ETHNOGRAFISCHE ONLINE-ARCHIVE ALS PARÖMIO-LOGISCHES WERKZEUG. NUTZUNGSSZENARIEN DES WOSSIDLO DIGITAL ARCHIVE (WOSSIDIA)

Abstract: The article aims to demonstrate the connectivity of an online folklore archive to proverbial research and proverb databases. The case study is WossiDiA, which presents the exensive (Low German language based) regional ethnographic collection of Mecklenburg's field-worker Richard Wossidlo (1859-1939). The challenge of the WossiDiA-project was to model the idiosyncratic nature of this tradition archive. About two million resarch notes, mostly small-scale information, are highly networked by hypergraphs. Some examples show how to extract proverbs from the digital archive. Variant relationships, social data of the informants, and other aspects of knowledge production, can also be included. Experiences within the transatlantic Big Data-project ISEBEL (Intelligent Search Engine for Belief Legends), in which WossiDiA actually participates, provide some ideas on how WossiDiA can be connected to proverbial databases.

Keywords: Computational Folkloristics, tradition archives, harvesting, Big Data, German, Mecklenburgian Folklore, computer-linguistics, dialect, proverbial corpora, proverbs, research notes

1. Einleitung

Während sich im Bibliotheksbereich die Digitalisierung gedruckten Schrifttums vergleichsweise rasant entwickelte und "eborn"-verfasste Werke rasch um sich griffen, vermisst man eine vergleichbare Dynamik im Archiv- und Sammlungswesen. Zwar ist es inzwischen üblich, dass verwahrende Institutionen ihre Geschichte, Programmatik und Findmittel (Repertorien) im Internet über Portale sichtbar machen; die archivalischen Quellen selbst werden jedoch dem Online-Nutzer in der Regel vorenthalten, mit Ausnahme von Einzeldokumenten, die als "Schätze" werkewirksam in Szene gesetzt werden. Hierfür gibt es viele Gründe, von dem unterschiedlichen Selbstverständnis zwischen Bibliotheken, Museen, Archiven und Forschungssammlungen, über Urheber-,

PROVERBIUM 36 (2019)

Persönlichkeits- und Nutzungsrechte (wie Sperrfristen) bis hin zur Trägerschaft und Finanzierbarkeit. Es gibt eine Fülle von Archivgattungen, und jedes Archiv besitzt seine unverwechselbare Individualität.

Dies gilt mindestens ebenso für wissenschaftliche Sammlungen, deren Begründer ihre Werkstätten im Hinblick auf forschungspraktische Anliegen ertüftelten - wie der volkskundliche Privatgelehrte und Gymnasialprofessor Richard Wossidlo (1859-1939) aus Mecklenburg, der das per extensiver Feldforschung und virtuoser Exzerptkultur zusammengetragene Wissen seines Sammelunternehmens durch ein ausgeklügeltes Zettelkastensystem tief verschlagwortete, Begriffe hierarchisierte und vielfach vernetzte. Strukturbildungsregeln für generiertes Wissen, welche sich das Bibliotheks- und Verlagssystem schon früh durch Normdaten schuf, waren daher in Forschungssammlungen erst werdender Fächer, wie der Volkskunde, nicht verbindlich und setzten sich auch später allenfalls graduell durch. Forschungssammlungen im Übergang zu sich institutionalisierenden Fachwissenschaften sind daher in hohem Maße idiosynkratisch. Diese zu digitalisieren erfordert zum einen anzupassende Workflows mit hohem Vorbereitungsaufwand. Weit schwerer wiegt jedoch wegen des idiosynkratischen Charakters vieler Forschungssammlungen das Problem ihrer Onlinestellung. Denn die Transformation eines analogen in ein virtuelles Archiv ist so zu gestalten, dass die Dokumente - zumindest für Forschungswecke lesbar sind, wobei ein Archivar, der in Praxis auch intuitives Wissen erworben hat, nicht maschinell ersetzt werden kann.

Während gedruckte Editionen archivalische Quellen für den Leser kommentieren, stellt dies im Falle der Onlinestellung einer ganzen Forschungssammlung, die über einen hohen Anteil bislang unpublizierter Quellen verfügt, eine Herausforderung dar. Im Projekt WossiDiA – the Digital Wossidlo Archive,¹ das von 2010 bis 2014 sowie von 2017 bis 2020 im Rahmen des transatlantischen Harvesting-Projekts ISEBEL – Intelligent Search Engine for Belief Legends² durch die Deutsche Forschungsgemeinschaft gefördert wird, wurde der Weg, einmal sämtliche Dokumente einer Forschungssammlung ins Netz zu stellen, pionierhaft umgesetzt. Der Nutzen dieses Ansatzes der Digitalisierung und Onlinestellung of the entire material of a collector or collection network (Ilyeflavi 2018: 218) in Verbindung mit dem Ziel,

damit auch gattungsspezifische Datenbanken zu beliefern und direkten Korpuszugang zu *Folklore Archives* zu ermöglichen, soll im Folgenden für die Parömiologie und Phaseologie beleuchtet werden.

Eine wesentliche Voraussetzung dieses recht radikal anmutenden Versuchs war die von Richard Wossidlo geleistete Verschlagwortungspraxis, die grundsätzlich rekonstruierbar ist. Es ist Siegfried Neumann zu verdanken, dass er die Sprichwort- (und im Besonderen die Sagwort-) Überlieferung Mecklenburgs erforschte und der internationalen Parömiologie durch zahlreiche Editionen, Aufsätze und Lexikoneinträge (besonders für die "Enzyklopädie des Märchens") zugänglich machte.³ Ergänzt durch eigene Erhebungen, bot für ihn die Sammlung Wossidlos ein unerschöpfliches Quellenreservoir, die von ihm zu dechiffrieren war. Ihm, dem lieben Kollegen und vormaligen Leiter der Wossidlo-Forschungstelle, dessen Amt ich nachfolgen durfte, ist dieser Beitrag anlässlich seines 85. Geburtstages gewidmet.

2. Elektronische Parömiographie und Computational Folkloristics

Bieten digitale Archive schon per se einen Mehrwert, indem man sie wegsparend rund um die Uhr und (im Falle öffentlich geförderter Projekte) unentgeltlich konsultieren kann, liegt doch ihr tieferer added value in ihrem Einsatz als Forschungsmittel für innovative Navigationsformen, Retrieval- und Data Mining-Szenarien sowie für (z. B. geobasierte) Visualisierungen. In der Wissenschaftsgeschichte führte die Entwicklung neuer Werkzeuge bekanntlich zu neuen Forschungspraktiken, die neue Sichtweisen ermöglichten und schließlich Theorien revidierten. Anliegen und erste Ergebnisse der noch jungen Teildisziplin der sog. Computational Folkloristics haben in den USA bahnbrechend Tim Tangherlini von der University of California (Abello/Broadwell/ Tangerlini 2012) und in Europa Theo Meder vom Meertens-Instituut (Meder 2014) formuliert. Dabei reichen die Ansätze von der semiautomatischen Klassifizierung von Volkserzählungen (Meder 2016) über die geobasierte Auswertung von Repertoireanalysen (Tangherlini 2010) und von Erzählelementen.wie Handlungsorten (Broadwell/Tangherlini 2016), bis zum Topic Modeling, bei dem semantisch verwandte Wörter aus großen Textkorpora herausgefiltert werden. Um das Retrieval von Erzählmotiven über thematische Beziehungen zu beschleunigen, wurde mit dem *Meertens online finder (MOMFER)*⁴ eine neuartige Suchmaschine geschaffen. Ein Ziel ist es ebenso, durch künstliche Intelligenz gestützte Übersetzungsverfahrungen für multilinguale Präsentationen von Erzähldatenbanken nutzbar zu machen.

Die heutige, durch Markupsprachen wie XML fortentwickelte Technik ermöglicht es, korpusbasiert, das heißt aus großen Textmengen, nicht nur Einzelwörter zu erkennen (sog. *Tokenisierung*), wovon die Verschlagwortung für Suchzwecke (keywording) profitiert, sondern überdies Regelmäßigkeiten und Ähnlichkeiten des Auftretens bestimmter Elemente zu markieren (pattern recognition).

Unabdingbar für die Erforschung von plot grammars ist die Identifizierbarkeit regelmäßig aufeinanderfolgender Handlungselemente. Für die computative Folkloristik und Narratologie steht Propps morphologisches Modell für russische Volksmärchen (Syntagma von 31 Funktionen) und das in es integrierte Rollenmodell (7 dramatis personae und mit ihnen korrelierende Funktionen) Pate, für dessen automatisierten Einsatz bereits mehrfach Datenmodelle entwickelt und implementiert wurden (Yarlott/Finlayson 2016). Ziel solcher Arbeiten im weiteren Sinne ist es, per *machine learning* ein wirkmächtiges computatives Werkzeug für die künstliche Erkennung und Generierung formaler narrativer Repräsentationen schaffen zu können. Die Identifizierung von Morphemen und ihrer Verkettungen ist längst ein wesentlicher Teilbereich der Computerlinguistik, die sich dies für semantische Analysen, die Generierung neuer Texte und für multilinguale Präsentationen zunutze macht.

Für die Parömiologie und Phraseologie spielt die Identifizierbarkeit syntagmatischer Relationen im Hinblick auf die Häufigkeit der Verbindung mehrerer Wörter eine entscheidende Rolle. Indem die Festigkeit und Varianz von Mehrwortverbindungen untersucht wird, kann Sprache insgesamt erforscht werden (Steyer 2012: 8). Die Untersuchung der Festigkeit bestimmter Wortverbindungen, sog. Phraseologismen, in Bezug auf das gesamte Sprachsystem, entwickelte sich zu einem derzeitig sehr nachgefragten Teilbereich der Linguistik, der Phraseologie. Die kulturwissenschaftliche Relevanz solcher Festigkeiten stand hingegen im Fokus des volkskundlichen Sammelinteresses. Der

Sprichwortforschung volkskundlicher Provenienz ist insbesondere der an stabile Mehrwortverbindungen oft geknüpften Bildhaftigkeit und ihrer Bedeutungsfunktionen im Alltag, in Abhängigkeit vom Gebrauch, gelegen. Sprichwörtliche Redensarten und durch ihre Satzwertigkeit davon abzugrenzende Sprichwörter, stellen in der Folkloristik eine von mehreren kulturellen Ausdrucksformen dar, wurden also im Kontext "volksläufiger" Überlieferungen, wie Sagen, Märchen, "Sitten und Gebräuchen" etc. aufgesammelt. Somit wurden auch Übergänge bildhaften Sprachgebrauchs, Sprichwörter und Redensarten als Teil "volkssprachlichen Stils" (Mieder 2007) und "volkstümlicher" Ausdruckskunst erfasst.

Im Gefolge der Disziplinentwicklung erwies sich dann rasch die besondere Komplexität und Aufgabenvielfalt der Sprichwortforschung von der weltweiten Dokumentation bildhaften oder indirekten Ausdrückens in den Sprachen bis zur Theoretisierung der Formenvielfalt kreativer Wortschatzverbindungen, sodass die Übergänge zur Sprachwissenschaft, besonders in der Sprichwortsemantik und Sprichwortpragmatik, fließend geworden sind. In deutschsprachigen Ländern haben Sprach- und Kulturwissenschaft. Dialektologie/Regionalsprachenforschung und Volkskunde/Europäische Ethnologie trotz vielbeschworener Interdisziplinarität ihre gegenseitige Anschlussfähigkeit verloren, während Parömiologie und Phraseologie eine wegweisend fruchtbare Ehe eingingen; Proverbium legt davon eindrücklich Zeugnis ab. Grob fahrlässig wird im deutschsprachigen Vielnamenfach der Europäischen Ethnologie/Kulturanthropologie/Volkskunde etc. die Erzählforschung und somit auch die Parömiologie behandelt. Spätestens seit dem Abschluss des akademiegeförderten Langzeitprojekts der "Enzyklopädie des Märchens" ist es um die deutschsprachige Erzählforschung still geworden. Auch in skandinavischen Ländern, den Pionieren der Erzählforschung, ist teils Ähnliches zu beobachten. Das wirkt sich unmittelbar auf den digital way aus, der für die fachspezifischen Sammlungen allenfalls zögerlich beschritten wird. Denn für deren Entwicklung und damit auch Pflege sind die Fachwissenschaften in erster Linie selber verantwortlich. Verschieben sich aber die Forschungsprogramme und disziplinären Interessen, gilt das Sprichwort: "Aus den Augen, aus dem Sinn".

Als Mediator entfremdeter Wissenschaftsdisziplinen könnten die *Digital Humanities* dienen. *Tradition Archives* fungieren als Stabilisatoren disziplinärer Identität. Ist ein Fach bedroht, wird die Fachgeschichte bemüht, und man erinnert sie wieder an jene Säulen, auf denen man ruht. In einer sammlungsintensiven Wissenschaft, wie der Volkskunde, hilft dann der Verweis auf "ihre" Museen und Folklorearchive aus der Erklärungsnot. Die digitale Verfügbarmachung von Folklorearchiven und damit auch von Sprichwörtern und Redensarten einer Sprachgemeinschaft ist daher für kleinere Fächer auch von wissenschaftsstrategischer Bedeutung.

Wie steht es nun mit dem *state of the art* der ektronischen Parämiographie,⁵ was kann die noch sehr junge <u>Computional Folkloristics</u> zu ihr beitragen, und wie können sich beide gegenseitig befruchten? Suchanliegen können zunächst zeit- und platzsparend durch "Klassiker", per kostengünstiger CD etwa durch Lutz Röhrichs *Lexikon der sprichwörtlichen Redensarten* oder per *free access* durch Wanders *Deutsches Sprichwort-Lexicon*, das im Rahmen des Trierer Wörterbuchnetzes präsentiert wird. Gleichwohl bringen die Druckwerke letztlich denselben Erkenntnisgewinn. Anders verhält es sich mit Datenbanken zu Analysezwecken.

Ein Meilenstein der Vernetzungsarbeit von Sprichwortausgaben aus aller Welt über key words ist die noch immer stabile Online-Sprichwortdatenbank The Matti Kuusi International Database of Proverbs,6 die der berühmte finnische Parömiologie mit seiner Tochter Outi Lauhakangas auf Basis des von ihm erstellten internationalen Typenverzeichnisses (Kuusi 1972) begann, und die von ihr fortgeführt wurde (Lauhakangas 2001).

Als Pionierin des Aufbaus hypermedial konzipierter parömiographischer Onlinedatenbanken im deutschsprachigen Bereich gilt Kathrin Steyer vom Institut für Deutsche Sprache (IDS)⁷, über deren Forschungsfortschritt *Proverbium* regelmäßig berichtet.⁸ Im grenzüberschreitenden Forschungsverbund gilt ihr Spezialgebiet der computativen, korpusempirischen Sprichwortforschung, die auf paradigmatische und syntagmatische Beziehungen von Sprichwörtern und deren Äquivalenzen im Sprachsystem zielt (Ďurčo/Steyer/Hein 2015: 11).

Word patterns, festgefügte Wortverbindungen, können heute aus großen Textmengen mit Hilfe spezieller Algorithmen heraus-

gefiltert werden. Maschinenlesbare Basis für die deutsche Gegenwartssprache bildet das seit 1964 vom IDS aufgebaute Deutsche Referenzkorpus (DeReKo), eine inzwischen über 42 Milliarden (!) Wörter umfassende Sammlung elektronischer Korpora. Aus den Befunden über zentrale Muster des Sprachgebrauchs können Lexika zum rezenten Sprichwortgebauch generiert werden (Stever 2013). Für Lernzwecke (Fremdspachenlernen, Erwerb interkultureller Kompetenzen etc.) wurde von 2008 bis 2010, gefördert durch ein EU-Projekt für Lebenslanges Lernen, unter Federführung der Universität Maribor (Slowenien) die Datenbank Sprich-Wort geschaffen, mit der für die Sprachen Deutsch, Slowenisch, Slowakisch, Tschechisch und Ungarisch Sprichwort-Äquivalente über das Internet frei zugänglich sichtbar gemacht werden.9 Der deutsche Teil dieser Plattform ist in OWID (Online-Wortschatz-Informationssystem Deutsch), das Portal für wissenschaftliche, korpusbasierte Lexikographie des IDS,¹⁰ integriert. Die Einträge sind alphabetisch nach dem Initium der Sprichwörter (in ihrer häufigsten satzwertigen Form) gelistet, die Satzphraseme werden in ihre Basiskomponenten zergliedert und dienen der bestandsübergreifenden Keyword-Suche. Neben Bedeutungsangaben enthalten die Suchergebnisse Angaben über Gebrauchsbesonderheiten. Beeindruckend sind die hierfür beigebrachten Belege von Texten ausgewählter Zeitungen, die im DeReKo stark vertreten sind. Die Plattform soll weiter verbessert werden, sodass man künftig auch über thematische Schlüsselwörter fündig wird. Auch soll ein direkter Zugang zu den Korpusbelegen geschaffen werden.

Da der *DeReKo* die Gegenwartssprache erfasst, werden, so Steyer, inzwischen Überlegungen angestellt, wie an diachrone Informationen angedockt werden kann. Andockstationen wären hierfür sicherlich maschinenlesbare parömiographische Editionen älterer Zeit, die jedoch ihre Datengrundlage mit unterschiedlicher Schärfe sichtbar machen. Wie es auch sei, speist sich die elektronische korpusempirische Sprichwortforschung bislang aus dem Gedruckten, kaum jedoch aus der archivalischen Überlieferung, und auch hier – denkt man an die im *DeReKo* archivierten Zeitungstexte, die als Archiv sui generis angesehen werden können – allenfalls aus publizierten Textsorten, nicht jedoch aus unpubliziert gebliebenen archivalischen Quellen. Editionen mögen den Anschein erwecken, dass sie erschöpfend seien. Wer die ihnen zugrundeliegenden Sammlungen konsultiert, kann sich leicht vom

Gegenteil überzeugen. Autoritative Formate, wie Lexika, sind von unschätzbarem Wert, zurren jedoch das Wissen auf einem bestimmten Gebiet für lange Zeit fest, so nicht neue Erhebungen, deren Auswertung und Repräsentation von der Wandelbarkeit des früher einmal Dargestellten zeugen.

Besucht man die volkskundlichen Archive zur immateriellen Überlieferung, wird man feststellen, wie (notwendig) ausschnitthaft, weil unendlich mühsam, sich das Geschäft des Edierens vollzieht. Archive sind, genauer betrachtet, keine Wissensspeicher, sondern Aufbewahrungsorte für Informationen, die aus dem Leben gerissen sind. Der Archivnutzer kann sie nur dadurch wiederbeleben, indem er sie in seinen eigenen Wissenshorizont einbettet, also subjektiv bewertet.

Zu Wissen gerinnen Informationen nur, wenn damit ein bestimmter Nutzen bezweckt werden kann. Die semantische Dimension von Daten wird erst akut, wenn sie zweckbezogen zu Informationen interpretiert werden. Wissen ergibt sich erst aus begründeten, miteinander in Beziehung gesetzten Informationen (Fuchs-Kittowski 2010: 37). Und dieser Nutzen weicht nicht unerheblich vom Anliegen derjenigen ab, die einst ihre Sammlungen zur kulturellen Überlieferung generierten. Für die Erhellung der Wissensproduktion, so auch über Sprichwörter und Redensarten, ist es daher unabdingbar aufzuzeigen, wie Lexika, Editionen u.a.m. aus den ihnen zugrundeliegenden archivierten Sammlungen entstanden sind - was nicht allein durch theoriefundierte Postulate, sondern vielmehr durch detektivische Spurensuche erreicht werden kann. Bis hin zum rezenten Sprichwortgebrauch stellen Folklorearchive keine geringe Schaltstelle im Wissenskreislauf darf. Mit anderen Worten: Nicht nur ready made folklore, sondern folklore in making ist sichtbar zu machen. Für die Parämiographie könnte eine so verstandene Computational Folkloristics dazu beitragen, die Entwicklung und Änderung von Feldforschungsstilen. Ordnungsmustern, Archivierungs- und Repräsentationspraxen zu erhellen (Schmitt/Tangherlini 2018: 187).

3. Collection-centered approach: Richard Wossidlos volkskundliche Sammlung und ihre Transformation zum Online-Archiv WossiDiA

Die Ethnografie volkskundlicher Provenienz entwickelte sich als sammelnde, Überlieferungsforschung treibende Wissen-

schaft. Sie klassifizierte und beschrieb Mündlichkeitskulturen und damit ein Gebiet, auf dem die Philologie bislang ungeübt war, da letztere ihr Methodenarsenal an skripturalen Texten entwickelt hatte. Im Fokus landschaftsgebundener Volkskunden waren überlieferte regionalkulturelle und damit auch regionalsprachliche Muster, wobei die Dokumentation von Festigkeit auch deren Varianz einschloss. Da Volkskundler zugleich – wenn auch unterschiedlich intensiv und umfangreich – den arealen mundartlichen Sprachschatz miterfassten, waren Sprichwörter, sprichwörtliche Redensarten, Routineformeln, Wellerismen, Gemeinplätze und überhaupt festgeprägte Sätze bzw. feste Phrasen, "Reime und Sprüche", wesentlicher Teil ihrer Dokumentation.

Wertvoll waren sammelnde Volkskundler nicht nur, weil sie verstreute Artefakte der Anschauung willen zusammenzutrugen und bewahrten,¹² was Grundlage jeder Forschung ist, sondern weil ihre Daten per Feldforschung gewonnen wurden.

Volkskundler griffen - entweder selber oder mit Hilfe von Sammelhelfern ("Gewährsleuten") – mitten ins Leben, genauer gesagt, in die Erinnerung besonders älterer Informanten. Diese wurden von ihnen bevorzugt aufgesucht, weil es galt, die in der Hochindustrialisierung schwindenden Überlieferungen noch festzuhalten ("Rettet, was noch zu retten ist!"). Die mitpraktizierte Sprachsammlung war zwar insofern ebenfalls rückwärtsgerichtet, weil bei der "Jagd" kaum mehr übliche Wörter und Phraseologismen präferiert wurden. Doch ist ein einmal angeeigneter Sprachgebrauch wohl in geringerem Grad von der Verblassung bedroht, als etwa die Erinnerung an Einzelheiten eines längst nicht mehr praktizierten Brauchs. Vor allem konnten die Informanten als Sprecher von Sprichwörtern aus eigener Anschauung beobachtet werden. Daher berücksichtigen die per Feldforschung generierten Sammlungen "volkstümlicher Überlieferungen" den noch gesprochenen Usus von Sprichwörtern und Redensarten, der in Abhängigkeit von der Qualität dokumentierter Metadaten rekonstruiert werden kann. Bei genauerem Hinsehen wird man entdecken, dass diese nicht kontextlos akkumuliert wurden.

Richard Wossidlos Sammelunternehmen, das der mecklenburgische, in Waren an der Müritz als Gymnasiallehrer tätige Privatgelehrte im Zeitraum von 1884 bis kurz vor seinem Tode im Jahre 1939 fast unentwegt am Leben hielt, ist für die Parömiologie und Phraseologie aus mehreren Gründen besonders ergiebig:¹³

Erstens war Wossidlo gelehrter, philologisch trainierter Sammler, der in seinen Studienjahren in Rostock, Leipzig und Berlin so bedeutsame Gelehrte wie Richard Förster, Carl Robert, Karl Müllenhoff, Wilhelm Scherer, Theodor Mommsen u.a.m. gehört hatte (Schmitt 2011: 88 ff.). Die Generation dieser Gelehrten revidierte bereits das Grimmsche Vorbild, indem sie auf die Autopsie, die eigene Überprüfung von Befunden, Wert legte und vor vorschnellen und sprunghaften Interpretationen warnte.

Zweitens war der Volkskundler zugleich in hohem Maße Sprachsammler, der nicht einfach Wörter auflistete, sondern diese nach semantischen und pragmatischen Relationen gruppierte. Aufgewachsen im heimischen Dialekt und überzeugt von der sprachlichen Kreativität bildungsferner niederer sozialer Schichten, richtete Wossidlo seine Aufmerksamkeit auf den bildhaften Sprachgebrauch, daher auch auf Sprichwörter und sprichwörtliche Redensarten und überhaupt auf den "volkssprachlichen" Stil. Bemerkenswert ist, dass er ganz im modernen Sinne die pragmatische Funktion tradierter, also festgefügter Wortverbindungen erfasste. Phraseme werden von ihm vielfach mit situativen Verwendungen notiert, sodass sprachliche zu außersprachlichen Faktoren in Beziehung gesetzt werden (Schmitt 2014: 271).

Bei der Umsetzung des nicht unbescheidenen Wunsches, eine Synonymik des Niederdeutschen zu erschaffen, wurden insofern auch Regelmäßigkeiten situativen Sprachgebrauchs registriert. Die ersten Jahre seiner Sammelarbeit, die Zeit von 1884 bis 1895, waren noch ganz der Inventarisierung des mundartlichen Sprachschatzes gewidmet. Seine ab 1891 im Auftrag des Vereins für mecklenburgische Geschichte und Altertumskunde einsetzende systematische "volkspoetische" Sammelarbeit hat jenes linguistische Erkenntnisinteresse, das im Sinne des damaligen Weltbildgedankens der Erhellung des Empfindens und Denkens seiner Informanten diente (Schmitt 2014: 266-271), niemals verlassen. Zeugnis legen hiervon seine "Mecklenburgischen Volksüberlieferungen" ab. So findet sich in Bd. 2 der Reihe, den Tieren im Munde des Volkes, ein längerer Abschnitt über apologische Sprichwörter (Wossidlo 1899: 29-42).

Drittens war Wossidlo in seinen beiden letzten Lebensjahrzehnten unmittelbar in die Lexikographie verwoben. Seine Sammlung bildet weithin die Quellengrundlage für das großlandschaftliche *Mecklenburgische Wörterbuch* (MWB, 1942-1992). Hierfür leistete er eine ganz wesentliche Vermittlungsarbeit, indem er seine sachsystematisch geordnete Sammlung und den wortfeldartig gruppierten Sprachschatz für das Mundartwörterbuch in die alphabetische Ordnung übertrug.

Viertens galt sein effektiv organisiertes Sammelunternehmen aufgrund der Kombination mehrerer Erhebungsmethoden für damalige Verhältnisse als mustergültig. Vera Deißner bewertet Wossidlo als seinerzeit angesehendsten Fachmann auf dem Gebiet der Forschungstechnik (Deißner 1997: 221), welcher der sich formierenden scientific community über seine praktischen Erfahrungen Rechenschaft ablegte (Wossidlo 1906). Seine Methoden- Kombination beruht auf folgenden Faktoren: der eigenen Feldforschung, die zumeist in Form von Mehrtagesreisen durchgeführt wurde; auf der Schaffung eines vergleichsweise stabilen Netzwerks von Sammelhelfern, die in seinem Auftrag ihr lokales Umfeld ethnografierten und die Erträge an Wossidlo sandten; und last but not least auf der Kollektion ethnografischer und lexikografischer Parallelen per ausgezogener Schriften aus dem gesamten deutschsprachigen Raum.

Anlässlich der Gründung des Folkloristischen Forscherbundes bezeichnete Kaarle Krohn, nachdem er in Waren die Sammlung besichtigt hatte, Wossidlo sogar als "größten Sammler deutscher Volkskunde" (Krohn 1907) und stellte ihn in eine Reihe mit dem dänischen Lehrer Evald Tang Kristensen und dem dänischen Pfarrer Jakob Hurt (Krohn 1907: 1). Archer Taylor suchte Wossidlo im Sommer 1925 auf, nachdem er die finnischen und estnischen Kollektaneen in Augenschein genommen hatte. Auch andere bedeutsame Gelehrte, wie Walter Anderson, statteten der Sammlung, die damals noch nicht institutionalisiert war, einen Besuch ab.

Fünftens wird die Sammlung durch ein ausgeklügeltes Zettelkastensystem repräsentiert, in dem die verschiedenen Bestandsgruppen, also eigene Feldforschungsnotizen, Korrespondentenbelege, extrahierte Publizistik und Umfragen tiefenindexiert und thesaurusartig miteinander verlinkt sind. Dieses System wurde recht früh entwickelt und – bedenkt man die schier unend-

liche Fülle von Aufzeichnungen und die zeitliche Ausdehnung des Sammelunternehmens – treu bis zum Lebensende (in praxi freilich nicht immer konsequent) durchgehalten. Fast könnte man von einem selber auferlegten Algorithmus sprechen (vgl. Schmitt 2019). Für die Datenerhebung und Erschließung entwickelte er ein eigenes Regelwerk mit ausufernden, zumeist lateinischen Abkürzungen, das zeitsparend eingesetzt wurde. Felderträge wurden nach morphologischen Kriterien, typischerweise nach Motiven, zerpflückt, und übereinstimmende Inhalts- und Formmuster zu Variantenreihen gebündelt. Nicht selten werden auch Aspekte des Kommunikationsprozesses, so Bewertungen von Informanten zu ihren Äußerungen, oder etwa Aspekte der Herstellung von Objekten, von Werkzeugen, Brauchrequisiten etc., miterfasst

Typisch ist für Wossidlo die Form verdichteter Notation auf kleinformatigen Zetteln (Sprichwörter und Redensarten vertragen sich besser mit dem Zettelformat, doch ist weiterer Platz für Bedeutungs- und Gebrauchsangaben in Betracht zu ziehen). Wossidlos Entscheidung, die Überlieferungen einer Landschaft in toto, und dazu den in ihr (schwindenden) Wortschatzgebrauch aufzunehmen, zwang ihn zur Zeitersparnis. Ausführliche Niederschriften konnten somit nicht angefertigt werden. Notiert wurde daher bei einer Ausdrucksform, die bereits an anderer Stelle beschrieben wurde, nur Abweichungen von einmal konzipierten Mustern sowie das Vorkommen an sich (per Metadaten: Ort und Zeit der Aufnahme, Name und ggf. Herkunftsort des Erzählers/der Erzählerin, ggf. Beruf und Alter des Informanten gesondert per Tagebuchnotizen).

Sein Feldforschungsstil und die Form seines *Note-Taking* erinnern an die Tradition gelehrter Exzerptkultur, bei der ausgewählte Passagen aus Druckschriften "gezogen" und in Merkbücher übertragen wurden, wobei sich im Feld die Leichtigkeit von Zetteln bezahlt machte. Die Methode des Exzerpierens (*methodus excerpendi*) verlangt das Abstrahieren vom Gelesenen und lenkt den Prozess der Materialanordnung (Te Heesen 2006: 27). Wossidlo übte sie ohnehin regelmäßig beim Durchsuchen gedruckter Quellen und Kommentare. Wir können heute leicht ganze Bücher kopieren, Informantengespräche elektronisch aufzeichnen und durch spätere, iterative Bearbeitung ausschnitthaft Tiefen ausloten. Man bedenke jedoch, wie viele Informanten,

wie viele kulturelle Ausdrucksformen, wie viele Wörter und deren sprachliche Muster, wie viele Befunde anderer hier ein einzelner in seinem Leben erstellte, verschlagwortete, anordnete und miteinander verkettete.

Zettel stellen ein dynamisches Wissenssystem dar. Man kann die beweglichen kleinformatigen Papiere umsortieren und für verschiedene Zugangsarten (Inhalt, Motiv, Ort, Sprachliches etc.) duplizieren. Man kann mit ihrer Hilfe gut referenzieren. Sie zwingen zu Auslassungen, damit der Informationsgehalt dicht gepackt werden kann. Durch die "vererbende" Verschlagwortung von Befunden mit weiterem und engerem Begriffsinhalt können Redundanzen (physisch durch das box-in-the-box-in-the-box-Prinzip: Zettel in Konvolut in Kasten in Regaleinheit) vermieden werden, weshalb Zettelinhalte nur gemeinsam mit ihrer Stellung im hierarchisierten Begriffsinventar dekodiert werden können.

Zettelkastensysteme können als mediale Vorform computativer Arbeitsweise betrachtet werden. Dies war einer der Gründe, weshalb sich Mitarbeiter des Rostocker Lehrstuhls für Datenbank- und Informationssysteme herausgefordert sahen, das dynamische Zettelkastensystem und seine sammlerspezifischen Eigentümlichkeiten für eine onlinefähige Graphdatenbank zu modellieren. Schon der bloße Digitalisierungsworkflow war wegen der Fragilität und des variierenden Formats vieler Zettelbelege und ihrer hohen Stückzahl (produziert wurden über zwei Millionen Scans) für das Projektteam belastend, wiewohl eine Berliner Firma als Dienstleister beauftragt wurde. Zudem wurden die Farbscans auf 36mm-Film ausbelichtet und im "Oberrieder Stollen", dem zentralen Bergungsort für verfilmtes ("national wertvolles") Kulturgut, eingelagert, welche Kosten anteilig das Bundesamt für Bevölkerungsschutz und Katastrophenhilfe (BBK) übernahm. Um aus Sicht der Informatikwissenschaften die Kleinteiligkeit, Vielfalt und Vielzahl der individuellen Zettelbelege, ihren lückenhaften, vielfach erst durch anreichernde Verbindungen zu erschließenden Informationsgehalt und den Verweischarakter der Sammlung bestmöglichst erfassen zu können, entschied man sich für die Implementierung sog. Hypergraphen.

Graphen sind ideal für Repräsentationen, die nicht nur Objekte (sog. "Knoten"), sondern die Beziehungen zwischen ihnen

(sog. "Kanten") erfassen (Meyer/Schering/Schmitt 2014: 72 ff.). Sie bieten sich daher ebenso für die Modellierung von Sätzen an, indem Prädikate durch Kanten sowie Subjekte und Objekte durch Knoten repräsentiert werden. In sog. "Hyperkanten" können Objekte mengenartig zusammengefasst werden, z.B. mehrere *keywords*. Zudem können Knoten typisiert und Kanten attribuiert werden.

Hängt man mehrere Stichwörter an eine einzige "Inhaltskante", wird, bezogen auf einen Beleg und seine Einbettung in das Datengeflecht, das Beisammensein von Stich- und Schlagwörtern miterfasst, also die besondere Verbindung eines keywords mit anderen, und zwar mit weiterreichenden Referenzen auf Bedeutungsgehalte, Quellen und ihre Generierung. Auf ähnliche Weise könnten Phraseme gewinnbringend markiert werden. Was bruchstückhafte bzw. verschwommene (fuzzy) Informationen anbelangt, können Lücken im jeweiligen Bezugsfeld erkennbar gemacht, im weiteren Verarbeitungsprozess mitgenommen und per sog. Matching ggf. ganz oder teilweise geschlossen werden.

"Hypergraphen" bieten dazu ein mehrdimensionales Modell, in dem theoretisch unendlich viele Hyperkanten angedockt werden können, heutige Speicherkapazitäten und Verarbeitungsgeschwindigkeiten setzen auch der Praxis kaum Grenzen. In WossiDiA sind derzeit nicht weniger als 74 Hyperkantentypen definiert, die alle zueinander in Beziehung gesetzt werden können. Ein häufig benutzter Hyperkantentyp lautet: erzählt mündlich, ein alternativer dazu: erzählt schriftlich. Ist das Medium unbekannt, wählt man den verallgemeinernden Hyperkantentyp erzählt. Dem Hyperkantentyp erzählt mündlich können in Wossi-DiA etwa folgende Knoten zugeordnet werden: Notationsquelle (Zettelbeleg Wossidlos; Seite der Korrespondenz eines Sammelhelfers; Wörterbuchzettel; Transkription; Sprache), Aufnahmedaten (Ort und Zeit der Aufnahme: Name, Herkunftsort und Rolle, d.h. Beruf, des Erzählers), Wortabkürzung, Schlagwort u.a.m. Der Hyperkantentyp "erzählt schriftlich" gilt für das Korrespondentennetzwerk Wossidlos, das über tausend verschiedene Personen umfasst (die Zahl der Sammelhelfer im eigentlichen Sinne, die also über einen längeren Zeitraum schriftlich zuarbeiten, ist freilich geringer). Da die Sammelhelfer vielfach auch mündliche Quellen in Verbindung mit Namen und ggf. Orten angeben,

wurde hierfür der besondere Knotentyp "Informant des Erzählers" entwickelt.

Literaturexzerpte, hierfür verwendete idiosynkratische Abkürzungen, die dazugehörigen bibliographischen Angaben sowie die Referenznummer für Verbundkataloge (und damit auch der Standort eines ausgezogenen Exemplars in der noch erhaltenen Gelehrtenbibliothek Wossidlos) können ebenso eingetragen werden. Dargestellt werden können ebenso ganze Verarbeitungsketten, vom Feldforschungszettel oder einer Korrespondenzseite über in Reinschrift übertragene Belege für Publikationszwecke und (mehrsprachig ausführbare) Transkriptionen bis zur gedruckten Seite (sofern publiziert wurde und das Publizierte, wie die Editionen Wossidlos und ausgewählte sammlungsbezogene Ausgaben Siegfried Neumanns) in WossiDiA eingepflegt wurden.

Auch ohne Informatikkenntnisse mag sich ausmalen, dass ein solches Modell geeignet ist, die Idiosynkrasie einer solchen "Zettelwerkstatt" (Schmitt 2015:197) abzubilden.

Die inzwischen in der WossiDiA-Datenbank erprobte Weiterentwicklung von Hypergraphen (Meyer/Schering/Heuer 2017) ermöglicht es, mit vergleichsweise geringem Implementierungsaufwand proprietäre Eigenschaften zu modellieren. Diesen, man möchte fast meinen, "Quantensprung" für die Darstellung von Vernetzungsstrukturen, könnte sich ebenso die elektronische Parömiographie zu Nutze machen. Kleinteiligkeit von Mehrwortverbindungen bildet hier den Ausgangspunkt, die freilich einen hohen Vernetzungsgrad bedingt. Die Vielschichtigkeit der Beziehungen des Sprichworts – bedenkt man allein seine weite Verbreitung, seine Polysituativität, Polyfunktionalität und Polysemantizität (Mieder 2014: 1101) – wäre prägnanter darstellbar.

4. Suchszenarien und vernetzte Darstellung von Sprichwörtern in WossiDiA

In WossiDiA kann man zum einen die Bestände in der Reihenfolge ihrer Anordnung wie im "realen" Archiv "durchblättern". Zum anderen kann man per "Hypergraphsuche", gepaart mit einem Filtermechanismus für Bestandsgruppen und Metatypen (facility search), ein Schlag- oder Stichwort, einen Personenoder Ortsnamen, einen Buchtitel, eine Abkürzung u.a.m. suchen. Im Folgenden sollen einige stichwortbezogene Retrieval-

Ergebnisse und ihr Verknüpfungsbild erläutert werden. Beispiel eines Suchszenariums seien im Folgenden Sprichwörter für "krank" und "Krankheit". Da Sprichwortbefunde von Wossidlo im semantischen und pragmatischen Zugang verzettelt wurden, wird man schnell über die gattungsspezifische Filtersuche fündig. Neben Sprichwörterbündeln zu Arbeit, Kleidung, Geld, Tanz, Tod, Umgangsformen von Menschen untereinander, Eigenschaften wie klug, dumm oder verrückt, gelangt man ohne größeren Zeitaufwand zum Konvolut mit der Bezeichnung "Gesundheit, Kranhkeit: Sprw. [für Sprichwort, aufgelöst per Verweis]". 14

Das physisch aus blauem Schulheftkarton gefertigte Konvolut und seine Etikettierung wird mit angezeigt. Nun kann man sich in aller Ruhe, ohne den Druck archivalischer Schließzeiten, die dort enthaltenen 64 Belege anschauen. Auf den ersten Blick erscheinen die zumeist handschriftlichen Aufzeichnungen kryptisch. Geduld sollte man mitbringen, auch um die Erklärungen des digitalen Systems über seine Handhabung durchzulesen. Die Usability ist allerdings noch zu verbessern und nutzerfreundlicher zu gestalten. Vieles ist selberklärend und kann durch das spielerische Verfolgen von Verlinkungen leicht erlernt werden. Mit Vorwissen ausgestattete Forscher werden sich noch schneller zurechtfinden. Nach der selbsttätigen Einarbeitungszeit staunt man dann doch über die Logik des Wossidlo'schen Note-Taking und des Zettelkastensystems.

Das Konvolut unseres Beispiels enthält Feldforschungsbelege. Exzerpte aus dem Gedruckten zum gesuchten thematischen Schlagwort findet man in gesonderten Bündeln. Beim Durchblättern stoßen wir etwa auf einen Beleg mit dem vollständig notierten Satzphrasem: gnarren wagens gahn am längsten (siehe Abb. 1). Es zeigt einen von Wossidlo im Feld erstellten Zettel. Während Sprichwörter und Redensarten in WossiDiA bislang über den sammlerspezifischen, von Wossidlo erstellten und im WossiDiA-Projekt überarbeiteten Index gefunden werden können, wurde beispielhaft folgender Zettel mit Hilfe des hypergraphbasierten Eingabewerkzeuges aufgelöst, verschlagwortet und vernetzt. Wossidlos Indexierung per Eintrag in der oberen rechten Ecke ("krank") erschließt sich navigierend über die baumartige Anlage der Sammlung.

Neu eingegeben wurden mit Hilfe der Hyperkante "erzählt mündlich" die unten auf dem Zettel vermerkten Aufnahmedaten: Aufnahmeort ist das am Fluss Warnow gelegene kleine Dorf Warnow bei Bützow in Mecklenburg. Informantin ist "Fräulein" Ida Rumpf, Wossidlo notierte ihre Aussage am 18. August 1935 vor Ort. Das Sprichwort wurde nicht erfragt, sondern "ss" [= sua sponte], auf eigenen Antrieb beiläufig von der Informantin mitgeteilt. Per "Inhalts-Kante" wurden die Wörter gnarren (als lautmalerischer Anker des Sprichworts) und Wagen vergeben. Als Bedeutungsgehalt wurde in dieselbe Hyperkante das neu vergebene Schlagwort "alt werden" gepackt. Angehangen wurden bilingual (in hochdeutscher Standardsprache und in ostniederdeutschem Dialekt) erstellte Transkriptionen.

Klickt man auf den Namen der Informantin, erhält man sämtliche, von ihr beigetragenen Informationen. Da Wossidlo seine Sammlung gattungsspezifisch (wenn auch nicht immer konsequent) verschlagwortete, erhält man eine Vorstellung vom "Repertoir". Die vorliegende Informantin ist jedoch keine eigentliche Erzählerin. Auch korrespondierte sie mit dem Hauptsammler, wodurch man mehr über sie erfahren kann. Per Klick auf das Aufnahmedatum enthüllt sich (perspektivisch, da längst noch nicht alle Aufnahmedaten eingetragen sind) die Verkettung mit allen weiteren, an diesem Tage aufgenommenen Feldforschungsdaten, sodass ebenso der Verlauf von *field trips* rekonstruiert werden kann (wofür Wossidlos skizzenhaft geführten Feldforschungstagebücher weitere Anreicherungen liefen). Analoges gilt für den Aufnahmeort und für die vergebenen *Keywords*.

Klickt man in der Linkliste auf das Wort gnarren (siehe Abb.1), gelangt man zu den Belegen, aus denen das Mecklenburgische Wörterbuch generiert wurde, hier zu den von Wossidlo aus seiner Sammlung übertragenen Wörterbuchzetteln (Bestand "MWW"). Im Folgenden soll gezeigt werden, auf welche Weise sich der Eintrag in das großlandschaftliche Wörterbuch und der damit verbundene Reduktionsprozess vollzog, und wie gewinnbringend es ist, die zugrunde liegenden Wörterbuchzettel und ihre "Urbelege" zu befragen. Das Verb gnarren steht unter dem Lemma Gnarr (MWB 3, Sp. 205-206), erklärt als wimmern und quäken ... besonders von weinerlichen Kindern. Unter den zahlreichen phraseologischen Belegen findet man sprw. von

kränkelnden Menschen: gnarren Wagens fäuhren am längsten HARed. Das diatopische, wortgeographisch angelegte Wörterbuch zitiert nur den Ort, nicht jedoch die Zeit der Aufnahme. Vor allem ist der Sprecher samt seiner Sozialdaten verschwunden.

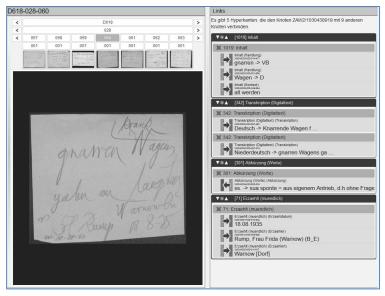


Abb. 1: Zettelansicht eines von Wossidlo notierten Sprichwortes. Darüber die Topologie, rechts die Verlinkung per fünf Hyperkanten mit neun Knoten

Bei der Selektion wurde die Aussage der weiblichen Sprecherin nicht aufgenommen, sondern die eines Büdners aus Redefin (das der Region Hagenow [HA] zugeordnet ist, abgekürzt als "HARed"). Sowohl auf dem thematisch verzettelten "Urbeleg" vom 15. August 1919 (siehe Abb. 2) als auch auf dem davon abgeleiteten Wörterbuchzettel (siehe Abb. 3) wird der Verwendungskontext genauer erklärt (*Wenn Frauen klagen*, 15 von kränkelnden Frauen 16). Hinter dieser miterwähnten Erklärung des Sprichwortes verbirgt sich ein misogynes Einstellungsmuster, das die Wörterbuchkonstrukteue womöglich absichtlich unter den Tisch fallen ließen.

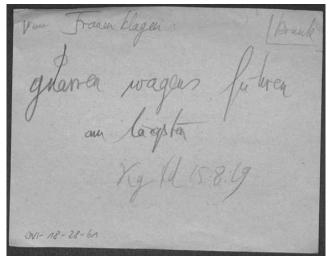


Abb. 2: Sprachlicher Feldforschungsbeleg Wossidlos mit dem topic keyword "krank"

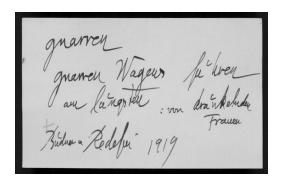


Abb. 3: Davon abgeleiteter Wörterbuchbeleg mit Alphabetisiertem Zugang

Den älteren mündlichen Beleg bevorzugte man bei der Auswahl vor dem jüngeren und den unteren (indigener anmutenden) sozialen Stand vor dem höheren, den Büdner gegenüber dem gebildeteren "Fräulein", das von sich in geübter Schrift meint: *Ich kann wohl sehr gut plattdeutsch schreiben und lesen.* ¹⁷ Bemerkenswert ist, dass auf dem Beleg der Informantin kein misogyner Bedeutungsgehalt notiert wurde. Hat sich der 1919 notier-

te semantische Gehalt des Sprichwortes bis 1935 gewandelt? Wie alt ist das "Fräulein", handelt es sich um eine Erinnerungslücke? Die Beleglage ist zu dünn, aber es werden bezugsreiche Fragen provoziert.

Verwandtes findet sich im größeren sprachlichen Formenspektrum in den übergeordneten, mit "krank, Krankheit" etikettierten Kästen. Diese umfassen an Exzerpten (vielfach aus der klein- wie großräumlichen Idiotika-Literatur) insgesamt 2.676 Zettel (zugänglich über 65 semantisch-pragmatische Verzeichniseinheiten) und an Feldforschungsbelegen 1.436 Belege (mit 46 Verzeichniseinheiten). 18 Beim gefilterten "Durchblättern" findet man weiter gefasste bedeutungsähnliche Festwortverbindungen und stößt z.B. auf folgendes, per onomatopoetischem Verb geprägtes Satzphrasem: dat knakkt wol äwerst dat brekkt nig mit der Bedeutungserklärung: stirbt nicht gleich wenn man krank ist (siehe Abb. 4). Er ist Dähnerts Platt-Deutschem Wörter-Buch (Dähnert 1781) entnommen und stammt aus der frühen "sprachlichen" Sammelzeit Wossidlos, weshalb man wohl die Seitenzahl vermisst. Mit einer online gestellten digitalisierten Ausgabe dieses basishaften Wörterbuches ließe sich dieser Zettel per Normdatensatz der Pica-Produktionsnummer (PPN) leicht verlinken.

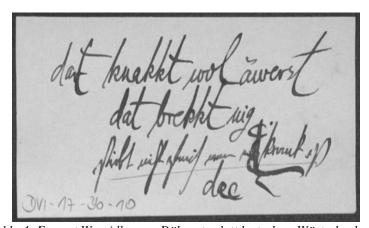


Abb. 4: Exzerpt Wossidlos aus Dähnerts plattdeutschem Wörterbuch

Was die Dekodierung der von Wossidlo verwendeten Literaturabkürzungen anbelangt (im Beispiel: "dae") wurden inzwi-

schen erhebliche Fortschritte gemacht. Die Datenbank verzeichnet aktuell 3.230 Literaturabkürzungen, weitere werden im Verlauf des Erschließungsprozesses aufgenommen. Sie werden sukzessive mit den vollständigen bibliographischen Angaben incl. der PPN-Nummer verknüpft. Da die Wossidlo-Bibliothek noch erhalten ist, können solche Abkürzungen und ihr Auflösungsversuch per exemplarischer Autopsie verifiziert werden. Bedenkt man, dass etwa zwei Drittel der sachsystematisch geordneten Zettelsammlung Wossidlos (ZAW) aus verschlagworteten Exzerpten besteht, soll diese exorbitante Erschließungsleistung des Sammlers zugänglich gemacht werden. Da nicht wenige der älteren Werke inzwischen online gestellt wurden – ein Prozess, der sich lawinenartig fortsetzt –, kann man von WossiDiA aus direkt in die referenzierten Publikationen hineingelangen. Diese thematische Indexierung könnte auch in umgekehrter Richtung genutzt werden. So wurden von Wossidlo sämtliche Werke von Fritz Reuter filigran auf Sprichwörter und Redensarten hin ausgezogen und verschlagwortet, was sich online gestellte Reuter-Ausgaben zunutze machen könnten.

Abschließend noch ein Beispiel zur Sagwortüberlieferung Mecklenburgs, die Siegfried Neumann in dem Band Sprichwörtliches aus Mecklenburg in mühsamer Kärrnerarbeit ediert hat. WossiDiA bietet die Möglichkeit, die dort als "hs" (handschriftlich) mit Ort- und Jahresangabe markierten Belege zu identifizieren. Unter den Sagwörtern mit "Sperling" findet sich der Eintrag: Dat gifft 'n grot Lock", säd' de Lüning, dor süll hei 'n Gosei leggen (Neumann 1996: 404, Nr. U70). Zitiert werden Belegstellen aus publizierten Quellen sowie Feldforschungserträge. Letztere werden per Ort und Jahreszahl der Aufnahme ausgewiesen (z.B. hs. Rostock 1898). Über die Rückverfolgung des Zettels können der Beiträgername und verwandte Einträge ermittelt werden.

Ist der Inhalt des Zettels über den Sperling des Kastens "Apologische Sprichwörter, Tiere" aus Abb. 5 kaum lesbar, so ist doch der Verweis aufgrund seiner eindeutigen Syntax (im Vergleich zu anderen Verweistypen) dechiffrierbar. Er führt zu einem akribischen Beiträger mit gut lesbarer Handschrift, den Rostocker Archivar Ludwig Krause (1863-1924). Ein Blick auf den Ausschnitt der betreffenden Korrespondenzseite (siehe Abb. 6) zeigt, dass Krause das Sprichwort dem damaligen christlichen

Erziehungsdiskurs entnommen hat. Es wurde vermutlich aus der Zeitschrift Weißes Kreuz. Zeitschrift zur Förderung sittlicher Reinheit unter Männern und jungen Männern exzerpiert. Beglaubigt wird, dass es sich hierbei um eine landläufige Redensart handelt. Für eine Autopsie müsste man sich das Heft besorgen, das Erscheinungsdatum (20.8.1898) und der Verfasser des Beitrages werden genannt. Der Screen zeigt die Ansicht der Beiträgerseite im für den Nutzer nicht öffentlichen Eingabemodul.

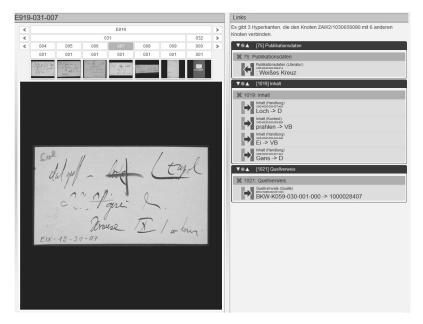


Abb. 5: Zettelnotiz Wossidlos, die auf eine Korrespondenzseite referenziert, mit Links¹⁹

All diese Beispiele sind zufällig gewählt und demonstrieren die Möglichkeiten des in *WossiDiA* implementierten Hypergraph-Modells als detektivisches Forschungswerkzeug. Deutlich wurde dabei ebenso, wie lohnenswert für die Sprichwortforschung ein direkter Korpuszugang zu den Quellen regionalsprachlicher Wörterbücher ist, was insbesondere für das mit volkskulturellen Belegen prall gefüllte *Mecklenburgische Wörter terbuch* zutrifft (Schmitt 2014: 257 ff.). Erleichtert würden solche Recherchemöglich-



Abb. 6: Stelle der referenzierten Beiträgerseite aus Abb. 5, aufgerufen im Eingabemodul. Links sind drei Hyperkanten sichtbar.

keiten, indem auch das Wörterbuch selbst im Kontext des Trierer Wörterbuchverbundes für freien Zugang digitalisiert und mit WossiDiA vernetzt würde, damit man von den Zettelbelegen direkt in das Wörterbuch und vice versa gelangen kann. Dadurch würde ebenso ein hochdeutscher und keyword-gestützter Zugang zu dem siebenbändigen Mundartwörterbuch ermöglicht. Die mithin wesentlichste Voraussetzung wurde hierfür bereits durch WossiDiA geschaffen, das direkten Korpuszugang zu den Zettelkästen des Wörterbuches gewährt: Sowohl Wossidlos, von ihm umgegossene Wörterbuchbelege ("MWW" = 383.953 Zettel) als auch der davon abgeleitete und erweiterte Teuchert'sche Zettelkatalog ("MWT" = 394.984 Zettel) sind in WossiDiA integriert und sachsystematisch (bislang auf den oberen und mittleren Ebenen des Thesaurus) vernetzt.

5. WossiDiA im Verbund mit gattungsspezifischen Folklorearchiven

Eine Voraussetzung für die komparatistische Erzählforschung besteht in der grenzüberschreitenden Klassifikation von Überlieferungen. Für Märchen und Schwänke konnten nationale und regionale Erzählkataloge aufgrund des internationalen Typenindexes von Aarne, Thompson und Uther (Uther 2004) anschlussfähig gemacht werden, für Sprichwörter wurde bereits auf Kuusis Beitrag hingewiesen. Sagen sind, trotz des Wandercha-

rakters vieler Motive, nur schwer zu vergleichende Mikrotraditionen. Vielfach bloß durch kleine Erzählkerne beschreibbar, bilden sie kaum vorhersehbare Kombinationen narrativer Elemente, weshalb ihre Katalogisierung in den 1960er Jahren stecken geblieben ist (Schmitt/Tangherlini 2018: 200).

In dem *Big-Data-*Projekt ISEBEL haben sich die *Nederlandse Volksverhalenbank*, der auf Basis der Sammlung Evald Tang Kristensens erstellte *Danish Folklore Data Nexus* und *WossidiA* mit dem Ziel vernetzt, ihre jeweiligen Sagenbestände für Forschungszwecke übergreifend durchsuchbar zu machen. Hierfür wird derzeit am Meertens Instituut unter der Leitung von Theo Meder ein gemeinsamer *Harvester* entwickelt. Ein wesentliches Ziel ist die Schaffung eines multilingualen Zugangs. Mit Hilfe von *Google translate* sollen (im Hintergrund, für den Nutzer unsichtbare) *dirty translations* für Suchzwecke geschaffen werden, ein Forschungsfeld, dem sich besonders Tim Tangherlini verschrieben hat.

Die Datenbanken sind unterschiedlich strukturiert, wobei die Ähnlichkeiten zwischen der dänischen und der nordostdeutschen Sammlung am größten sind, auch wenn Kristensens und Wossidlos Feldforschungsstile differieren (Schmitt/Tangherlini 2018: 187-195). Und das ist gut so – kommt es doch ebenso darauf an zu zeigen, wie verschiedene Datenbanken, trotz ihrer jeweiligen Individualität, "abgegrast" werden können. Ihre Struktureigentümlichkeiten sollen bestehen bleiben, es kommt nur darauf an, Brücken zu bauen. Dies geschieht in WossiDiA über die Implementierung einer OAI-Schnittstelle in Anlehnung an den (leider eingeschränkten) Regelsatz der Open Archives Initiative. Ein Prototyp, der ein Minimum an Anschlussmöglichkeiten garantiert, läuft bereits (Meder 2018). Im Verlauf des Projekts sollen verfeinerte Vernetzungsmöglichkeiten entwickelt werden.

WossiDiA ist über eine Vielzahl von Metadatentypen in verschiedene Richtungen anschließbar. Korpusbasierte Auswertungsverfahren erfordern darüber hinaus maschinenlesbare Texte. WossiDiA offeriert derzeit 1.852.240 Digitalisate, weitere ca. 200.000 Scans anreichender Bestandsgruppen harren noch der Einarbeitung in die Datenbank. Daher sind erstens alle früheren maschinenschriftlichen Übertragungsarbeiten (ca. 30.000 ausgefüllte Formulare) und zweitens alle bislang edierten Texte (so-

weit Nutzungsrechte einholbar sind) über Optische Zeichenerkennung (OCR) anschlussfähig zu machen.

Zum anderen muss vor allem die künftige Erschließungsarbeit mit Hilfe des neu geschaffenen hypergraphbasierten Eingabewerkzeuges vorangetrieben werden. WossiDiA punktet mit dem direkten Zugang zu seinen handschriftlichen Korpora und legt Zeugnis von der Vielfalt kultureller und sprachlicher Ausdrucksformen einer Landschaft ab. Der Vorteil der beiden anderen Datenbanken besteht darin, dass diese bislang weit mehr maschinenlesbare Texte vorhalten können. Handschriften spielen hier eine geringere Rolle. So publizierte Kristensen selber noch einen Großteil seiner Sammlung. Da Wossidlo jedoch ebenso publizierte Quellen gattungs- und motivspezifisch verschlagwortete, also ältere und vor allem auch weniger bekannte Werke auszog, die heute zunehmend auf durchsuchbare Weise online gestellt werden, wird sich die maschinenlesbare Textbasis von WossiDiA vergleichsweise rasch entwickeln können.

6. Fazit

Ein vergleichbares *Harvesting*, mit dem im Projekt *ISEBEL* derzeit experimentiert wird, könnte per gesondertem Projekt zwischen Sprichwort-Datenbanken und online gestellten Folklorearchiven in Gang gesetzt werden; und zwar ebenso grenzüberschreitend zur Erzeugung von Multilingualität. *WossiDiA* wäre hierfür ein "intelligenter" Partner, dessen mit Hypergraphen modellierte digitale Archivlösung beispielhaft demonstrieren kann, wie der phraseologische Anteil einer landschaftlichen Sammlung in hochvernetzter Form für die Bedürfnisse der Sprichwortforschung extrahiert werden kann. Mit einem solchen Projekt könnte sich die Parömiologie Korpuszugang zu handschriftlichen Dokumenten verschaffen und dabei überdies Einblicke in die Werkstätten von Sprichwortsammlern erhalten. Erreicht würde damit eine neue epistemische Qualität des Korpuszugangs.

Folklore-Archive können sich aufgrund der in ihnen enthaltenen Vielfalt kultureller und sprachlicher Ausdrucksformen für mehrere Disziplinen öffnen. Ihre verschiedenen Bereiche sind in der Regel unterschiedlich tief erschlossen, je nachdem, welchen Erkenntnisgewinn sie welchem Fachgebiet versprochen haben und neu in Aussicht stellen. Die digitale Transformation solcher Sammlungen bedeutet nicht, dass die Erschließungsarbeit endet.

Ganz im Gegenteil wird man aufgefordert, tradition archives mit den Möglichkeiten digitaler Werkzeuge noch gründlicher als bisher zu beschreiben, dabei jedoch frühere Erschließungsstände mitzunehmen.

Ohne Anpassung an die Bedürfnisse der Parömiographie und Parömiologie, durch einfachen Anschluss der Folklore-Archive an Sprichwort-Datenbanken, durch simples *Big Data*, bleiben die weiterreichenden Möglichkeiten des *Harvesting* unausgeschöpft. Für die Folklorearchive bietet sich dadurch die Chance, "abzuerntende" Sammlungsbereiche tiefer zu erschließen bzw. früher einmal begonnene, dann infolge des Wechsels von Forschungsprogrammen liegen gebliebene Arbeiten unter neuen Vorzeichen wieder aufzunehmen. Und virtuelle Sprichwortarchive können mit Datenbeständen kommunizieren, die sich auf "authentischere" (handschriftliche, mundartliche) Quellen beziehen. Alles keine Kleinigkeit, aber für computativ enthusiasmierte Parömiologen und Folkloristen keine unüberwindliche Hürde.

Notes

- ¹ https://apps.wossidia.de/webapp/run, 15. März 2019.
- ² https://search.isebel.eu, 15. März 2019 (Beta-Version).
- ³ Siehe die Auswahlbibliographien in Schmitt 1999: 411-424 und Neumann 2018: 253-265.
 - ⁴ www.momfer.ml (15. März 2019).
- ⁵ Für die folgenden Informationen danke ich Frau Dr. Kathrin Steyer vom Institut für Deutsche Sprache (IDS), E-Mail vom 9. November 2018.
 - ⁶ http://lauhakan.home.cern.ch/lauhakan/cerp.html, 15. März 2019.
 - ⁷ Ihr verdanke ich die folgenden Informationen.
- ⁸ Siehe Proverbium 30 (2013), 462-470; 32 (2015): 445-449 [bitte ggf. ergänzen, da die letzten Proverbiums-Bd. noch unterwegs nach Deutschland sind].
 - ⁹ http://sprichwort-plattform.org/sp/Sprichwörter, 15. März 2019.
 - ¹⁰ http://www.owid.de/wb/sprw/start.html, 15. März 2019.
 - 11 Steyer (wie Anm. 3). Weiteres
 - wurde hierüber nicht mitgeteilt.
- ¹² Vgl. Manfred Sommer: Sammeln. Ein philosophischer Versuch, Frankfurt am Main: Suhrkamp Verlag 2002, S. 8.
- ¹³ Auf die nähere Darstellung von Biographie und Lebenswerk muss an dieser Stelle verzichtet werden. Näheres siehe Gratopp 1935; Neumann 1994: 9-28; Gundlach 1992; Neumann 2014 (dort weiterführende Literatur).
 - 14 ZAW-D616-028.

- 15 ZAW D618-028-061.
- ¹⁶ MWW-G031-018-053-021.
- 17 Brief vom 2. September 1935 (Warnow) an Richard Wossidlo; BKW-R052-001-001-001.
 - ¹⁸ ZAW D16, D17 und D18.
 - 19 BKW K059-030-001.

References:

- Abello, James, Peter Broadwell and Timothy R. Tangherlini (2012): Computational Folkloristics. In: *Communications of the ACM* 55, No. 7: 60-70.
- Broadwell, Peter M. and Timothy R. Tangherlini (2016): WitchHunter. Tools for the GeoSemantic Exploration of a Danish Folklore Corpus. In: *Journal of American Folklore* 129, No. 511: 14-42.
- Dähnert, Johann Carl (1781): Platt-Deutsches Wörter-Buch nach der alten und neuen Pommerschen und Rügischen Mundart. Stralsund. Struck.
- Deißner, Vera (1997): Die Volkskunde und ihre Methoden. Perspektiven auf die Geschichte einer "tastend-schreitenden Wissenschaft" bis 1945 (= Studien zur Volkskultur in Rheinland-Pfalz 21). Mainz. Gesellschaft für Volkskunde in Rheinland-Pfalz.
- Ďurčo, Peter, Kathrin Steyer und Katrin Hein (2015): Sprichwörter im Gebrauch. Trnava. Univerzita sv. Cyrila a Metoda v Trnave (Reprint).
- Fuchs-Kittowski, Klaus (2010): Wissens-Ko-Produktion Organisationsinformatik. In: ders. u.a. (Hrsg.): *Organisationsinformatik und Digitale Bibliothek in der Wissenschaft*. 2. Aufl. Berlin. Gesellschaft für Wissenschaftsforschung [www.wissensforschung.de/JB00_9-88.pdf, 15.3.2019].
- Gratopp, Karl (1935): Richard Wossidlo. Wesen und Werk. Neumünster. Wachholtz.
- Gundlach, Jürgen (1992): Das mecklenburgische Wörterbuch von Richard Wossidlo und Hermann Teuchert. Seine Geschichte und seine Aussage. In: Niederdeutsches Jahrbuch 115: 145-158.
- Ilyefalvi, Emese (2018): The Theoretical, Methodological and Technical Issues of Digital Folklore Databases and Computational Folkloristics. Acta Ethnographica Hungarica 63(1): 209-258.
- Krohn, Kaarle (1907): Erste Mitteilung des Folkloristischen Forscherbundes "FF". Helsinki, 23. November 1907: 1-5.
- Kuusi, Matti (1972): Towards an International Type-System of Proverbs (= FFC 211). Helsinki. Suomalainen Tiedeakatemia, Academia Scientiarum Fennica.
- Lauhakangas, Outi (2001): The Matti Kuusi International Type System of Proverbs (= FFC 275). Helsinki. Suomalainen Tiedeakatemia, Academia Scientiarum Fennica.
- Meder, Theo (2014): The Folktale Database as a Digital Heritage Archive and as a Research Instrument. In: Meyer, H. u.a. (Hrsg.): Corpora ethnographica online. Strategien der Digitalisierung kultureller Archive und ihrer Präsentation im Internet (= Rostocker Beiträge zur Volkskunde und Kulturgeschichte 5). Münster/New York. Waxmann Verlag:119-128.

- Meder, Theo (2016): Automatic Enrichment and Classification of Folktales in the Dutch Folktale Database. In: *Journal of American Folklore* 129, No. 511: 76-94.
- Meder, Theo (2018): Harvesting International Folktale Data with ISEBEL, presented at the SIEF workgroup on Archives in Vilnius, Lithuania, 18 October 2018 [http://www.theomeder.nl/VilniusISEBEL.pdf, 15. März 2919].
- Meyer, Holger, Christoph Schmitt and Alf-Christian Schering (2014): WossiDiA
 The Digital Wossidlo Archive. In: Holger, M. u.a. (Hrsg.): Corpora ethnographica online. Strategien der Digitalisierung kultureller Archive und ihrer Präsentation im Internet. Münster/New York. Waxmann-Verlag: 61-84; kostenfrei downloadbare e-book-Version siehe: [http://rosdok.unirostock.de/resolve/id/rosdok_document_0000005031]
- Meyer, Holger, Alf-Christian Schering, and Andreas Heuer (2017): The Hydra.PowerGraph System. Building Digital Archives with Directed and Typed Hypergraphs. *Datenbank Spektrum* 17: 113-129.
- Mieder, Wolfgang (2007): Sprichwort. Enzyklopädie des Märchens, ed. Rolf Wilhelm Brednich et al. Vol. 12. Berlin/Boston. De Gruvter: col. 1099-1115.
- Neumann, Siegfried (1994): Richard Wossidlo und das Wossidlo-Archiv in Rostock (Kleine Schriften 2). Rostock. Wossidlo-Archiv.
- Neumann, Siegfried (1996): Sprichwörtliches aus Mecklenburg. Andekdotensprüche, Antisprichwörter, apologische Sprichwörter, Beispielsprichwörter, erzählende Sprichwörter, Sagte-Sprichwörter, Sagwörter, Schwanksprüche, Wellerismen, Zitatensprichwörter. Göttingen. Verlag Otto Schwarz.
- Neumann, Siegfried (2014): Wossidlo, Richard. In: Enzyklopädie des Märchens. Handwörterbuch zur historischen und vergleichenden Erzählforschung, hrsg. von R. W. Brednich u.a. Bd. 14. Berlin/Boston. De Gruyter, Sp. 1015-1018.
- Neumann, Siegfried (2018): Erzählwelten. Fakten und Fiktionen im mündlichen und literarischen Erzählen. Beiträge zur volkskundlichen Erzählforschung. Berlin/New York. Waxmann-Verlag.
- Schmitt, Christoph (1999): *Homo narrans. Studien zur populären Erzählkultur.* Festschrift für Siegfried Neumann zum 65. Geburtstag (Rostocker Beiträge zur Volkskunde und Kulturgeschichte 1). Münster u.a.: Waxmann-Verlag.
- Schmitt, Christoph (2011): Richard Wossidlo und die Genese der Volkskunde in ihrem Verhältnis zur Philologie. In: Wissen im Wandel. Disziplinengeschichte im 19. Jh. Referate der interdisziplinären Ringvorlesung des Arbeitskreises "Rostocker Universitäts- und Wissenschaftsgeschichte" im Wintersemester 2007/08, hrsg. von G. Boeck und H.-U. Lammel (= Rostocker Studien zur Universitätsgeschichte 12). Rostock. Universität Rostock: 77-104.
- Schmitt, Christoph (2014): Szenarien semantischer Vernetzung zwischen regionalethnographischen und dialektlexikographischen Korpora im Online-Projekt "WossiDiA". In: Bühler, R., R. Bürkle und N. K. Leonhardt (Hrsg.): Sprachkultur Regionalkultur. Neue Felder kulturwissenschaftlicher Dialektforschung (= Studien und Materialien des Ludwig-Uhland-Instituts der Universität Tübingen, 49). Tübingen. Tübinger Vereinigung für Volkskunde: 255-286.
- Schmitt, Christoph (2015): Zettelwerkstatt. Feldforschungsbasierte Wissenszirkulation um 1900 und die Praxis papierner Gelehrtenmaschinen am Fallbei-

- spiel des "Volksforschers" Richard Wossidlo. In: *Volkskunde in Sachsen* 27, hrsg. v. Institut für Sächsische Geschichte und Volkskunde e. V. Dresden. Thelem: 7-47.
- Schmitt, Christoph (2018): Folklore Archives Online. Zur Sichtbarmachung, Auswertbarkeit und Interoperabilität einer dänischen und einer nordostdeutschen Sammlung. In: *Jahrbuch für Europäische Ethnologie* (Dritte Folge) 13 (Sonderband Dänemark): 181-204.
- Schmitt, Christoph (2019): From Idiosyncratic 'Paper Slip Machines' to Digital Folklore Archives. The Case von WossiDiA. In: P. Hakamies and A. Heimo (edd.): *Folkloristics in the Digital Age* (= FFC 316). Helsinki. Suomalainen Tiedeakatemia, Academia Scientiarum Fennica (im Druck).
- Steyer, Kathrin (2012): Vorwort. In: dies. (Hrsg.): Sprichwörter mulitlingual. Theoretische, empirische und angewandte Aspekte der modernen Parömiologie (= Studien zur Deutschen Sprache 60). Tübingen. Narr Francke Attempto Verlag, 7-9.
- Steyer, Kathrin (2013): *Usuelle Wortverbindungen: Zentrale Muster des Sprachgebrauchs aus korpusanalytischer Sicht* (= Studien zur Deutschen Sprache 65). Tübingen. Narr Francke Attempto Verlag.
- Tangherlini, Timothy R. (2010): Legendary Performances. Folklore, Repertoire and Mapping. In: *Ethnologia Europaea*. *Journal of European Ethnology* 40, 2: 103-115.
- Tangherlini, Timothy R. (2013): The Folklore Macroscope. Challenges for a Computational Folkloristics. In: *Western Folklore* 72, 1 (2013): 7-27.
- Tangherlini, Timothy R. (ed., 2016): Big Folklore. A Special Issue on Computational Folkloristics. In: *Journal of American Folklore* 129, No. 511.
- Te Heesen, Anke (2006): Der Zeitungsausschnitt. Ein Papierobjekt der Moderne. Frankfurt. Fischer Taschenbuch Verlag.
- Uther, Hans Jörg (2004): *The Types of International Folktales. A Classification and Bibliography*. Vol. 1-3 (FFC 284-286). Helsinki. Suomalainen Tiedeakatemia, Academia Scientiarum Fennica.
- Wossidlo, Richard (1899): *Die Tiere im Munde des Volkes* (= Mecklenburgische Volksüberlieferungen, Bd. 2). Wismar. Hinstorf'sche Hofbuchhandlung Verlagsconto.
- Wossidlo, Richard (1906): Über die Technik des Sammelns volkstümlicher Überlieferungen. In: Zeitschrift des Vereins für Volkskunde 16 (1906): 1-24.
- Wossidlo, Richard und Herman Teuchert. *Mecklenburgisches Wörterbuch*. Bd. 1 hrsg. von Prof. Dr. h. c. Richard Wossidlo und Univ.-Professor Dr. Hermann Teuchert. Neumünster 1942; Bd. 2–7 aus den Sammlungen Richard Wossidlos und aus eigenen Ergänzungen bearb. u. hrsg. von Hermann Teuchert. Berlin/Neumünster 1957–1992 (ab Bd. 6, 1976, bearb. unter der Leitung von Jürgen Gundlach) [*Zitiert:* MWB].
- Yarlott, W. Victor H. and Mark A. Finlayson (2016): ProppML: A Complete Annotation Scheme for Proppian Morphologies. In: Miller, B. u.a.: (Hrsg.): 7th Workshop on Computational Models of Narrative (CMN 2016). Saarbrücken/Wadern. Leibnitz-Zentrum für Informatik/Schloss Dagstuhl. Article No. 8: 1-19 [http://drops.dagstuhl.de/opus/volltexte/ 2016/6709, 15. März 2019]

Christoph Schmitt
Wossidlo-Forschungsstelle für
Europäische Ethnologie/Volkskunde
Universität Rostock
Am Reifergraben 4
18055 Rostock
Germany

E-Mail: christoph.schmitt@uni-rostock.de