

Hauke Bartels

Sorbisch lernen digital: Sprachressourcen im Internet Zehn Jahre niedersorbisch.de als Sprachportal

Einleitung

Die Sorben sind eine autochthone nationale Minderheit in Deutschland. Das Sorbische bezeichnet historisch die Gesamtheit aller westslawisch-sorbischen Dialekte, die im einst weit größeren Siedlungsgebiet der Sorben Verwendung fanden. Heute steht es auch als Oberbegriff für die sorbischen Schriftsprachen Niedersorbisch und Obersorbisch (sowie weitere Varietäten). Beide Einzelsprachen sind in ihrer Existenz bedroht, das Niedersorbische deutlich stärker als das Obersorbische.¹ Mit Blick unter anderem auf den Minderheitenstatus der Sorben, die Größe (d. h. die relative „Kleinheit“) der beiden sorbischen Sprachgemeinschaften und den soziolinguistischen Status des Ober- bzw. Niedersorbischen ist häufig – je nach Perspektive – auch von Minderheitensprachen oder Kleinsprachen die Rede. Darin kommt über die Bestandteile „Klein-“ bzw. „Minderheiten-“ ein Aspekt zum Tragen, der gerade im Kontext der Digitalisierung von Bedeutung ist: Kleinheit und Minderheitenstatus als Ursache für einen Mangel an Informationen und Ressourcen. In BARTELS (2015: 24; 31) wurde dies mit Blick auf das Niedersorbische als „Kleinsprachen-Dilemma“ beschrieben:

Das Niedersorbische wurde von einem Großteil der Sprecher mittlerer und jüngerer Generation nicht mehr als Erstsprache (Muttersprache) erlernt. Der aktive Wortschatz ist oft thematisch stark beschränkt und Unsicherheiten im eigenen Sprachgebrauch sind weit verbreitet. Der Bedarf an Sprachberatung und zuverlässigen Hilfsmitteln wie Wörterbüchern ist daher hoch.

In bedrohten Kleinsprachen wie dem Niedersorbischen steht diesem erhöhten Beratungsbedarf jedoch ein Mangel an Informationsquellen gegenüber: eine abnehmende Anzahl zuverlässiger Informanten, wenig Drucktexte, die geringe Häufigkeit geeigneter Kommunikationssituationen, d. h. ein insgesamt unbefriedigender sprachlicher „Input“ als Korrektiv für den eigenen Sprachgebrauch. Und auch die Datengrundlage für sprachwissenschaftliche Forschung ist im Vergleich zu „großen“ Sprachen eher klein.

Die einzige Möglichkeit, diesem „Kleinsprachen-Dilemma“ zu begegnen, ist ein konsequentes und systematisches Sammeln aller verfügbaren Informationen sowie eine effektive Aufbereitung und Vernetzung derselben, um ihre breite Nutzbarkeit zu ermöglichen. [Im Original als niedersorbischer und deutscher Text.]

Der skizzierte Ausweg aus dem Dilemma lässt sich nur mit den Möglichkeiten der Digitalisierung gehen. Die folgenden Abschnitte zeigen, welche Schritte in der vergangenen Dekade bereits umgesetzt und welche Möglichkeiten dabei genutzt wurden. Allein für das Niedersorbische wird es mindestens zehn weitere Jahre brauchen, um einen optima-

¹ Mehr Informationen finden sich in MENZEL/POHONTSCH (2020).

len Stand der Aufbereitung, Vernetzung und Zugänglichmachung der wichtigsten sprachlichen Informationen zu erreichen. Insofern wird hier von einem Zwischenstand berichtet. Das heutige Sprachportal niedersorbisch.de umfasst sehr unterschiedliche Ressourcen, die nicht nur der Sprachförderung, sondern auch der Sprachdokumentation dienen. Im Folgenden liegt der Schwerpunkt aber auf den Angeboten, die sich auch zum autodidaktischen Lernen eignen. Außerdem geht es – auch mit Blick auf das im Untertitel genannte Jubiläum – vor allem um das Niedersorbische. Weiter unten findet sich dennoch ein kurzer Abschnitt über entsprechende Angebote zum Obersorbischen.

Digitalisierung als Teil der Lösung: Das Sprachportal niedersorbisch.de

Die *Internetseite* niedersorbisch.de (alternativ: doloserski.de)² ging im Jahr 2003 online und diente zunächst allein zur Veröffentlichung einer ersten Version des aktiven Deutsch-Niedersorbischen Online-Wörterbuchs (DNW).³ Diese umfasste gerade einmal 800 Artikel und wurde in den Folgejahren beständig ausgebaut. Grundlage dafür war ein auf traditionelle Weise erstelltes WORD-Manuskript, das seit 2001 hauptsächlich vom sorbischen Lexikografen Manfred Starosta erarbeitet⁴ und stetig erweitert wurde. Diese Form der schrittweisen aber schnellen Veröffentlichung bei laufender Bearbeitung und Ergänzung war mit Blick auf die Sprachsituation wichtig und nur im digitalen Medium möglich. (Mehr zum DNW siehe den folgenden Abschnitt.)

Ein wichtiger Schritt zum Ausbau der Internetseite zu einem *Sprachportal* wurde 2012 getan, sodass das Jahr 2022 ein 10-jähriges Jubiläum markiert. Bereits einige Jahre zuvor war mit der Retrodigitalisierung der vier wichtigsten niedersorbisch-deutschen Wörterbücher begonnen worden, und vor nunmehr zehn Jahren konnte als Ergebnis eine einheitliche Internetversion dieser Wörterbücher veröffentlicht werden. Seitdem ist die bis dahin in gedruckten Wörterbüchern erfasste niedersorbische Lexik über einen zentralen Zugang und ebenso leicht wie gut recherchierbar frei zugänglich. Und was für die weitere Entwicklung von Sprachressourcen wichtig ist: in vollständig computerlesbarer Form.

Im Folgenden werden diese beiden sowie weitere, in den Folgejahren ergänzte Sprachressourcen⁵ kurz vorgestellt, mit dem bereits erwähnten Fokus auf ihrem Nutzen für das autodidaktische Sprachenlernen. Außerdem sollte in der Darstellung deutlich werden, welche Möglichkeiten die Digitalisierung bietet, ohne dass hier alle relevanten Aspekte dieser Teilthematik angesprochen werden können. Digitalisierung wird aber als Teil der Lösung bestimmter Probleme von Kleinsprachen betrachtet.

² Siehe <https://www.doloserski.de/> oder <https://www.niedersorbisch.de/>.

³ Siehe <https://www.doloserski.de/dnw/>.

⁴ Ein weiterer muttersprachlicher Ko-Autor war Erwin Hannusch. Mehr Informationen finden sich auf der Internetseite des DNW.

⁵ Nahezu alle im Artikel erwähnten Ressourcen wurden von Projektgruppen, d. h. von mehreren Personen erarbeitet. Dies wird auf den jeweiligen Portalseiten dargestellt. Dort finden sich auch die Hinweise auf finanzielle Förderung der einzelnen Vorhaben. – Von der weiter unten abgebildeten Einstiegseite zum Portal (Screenshot) abgesehen, wird im Folgenden bei der Darstellung der einzelnen Portalangebote auf eine bildliche Illustration verzichtet, da alle Beispiele im Internet nachvollzogen werden können.

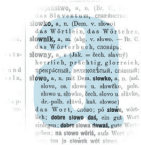
NIEDERSORBISCH.DE

Die Cottbuser Zweigstelle des Sorbischen Instituts stellt an dieser Stelle verschiedene seit 2003 erarbeitete niedersorbische Sprachressourcen bereit. Die Seite bietet damit umfassende Informationen zum Niedersorbischen.



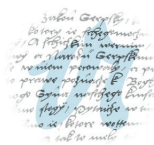
DEUTSCH-NIEDERSORBISCHES WÖRTERBUCH

Aktives Wörterbuch mit mehr als 80000 Artikel (in ständigem Ausbau)



NIEDERSORBISCH-DEUTSCHE WÖRTERBÜCHER

Vier retrodigitalisierte Wörterbücher mit einheitlichem Zugang



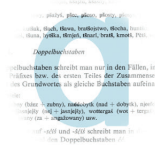
NIEDERSORBISCHES TEXTKORPUS

Suche in niedersorbischen Texten



MUTTERSPRACHLICHES NIEDERSORBISCH

Mehr als 100 Stunden Tonaufnahmen des dialektalen Niedersorbischen samt orthografischer Transkription



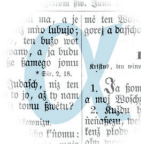
NIEDERSORBISCHE RECHTSCHREIBUNG

Aktuelle Regelungen und ein Modul zur automatischen Rechtschreibkontrolle



NIEDERSORBISCHE REDEWENDUNGEN UND SPRICHWÖRTER

Sammlung von Phraseologismen aus diversen Wörterbüchern, bearbeitet und angereichert



NIEDERSORBISCHE BIBEL VON 1868

Digitale Ausgabe in originaler und modernisierter Schreibung, mit Suchfunktionen



NIEDERSORBISCHE AUSSPRACHE

Thematisch geordnete und zum Selbstlernen geeignete Darstellung phonetischer Besonderheiten



NIEDERSORBISCHE NAMEN

Sprach- und sachbezogene Informationen zu Personen- und Ortsnamen



NIEDERSORBISCHE DIGITALE BIBLIOTHEK

Digitale Leseanfänge historischer nieder-sorbischer Drucke



Serbski Sorbisches
institut Institut

IMPRESSUM, DATENSCHUTZ

Screenshot des Portals niedersorbisch.de (deutschsprachige Version) [29.06.2022].

Leider gibt es bisher keine wissenschaftlichen Untersuchungen zum Ausmaß und zur Bedeutung des autodidaktischen Lernens des Nieder- oder Obersorbischen. Vor allem beim Niedersorbischen, wo nur sehr wenige Möglichkeiten des Spracherwerbs außerhalb des Bildungssystems (und damit jenseits des gesteuerten Spracherwerbs) existieren, besteht aber der deutliche Eindruck, dass ein bei Lernenden leider nur in Ausnahmefällen festzustellendes hohes Kompetenzniveau sehr stark vom persönlichen Engagement und von der Bereitschaft zur individuellen Weiterbildung abhängt. Daher ist anzunehmen, dass zumindest für diesen Personenkreis, der in der Lage und willens ist, die eigene Sprachkompetenz selbständig zu erweitern, die verschiedenen Informationsangebote des Sprachportals von großer Relevanz sind. Neben den Inhalten ist dabei auch die fast durchgängige (deutsch-niedersorbische) Zweisprachigkeit des gesamten Angebots von Bedeutung. Durch die deutsche Benutzerführung werden des Niedersorbischen (noch nicht hinreichend) mächtige Nutzende einbezogen. Die niedersorbische

Variante schafft gewissermaßen einen zusätzlichen rezeptiven Sprachraum, in dem zumindest das Leseverstehen auch von Texten mit nicht alltäglicher Lexik gefördert wird.⁶

Das Deutsch-Niedersorbische Wörterbuch (DNW)

Angesichts der bereits erwähnten starken Gefährdung der niedersorbischen Sprache war und ist der Bedarf nach einem aktiven Wörterbuch,⁷ das ausreichende Informationen zur korrekten Sprachproduktion enthält, sehr groß. Alle Bemühungen um eine Revitalisierung des Niedersorbischen werden langfristig ins Leere laufen, wenn es nicht gelingt, den Wortschatz der Sprache zu dokumentieren und Lernenden auf geeignete Weise zur Verfügung zu stellen. Daher wurde um die Jahrtausendwende mit den Arbeiten an einem solchen Wörterbuch begonnen. Sie zielten zunächst auf ein aktuelles, dem Umfang nach überschaubares deutsch-niedersorbisches Schulwörterbuch, das auf traditionelle Weise im Druck erscheinen sollte.⁸

Allerdings sprachen mehrere Gründe zum einen für einen Wechsel ins Internet und damit für die Erarbeitung eines Online-Wörterbuchs, zum anderen für die Aufgabe der ursprünglichen Beschränkungen: (a) Eine digitale Erscheinungsweise über das Internet bietet erheblich bessere Darstellungsmöglichkeiten und längerfristige Ausbauoptionen; sie beseitigt zudem (b) ein für Druckpublikationen stets bestehendes Raumproblem und ermöglicht so die Platzierung zusätzlicher Informationen (etwa zahlreicher Beispielsätze). Dieser neue Entwicklungsrahmen erlaubte es überdies, (c) möglichst alle verfügbaren Informationen zu sammeln und soweit wie möglich einzubeziehen – auf diese Weise konnte vor allem das in den Anfangsjahren noch greifbare Sprachwissen älterer Muttersprachler genutzt werden. Und schließlich wurde wegen des schon zu Beginn der Bearbeitung dringenden Bedarfs (d) eine sukzessive Erscheinungsweise gewählt, die bereits erarbeitete Informationen möglichst schnell an die Öffentlichkeit brachte.

Auf diese Weise entstand über die Jahre das bisher umfangreichste sorbische Wörterbuch überhaupt, das eine herausragende Rolle als Instrument der Sprachförderung spielt.⁹ Es wird auch in den kommenden Jahren weiter bearbeitet, ergänzt sowie funkti-

⁶ Als weiteres wichtiges Angebot neben den Ressourcen des niedersorbischen Sprachportals gibt es einen Internet-Sprachkurs: Sorbisch online lernen (2015–2022). Dieser umfasst Obersorbisch und Niedersorbisch und richtet sich ausdrücklich an autodidaktisch Lernende. Für Obersorbisch liegt außerdem ein Lehrbuch vor, das sich ebenfalls speziell an Selbstlerner richtet und dem eine CD mit vertonten Texten und Übungen beiliegt: ŠOLČINA/WORNAR (2012).

⁷ Der Unterschied zwischen einem tatsächlich als „aktives“ Wörterbuch konzipierten lexikalischen Nachschlagewerk und traditionellen zweisprachigen Wörterbüchern kann illustriert werden durch einen direkten Vergleich zwischen dem Artikel *erfüllen* aus dem 1990 erschienenen „Deutsch-niedersorbischen Schulwörterbuch“ (JANNASCH 1990) und dem entsprechenden Artikel des DNW. Im o. g. Schulwörterbuch, das sich gleichwohl als „Handwörterbuch für Lernende und Schreibende“ verstand, sieht der vollständige Artikel wie folgt aus: „erfüllen spoŋi|š 79 p, dopoŋi|š 79 p, spoŋj|owaš 62, dopoŋj|owaš 62; sich ~ se dopoŋi|š 79 p, se dopoŋj|owaš 62“. Eine einfache Auflistung (in bestimmten Kontexten mehr oder weniger) möglicher Äquivalente bietet aber gerade Lernenden keine hinreichenden Informationen für den aktiven Sprachgebrauch.

⁸ Ein solches erschien auf Grundlage des erweiterten DNW, siehe DNSW (2012).

⁹ Eine genauere Beschreibung des DNW kann an dieser Stelle nicht erfolgen. Einen groben Einblick in die zugrundeliegende Konzeption bietet BARTELS (2010).

onal erweitert: So sollen demnächst die noch aus der Erstentwicklung von 2003 stammenden einfachen Suchmöglichkeiten über das Deutsche, die sich im DNW selbst noch auf die deutschen Lemmata beschränken, erweitert werden. Bereits jetzt gibt es einen zusätzlichen und komfortablen Zugang zu den im DNW enthaltenen niedersorbischen Wörtern über eine zentrale sorbische Suche (siehe den Abschnitt zur Orthografie). Zwischen 2015 und 2017 wurde es zudem um Tonaufnahmen zu zahlreichen niedersorbischen Übersetzungsäquivalenten ergänzt (siehe den Abschnitt zu Aussprache und Orthoepie).

Die typische Aufgabe, die mit Hilfe des DNW lösbar sein soll, ist eine sprachlich angemessene Übertragung eines deutschen Ausdrucks oder einer auf Deutsch formulierten Äußerung ins Niedersorbische. Die grundlegende Kenntnis der grammatischen Strukturen des Niedersorbischen vorausgesetzt, soll das DNW für den lexikalischen Bereich die nötigen Informationen bereitstellen. Dieses Ziel ist hochgesteckt und wird sich nicht vollumfänglich und nicht in allen Artikeln gleich gut erreichen lassen; sich ihm möglichst weit anzunähern bestimmt aber die Arbeit an diesem Wörterbuch.

Niedersorbisch-deutsche Wörterbücher

Wie bereits erwähnt war ein entscheidender Schritt zum Ausbau der Internetseite niedersorbisch.de in Richtung Sprachportal die Veröffentlichung einer einheitlichen Nutzerschnittstelle für vier retrodigitalisierte niedersorbisch-deutsche Wörterbücher.¹⁰ In den Jahren zuvor waren diese in einem mehrstufigen Verfahren digitalisiert, analysiert und feingranular in ein einheitliches XML-Format gebracht worden (SZCZEPAŃSKI 2012). Diese zweite lexikalische Ressource auf dem Portal bietet einen einfachen Zugang zu reichhaltigen Informationen aus mehr als 150 Jahren lexikografischer Tradition.¹¹

Die Suche ist sowohl in einzelnen als auch in allen Wörterbüchern gleichzeitig möglich. Wegen der in ihnen enthaltenen historischen Schreibungen und ggf. nur ungenauer Kenntnisse davon kann beim Suchen auf Diakritika verzichtet werden. Zusätzlich ist es möglich, über die Suche nach aktuellen Schreibungen (Option „modernisiert“) auch Einträge in alter Schreibweise zu finden. Ebenso kann nach Wortteilen, Wortanfängen usw. gesucht werden. Für die gezielte Suche nach bekannten historischen Schreibungen gibt es eine Eingabehilfe für „Sonderzeichen“.

Unter den vier Wörterbüchern nimmt das von STAROSTA (1999) eine Sonderstellung ein, da es sich hierbei um das in dieser Sprachrichtung aktuellste Nachschlagewerk handelt, das zudem offiziell als Lehr- und Hilfsmittel für den Unterricht und für Prüfungen an Schulen zugelassen ist. Bei Suchanfragen über alle Wörterbücher (als Standardfall) werden die Ergebnisse aus diesem Werk an erster Stelle angezeigt. Mit Blick auf den Sonderstatus dieses Wörterbuchs wurden einige Jahre später auch nur hier wichtige, seit dem Erscheinen vor mehr als zwanzig Jahren beschlossene orthografische Änderungen eingearbeitet. So wird zum Beispiel im Artikel *pjakaŕ* (Bäcker) die neue Schreibweise der dort ebenfalls erwähnten movierten Bildung *pjakaŕka* (Bäckerin) als „aktuelle

¹⁰ Es handelt sich um ZWAHR (1847), MUCKE (1911–1928), ŠWJELA (1961) und STAROSTA (1999).

¹¹ Zugleich ist hier ein wichtiger vorbereitender Schritt für die historische Lexikografie des Niedersorbischen getan (siehe BARTELS 2013, 2022).

Schreibweise“ angeführt (alte Schreibweise: *pjakarka* mit <r> statt <ř>).¹² Allerdings wurde diese (nicht vollumfängliche) Aktualisierung so vorgenommen, dass die Originalinformationen erhalten bleiben, da die gemeinsame Seite der vier niedersorbisch-deutschen Wörterbücher ja zugleich der Dokumentation der bisherigen lexikografischen Tradition dient.

Unter didaktischen Gesichtspunkten handelt es sich bei diesen vier Wörterbüchern primär um Nachschlagewerke zum Textverstehen, auch von historischen niedersorbischen Texten. Für sprachgeschichtlich Interessierte und fortgeschrittene Lernende bieten die in den Werken enthaltenen Informationen aber auch eine Möglichkeit zur Erweiterung des eigenen Wortschatzes, etwa zum Zweck einer größeren stilistischen Variation. Dies setzt jedoch eine hohe Bewertungskompetenz mit Blick auf die besonders in den drei ältesten Wörterbüchern oft veralteten Angaben voraus. Daher ist allen Artikeln der Hinweis beigefügt. „Es handelt sich um ein älteres Wörterbuch; die Informationen können teilweise veraltet sein.“

Es liegt hier generell eine gewisse Gefahr der Retrodigitalisierung historischer Wissensbestände, wenn diese nicht hinreichend explizit gekennzeichnet sind. Schon eine moderne Präsentationsform verschleiert die zeitliche Distanz zwischen Entstehungs- und Nutzungszeit. Bei den zugrundeliegenden Druckwerken rufen dagegen allein die physische Gestalt oder das Druckbild (etwa Frakturschrift) sofort zur Vorsicht auf. Noch problematischer kann sich eine lexikografisch unreflektierte automatische Verarbeitung von Wörterbuchinformationen auswirken, die sich in manchen stark informationstechnisch dominierten Projekten findet, so etwa, wenn die sprachlich gerichtete (Ausgangssprache-Zielsprache) Informationsanordnung aus zweisprachigen Wörterbüchern einfach „umgedreht“ wird, ohne diesen Verarbeitungsschritt zu kennzeichnen und vor seinen Folgen gerade auch für Sprachlernende zu warnen. Insofern erfordert die Nutzung digitaler lexikalischer Nachschlagewerke eine noch größere kritische Verwendungskompetenz als schon traditionelle Wörterbücher. Abgesehen von der Verantwortung der Anbieter derartiger Informationen¹³ besteht hier sicher ein Auftrag für das allgemeine Bildungssystem mit Blick auf eine letztendlich sehr breit zu verstehende „digital literacy“¹⁴ (SCHMITT 2021: 7).

Die Nutzungsmöglichkeiten der vier Wörterbücher konnten 2014 durch einen (aufgrund seiner Konzeption im obigen Sinne unkritischen) „deutschen Zugang“ noch erweitert werden. Die aufwendig annotierten Textdigitalisate enthalten mehr als 580 000 deutsche Wortformen – als Übersetzungsäquivalente, aber auch in Bedeutungsbeschreibungen, Kommentaren usw. Diese wurden durch Lemmatisierung und orthografische Normalisierung (heutige Schreibung, z. B. *Karussell* statt *Carouffel*) indiziert, sodass

¹² Es handelt sich bei dieser Online-Version des Wörterbuchs um das einzige derzeit verfügbare orthografisch teilweise aktualisierte niedersorbisch-deutsche Nachschlagewerk. Eine vollständig orthografisch aktualisierte Neuauflage der Druckfassung ist in Vorbereitung.

¹³ Dieser hoffen wir auf dem Portal nachgekommen zu sein: Die originale Gerichtetheit der Information bleibt stets erhalten, die Disclaimer geben explizite Hinweise auf den Status der Wörterbücher, und Aktualisierungen sind deutlich gekennzeichnet und überschreiben nicht die ursprünglichen Informationen.

¹⁴ In SCHMITT (2021: 7) heißt es dazu: „Es geht um ein neues Verständnis des Digitalen und ein damit zusammenhängendes neues Selbstverständnis der Anwender. Bildung (respektive »High-status«-Verhalten) der Zukunft wird zu einem beträchtlichen Teil auf dem kritischen Umgang mit den Bildschirmen beruhen.“

seitdem auch eine Suche über deutsche Wörter möglich ist. Dies stellt einen weiteren großen Vorteil gegenüber den alten Druckausgaben dar und gewährt zudem einen Zugang zu teilweise in den Quellen „versteckten“ sprach- wie kulturwissenschaftlichen Informationen, was vor allem im tendenziell „enzyklopädischen“ Wörterbuch von Mucke von Bedeutung ist.¹⁵ Außerdem eröffnet sich durch diese Erweiterung ein zusätzlicher explorativer Zugang: Dadurch, dass nahezu alle in den Wörterbüchern vorkommenden deutschen Wörter auch als Links fungieren, lässt sich der dokumentierte Wortschatz z. B. auf der Suche nach einem passenden Ausdruck gezielt durchstöbern, was in den Druckfassungen völlig unmöglich wäre. So führt etwa die Suche nach *gnás* (u. a.) zum entsprechenden Eintrag in STAROSTA (1999) mit der ersten Bedeutung ‚treiben‘. Über dieses Wort als Link kommt man dann aber zu einer Liste mit 172 Wörterbucheinträgen, in denen *treiben* auftaucht, als Äquivalent, Teil der Bedeutungsbeschreibung usw. Und da auch bereits viele sorbische Wörter als Links wirken (eine vollständige Indexierung steht aus), ergeben sich mit der nötigen Mühe interessante Einblicke in den dokumentierten Wortschatz.

Eine weitere Ergänzung, die auf den retrodigitalisierten Wörterbüchern basiert, ist die Seite „Niedersorbische Redewendungen und Sprichwörter / Dolnosěrbske rěcne wobroty a pśisłowa“, die ebenfalls 2014 online ging.¹⁶ Hier wurden die in den Quellen in großer Menge enthaltenen aber teilweise schwer auffindbaren idiomatischen Wendungen zusammengeführt und über eine zusätzliche XML-Modellierung so aufbereitet, dass sie relativ einfach zugänglich sind. Es handelt sich noch nicht um ein lexikografisch ausgearbeitetes phraseologisches Wörterbuch, sondern zunächst um eine systematisierende Materialsammlung. Diese kann aber nicht nur als Grundlage für eine spätere wissenschaftlich-lexikografische Auswertung dienen, sondern stellt schon jetzt in diesem speziellen Bereich eine Hilfe auch für Lernende dar.

Aussprache und Orthoepie

Für das Erlernen der Lexik einer Sprache ist die Aussprache der Wörter von besonderer Bedeutung. Jedoch existieren für Niedersorbisch weder eine allgemeingültige Orthoepie noch ein entsprechendes Aussprachewörterbuch. Bis etwa 2015 standen Lernenden daher lediglich textuelle Aussprachehinweise in gängigen Grammatiken, Lehrbüchern und Wörterbüchern (JANAŠ 1984, STAROSTA 1991; 1999) zu Verfügung. Ein Lehrbuch (HANNUSCH 1998, 32022) enthält zusätzlich zwei CDs zur lautlichen Illustration der dort vertretenen Ausspracheregeln. Diese insgesamt dennoch spärlichen Informationen sind zumindest mit Blick auf autodidaktisches Lernen häufig unzureichend, zumal viele Lernende wegen der schwierigen soziolinguistischen Situation kaum Kontakt mit voll-

¹⁵ So gelangt man z. B. bei der Suche nach „Ehemann“ auch zum Stichwort *witawa* bei MUCKE (1911–28), wo es in einer Bedeutungserklärung heißt „das Geschenk, welches beim Einzug einer neuen Herrschaft den Dienstboten oder beim Einzug des jungen Ehemannes oder einer jungen Frau ihrer neuen Hausverwandtschaft gegeben wird“.

¹⁶ Siehe <https://www.dolnosěrbski.de/nrs/>.

kompetent und authentisch¹⁷ Sprechenden haben. Daher wurde das Sprachportal seit 2015 um zwei Ressourcen zur Aussprache ergänzt.

Zum einen wurde das DNW bis 2017 um insgesamt 44 000 Tonaufnahmen zu im Wörterbuch auftretenden niedersorbischen Äquivalenten erweitert, um die Aussprache der Wörter zu illustrieren.¹⁸

Auf dieser Maßnahme aufbauend wurde zum anderen eine neue Seite mit dem Titel „Niedersorbische Aussprache / Dolnosěrbske wugronjenje“ erarbeitet, die eine ausführliche, systematische und zum Selbstlernen geeignete Darstellung phonetischer Besonderheiten des Niedersorbischen im Kontrast zum Deutschen bietet.¹⁹ Die Notwendigkeit einer solchen Seite war während der Tonaufnahmen für das DNW erneut deutlich geworden, da zahlreiche der zunächst für das Vorhaben ausgewählten Sprecher:innen trotz vergleichsweise guter Sprachkenntnisse unterschiedliche, teilweise zahlreiche Probleme mit der niedersorbischen Phonetik hatten (KAULFÜRST 2019: 3–4, 38–40).

Auf der 2019 bereitgestellten Seite ist der Zugang zu den Aussprache-Informationen über drei Wege möglich: (a) hauptsächlich über die den Lernenden in aller Regel am meisten vertrauten *Buchstaben*, daneben aber auch (b) über wichtige *Themen* oder (c) über die im Niedersorbischen vorkommenden *Laute*. Eine zusätzliche Seite informiert über die in den Beschreibungen verwendeten Fachbegriffe. Der thematische Zugang stellt soweit wie möglich die Zusammenhänge zwischen den einzelnen Darstellungen aus Buchstaben- oder Laut-Perspektive dar und informiert über grundlegende Eigenschaften der niedersorbischen Phonetik und Phonem-Grafem-Beziehungen. So wird etwa reagiert auf das aufgrund mangelhaften Schulunterrichts verbreitete Problem der „buchstabengetreuen Aussprache“ (vgl. Fußnote 17), indem über den bedeutsamen Unterschied zwischen Schreibung und Aussprache aufgeklärt wird. Weitere Themen sind zum Beispiel die orthografische Bezeichnung der Palatalität, Assimilationsphänomene, die Bedeutung von Silbengrenzen oder die wichtige Unterscheidung von Fremd- und Lehnwörtern. Ergänzend gibt es ein Kapitel zu einigen generell für Niedersorbisch-Lernende mit deutschsprachigem Hintergrund bedeutsamen Kontrasten bei der Aussprache (z. B. mit Blick auf die Vokallänge oder die Palatalisierung von Konsonanten). Alle dargestellten lautlichen Phänomene werden durch geeignete Tonbeispiele illustriert.

¹⁷ Bis mindestens zu Beginn der 1990er-Jahre war die unter anderem in Bezug auf die Aussprachepraxis insgesamt sehr heterogene niedersorbische Sprachgemeinschaft von zwei starken gegensätzlichen Polen geprägt, und zwar einerseits von den Muttersprachlern verschiedener niedersorbischer Dialekte. Diese Gruppe rekrutierte sich fast ausschließlich aus der älteren und ältesten Generation der teilweise weit vor 1950 Geborenen. Andererseits von einer mittleren und jüngeren Sprechergeneration, die ihre Sprachkompetenz zum großen Teil gesteuert über das regionale Bildungssystem erworben hatte. In dieser Gruppe dominieren Semi-Sprecher. Ein häufiges Phänomen ist dort unvollständiger Spracherwerb. Die Aussprache prägte oft eine sehr stark an der Schrift orientierte und auf phonetische Besonderheiten des Niedersorbischen wenig achtende sowie durch das deutsche bzw. obersorbische phonetische System beeinflusste „buchstabengetreue Aussprache“.

¹⁸ Eine vollständige Vertonung des Wörterbuchs war aus Kostengründen nicht möglich. Ebenso kam aufgrund der besonderen soziolinguistischen Situation eine Arbeit mit professionellen, auf Grundlage einer anerkannten Orthoepie gründlich ausgebildeten Sprechern nicht in Frage. Diese und andere Umstände bedingten im Projekt generell ein stark pragmatisches Vorgehen. Für ein Beispiel siehe <https://www.dolnosěrbski.de/dnw/woerterbuch/aussprache>. Mehr Informationen zum Projekt siehe KAULFÜRST (2019).

¹⁹ Siehe <https://www.dolnosěrbski.de/wugronjenje/>.

Die auf dieser Internetseite vermittelten Ausspracheempfehlungen orientieren sich an dem ersten umfassenden Versuch, eine Orthoepie der niedersorbischen Schriftsprache zu entwerfen (KAULFÜRST 2019).²⁰ In diese Studie wurde auch die Lautung der niedersorbischen Dialekte einbezogen, wie sie u. a. auf der ebenfalls über das Sprachportal zugänglichen Seite „Muttersprachliches Niedersorbisch / Maminorěčna dolnosorbšćina“ dokumentiert ist.²¹

Orthografie

Das Portal beinhaltet auch eine Seite zur niedersorbischen Orthografie („Niedersorbische Rechtschreibung / Dolnoserbski pšawopis“).²² Die Grundlage der Informationen bildet einerseits eine retrodigitalisierte Fassung des gültigen Regelwerks (STAROSTA 1982), das in dieser Internet-Ausgabe bereits an wichtigen Stellen ergänzt bzw. aktualisiert wurde im Sinne neuerer Entscheidungen der Niedersorbischen Sprachkommission.²³ Eine grundlegende Überarbeitung des Regelwerks ist aktuell in Arbeit. Ergänzend bietet die Seite eine Zusammenstellung der Beschlüsse der Sprachkommission, wahlweise thematisch oder auch chronologisch.

Auch die 2016 in erster Version veröffentlichte und seitdem ständig erweiterte automatische Rechtschreibkontrolle für Niedersorbisch²⁴ steht hier zum Download bereit. Die aktuelle Version (Anfang 2022) umfasst mehr als 90 000 Lexeme mit etwa 3,7 Mio. Wortformen.

²⁰ Diese Arbeit, für die alle o. g. sowie weitere Quellen mit Aussprachehinweisen ausgewertet wurden, bildet zugleich die Grundlage für weitere notwendige Forschungen zur niedersorbischen Phonetik/Phonologie und Orthoepie. Diese finden seit 2019 im Rahmen der Entwicklung einer Vorlesefunktion für die beiden sorbischen Schriftsprachen statt. Ergebnisse aus diesem Forschungsprojekt können in Zukunft auch zu Revisionen der derzeitigen Empfehlungen führen. Für eine Einführung in dieses Vorhaben siehe MĚŠKANK/SCHMIEDEL (2022). Erste Forschungsergebnisse aus dem Projekt präsentieren JOCZ/MĚŠKANK (2022).

²¹ Siehe <https://www.dolnoserbski.de/dobes/>. Auf der Seite finden sich etwa 100 Stunden Tonaufnahmen verschiedener niedersorbischer Dialekte. Da die Aufnahmen nicht nur geografisch zugeordnet und thematisch verschlagwortet, sondern zudem vollständig transkribiert und die Transkripte zusätzlich ins Deutsche übersetzt wurden, ist ein mehrfacher und relativ komfortabler Zugang zu den Tondokumenten möglich. Lernende, die sich stärker an der dialektalen Praxis orientieren möchten, finden hier eine geeignete Inspirationsquelle.

²² Siehe <https://www.dolnoserbski.de/ortografija/>.

²³ Informationen zur Arbeit der Sprachkommission finden sich auf der Internetseite.

²⁴ An dieser Stelle sei erwähnt, dass im Niedersorbischen wie generell in (digital unterversorgten) Kleinsprachen Ressourcen, die für große Sprachgemeinschaften mittlerweile selbstverständlich sind, in Eigenentwicklung hergestellt werden müssen. Denn es existiert weder eine differenzierte und unterstützende universitäre Infrastruktur, noch ein hinreichendes privatwirtschaftliches Interesse. Daher muss zum Beispiel auch die in Fußnote 20 erwähnte Vorlesefunktion (ein Text-to-speech-System auf Grundlage notwendiger phonetischer Forschungen) für beide sorbische Schriftsprachen in Eigenentwicklung erfolgen. Dieses Phänomen stellt – vor allem in Kombination mit dem erwähnten Mangel an spezifisch qualifiziertem Personal – zugleich einen Bereich dar, in dem Digitalisierung kleine Sprachgemeinschaften vor große Probleme stellt. Dies betrifft zum Beispiel auch Verfahren zur automatischen Spracherkennung und Übersetzung. Die Entwicklung wird zusätzlich dadurch erschwert, dass für große Sprachen entwickelte Methoden häufig nicht einfach übertragen werden können, da sie auf großen Datenmengen aufbauen, die für Kleinsprachen nicht existieren. Vgl. aber Fußnote 45.

Außerdem wurden für diese Portalseite die verfügbaren Informationen zur Schreibung des aktuellen Wortschatzes zusammengeführt. Dadurch konnte erstmals für das Niedersorbische ein umfassendes Rechtschreibwörterbuch bereitgestellt werden.²⁵ Datenbestand und Suchmechanismen sind so gestaltet, dass ein gesuchter niedersorbischer Ausdruck (einschließlich von Mehrwortausdrücken) auch bei nur sehr ungenauer Kenntnis der Schreibweise gut gefunden werden kann. Das System verweist zu diesem Zweck auf grafemisch ähnliche Wörter. Bei der Abfrage werden auch Flexionsformen niedersorbischer Wörter erkannt (etwa solche, die Lernende in Texten vorfinden), und es wird auf ihre Grundformen verwiesen. Umgekehrt erlaubt eine zusätzliche Suchoption auch das Finden von Ausdrücken mit Flexionsformen bei einer Suche nach der Grundform. Die Abfrageergebnisse erscheinen in Form von Wortlisten. So erhält man zum Beispiel bei einer Suche nach *nan* (Vater) eine Liste mit (derzeit) insgesamt 29 Ein- und Mehrwortausdrücken, darunter auch solchen wie *mama starego nana* (Urgroßmutter), die nur eine Flexionsform des Suchwortes enthalten. Von diesen Suchergebnissen gelangt man dann bei Bedarf in die Artikel der jeweiligen Wörterbücher, wo häufig weitere Informationen zur Bedeutung, Flexion usw. zu finden sind. Weitere komplexe Abfragen sind möglich. Die Seite enthält dazu Hilfestellungen.

Die Suchfunktion zu den im Rechtschreibwörterbuch gebotenen Informationen, die im Wesentlichen auf den Informationen des DNW und des Niedersorbisch-deutschen Wörterbuchs von STAROSTA (1999) basieren und über die Trefferlisten weiter zu diesen Quellen führen, ist seit einiger Zeit zentral auf der Portalseite platziert und bietet damit auch einen komfortablen sorbischsprachigen Zugang zum DNW.

Eigennamen

In stark flektierenden Sprachen wie den beiden sorbischen ist die Grammatik von Eigennamen für Lernende durchaus eine Herausforderung. Gleichzeitig enthalten Wörterbücher traditionell Eigennamen häufig nur unvollständig, beschränken sich z.B. auf Ortsnamen. Auch bei der zuvor erwähnten Entwicklung der automatischen Rechtschreibkontrolle wurde schnell sichtbar, dass es sinnvoll wäre, sorbische Eigennamen einzubeziehen. Aus diesen und weiteren Gründen fiel die Entscheidung, auf dem Sprachportal den Namen eine eigenständige Ressource zu widmen, einen Informationsservice zu „niedersorbischen Namen“ („Dolnoserbske mjenja“), dessen erste Version Anfang 2020 veröffentlicht wurde.²⁶

Den Kern des dort aktuell präsentierten Bestandes an Ortsnamen bilden Siedlungsnamen (einschließlich Namen von Verwaltungseinheiten) des amtlichen sorbischen Siedlungsgebietes in Brandenburg. Bereits ergänzt wurden die Namen weiterer brandenburgischer Siedlungen außerhalb des Siedlungsgebiets sowie belegte niedersorbische Ortsnamen in Sachsen, sodass bisher (Stand Mitte 2022) gut 1000 Siedlungen bzw. Verwaltungseinheiten erfasst sind. Die Seite enthält derzeit außerdem Informationen zu mehr als 2600 Familiennamen und etwa 120 Vornamen. Genaue Informationen zum Bestand, zu den ausgewerteten Quellen usw. sind auf der Seite selbst zu finden.

Zu den Eigennamen werden jeweils umfassende Informationen angeboten, Auswahl und Darstellung variieren je nach Namentyp. Bei Ortsnamen finden sich neben eher

²⁵ Siehe <https://www.dolnoserbski.de/ortografija/slownik>.

²⁶ Siehe <https://www.dolnoserbski.de/mjenja/>; vgl. auch SZCZEPAŃSKA (in Vorbereitung).

enzyklopädischen Angaben (u. a. Geo-Koordinaten und eine kurze Definition/Angabe zur Zugehörigkeit zum sorbischen Siedlungsgebiet) sowohl Informationen, die für das Sprachverstehen (aktuelle, alternative und veraltete niedersorbische und deutsche Namen und Namensformen) relevant sind als auch solche, die eine aktive Verwendung befördern: Derivate (Bewohner/in, Adjektiv) und die zugehörigen Flexionsangaben in Form einer fallspezifischen Tabelle.

Die Einträge zu den Personennamen beinhalten jeweils die Hauptform des Namens (z. B. *Kóńcař*, *Měto* – daneben auch als Kurzform von *Mjertyn*), bei Familiennamen zusätzlich die Namensformen für Frauen (*Kóńcarjowa*), für die Familie als Ganze (*Kóńcarjojc*, *Kóńcarjowe*) sowie das possessive Adjektiv (*Kóńcarjowy*). Bei Vornamen werden ggf. auch alternative und veraltete Namensformen, Verkürzungen und Verniedlichungen, Diminutiva (beim obigen Beispiel *Měčo*, *Mječo*, *Mětko*) und Augmentativa angeführt. Auch hier finden sich neben evtl. relevanten enzyklopädischen Informationen (etwa einem Verweis auf eine(n) prominente(n) Namensträger(in)) zusätzlich die vom jeweiligen Namen gebildeten Adjektive sowie spezifische Flexionstabellen.

Für eine effektive Suche nach Namen bzw. Namensformen ist auch die Einbeziehung der entsprechenden deutschen Namen und Namensvarianten von Bedeutung, für obiges Beispiel etwa *Konzak*, *Konzack*, *Konzag*, *Kointzag*, *Kanzack*, *Keinzag*, *Kinzack*, *Kühnzack*, *Koinzar*, *Kointzar* bzw. *Martin*. So können die aktuellen und grammatisch korrekten niedersorbischen Formen auch bei nur ungenauer Kenntnis des Namens – in welcher Form auch immer – auf verschiedenen Suchpfaden und mit hoher Wahrscheinlichkeit gefunden werden.²⁷ Hier zeigen sich erneut deutlich die Vorteile einer digitalen Datenaufbereitung und Präsentation der zusammengestellten Informationen.

Dasselbe gilt für eine seit 2021 zunächst testweise vorgenommene Erweiterung des „Eigennamen-Informationsservice“ um Informationen zu Namenbedeutungen. Im sorbischen Kontext sind Eigennamen nicht nur als Wörter – und damit für den Spracherwerb – von Bedeutung, sondern auch als hör- oder sichtbarer Teil des sorbischen Kulturerbes der Region, der bikulturell geprägten Lausitzer „Namenlandschaft“. Daher läuft seit 2020 ein Versuch, auch etymologische Informationen zu den sorbischen bzw. auf das Sorbische zurückgehenden deutschen Namen anzubieten und zu vermitteln – Letzteres explizit auch mit dem Ziel, das Wissen über die sorbische Herkunft *vieler* in der Lausitz vorkommender Namen zu stärken. Dabei bezieht sich der verwendete Begriff „Versuch“ im Sinne von Experiment vor allem auf das Vorgehen bei der Zusammenstellung, Formulierung und Vermittlung der Bedeutungserklärungen auf innovative modulare Weise (ZSCHIESCHANG 2021). Bis dato wurde so die Etymologie von etwa 150 Namen ergänzt, ein kleiner Teil dessen, was längerfristig angestrebt wird.

Es ist vorgesehen, die Eigennamen-Seite in Zukunft um weitere Datenbestände zu erweitern. Eine systematische Auswertung des Textkorpus (siehe unten) steht noch aus. Vorgesehen ist auch, stärker explorative Zugänge zu den Daten zu ermöglichen („blättern“ und „stöbern“ wie in Druckpublikationen sowie eine kartenbasierte Präsentation).

²⁷ Die Seite enthält insgesamt mehr als 18 000 niedersorbische Namen und davon abgeleitete Formen, ein Großteil davon mit vollständigen Flexionstabellen. Der Datenbestand erfasst zudem über 10 000 deutsche Namen als Entsprechungen der niedersorbischen.

Texte: Korpusuche und Digitale Editionen

Die bisher vorgestellten Ressourcen boten Informationen zur niedersorbischen Lexik, wie sie in älteren wie aktuellen Wörterbüchern beschrieben ist, entweder in niedersorbisch-deutscher oder in deutsch-niedersorbischer Perspektive, aber mit Zugängen auch jenseits dieser traditionellen Ausrichtung.²⁸ Diese Informationen wurden ergänzt durch Tonaufnahmen sowie durch Erläuterungen zur Phonetik und Orthoepie sowie zur Orthografie. Und zu diesen *Wörtern* zählen Eigennamen, die auf einer gesonderten Seite auf besondere Weise beschrieben werden.

Ergänzend werden aber auch ganze *Texte* auf dem Sprachportal bereitgestellt. Und auch diesen kann man sich zunächst von der Wortseite nähern – das ist die traditionelle korpuslinguistische Vorgehensweise: über Suchwörter zu Konkordanzen (KWIC: Key Word in Context). Einen solchen Zugang gibt es auf dem Sprachportal seit 2010. Er basiert auf dem alten niedersorbischen Textkorpus, das aus reinen Volltexten besteht, die nicht weiter aufbereitet sind. Dieser Zugang besteht weiterhin,²⁹ wird aber schrittweise durch ein neues Textkorpus ersetzt. Denn in den letzten Jahren wurden große Anstrengungen unternommen, um das niedersorbische Textkorpus sowohl quantitativ als auch qualitativ auf eine neue Grundlage zu stellen (siehe BARTELS 2020).

Die Korpustexte wurden so aufbereitet, dass sie gleichzeitig verschiedenen Zwecken dienen können: Zum einen gewährleisten Textqualität und umfangreiche Annotationen (bisher v. a. Textstruktur, Normalisierung und Lemmatisierung) eine zuverlässige und komfortable Suche.³⁰ Komfort bedeutet hier insbesondere, dass für eine effektive Suche keine Spezialkenntnisse etwa zur sorbischen Flexion oder zu historischen Schreibvarianten³¹ notwendig sind. Zum anderen ermöglicht die Aufbereitung zugleich eine Textausgabe im Rahmen einer Digitalen Bibliothek und stellt damit Lesefassungen zur Verfügung, die über traditionelle Zugänge zu Korpustexten weit hinausgehen.³²

Die erste auf dem Sprachportal verwirklichte und ebenfalls bereits auf einem Korpustext gründende digitale Edition betraf die niedersorbische Bibel-Gesamtausgabe von 1868. Es handelt sich hierbei um einen Text von hoher sprachlicher Qualität, der bis zur digitalen Neu-Edition nie wieder aufgelegt worden war und daher – trotz seiner herausragenden Bedeutung für eine sorbischsprachige christlich-religiöse Praxis – nur noch antiquarisch und häufig schlecht erhalten verfügbar war. Daher eignete sich dieser umfangreiche Text für einen Prototyp einer korpustextbasierten Digitalen Bibliothek, der 2018 zum 150. Jubiläum der Originalausgabe veröffentlicht wurde.³³ Abgesehen davon, dass auf diese Weise der Text überhaupt wieder leicht zugänglich wurde,³⁴ konnten auch

²⁸ Einsprachige Wörterbücher des Niedersorbischen existieren nicht.

²⁹ Siehe <https://www.dolnoserbki.de/korpus/standard/pytanje>.

³⁰ Siehe <https://www.dolnoserbki.de/korpus/komfort/pytanje/>.

³¹ Es handelt sich um ein historisches Korpus mit Texten seit dem 16. Jahrhundert.

³² Die Korpustexte basieren auf einer hochwertigen Primärdigitalisierung per Double-Keying-Verfahren und wurden zusätzlich einigen weiteren Korrekturen unterzogen. Gleichwohl sind Fehler möglich.

³³ Siehe <https://www.dolnoserbki.de/biblija/>. Hier finden sich auch genaue Informationen zum Text, zu den Hintergründen der digitalen Edition und zur Textbearbeitung.

³⁴ Die Textausgabe wurde in diesem Fall noch einer besonderen Qualitätsprüfung unterzogen, und zwar in Kooperation mit dem Verein zur Förderung der wendischen Sprache in der Kirche e. V. (Spěchowanske towaristwo za serbsku rěc w cerkwi z. t.). Parallel erschien auch eine CD mit der digitalen Edition.

hier die Vorteile des digitalen Mediums genutzt werden: Neben einem Zugang über die bekannten Textstrukturelemente (Buch, Kapitel, Vers) ist der Text auf vielfältige Weise durchsuchbar. Dieser wird außerdem in drei verschiedenen Varianten angeboten, die in ihrer Gesamtheit einerseits historische Originaltreue, andererseits eine leichtere Erschließbarkeit für das heutige Lesepublikum gewährleisten. So kann einerseits der ursprüngliche Text in einer nahezu originalgetreuen Frakturschrift (wie bei der Ausgabe 1868) oder alternativ als Transliteration in lateinischer Schrift angezeigt werden. Die Fassung in lateinischer Schrift ist dann zusätzlich auch noch in einer orthografisch modernisierten Fassung verfügbar.

Wie in BARTELS (2020) dargestellt, wurde in den Folgejahren das Verfahren zur Aufbereitung von Korpustexten auf eine Weise fortentwickelt und angewandt, die es erlaubte, auch eine größere Anzahl von Korpustexten im Rahmen einer Digitalen Bibliothek bereitzustellen – eine erste Version ging im Januar 2022 online³⁵ und enthält neben den Ausgaben periodischer Schriften³⁶ auch etwa 200 teilweise sehr umfangreiche Einzeltexte und damit bereits einen Großteil des niedersorbischen Schrifttums bis zum Ende des Zweiten Weltkriegs 1945. Die Digitale Bibliothek stellt freilich keine philologische Neuausgabe dar, basiert aber auf hochwertigen Korpustexten, sodass sie relativ zuverlässige und gut handhabbare Lesefassungen bietet. Alle Texte stehen zusätzlich in den Formaten PDF, TXT und ePUB zum Download bereit. Generell wird eine Originalfassung in lateinischer Transliteration angeboten, analog zur „mittleren“ Textfassung der Bibel-Edition. Die beiden im Falle der orthografisch wie typografisch relativ einheitlichen³⁷ Bibel von 1868 vorhandenen anderen Varianten (Original in Fraktur und orthografisch modernisiert) sind in der Digitalen Bibliothek wegen der historischen Tiefe (Texte, die zwischen 1706 und 1936 gedruckt wurden) und der damit verbundenen sehr großen Varianz sowohl der Schreibung als auch der verwendeten Drucktypen nur teilweise und längerfristig umzusetzen. Eine modernisierte Textvariante wurde zunächst versuchsweise für gut ein Dutzend Bücher umgesetzt, die sich hinsichtlich der o. g. Kriterien weitgehend an der Bibel orientierten.

Auf der Seite der Digitalen Bibliothek sind die Texte über eine Katalogsuche zugänglich. Zusätzlich ist dort eine Auswahl über Autorennamen und Kategorien wie Lyrik, Kirchenlied, Zeitschrift usw. möglich. Da den Lesefassungen Korpustexte zugrunde liegen, gibt es einen alternativen Zugang über die Korpusseite und damit über die bereits erwähnte „Komfortsuche“. Auch von dort gelangt man aus der Textsuche heraus (über die Suchergebnisse) zu den jeweiligen Lesetexten der Digitalen Bibliothek.

Inwieweit sind die Texte dieser Digitalen Bibliothek noch heute mit Blick auf den Sorbischerwerb relevant? Sicherlich ist der Hauptzweck der Korpustextaufbereitung ein anderer: Das Korpus dient primär der Sprachdokumentation sowie als Grundlage für diverse textauswertende Forschungen. Angesichts des in der Einleitung thematisierten Mangels an Drucktexten können jedoch zumindest die im 20. Jahrhundert erschienenen

³⁵ Siehe <https://www.dolnoserbski.de/biblioteka/>.

³⁶ Vor allem dieser Teil des Schrifttums hat einen relativ großen Umfang. Verfügbar sind alle 84 Jahrgänge des *Bramborski Casnik* (eine Wochenzeitung, erschienen zwischen 1848 und 1933 unter verschiedenen Namen), alle 10 Jahrgänge der evangelischen Zeitschrift *Wósadnik* (1904–13) sowie 24 Ausgaben des Jahreskalenders *Pratyja*.

³⁷ Unterschiede bestehen vor allem zwischen Altem und Neuem Testament. Die aufwendige Transformation besorgte Fabian Kaulfürst. Mehr Informationen finden sich auf <https://www.dolnoserbski.de/biblija/info>.

Werke, darüber hinaus aber auch viele der seit etwa den 1880er-Jahren veröffentlichten Texte fortgeschrittenen Lernenden als zusätzlicher Input helfen, durch Lesen ihre Sprachkenntnisse zu festigen und ihren Wortschatz zu erweitern. Zweifellos können die ausschließlich älteren Texte der Digitalen Bibliothek – neuere können dort aus rechtlichen Gründen nicht veröffentlicht werden – das Fehlen einer thematisch wie stilistisch usw. breit gefächerten modernen Literatur nicht kompensieren. Angesichts des eingangs erwähnten „Kleinsprachen-Dilemmas“ bleibt aber nur, das Vorhandene möglichst umfassend und barrierearm zur Nutzung bereitzustellen.

Obersorbisch

Ein den dargestellten niedersorbischen Ressourcen gleichwertiges Angebot für das Obersorbische gibt es bisher nicht. Vor allem mangelt es hier noch an einem aktiven, d.h. mit dem Fokus auf einer Förderung des aktiven Sprachgebrauchs konzipierten Wörterbuch und ähnlichen Hilfsmitteln. Ein obersorbisches Sprachportal befindet sich im Aufbau.³⁸ Die Seite besteht seit 2017 und enthält bislang (Stand Mitte 2022) zwei Angebote: Eine retrodigitalisierte Fassung des materialreichen, jedoch konzeptionell eher traditionellen Deutsch-obersorbischen Wörterbuchs von 1989³⁹ sowie die sogenannten „Rěčne kućiki / Sprachecken“.⁴⁰ Dabei handelt es sich um eine in digitaler schriftlicher Form aufbereitete und daher gut recherchierbare Sammlung von etwa 730 Radiobeiträgen, die grob als „Sprachberatung“ charakterisiert werden können. Ein Zugang zu den einzelnen Artikeln ist u. a. über die behandelten Wörter, über Themen (z. B. Wortbildung) oder auch bestimmte Schlagwörter wie *požconka* (Lehnwort) oder *přenjeseny woznam* (übertragene Bedeutung) möglich.

Das oben genannte Wörterbuch ist zusätzlich mit weiteren lexikalischen Ressourcen auf der Internetseite *soblex*⁴¹ verknüpft. Diese Seite ist auf Initiative von Privatpersonen entstanden, aber mittlerweile ein Gemeinschaftsprojekt verschiedener Personen und Institutionen, in den letzten Jahren unter maßgeblicher Beteiligung des Sorbischen Instituts. Den Datenbestand bildete zunächst eine digitale Fassung des Verlagsmanuskripts des obersorbischen Rechtschreibwörterbuchs (VÖLKELEL 2005).⁴² Dieser Kernbestand wurde anschließend durch eine ganze Reihe weiterer Quellen⁴³ erweitert, sodass *soblex* mittlerweile eine sehr reichhaltige lexikalische Materialsammlung darstellt, die mittels gut ausgebauter Suchmechanismen auf Anhieb viele Informationen anbietet. Vor allem ein als aktives Wörterbuch konzipiertes Deutsch-obersorbisches lexikalisches Informationssystem kann dadurch freilich nicht ersetzt werden.

³⁸ Siehe <https://www.hornjoserbsce.de/> oder <https://www.obersorbisch.de/>.

³⁹ Siehe <https://www.hornjoserbsce.de/dow/>. Für die Druckausgabe siehe DOW (1989/1991).

⁴⁰ Siehe <https://www.hornjoserbsce.de/kuciki/>.

⁴¹ Siehe <https://www.soblex.de/> [27.06.2022].

⁴² In einem späteren Projekt am Sorbischen Institut (2016) wurden zahlreiche Änderungen der 6. Auflage (VÖLKELEL 2014) eingearbeitet.

⁴³ Die wichtigsten Erweiterungen waren: 2017/18 um Terminologie aus einem neuen, in Arbeit befindlichen Deutsch-obersorbischen Wörterbuch sowie um die Lexik einer 2006 erschienenen Erweiterung des DOW (DOW-NL 2006); 2019/20 um sorbische Vornamen (MEŠKANK 2017), neue Lexik aus den oben genannten „Sprachecken“ und um obersorbische Exonyme nach einer Sammlung des Sorbischen Instituts.

Unter den im Laufe der Zeit auf *soblex* ergänzten Quellen nimmt der Sprachratgeber „Kak to jenož praju?“ (HAJDUK-VELJKOVIĆ 2017) eine besondere Stellung ein, da auf diese Weise eine stärker didaktisch und auf die Erweiterung des aktiven Sprachvermögens ausgerichtete Publikation einbezogen wurde. Es handelt sich um eine Sammlung bedeutungsähnlicher Wörter mit Anwendungsbeispielen sowie von Redewendungen und damit um die derzeit am stärksten „aktive“ Komponente der Seite.

Die Bedeutung von *soblex* wird zusätzlich erhöht durch einen integrierten morphologischen Generator, der für alle in der Datensammlung enthaltenen Wörter vollständige und jeweils spezifische Flexionsparadigmen anzeigt. Auch die Silbentrennungsmöglichkeiten sind angezeigt, was für das Niedersorbische zurzeit nicht möglich ist. Und sie lassen sich zudem für beliebige obersorbische Wörter generieren.⁴⁴ Eine Besonderheit stellt außerdem eine „Textanalyse“ genannte Hilfe zur sorbisch-deutschen Übersetzung dar: ein obersorbischer Text wird hier in seine Bestandteile zerlegt, für die dann die im System hinterlegten Übersetzungsäquivalente (einschließlich grammatischer Angaben im Falle von Flexionsformen im Eingabetext) vorgeschlagen werden. Dieses Ergebnis kann als Grundlage für eine deutsche Übersetzung des analysierten Textes dienen.⁴⁵

Ebenfalls über diese Seite als Download zugänglich sind Tools für die automatische Rechtschreibkontrolle sowie zusätzlich zur Silbentrennung.

Ausblick und Resümee

Der weitere Ausbau der beiden Sprachportale wird einerseits in einer schrittweisen Erweiterung bestehender sowie in der Entwicklung neuer digitaler sprachlicher Ressourcen bestehen. Vor allem beim Obersorbischen spielen dabei neben der (laufenden) Erarbeitung eines aktiven Wörterbuchs die Retrodigitalisierung weiterer Quellen und deren Einbindung in das Portal eine große Rolle. Für beide Portale ist in nächster Zeit ein neues Modul geplant, das zügig neue Lexik bereitstellen soll, die sich aus der Auswertung des aktuellen sorbischen Schrifttums (BARTELS u. a. 2021), aus Übersetzungsvorhaben oder der ständigen Korpusplanung (Bildung neuer Terminologie usw.) speist.

Andererseits wird in den nächsten Jahren ein Schwerpunkt auf der weiteren Erschließung und Vernetzung bereits verfügbarer Ressourcen liegen. Wie dargestellt enthalten schon jetzt einige der Ressourcen quellenübergreifende und über den Weg zusätzlicher Datenmodellierung geschaffene erweiterte Zugänge zu den verfügbaren Informationen. Dies erhöht nicht nur den Komfort bei der Suche und baut Barrieren ab, sondern sorgt dafür, dass die vorhandenen und bereits erschlossenen Daten möglichst gut zugänglich und damit nutzbar sind.

Auch die Präsentation schon vorhandener Daten kann noch verbessert werden: Neben den vorgestellten Suchfunktionen sollen mehr explorative Zugänge geschaffen werden, die das ermöglichen, was man mit Blick auf Druckpublikationen oder traditionelle Wissensbestände mit Begriffen wie „stöbern“ oder „sich umschaun“ umschreiben

⁴⁴ Siehe <https://soblex.de/?cmd=hyph.show&>. Die Silbentrennmuster wurden 2018/19 in einem Projekt am Sorbischen Institut entwickelt.

⁴⁵ An einem System zur automatischen Übersetzung sorbischer Texte wird seit einiger Zeit am WITAJ-Sprachzentrum (siehe <https://www.witaj-sprachzentrum.de/> [27.06.2022]) gearbeitet: siehe <https://sotra.app/> [27.06.2022]. Auch hier schreitet die Arbeit trotz großer Herausforderungen erkennbar voran.

würde. Dies im digitalen Medium zu realisieren ist zwar nicht trivial. Es ist aber sinnvoll, um zum einen alten, noch dem Druckzeitalter entstammenden Rezeptionsgewohnheiten entgegenzukommen, und um zum anderen zusätzliche Möglichkeiten zu nutzen – wie etwa die Bereitstellung eines kartenbasierten Zugangs zu einzelnen Ortsnamen. Schließlich soll die trotz des benannten grundsätzlichen Informationsmangels dennoch vorhandene relative Reichhaltigkeit der angebotenen Informationen gut sichtbar sein, auch als Motivation, immer weiter zu lernen und Neues zu entdecken.

Aus der geschilderten Vorgehensweise und der groben Beschreibung der bislang erarbeiteten Ressourcen sollte deutlich geworden sein, dass die Digitalisierung gerade Kleinsprachen neue Möglichkeiten eröffnet, auch um mit dem dort typischen Informationsmangel umzugehen. Dieser kann zwar durch Digitalisierung nicht grundsätzlich behoben, wohl aber teilweise kompensiert werden. Eine kritische und reflektierte Herangehensweise vorausgesetzt, kann der Einsatz digitaler Methoden von großem Vorteil sein.

Freilich ist es für kleine Sprachgemeinschaften häufig schwierig, die Möglichkeiten der Digitalisierung durchgängig und intensiv zu nutzen, da ja zusätzlich zur fachlichen (z. B. linguistischen) Erarbeitung von Informationen auch noch die technischen Prozesse bedient werden müssen. Zudem müssen einmal zusammengestellte Datenbestände inhaltlich gepflegt und einschließlich aller Komponenten zur Bearbeitung und Präsentation (Software, Internetseiten usw.) beständig aktualisiert werden. Da das „Kleinsprachen-Dilemma“ in der Regel auch die Verfügbarkeit entsprechend qualifizierten Personals betrifft, stellt dies eine dauerhafte Herausforderung dar und führt dazu, dass wichtige Entwicklungsschritte gar nicht, nur langsam oder in einer durch Kompromisse gekennzeichneten Form realisiert werden können.

Trotz allem stellt die Digitalisierung im beschriebenen Tätigkeitsbereich einen Teil der Lösung, nicht des Problems dar. Denn der Einsatz digitaler Methoden ermöglicht eine optimale Bereitstellung knapper sprachlicher Ressourcen und unterstützt damit Maßnahmen zu Spracherhalt und -revitalisierung. Und mit Blick auf den bisherigen Entwicklungsstand beim Niedersorbischen lässt sich wohl mittlerweile sagen: Weder das Sprachlernen Erwachsener noch die Fortbildung von Lehrkräften sollten am Informationsmangel scheitern.

Literaturverzeichnis

- BARTELS, Hauke 2010: The German-Lower Sorbian Online Dictionary, in: DYKSTRA, Anne; SCHOONHEIM, Tanneke (Hgg.), Proceedings of the XIV Euralex International Congress (Leeuwarden, 6–10 July 2010). Ljouwert, S. 1450–1462.
- BARTELS, Hauke 2013: Zur Konzeption eines historisch-dokumentierenden Wortschatz-Informationssystems des Niedersorbischen. Pläne zur Behebung eines drängenden Forschungsdesiderats, in: KEMPGEN, Sebastian; WINGENDER, Monika; FRANZ, Norbert; JAKIŠA, Miranda (Hgg.), Deutsche Beiträge zum 15. Internationalen Slavistenkongress, Minsk 2013. München-Berlin-Washington/D.C., S. 37–46.
- BARTELS, Hauke 2015: Słowniki – spěchowanjje rěcy – nowe medije, in: Schulz, Jana; Ćorić, Božo (Hgg.), Słowjanske spisowne rěče a medije. Mjezynarodna konferenca Komisije za słowjanske spisowne rěče při Mjezynarodnym komiteju slawistow, Budyšin 18.–21.6.2014. Budyšin, S. 24–31.

- BARTELS, Hauke 2020: Das niedersorbische Globalkorpus als Ziel einer ganzheitlichen Konzeption zum Aufbau von Textkorpora, in: *Lětopis* 67/2, S. 3–44.
- BARTELS, Hauke 2022: The Long Road to a Historical Dictionary of Lower Sorbian. Towards a Lexical Information System, in: KLOSA-KÜCKELHAUS, Annette; ENGELBERG, Stefan; MÖHRS, Christine; STORJOHANN, Petra (Hgg.), *Dictionaries and Society, Proceedings of the XX EURALEX International Congress*. Mannheim, S. 540–548.
- BARTELS, Hauke; KAULFÜRST, Fabian; SZCZEPAŃSKI, Marcin; WÖLKE, Sonja 2021: Das Monitoring des sorbischen Schrifttums. Grundlagen und erster Jahresbericht eines neuen Forschungsvorhabens, in: *Lětopis* 68/1, S. 77–136.
- DNSW 2012 – HANNUSCH, Erwin; POHONTSCH, Anja; STAROSTA, Manfred: *Deutsch-niedersorbisches Schulwörterbuch*. Bautzen.
- DNW 2003–2022 – STAROSTA, Manfred; HANNUSCH, Erwin; BARTELS, Hauke (unter Mitarbeit von KAULFÜRST, Fabian): *Deutsch-niedersorbisches Wörterbuch*. URL: <https://www.dolnoserbski.de/dnw/>.
- DOW 1989/1991 – *Deutsch-obersorbisches Wörterbuch*. Begründet von Rudolf JENTSCH. Verfaßt von Helmut JENTSCH, Siegfried MICHALK und Irene ŠĚRAK unter Mitarbeit von Georg MIRTSCHINK, 2 Bände. Bautzen.
- DOW-NL 2006 – JENTSCH, Helmut; POHONTSCH, Anja; SCHULZ, Jana: *Deutsch-obersorbisches Wörterbuch neuer Lexik*. Bautzen.
- HAJDUK-VELJKOVIĆ, Lubina 2017: *Kak to jenož praju? Příručka za bohatši słowoskład za wyšu šulu a gymnazij*. Bautzen.
- HANNUSCH, Erwin 1998/2022: *Niedersorbisch praktisch und verständlich. Das Lehrbuch zum Lernen und Auffrischen*. Bautzen.
- JANAŠ, Pěťš 1984: *Niedersorbische Grammatik für den Schulgebrauch*, 2. Auflage. Bautzen.
- JANNASCH, Peter [JANAŠ, Pěťš] 1990: *Wörterbuch Deutsch-Niedersorbisch*. Bautzen.
- JOCZ, Lechosław; MĚŠKANK, Jan 2022: Beiträge zur sorbischen Phonetik und Phonologie Teil 1 – Zum Lautsystem des Niedersorbischen: Distribution von [ɛ] und [e], in: *Lětopis* 69/1, S. 54–68.
- KAULFÜRST, Fabian 2019: Pśinosk k dolnoserbskej ortoepiji na zakłaže projekta awdijowych datajow za nimsko-dolnoserbski internetowy słownik, in: *Lětopis* 66/1, S. 3–41.
- MENZEL, Thomas; POHONTSCH, Anja 2020: Sorbisch, in: BEYER, Rahel; PLEWNIA, Albrecht (Hgg.), *Handbuch der Sprachminderheiten in Deutschland*. Tübingen, S. 227–269.
- MĚŠKANK, Jan; SCHMIEDEL, Astrid 2022: Zur Vorlesefunktion für die nieder- und obersorbische Schriftsprache – Einführung und Überblick, in: *Lětopis* 69/1, S. 131–142.
- MĚŠKANK, Timo 2017: *Serbske předmjena. Serbske pśedmjenska. Sorbische/wendische Vornamen*. Budyšin.
- MUCKE, Ernst 1911–15/1926, 1928: *Wörterbuch der nieder-wendischen Sprache und ihrer Dialekte*. Bd. 1: St. Petersburg 1911–15, Prag 1926; Bd. 2/3: Prag 1928. (Foto-mechanischer Neudruck. Bautzen 2008)
- SCHMITT, Peter 2021: *Postdigital. Medienkritik im 21. Jahrhundert*. Hamburg.

- SORBISCH ONLINE LERNEN 2015–2022: URL <https://sprachkurs.sorbischlernen.de> [27.06.2022]; <https://tu-dresden.de/gsw/slk/lsk/die-einrichtung/E-Learning-LSK-MSZ/sorbisch-online> [27.06.2022].
- ŠOLĆINA, Jana; WORNAR, Edward 2012: Obersorbisch. Ein Sprachkurs im Selbststudium für Unerschrockene / Hornjoserbšćina. Rěčny kurs za samostudij za njebojazychnych. 3., přehladany nakład. Budyšin.
- STAROSTA, Manfred 1982: Niedersorbische Orthographie und Interpunktion. Regeln (2., bearbeitete Auflage). Bautzen/Budyšin.
- STAROSTA, Manfred 1991: Niedersorbisch schnell und intensiv. Lehrbuch für Fortgeschrittene und sorabistisch Interessierte. Teil 1. Bautzen.
- STAROSTA, Manfred 1999: Niedersorbisch-deutsches Wörterbuch. Bautzen.
- SZCZEPAŃSKA, Joanna (in Vorbereitung): Dolnosěrbske mjenja | Niedersorbische Namen – serwis internetowy poświęcony dolnołużyckim nazwom własnym. Popularyzacja wiedzy na temat języka mniejszościowego w dwujęzycznym środowisku. Erscheint in: *Onomastica* (<https://onomastica.ijp.pan.pl/index.php/ONOM>).
- SZCZEPAŃSKI, Marcin 2012: Modelowanie danych do słownika elektronicznego na przykładzie artykułów słownika A. Muki, in: MOTORNYJ, Volodymyr; SCHOLZE, Dietrich (Hgg.), *Pytannja sorabistyky: materialy XIII Mižnarodnogo sorabistyčnogo seminary 2010*. L'viv, S. 26–61.
- ŠWJELA, Bogumił 1961: Dolnosěrbski-němski słownik. Budyšin.
- Völkel, Pawoł 2005: *Prawopisny słownik hornjoserbskeje rěče, wobdźěłał Timo Meškank. 5., wobdźěłany a sylnje rozšěrjeny nakład*. Bautzen.
- Völkel, Pawoł 2014: *Prawopisny słownik hornjoserbskeje rěče, wobdźěłał Timo Meškank. 6., přehladany nakład*. Bautzen.
- ZÖFGEN, Ekkehard 1991: *Bilingual Learner's Dictionaries*, in: HAUSMANN, Franz Josef; REICHMANN, Oskar; WIEGAND, Herbert Ernst; ZGUSTA, Ladislav (Hgg.), *Wörterbücher. Ein internationales Handbuch zur Lexikographie. Dritter Teilband*. Berlin, New York, S. 2888–2903.
- ZSCHIESCHANG, Christian 2021: *Onomastischer Wissenstransfer am Sorbischen Institut. Zwei neue Projekte*, in: *Onomastica Lipsiensia* 14, S. 641–656.
- ZWAHR, J. G. 1847: *Niederlausitz-wendisch-deutsches Handwörterbuch*. Spremberg (Fotomechanischer Neudruck. Bautzen 1989).