziel homogenisiert werden können.« Ebenda, S. 17.

21 Berlin-Brandenburgische Akademie der Wissenschaften: Digitales Wörterbuch der deutschen Sprache (DWDS), Referenzkorpora, <<https://www.dwds.de/d/k-referenz>>, 19.4.2020.

22 Vgl. Daniel Bunčić: Über den Nutzen von Google Books & Co. für (nicht nur slavistische) Begriffsgeschichten. In: Irina Podtergera (Hg.): Schnittpunkt Slavistik. Ost und West im wissenschaftlichen Dialog. Festgabe für Helmut Keipert zum 70. Geburtstag. Göttingen 2012, S. 393–407, hier: S. 403f. Bunčić führt alternative Digitalisierungsprojekte, wie z. B. die HathiTrust Digital Library, <<https://www.hathitrust.org/>>, mit derzeit mehr als 17 Millionen digitalisierten Bänden an, die mittlerweile über erweiterte Suchfunktionen, jedoch nicht über vergleichbare Visualisierungsmöglichkeiten verfügt.

23 Vgl. Hartmann: Deutsche Sprachgeschichte, S. 256f. Auf die Nutzung verweisen auch: Gibbs, Owens: The Hermeneutics.

zu allen n -Grammen (Wörter und Phrasen) des gleichen Jahres darstellt.²⁴ Die Visualisierungen basieren allerdings ausschließlich auf den entweder 2009 oder 2012 zusammengestellten Datensätzen und geben nicht den aktuellen Stand der Google Books wieder.²⁵ Im Unterschied dazu, stammen die Einzelbelege unterhalb des Diagramms, die automatisiert nach Zeitabschnitten gruppiert werden, nicht aus den 2009er- und 2012er-Datensätzen (bzw. Korpora), sondern sind mit einer Suche über die gesamten Angebote von Google Books verknüpft. Das bedeutet, dass beispielsweise bei einer Suche nach ›Auslandsdeutsche‹ die Jahresangaben in der Verlaufskurve nicht mit denen der Belege in Google Books übereinstimmen und somit auch ihre Prüfung nicht möglich ist.²⁶

Im DWDS können mittels Verlaufskurve die relative Worthäufigkeit (pro Million Token²⁷ im Korpus) ebenfalls auf einer Zeitachse angezeigt und Einzelbelege gelistet werden.²⁸ Hier stimmen die Daten für die Kurve und die Belege – Stichproben zufolge – überein. Aufgrund der Widersprüche beim Google Books Ngram Viewer ist ein direkter Vergleich der statistischen Ergebnisse beider Korpora-Angebote kaum möglich. Zudem weisen Quasthoff und Ecker auf die Schwierigkeit eines solchen Vergleiches hin, da die Messwerte von der Art, Zusammenstellung und Bearbeitung der Textkorpora sowie von den Messungen selbst abhängen. Für eine sinnvolle Gegenüberstellung der statistischen Daten müssten die Korpora in einem einheitlichen Verfahren erstellt worden sein.²⁹ Die Auswahl der Textdaten für den Google Books Ngram Viewer erfolgte jedoch nach anderen Kriterien als für das DWDS. Die Publikationen sind nach Angaben von Google Books in 40 weltweit verteilten Universitätsbibliotheken gescannt worden,³⁰ obgleich die Mehrzahl der Treffer zu Publikationen führte, die sich in Universitätsbibliotheken in den USA sowie der Bayerischen Staatsbibliothek in München befinden. Die Werke wurden bis 2012 digitalisiert, mittels OCR auch in Frakturschrift im Volltext durchsuchbar gemacht und online

-
- 24 Unter n -Grammen werden Ketten oder Zeichenfolgen verstanden, die aus einer bestimmten Zahl (n) an Einheiten, vorrangig Wörtern zusammengesetzt sind. Vgl. Hirschmann: Korpuslinguistik, S. 131. Zum Verfahren der Erstellung von Google-Books-Daten vgl. Eitan Adam Pechenick, Christopher M. Danforth, Peter Sheridan Dodds: Characterizing the Google Books Corpus: Strong Limits to Inferences of Socio-Cultural and Linguistic Evolution, In: PLOS ONE. 10 (10), <<https://doi.org/10.1371/journal.pone.0137041>>, 6.2.2020. Die Angabe der n -grams erfolgt prozentual. Die Daten der Korpora können für eigene Analysen direkt heruntergeladen werden und enthalten die konkreten n -gram-Angaben. Vgl. [Datsets]. Google Books Ngram Viewer, <<http://storage.googleapis.com/books/ngrams/books/datasetsv2.html>>, 26.4.2020; Vgl. Yuri Lin u. a.: Syntactic Annotations for the Google Books Ngram Corpus. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (2012), S. 169–174, hier: S. 169.
- 25 »The Ngram Viewer has 2009 and 2012 corpora, but Google Books doesn't work that way. When you're searching in Google Books, you're searching all the currently available books, so there may be some differences between what you see in Google Books and what you would expect to see given the Ngram Viewer chart.« [Information] Google Books Ngram Viewer, <<https://books.google.com/ngrams/info>>, 2.6.2016.
- 26 Vgl. Google Books Ngram Viewer: ›Auslandsdeutsche‹ 1830–2008 in German 2012, <https://books.google.com/ngrams/graph?content=Auslandsdeutsche&year_start=1830&year_end=2008&corpus=20&smoothin_g=0&share=&direct_url=t1%3B%2CAuslandsdeutsche%3B%2CCo#t1%3B%2CAuslandsdeutsche%3B%2CCo0>, 24.5.2020 (Jahresangaben mit Treffern: 1879, 1884, 1902) und die Verlinkung auf Belege von 1800 bis 1921, <https://www.google.com/search?q=auslandsdeutsche&tbm=bks&tbs=cdr:1,cd_min:1830,cd_max:1921&lr=lang_de>, 24.5.2020 (Jahresangaben mit Treffern: 1835, 1907, 1913).
- 27 Token ist die zählbare Primäreinheit (hier Wörter und Satzzeichen). Vgl. Haß-Zumkehr: Das Wort, S. 47f.
- 28 Berlin-Brandenburgische Akademie der Wissenschaften: Verlaufskurven im DWDS, <<https://www.dwds.de/d/plot>>, 27.4.2020.
- 29 Vgl. Uwe Quasthoff, Thomas Eckart: Vergleichbarkeit von sprachstatistischen Messungen. In: Travaux neuchâtelois de linguistique 55 (2011), S. 45–57, hier: S. 56.
- 30 Jean-Baptiste Michel u. a.: Quantitative Analysis of Culture. Using Millions of Digitized Books. In: Science 331 (2011), S. 176–182, hier: S. 176.

gestellt. Publikationen, in denen die Texterkennung nicht den Ansprüchen der Mitarbeiter*innen von Google genügte, sowie Reihen-/Serientitel wurden aussortiert.³¹ Im Vergleich zum DWDS bestand das Ziel hier nicht darin, systematisch einen Querschnitt an Literatur zusammenzustellen oder den Korpus in Bezug auf eine bestimmte linguistische Fragestellung hin zu erstellen. Vielmehr gingen die Ersteller*innen davon aus, dass durch die Zurverfügungstellung großer Mengen an Datenmaterial generell quantitative Untersuchungen linguistischer und kultureller Entwicklungen sowie die Analyse von Grammatik und Syntax ermöglicht würde.³² Es handelt sich nach Angaben von Lin et al. um circa 6 Prozent aller bis 2012 publizierten Bücher.³³ Die Metadaten zur Datierung und zum Ort der Publikation lieferten Bibliotheken und Verlage, die dem Google Books Team zufolge, um Informationen aus bibliografischen Datenbanken ergänzt wurden.³⁴ Die genaue Umsetzung bleibt offen, und die oben angeführten Beispiele weisen bereits darauf hin, dass die Metadaten, also die Informationen über die Primärdaten, offenbar fehlerhaft beziehungsweise unvollständig sind. Dass es sich um mehr als nur wenige Einzelfälle handelt, zeigen weitere Trefferanalysen. Neben den unterschiedlichen Werken diversen Datums, die unter einer bibliografischen Angabe vereint sind und Digitalisate, die Ungenauigkeiten in der Texterkennung aufweisen, werden auch diejenigen Werke den falschen Zeitabschnitten zugeordnet, die das Datum der Verlags- oder Körperschaftsgründung fälschlicherweise als Erscheinungsjahr aufweisen.³⁵ Die Metadaten, insbesondere die Datierungen der Texte, bilden gleichwohl die Grundlage für die Visualisierung der Zeitskala. Sie können das Ergebnis verfälschen und es somit in seiner Gesamtheit in Frage stellen. Welche Digitalisate aufgrund ihrer geringen OCR-Qualität aussortiert wurden und somit nicht im Graphen erscheinen, bleibt zudem intransparent. Dadurch ist auf die Angaben zum Erscheinen von Wörtern oder Phrasen im Bereich der Visualisierung grundsätzlich kein Verlass.³⁶ Die hohe Fehlerquote ist wohl auf die große Menge an Digitalisierungsdaten, die in einem relativ kurzen Zeitraum erstellt wurde, zurückzuführen. Eine bibliografisch korrekte ›Bearbeitung‹ der Digitalisate ist jedoch die unabdingbare Voraussetzung für Visualisierungen und Auswertungen.

Der Abgleich von Quellen und Graphen im DWDS zeigte, dass die Daten zwar sehr viel zuverlässiger sind. Die Treffer der Suche und die Verlaufskurve bieten allerdings kaum neue Erkenntnisse zum zeitlichen Erscheinen von Wörtern und Phrasen. Die wenigen Einzelbelege liegen innerhalb der zeitlichen Grenzen, die bereits aus der

31 [Information] Google Books Ngram Viewer, <<https://books.google.com/ngrams/info>>, 2.6.2016.

32 Vgl. Lin u. a.: Syntactic Annotations, S. 169.

33 Ebenda.

34 Metadaten sind Angaben über die eigentlichen, zusammengestellten Textdaten, die u. a. Auskunft über die Autorin / den Autor und die Entstehungszeit geben, die mit dem Digitalisat verbunden sind. Siehe auch: Hirschmann: Korpuslinguistik, S. 2; Zum Google Books Ngram Viewer heißt es, »Metadata describing the date and place of publication were provided by the libraries and publishers and supplemented with bibliographic databases.« Michel u. a.: Quantitative Analysis, S. 176.

35 Vgl. Nachrichten der Gesellschaft für Natur- und Völkerkunde Ostasiens/Hamburg: (1937–1939), Ausgaben 42–50 in Google Books, <https://books.google.de/books?id=Gio6A23mSpC&dq=Nachrichten+aus+der+Gesellschaft+/+Deutsche+Gesellschaft+für+Natur-+und+Völkerkunde+Ostasiens&hl=de&source=gbs_navlinks_s>, 11.5.2020. Ein weiteres Problem sind zweiseitige Drucke mit Zeilenumbrüchen, bei denen Wörter automatisiert zusammengefügt werden.

36 Zu den kritischen mathematisch-statistischen Punkten vgl. Pechenick, Danforth, Dodds: Characterizing the Google Books Corpus. Fehler in der Texterkennung und in Bezug auf die Metadaten weist auch Bunčić nach. Bunčić: Über den Nutzen, S. 402.

einschlägigen Literatur bekannt sind. Die Anzeige der Treffer erfolgt überwiegend als auszugsweiser Text, sodass sie kaum im Digitalisat selbst, sondern in den entsprechenden Print-Publikationen überprüft werden müssen. Aufgrund der Ausschnitthaftigkeit der DWDS-Ergebnisse bleibt mitunter offen, in welchem Kontext die Begriffe verwendet wurden.

ALTERNATIVEN UND ENTDECKUNGEN

Dennoch lassen sich neue Erkenntnisse gewinnen, indem statt der Visualisierungen, die das DWDS und der Google Books Ngram Viewer bieten, die Korpora primär als »Belegressource«³⁷ genutzt werden, da sie neue Ansätze und Entdeckungsmöglichkeiten bieten. Der Sprachwissenschaftler Daniel Bunčić sieht in der Google-Books-Suche durchaus einen Vorteil darin, dass man offen über verschiedenste Texte im Volltext hinweg recherchieren kann und nicht von vornherein wissen muss, wo der Begriff nachzuschlagen wäre.³⁸ Eine Suche nach der Phrase »Deutsche(n) im Ausland« führt beispielsweise zum *Universal-Lexikon der Gegenwart und Vergangenheit* von 1841.³⁹ Die Metadaten in Google Books sind hier korrekt angegeben. Das Lemma »Deutsche Sprache« beschäftigt sich mit dem Aufkommen sowie ihrem Sprachgebrauch und gibt eine »Übersicht der deutschen Dialekte« samt Würdigung vom schweizerischen über den schwäbischen und bairischen bis hin zu den Dialekten der Deutschen in Ungarn und Siebenbürgen: »In Siebenbürgen wird sehr gemischt gesprochen«, konstatiert der Autor, »weil aus den verschiedensten Theilen Deutschlands Colonien dahin gingen«.⁴⁰ Dieser Erkenntnis folgt die Einteilung in vier Hauptdialekte in Siebenbürgen: den Hermannstädter, Kronstädter, Bistritzer und bäuerischen Dialekt, die – leicht variierend – generell als »gemischt« und »unverständlich« charakterisiert wurden, während die Liv- und Estländer »unter allen Deutschen im Ausland ihre Sprache am reinsten u. unvermischtesten erhalten haben«.⁴¹ Von Seiten der Sprachwissenschaften zeigt sich somit ein frühzeitiges Interesse für deutschsprachige Bevölkerungsteile, die im Ausland leben; zugleich erfolgt, analog zu Strickers Ausführungen, eine Kategorisierung hinsichtlich sprachlicher »Reinheit« und Homogenität. Vornehmlich in bildungsbürgerlichen Kreisen entstand ein Interesse an deutschsprachigen Bevölkerungsteilen in der Welt und ihrer Einbeziehung in die nationale Idee im Zuge der deutschen Nationalbewegung des 19. Jahrhunderts, die auch im Kontext der europäischen Nationalisierungsprozesse zu verorten ist.⁴² Seit dem Germanistentag zu Frankfurt am Main und den Publikationen Strickers wurde die Bezeichnung »Deutsche im Ausland« häufiger in der Mehrzahl verwendet und umfasste Auswanderer und Auswanderinnen aus deutschen Ländern, die dauerhaft ins Ausland gegangen waren und oftmals eine andere Staatsangehörigkeit angenommen hatten, ebenso wie deutsche Staatsbürger*innen, die Ende des

37 Hirschmann: Korpuslinguistik, S. 14.

38 Vgl. Bunčić: Über den Nutzen, S. 395.

39 Verlinkung auf Belege von 1800 bis 1852 in: Google Books Ngram Viewer: »Deutschen im Ausland« 1800–1910 in German 2012, <https://books.google.com/ngrams/graph?content=Deutschen+im+Ausland&year_start=1800&year_end=1910&corpus=20&smoothing=0&share=&direct_url=t1%3B%2CDeutschen+im+Ausland%3B%2CCo>, 11.5.2020.

40 Lemma: »Deutsche Sprache«, in: Heinrich August Pierer: *Universal-Lexikon der Gegenwart und Vergangenheit* oder neuestes encyclopädisches Wörterbuch der Wissenschaften, Künste und Gewerbe. 8. D–Doit. 2., völlig umgearb. Aufl. Altenburg 1841, S. 262–268, hier: S. 265.

41 Ebenda, S. 266.

42 Vgl. Otto Dann: *Nation und Nationalismus in Deutschland, 1770–1990*. München 1993, S. 15f.

19. Jahrhunderts über einen längeren Zeitraum außerhalb des Deutschen Reiches, etwa in den Kolonien, lebten. Die Identifizierung der Ersteren erfolgte aufgrund ihrer Herkunft aus deutschsprachigen Ländern⁴³ und ihrer Sprach- beziehungsweise Dialektzugehörigkeit.⁴⁴ Die Suchergebnisse in Sprachdaten enthalten beide Auslegungen, müssten bei einer statistischen Auslegung jedoch unterteilt werden.

Ende des 19. und Anfang des 20. Jahrhunderts kamen weitere Begriffsvarianten hinzu, die auf das Konzept verwiesen, wie die kollektivierenden Bezeichnungen und Phrasen: das ›Ausland(s)deutschtum‹ beziehungsweise das ›Deutschtum im Ausland‹, das ›Deutschtum‹ etwa in Ungarn oder in den Vereinigten Staaten⁴⁵ und das ›transatlantische Deutschtum‹,⁴⁶ das explizit auf die europäische Perspektive hindeutet. Neben der Bezeichnung ›Ausland(s)deutsche‹ lassen sich zudem Kolonie-Begriffe⁴⁷ finden; spätestens ab 1917 kamen dem DWDS zufolge in dieser Bedeutung auch ›deutsche Volksgruppen‹⁴⁸ hinzu und etwa ab den frühen 1930er-Jahren die ›Volksdeutschen‹ als nationalsozialistischer Begriff. Letzterer verstärkte sprachlich die bevölkerungs- und machtpolitischen Ansprüche des Deutschen Reiches auf Grundlage der rassistischen Ideologie.⁴⁹

Die ›Deutschen im Ausland‹ standen zunächst für eine Tendenz, aus nationalistischer und völkischer Perspektive Menschen zu kategorisieren, sie der ›eigenen‹ Nation zuzuschreiben und durch sie zunächst sprachlich und mental gewissermaßen Zugriff auf Regionen und Landesteile in der Welt zu erhalten, die der Tradition der ›Kolonialphantasien‹ entsprechen, auf deren Erscheinen schon in vorkolonialer Zeit

43 Statistiken umfassen »im deutschen Sprachgebiet Mitteleuropas Geborene 1900«. Paul Langhans: Die Volkszahl der Deutschen. In: Alfred Geiser (Hg.): Deutsches Reich und Volk. Ein nationales Handbuch (Im Auftrag des Kyffhäuser-Verbandes der Vereine Deutscher Studenten und mit Unterstützung anderer nationaler Verbände). München 1906, S. 215–222, hier: S. 217.

44 Vgl. Naranch: *Inventing the Auslandsdeutsche*, S. 26f.

45 Vgl. Geiser: *Deutsches Reich*, S. VIII.

46 Wilhelm von Polenz: *Das Land der Zukunft oder Was können Deutschland u. Amerika voneinander lernen?* Berlin 1917, S. 327.

47 »Deutsche Kolonien an der Wolga [...] im Schwarzmeergebiet«, in: Otto Boelitz: *Das Grenz- und Auslandsdeutschtum: Seine Geschichte und seine Bedeutung*. München, Berlin 1926, S. VI.

48 Ein erster Beleg zu den deutschen Volksgruppen in Abgrenzung unter anderem zur tschechischen Arbeiterschaft ist laut DWDS in dem Periodikum *Die Grenzboten* 76 (1917) zu finden. Vgl. Korpusrecherche für »deutsche Volksgruppe«, aus dem Korpus *Historische Korpora (1473–1969)* des Digitalen Wörterbuchs der deutschen Sprache, <https://www.dwds.de/r?q=%22deutsche+Volksgruppe%22&corpus=dtaxl&date-start=1473&date-end=1969&subcorpus=dingler&subcorpus=dsdk&subcorpus=dae&subcorpus=dtak&subcorpus=gartenlaube&subcorpus=grenzboten&subcorpus=textberg_hist&format=full&sort=date_asc&limit=100,9.5.2020>. Vgl. zum »Projekt der nationalsozialistischen Volksgruppenbildung« Ingo Haar: *Vom »Volksgruppen-Paradigma« bis zum »Recht auf Heimat«. Exklusion und Inklusion als Deutungsmuster in den Diskursen über Zwangsmigrationen vor und nach 1945*. In: Jerzy Kochanowski, Maike Sach (Hgg.): *Die »Volksdeutschen« in Polen, Frankreich, Ungarn und der Tschechoslowakei. Mythos und Realität*. Osnabrück 2006, S. 17–39.

49 Schmitz-Berning dokumentiert Erwähnungen bereits ab 1930. Vgl. Cornelia Schmitz-Berning: *Vokabular des Nationalsozialismus*. 2., durchges. und überarb. Aufl. Berlin, New York 2007, S. 651. Haar datiert den Paradigmenwechsel von »Grenz- und Auslandsdeutschen« zu »Volksdeutschen« auf Juni 1938. Vgl. Haar: *Vom »Volksgruppen-Paradigma«*, S. 26; *Der Google Book Ngram Viewer zeigt den Begriff in der Bedeutung »Deutsche im Ausland« erstmals in: Albrecht Erich Günther: Was wir vom Nationalsozialismus erwarten: zwanzig Antworten*. Heilbronn 1932. Das DWDS listet das erste Erscheinen 1934, in einem Bericht über die »volksdeutsche Weihestunde unter Beteiligung des VDA und Hans Steinacher im Völkischen Beobachter auf. Vgl. DWDS-Wortverlaufskurve für »volksdeutsche«, erstellt durch das Digitale Wörterbuch der deutschen Sprache, <<https://www.dwds.de/r/plot?view=2&corpus=public&norm=abs&smooth=line&genres=0&grand=1&slice=1&prune=0&window=0&wbase=0&logavg=0&logscale=0&xrange=1600%3A2018&q1=volksdeutsche>>, 28.5.2020.

die Literaturwissenschaftlerin Susanne Zantop aufmerksam gemacht hat.⁵⁰ Mithilfe der sprachlichen Formen entstanden einerseits Vorstellungen von idealisierten Bevölkerungsgruppen, zugleich bildeten diese Zuschreibungen und Bekenntnisse die Grundlage zur Erweiterung der Bevölkerungsstatistiken zugunsten eines vermeintlichen »Deutschtums«.⁵¹

»Ausland(s)deutschtum« als Kollektivsingular suggeriert sprachlich Homogenitätsvorstellungen in mehrfacher Hinsicht. Auf die vermeintliche »Absolutheit« von Kollektivbegriffen wie »Geschichte« oder »Natur« verweist Jürgen Fohrmann.⁵² Sie scheint allumfassend zu sein und auch die »Gegenbegriffe« einzubeziehen, die sich zugleich durch Unterscheidungen ergeben. In Bezug auf »Ausland(s)deutschtum« ist eine solche Absolutheit insoweit erkennbar, als begrifflich von »einem Deutschtum« ausgegangen wurde, das im Deutschen Reich bestünde, sich durch Auswanderungen und Migrationsprozesse außerhalb des Reiches verbreite und weiter existieren würde. Zudem bezog der Begriff das »Nichtdeutschtum« im Ausland mit ein, ohne es als solches zu definieren, und fasste es ebenfalls – abgegrenzt und mitunter in seiner Gesamtheit als »fremd« und »feindlich« erklärt – als Einheit auf.

Der Oberbegriff erscheint in den Titeln zahlreicher Handbücher und Publikationen; in den Inhaltsverzeichnissen finden sich dann differenzierte Untergliederungen. Die einzelnen Unterbegriffe wie Siebenbürger Sachsen, Karpatendeutsche, Bessarabiendeutsche oder Donauschwaben verknüpfen das vermeintliche Deutschtum mit einzelnen Ländern und Regionen, in denen sich offenbar trotz national-homogener Vorstellung eine vermeintliche Spezifik des Deutschtums aufgrund geopolitischer »äußerer« Bedingungen entwickelt hätte. Die Bezeichnungen suggerieren weitere Kollektive aus Sicht der Akteure im Deutschen Reich. Dass eine solche Differenzierung in den kollektivierenden Begrifflichkeiten bezüglich der so genannten Deutschen im Ausland vorgenommen wurde, mag unter anderem daran liegen, dass die Thematik in den Unterricht einbezogen werden und gar ein universitäres Fach mit eigener Professur entstehen sollte.⁵³ Einflussreiche Akteure, die publizistisch gewissermaßen zur »Gliederung des Stoffes« beitrugen, waren Otto Boelitz (1876–1951), der – ursprünglich Lehrer – in den 1920er-Jahren als preußischer Kultusminister wirkte, sowie der Pädagoge und Schriftsteller Gottfried Fittbogen (1878–1941), dessen Interesse an den »Deutschen im Ausland« offenbar durch Reisen nach Siebenbürgen sowie ins Banat geweckt worden war.⁵⁴

50 Susanne Zantop hat herausgearbeitet, wie sich in verschiedenen Textsorten seit dem 18. Jahrhundert in Deutschland eine Art »kolonialistische Geisteshaltung« herauslesen lässt, indem sie in ihrer Analyse mehrdeutige Kolonialphantasien in den Blick nimmt. Susanne Zantop: *Kolonialphantasien im vorkolonialen Deutschland (1770–1870)*. Berlin 1999, S. 11.

51 Vgl. Wilhelm Winkler: *Statistisches Handbuch des gesamten Deutschtums*. Berlin 1927.

52 Die Kollektivbegriffe »Geschichte« oder »Natur« seien »absolut [...], weil ihre Extension alles zu umfassen sucht und weil sie damit alle Gegenbegriffe integrieren.« Jürgen Fohrmann: *Das andere der »Kultur«: Die »Kultur« der Kulturwissenschaften*. In: Georg Mein, Markus Rieger-Ladich (Hgg.): *Soziale Räume und kulturelle Praktiken. Über den strategischen Gebrauch von Medien (Kultur- und Medientheorie)*. Bielefeld 2004, S. 85–100, hier: S. 85f. Zu Kollektivsingularen vgl. Koselleck: *Begriffsgeschichten*, S. 67, S. 111.

53 Vgl. Anmerkungen zur Wissenschaft bei Gottfried Fittbogen (Hg.): *Was jeder Deutsche vom Grenz- und Auslandsdeutschum wissen muß*. München, Berlin 1924, S. 3.

54 Hans-Werner Retterath nimmt an, dass Fittbogens Südosteuropa-Reisen den Ausgangspunkt für sein Interesse am »Auslandsdeutschum« bildeten. Vgl. Hans-Werner Retterath: *Gottfried Fittbogen*. In: Michael Fahlbusch, Ingo Haar, Alexander Pinwinkler (Hgg.): *Handbuch der völkischen Wissenschaften. Akteure, Netzwerke, Forschungsprogramme*. Band 1. 2., grundl. erw. und überarb. Aufl. München 2017, S. 163–167, hier: S. 163; Fittbogen publizierte über das Banat und Ludwig Roth auch in den *Südost-Forschungen*. Einige seiner Beiträge sind gelistet unter: Dieter Rübsamen (Bearb.): *Südost-Forschungen* 1, 1936–20, 1961 –

Die Vielfalt der Bezeichnungen hat Vor- und Nachteile: Einerseits können sich zahlreiche Treffer für einen Begriff ergeben, die aber möglicherweise nicht der gesuchten Bedeutung entsprechen. Andererseits entgehen wichtige Quellen und Publikationen der Suchanfrage, da in ihnen die Thematik mit anderen Worten umschrieben wurde und ihre Zeichenketten nicht den Suchbegriffen entsprechen. Zuweilen, so eine Anmerkung von Bunčić, erweist sich die »Polysemie und Homonymie des untersuchten Begriffs«⁵⁵ als schwierig, da die Anzahl an Gesamt-Treffern sehr viel höher ist als die tatsächlich relevante Trefferzahl. So taucht etwa das Begriffspaar »deutsche Volksgruppe« in einem anderen Bedeutungszusammenhang schon 1862 auf. Der Kriminalist Friedrich Christian Benedikt Avé-Lallemant (1809–1892) verwendete es für ein Kapitel *Die Sprache deutscher Volksgruppen* seines umstrittenen linguistischen Abschnitts in seinem Werk *Das Deutsche Gaunerthum in seiner social-politischen, literarischen und linguistischen Ausbildung zu seinem heutigen Bestande*.⁵⁶ Avé-Lallemant ging auch auf die Formulierung »volksdeutsch« ein – wohl im Sinne Ernst Moritz Arndts als »gute, volkstümliche, deutsche Sprache«⁵⁷ –, in Abgrenzung zur jüdisch-deutschen Sprache. Er sah im »Volksdeutschen« die ursprüngliche, »originale« Sprache, die seiner Meinung nach fast unmerklich durch »exotische« Formen geprägt werde.⁵⁸

Der Bedeutungswandel oder auch Mehrdeutigkeiten bleiben von der automatisierten Texterkennung und Suche unerkannt. Diese Tatsache kann jedoch unter Umständen die Perspektive erweitern, wie ein weiteres Fundbeispiel mit Bezug zu Südosteuropa zeigt. »[Z]iemlich ungenaue Sammelbezeichnungen« nennt Gerhard Seewann die Gruppenbezeichnungen »Siebenbürger Sachsen« und »Donauschwaben« – und letztere noch dazu »historisch unzutreffend« und unglücklich.⁵⁹ Es handelte sich um eine Kollektivbezeichnung für die »deutschen Bevölkerungsteile, die seit Trianon (1920) politisch-administrativ auf die Nachfolgestaaten der Habsburgermonarchie aufgesplittert waren«.⁶⁰ Die Geschichte des Begriffes »Donauschwaben« haben Hans-Christian Petersen und Tobias Weger detailliert recherchiert und die Ersterwähnungen auf 1921/22 datiert; eingeführt wurde er von Robert Sieger (1864–1926), Hermann Rüdiger (1889–1946) und wohl Gerhard Gesemann (1888–1948).⁶¹ Dem Textkorpus des DWDS zufolge erschien die Bezeichnung erst ab 1930, wie Abb. 1 zeigt. Es fehlen also zentrale Texte, die historische Verweise auf den Begriff enthalten.

Doch bereits 1834 findet sich die Bezeichnung »Donauschwaben« laut Google Books in den *Jahrbüchern der Literatur* und zwar in einer Urkundenbesprechung zur Geschichte Obersachsens (mit korrekt angegebenen Metadaten). Die »heutigen

Magazine Stacks, <<https://magazinestacks.fordham.edu/mvst/sof.html>>, 9.5.2020. Vgl. Fittbogen (Hg.): Was jeder Deutsche.

55 Bunčić: Über den Nutzen, S. 402.

56 Friedrich Christian Benedikt Avé-Lallemant: *Das Deutsche Gaunerthum in seiner social-politischen, literarischen und linguistischen Ausbildung zu seinem heutigen Bestande*. Leipzig 1862, S. 91.

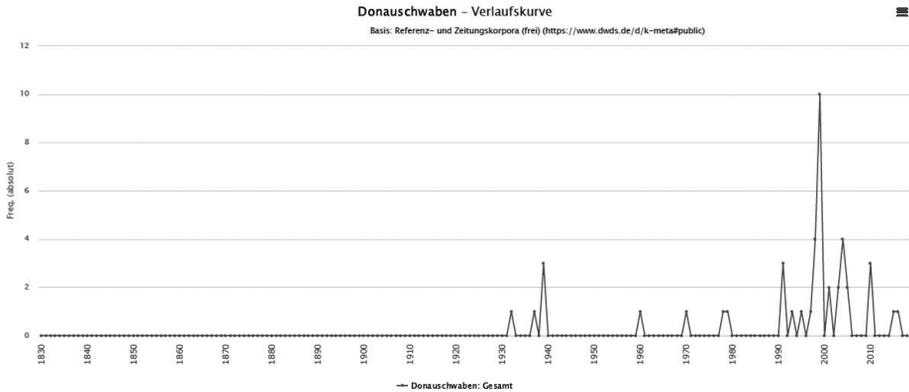
57 Schmitz-Berning: *Vokabular des Nationalsozialismus*, S. 650.

58 Avé-Lallemant: *Das Deutsche Gaunerthum*, S. 265.

59 Gerhard Seewann: Siebenbürger Sachse, Ungarndeutscher, Donauschwabe. Überlegungen zur Identitätsproblematik des Deutschtums in Südosteuropa. In: ders. (Hg.): *Minderheitenfragen in Südosteuropa. Beiträge der Internationalen Konferenz: The Minority Question in Historical Perspective 1900–1990*, Inter University Center Dubrovnik, 8.–14. April 1991. München 1992, S. 139–153, hier: S. 140f.

60 Ebenda, S. 141.

61 Vgl. Hans-Christian Petersen, Tobias Weger: Neue Begriffe, alte Eindeutigkeiten? Zur Konstruktion von »deutschen Volksgruppen« im östlichen Europa. In: *Jahrbuch des Bundesinstituts für Kultur und Geschichte der Deutschen im östlichen Europa* 25 (2017), S. 177–198, hier: S. 193f.



Digitales Wörterbuch der deutschen Sprache: DWDS-Wortverlaufskurve für »Donauschwaben«, <https://www.dwds.de/r/plot?view=2&corpus=public&norm=abs&smooth=line&genres=0&grand=1&slice=1&prune=0&window=0&wbase=0&logavg=0&logscale=0&xrange=1830%3A2018&q1=Donauschwaben>, 23.5.2020.

Donauschwaben«, heißt es dort, seien nicht zu verwechseln mit den Schwaben, die verschiedene Fürstenhäuser (unter anderem von Anhalt und Brandenburg) stellten.⁶²

Eine weitere Erwähnung, unter der Bezeichnung »Donau-Schwaben«, bezieht sich im Jahr 1838 auf die mögliche Benennung eines von acht Kreisen in Bayern, hier unter Bezugnahme auf die »neue bayerischen Reichs-Haupt-Eintheilung« von 1837.⁶³ Und auch Wilhelm Heinrich Riehl (1823–1897) erwähnt die Bezeichnung in seiner volkswissenschaftlichen Arbeit zu den Pfälzern, genau genommen in seiner Beschreibung der pfälzischen Küche. Er verweist ebenfalls auf die bayerischen Donauschwaben, deren »Röggel«, eine »Form des Weißbrodes«, rund und mit einem Kreuz verziert seien.⁶⁴ Anhand der Beispiele kommt auch zum Ausdruck, wie Seewann feststellt, dass die »Donau ein zu langer Fluß« ist,⁶⁵ um sie für eine geografische Abgrenzung in Betracht zu ziehen. Relativität und Perspektive bestimmen die Nutzung des Flusses als Charakterisierung. So ist ein Abschnitt der Donau für die bayerischen Regionen maßgebend gewesen, und möglicherweise bot dieser sich an, um einen engeren Bezug der Bevölkerungsteile im historischen Ungarn zum Deutschen Reich herzustellen. Allerdings ist nicht klar, ob die genannten »Erfinder« des noch heute verwendeten Donauschwaben-Begriffes ihn in seiner vorherigen Bedeutung überhaupt kannten. Wie verbreitet die bayerische Bezeichnung »Donauschwaben« vor den 1920er-Jahren war, lässt sich aus den Suchergebnissen nicht ohne Weiteres schließen. Für beide Suchsysteme gilt, dass aus den Ergebnissen kaum Rückschlüsse auf den tatsächlichen Sprachgebrauch im angegebenen Zeitraum erfolgen können. Die Häufigkeit des Erscheinens von Wörtern oder Redewendungen kann zwar visualisiert werden; eine

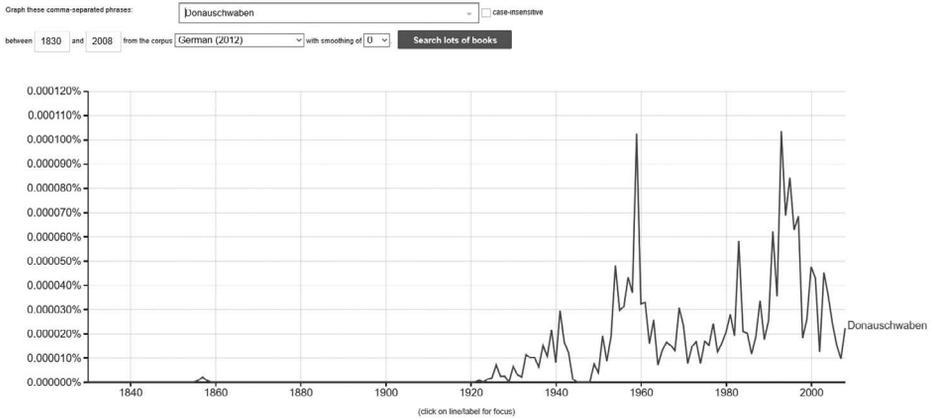
62 Besprechung des Directorium diplomaticum. oder chronologisch geordnete Auszüge von sämmtlichen über die Geschichte Obersachsens vorhandenen Urkunden (bearb. von Ludwig August Schultes). In: Jahrbücher der Literatur 68, Oktober, November, Dezember (1834), S. 214–232, hier: S. 224.

63 Vgl. Ueber Länder- und Staaten Ein- und Abtheilungen überhaupt und die neueste Reichs-Eintheilung Bayerns insbesondere sammt unmassgeblichen Bemerkungen über die Nothwendigkeit einer Reform und neuen Formation der Landgerichts-Bezirke Bayerns. Erlangen 1838, S. 56.

64 Wilhelm Heinrich Riehl: Die Pfälzer. Ein rheinisches Volksbild. Stuttgart, Augsburg 1857, S. 253. Weitere Erwähnungen Donauschwabens als Region um Ingolstadt finden sich in Johann Georg Kohl: Die Donau von ihrem Ursprunge bis Pesth. Triest 1854, S. 24.

65 Seewann: Siebenbürger Sachse, S. 140.

Google Books Ngram Viewer



Google Books Ngram Viewer: >Donauschwaben< 1830–2012 in German 2012, <https://books.google.com/ngrams/graph?content=Donauschwaben&year_start=1830&year_end=2008&corpus=20&smoothing=0&share=&direct_url=t1%3B%2CDonauschwaben%3B%2CCo>, 23.5.2020.

Aussage über ihre Popularität ist, zumindest mit dem Google Books Ngram Viewer, jedoch kaum möglich. Pechenick et al. (2015) sprechen in ihrem kritischen Beitrag zur Nutzung des englischsprachigen Bereichs der Google Books von einer »illusion of large-scale cultural popularity«, die vermittelt werde.⁶⁶ Sie weisen darauf hin, dass der Bestand nicht als Textkorpus zu charakterisieren sei, sondern eher als lexikalische Datenquelle, denn die Werke verschiedener Autoren und Autorinnen werden ungeachtet ihrer Rezeption in gleicher Weise erfasst und nicht gewichtet. Darüber hinaus zeigen sie anhand statistischer Untersuchungen, dass der Anteil wissenschaftlicher Literatur seit den 1980er-Jahren nachweislich zugenommen hat, und charakterisieren die Textauswahl als voreingenommen. Über die tatsächliche Nutzung der Wörter im Sprachgebrauch oder ihre Rezeption gäben diese Such- und Visualisierungsergebnisse daher kaum Aufschluss und sollten als Datenbestände für kulturgeschichtliche Fragestellungen nur mit Vorsicht in Betracht gezogen werden. In Bezug auf die zukünftige Korpora-Erstellung für den Google Books Ngram Viewer empfehlen Pechenick et al. deshalb, eine Unterscheidung zwischen wissenschaftlicher und populärer Literatur vorzunehmen.⁶⁷ Im DWDS ist eine solche Möglichkeit der Differenzierung umgesetzt. Die Verlaufskurve lässt sich nach Textsorten (Wissenschaft, Belletristik, Gebrauchsliteratur und Zeitung) getrennt darstellen. Generell gilt allerdings, dass die Repräsentativität der Korpora und damit auch der Suchergebnisse je nach Forschungsfrage unterschiedlich einzuschätzen ist. Die Ergebnisse sind abhängig vom Zeitpunkt der Suchanfragen und dem jeweiligen Stand der zugänglichen Textquellen beziehungsweise Korpora. Somit ist von einer Dynamik im Bereich der Texte, die zur Verfügung stehen, auszugehen. Die Abfragen sollten deshalb entsprechend zeitlich dokumentiert und zumindest stichprobenartig geprüft werden.

Die Auswahl an Texten und die mit beiden Projekten, DWDS und Google Books Ngram Viewer, vorgenommene Bestimmung der jeweiligen Sprache, in der die

⁶⁶ Pechenick, Danforth, Dodds: Characterizing the Google Books Corpus.

⁶⁷ Ebenda.