Lexikographie

Vom Papierfragebogen zur Online-Datenbank

Das Bayerische Wörterbuch auf dem Weg ins digitale Zeitalter.

VON EDITH BURKHART-FUNK UND URSULA WELSCH

DIE LEICHTIGKEIT und Unscheinbarkeit manches Wörterbuchartikels kann täuschen. Hinter ein paar Zeilen stecken oft stundenlange Exzerpier-, Sortierund Recherchearbeiten. Moderne wissenschaftliche Mundartwörterbücher sind rein empirisch, d. h. ausschließlich beleggestützt. Die Kompetenz des Bearbeiters, seine Introspektion, darf keine Rolle spielen.

Mehr Freude als das Zusammensuchen des Materials macht den Redaktoren die anschließende Interpretation: Aus den amorphen Rohdaten werden die unterschiedlichen Bedeutungen eines Wortes herausgehoben und in eine Struktur gebracht, die dem Artikel sein Gerüst gibt. Die Gliederung eines Artikels nach unterschiedlichen Bedeutungen spiegelt also die Strukturierung und Interpretation des Materials durch den Bearbeiter wider.

Im Folgenden geht es um das Aufbereiten und Herrichten der Sprachdaten für einen Artikel des Bayerischen Wörterbuchs, und zwar zum einen um die seit Jahren praktizierte "händische" Methode und zum anderen darum, wie die Digitalisierung des gesamten Materials und seine Einrichtung in einer Datenbank die Arbeit vereinfachen und beschleunigen kann.

Die Sammlungen des Bayerischen Wörterbuchs

Der Bestand des Bayerischen Wörterbuchs ist ziemlich heterogen. Gesammelt wird schon seit 1913. Es wurden literarische Quellen ausgewertet, zeitgenössische und ältere Mundartliteratur, aber auch Monographien und kleinere Darstellungen der bairischen Mundarten (wenn vom Dialekt die Rede ist, schreibt man "bairisch" mit "i") exzerpiert, ebenso wie Texte aus älteren Sprachstufen, da das Bayerische Wörterbuch auch den gesamten historischen Wortschatz darstellen

the same of the content of the conte

soll. Viele an ihrer Mundart Interessierte haben der Arbeitsstelle ihre privaten Sammlungen zur Verfügung gestellt. Ende der 1920er Jahre wurde auf so genannten "Kundfahrten" Material durch Befragung einheimischer Sprecher erhoben (Abb. 2). Der größte Teil des Bestandes basiert jedoch auf großangelegten Sammlungsaktionen mittels Fragebögen. Systematisch wurde in Altbayern und den angrenzenden bairischen Gebieten Schwabens und Frankens ein Netz von Sammlern aufgebaut, deren Anzahl zwischen ca. 200 und über 500 schwankte. In der schlechtesten Phase nach dem Ersten Weltkrieg war die Zahl auf nur noch 38 Sammler zurückgegangen. Heute haben wir etwa 430 Sammler.

In der ersten Periode bis zur Jahrhundertmitte fragte man ausgehend von Bedeutungen nach den Bezeichnungen. Auch übertragene Bedeutungen, Wendungen und Redensarten,

Abb. 1: Beispiel einer Wörterliste aus der Sammlung des Bayerischen Wörterbuchs. Sprichwörter etc., in denen ein Wort vorkommen kann, sowie Volkskundliches zu einem Begriff wurden erhoben. Ein Beispiel: "Welche Bezeichnungen verwenden Sie für das Schaf? Für das männliche, weibliche? Nennen Sie Schimpfwörter, Redensarten mit diesen Bezeichnungen, Bezeichnungen für unterschiedliche Altersstufen, unterschiedliche Lämmerrassen." Um die Jahrhundertmitte wurde das gesammelte Material gesichtet, um Schwachstellen und Lücken zu ermitteln. Von nun an konnte man umgekehrt vorgehen und fragte nicht mehr nach dem Wort für eine Sache, sondern in alphabetischer Reihenfolge gezielt nach dem Vorkommen, der Bedeutung, der Verwendung und der Aussprache von bestimmten Bezeichnungen. Die Sammler wurden gebeten, immer auch ein Satzbeispiel anzuführen. Das Material dieser sogenannten "Wörterlisten" (Abb. 1) ist die umfangreichste der Sammlungen, sie umfasst derzeit ca. 108.000 Listen mit je 60 Fragen, die wegen der regionalen sprachlichen Unterschiede freilich selten alle beantwortet wurden. Die Befragungen mit solchen "Wörterlisten" werden

bis zum Ende des Alphabets durchgeführt. Die Anzahl der Belege in den unterschiedlichen Sammlungen ist schwer zu schätzen, liegt aber auf jeden Fall bei mehreren Millionen. Aufgrund der mittlerweile sehr weit verbreiteten digitalen Technologien können Sammler die Fragebögen auch als Dateien einreichen. Ein Online-Eingabe-Tool ist in Planung – dann können die Fragen direkt in einem Bildschirmformular beantwortet werden.

Bearbeitung des Sprachmaterials vor dem Verfassen eines Wortartikels

Die Exzerpte der literarischen Quellen, die Sammlungen der Laien, der Kundfahrtenkatalog sowie die Ergebnisse der ersten schriftlichen Erhebungen sind in der alphabetisch geordneten Zentralkartei abgelegt (Abb. 3). Diese Belege sind sofort zur Hand, wenn es darum geht, einen bestimmten Wortartikel zu schreiben. Mehr Arbeit machen die Fragelisten. Aufgrund der Fragestellung weiß man, in welcher Liste ein bestimmtes Wort zu suchen ist. Nun kann man darangehen,

Abb. 2: Mundartgeographischer Fragebogen (1920er und 1930er Jahre), hier aus dem Landkreis Regen (Vorder- und Rückseite).

			Niederbayern
Det: Anvia	el Bez-Mint: Reg	eu greis	· Ndb.
53. Mi	undartgeographif der Wörterbuchk	der Frageb	
lichft um Mitteil Ihrer alteren ! waren Ihnen gu	ich-Ofterericifice Wört lung ber für die nachste Ortsmundart gebräuch großem Dant verpflich alb wieder zurückleiten	chend aufgeführt hlichen Entsprec htet, wenn Sie t	en Begriffe in
1. Genaue Auss	iprache des er in "ein fl	einer hund" 164	gloauri.
	r im Worte Schneib. ochen? (a gloanr Hund -		THE RESERVE OF THE PARTY OF THE
ober wie in "	rte "höher" das er gle Schneiber" ausgejproch einer Wenfch" er for	on? Lane for	or frain her , Efen
	noch feiner" eos ci		
	höner Mann" ov faf		
7et in "noch		The state of the s	
	r (Bfarr? Bfarl?)	Hara	
9er in Lehrer	, Schullehrer Offile	leava	
10 Fehler	Luia		
11. Teller	drein		
12. geftorben	gffloun		
18. geworfen	gevaaffa		
14. irgendwer	nlba		
15. irgendwas	26692		
16. irgendwie 🥏	his to Sieps de	straintry	the granaciette
17. irgendwann	in marus		

19. welche fi	piet?	of fine fel	V	
20. Hebamm	t fr	germans	see.	
21. Niemen	Rx	weer,		
	n, Gs. Sle			
23. Gurgel	(Ansiprache)	fixed me		
24. Stirne	tern	ficer mu	u (sheu)	
25. Kinn	afri			
	hel Hen (Schi	ibl, genaue Ausipe	adje des -b-) -s	e Oficer
27. Wiejenm	ohn Man	oublearus .		
28, bie Hühr	er (Musiprad	(10) & free	in	
29. fdpreädil	d) (Ausbrück	e dafür) Sefensy	6	
		r gordon I	ethe	
31. Anabe	Live			
32 fnarren	quan	eco, que nove	yazu	
33. Redmun	8 Raylo	en, queros		
	Amiase	V		
		17.14		Mark Colors
Datum		anino		1986
(Wetväh	römann: 🚜	Exerci della	minus the	der
		Jahre alt.)		
Warra b		iben: M. Brysis	hada de	11 44
Securit 6	en semblutten	went per ser appear	and the	ency nous
Die Rudfenl	oung bes ausg	gefüllten Fragebog	end fann amtfic	h mit Dienft.
wextzeichen	erfolgen, da li	t. MinEntschl. Nr.	. 16945 pom 21.	7. 1913 Gen-
bungen an l	die Wörterbud	hlommission ald D	ienstfache gu bel	handeln find.
Mar Alla Pro-	Der t			
**** DEC 1801		e der Biffenfdjaft		htommiffion,
	South the state of	II NO., Ludwigitr.	24/1, 3, 199.	

die entsprechenden Listen durchzugehen und zu exzerpieren. Bei z.T. über 500 Listen nimmt diese oft recht monotone und unangenehme Arbeit aber sehr viel Zeit in Anspruch. Sie ist auch kaum von Hilfskräften zu leisten, da das Material bereits beim Exzerpieren einer sprachwissenschaftlichen Interpretation bedarf. Man muss schon die bairischen Lautgesetze im Kopf haben, um sicher zu wissen, dass eine Lautform bema für "mit einem Fußboden belegen" zum Stichwort "bödnen" gehört. Da ein und dasselbe Wort aufgrund seiner unterschiedlichen Bedeutungen und Verwendungsweisen in den Listen häufig an mehreren Stellen abgefragt wurde, dauert das Exzerpieren oft wesentlich länger als das Verfassen des Artikels.

Nachdem die Wörterlisten und die Original-Fragebögen der verschiedenen Fragerunden mittlerweile komplett digital als Scans zur Verfügung stehen, können sie bei der Suche nach passenden Belegen bequem am Bildschirm durchgeblättert werden. Aktuell entwickelt die IT-Abteilung der Bayerischen Akademie der Wissenschaften ein Redaktionssystem, das es außerdem erlaubt, sich auf einzelne Fragen zu konzentrieren und die verschiedenen Antworten unmittelbar am Bildschirm miteinander zu vergleichen (Abb. 4). Mit einem einfachen Mausklick können sie markiert und z.B. nach unterschiedlichen Bedeutungen gruppiert werden. Damit wird die Arbeit des Vergleichens und Strukturierens der relevanten Belege sehr erleichtert und beschleunigt.

In diesem Redaktionssystem ist auch eine spezielle Suchmöglichkeit implementiert, die im Bestand der Dialektwörter unterschiedliche Schreib- und Lautungsvarianten findet, und zwar unabhängig von der Region, aus der diese stammen. Zudem können neben den bairischen Belegen auch diejenigen des Fränkischen und des Schwäbischen Wörterbuchs gefunden und auf

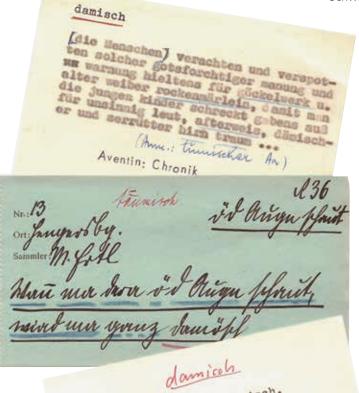
> diese Weise miteinander vernetzt werden. Das Fränkische Wörterbuch ist ebenfalls ein Projekt der Bayerischen Akademie der Wissenschaften.

Die Wörterbuchartikel

Die Wörterbuchartikel werden bislang mit einem Textverarbeitungsprogramm verfasst. Da sie in Zukunft online veröffentlicht werden sollen, und zwar zusammen mit dem ihnen zugrundeliegenden Belegmaterial, wird gerade mit einer Erfassungsmethode experimentiert, die es ermöglicht, die Artikel zugleich in Print- und in Online-Form zu produzieren. Die Texte werden strukturiert erfasst, um ihre verschiedenen Teile explizit zu kennzeichnen. Damit werden die unterschiedlichen Angaben, die in einem Artikel stecken, wie z.B. die Bedeutungen, die Verbreitungsgebiete, Phraseologismen, sachliche und volkskundliche Hinweise, die Etymologie und die unterschiedlichen Lautformen, sichtbar und recherchierbar gemacht (Abb. 5).

> Bislang sind vom Bayerischen Wörterbuch zwei Bände und drei Hefte des dritten Bands (von insgesamt 10 Bänden) in Buchform publiziert. Demnächst soll neben die

31



D' Leut maan, i 'is aba not wahr;

SLO

is glei wieds gar.

bin grod a weng damisch,

Altmann DEG

Abb. 3: Zettel aus der Zentralkartei: Belegbeispiele für das Wort "damisch".

Company Comp	to the way of the way	abe matyorer Parameter bewirkt sine UKO-Vermophings	two poundational un	the same included warm	car accretion many point	THE PERSON NAMED IN	r Bopen und Frage gesucht war	
# softstagend	# sufficient water and the state of the stat			Zveite Surter-Spatie	Dritte Surter-Spalte:	Vierbe Sortier-Bratte:	Fortis Serber Syste	
Application of the state of the	Applicate folia of the control of th	en).		WANTED AND ADDRESS OF THE PARTY	W. C. Branches	W. C. Stranger	Windshield	
Description and the City and severe 1000 Destinate severe severe who are a trained and a 1-1- required and another process of the City and severe and and	Description of the control of the co	rg .						
2 Decision when we want to be seen to be been and a to require at the Seen	2 government of the control of the c	Participant Control of the Control o		The second second	-1100000000000000000000000000000000000	1.20011190001	- 2-0-0000000	
province taken there are now white note the foreign and the contract process are now all the base and a to receive at the desiration of the contract contract of the contract process and the contract process are now and the contract process and the contract process are now and the contract process are contract process and the contract process an	province tolen The server when any the description are common to the server of the se	4						
The product water for the Upon has deem to deep the product of the Upon has deem to deep the product of the Upon has deem to deep the product of the Upon has deem to deep the product of	The parameter was not to the control to the parameter and the control to the cont	9. 3						
2 Ein solches Weißkraut! 2 Ein solches Weißkraut! 2 Ein solches Weißkraut! 3 solls Weißkraud 2 Ein solches Weißkraut! 3 solls Weißkraut! 4 solls Weißkraut! 5 solls Weißkraut! 6 solls Weißkraut! 6 solls Weißkraut!	2 Ein solches Weifikraut! 2 Ein solches Weifikraut! 3 sotts Weifskraud 3 sotts Weifskraud 4 so son Weifskraud 5 solches Weifikraut! 5 or solch Menkruni howen hand 6 solches Weifikraut! 6 solches Weifikraut! 6 solches Weifikraut! 7 solches Weifikraut! 7 solches Weifikraut! 8 solches Weifikraut!	de relevant mathietis Bintige eberfalk leden Bachalde lewen						
2 Ein polches Weiffkraut! or sells Weiskreisel howen brief Edward Schward Enragation	2 Ein solches Weilikraut! or solls Menkruist howner brief Einzelstein 2 Ein solches Weilikraut! or sölde Meiszgrand Lortres interest interest states (solth) 2 Ein solches Weilikraut! or söldes Meiszgrand Lortres (solth)	he ergab 5002 Theffer, Nachfelgend wurden daher nur die ersten 1	E aufgelistet. Sie können in der Er		-gohen durch die nachfol	genden, zentrierten Links		
2 Ein notchen Weillikrautt or selfs Wankreint housen brief Enragation	2 Ein volches Weilkraut! or selfs Menkennil housen brief. 2 Ein volches Weilkraut! or seololo Meifrance Lortres international and a second and the second a	che ergab 9022 Theffer, Nachfelgend worden daher nur die ersten 1	E aufgelistet. Sie können in der Er	potentiate ver- and actic		genden, zentrierten Link		
2 Ein solches Weithraut! Or scolds Meifigrand Lorores International States		hi wigh \$022 Teffer, hieldskynt weden debe ne zu esten 1 2. Ein nolchen Weifikraut!	I aufgelistet. Sie künnen in der Br	potentiate ver- and partic	nachuta 10			Schooled-Addess
		2. Ein solches Weiffikraut! 2. Ein solches Weiffikraut! 2. Ein solches Weiffikraut!	a significat. Six kirrows in der to	the Weigh	hround kround eighbround			Schward Andrea Romae of then Schward Schward Andrea

ative
browaposter brone Oper terms
Immergrants of or Commerciate Introductor of or Immerciate
grammatic notatiful() Cammo position
beledung-position out beledung schaesfred beledung
beledung-rainte re-la beledung schwedel dach Schweckreize ed mangeloden Schwedell
bedeating
herbriding position
rethretungungste beleg-mach 1968 blodghet-met back+CB
sedimbing-angular belog-mack 1958 Markylant-med Inscirct-MID
vertiredung ongobe belegimeach 1968 höutgkeitmiself, Bazzon GDH
verticerung pergode beleg mach 1966 houtgivetmerplas, begrit/hCP
adakappaka
beloggication belog-argular belog-test weens after is, training and belog-test
belegkentent tempe er des Velt auf die Alm belegkentent belegigselle belegiegen unt Teisenff at
landiren III landiren beleg-regon beleg-caste beleg-angabe beleg-position bedeutung-annante
bedruhing pooline

Printausgabe eine Online-Publikation treten. Natürlich werden auch die bisher erschienenen Bände online gestellt. Bis dahin ist es allerdings noch ein gutes Stück Weges: denn die nachträgliche Rückführung der Printfassung in eine Form, wie sie für die Online-Fassung neu erstellt wird, ist mit einigem Aufwand verbunden wie z. B. dem Einfügen der Links, der Auflösung von Abkürzungen und sonstigen Verdichtungen sowie der Nachstrukturierung der Inhalte in die Bestandteile, wie sie oben beschrieben sind.

Der interessierte Laie kann dann künftig sowohl in den Originalbelegen stöbern als auch sich durch die Interpretationen des Materials in Form der Artikel klicken. Es ist dann möglich, nach den Belegen einzelner Orte oder Regionen zu suchen oder etwa nach allen Belegen für einen bestimmten Begriff, ein Spiel, ein Werkzeug oder ein Tier, eine Pflanze.

Mit Hilfe der Digitalisierung wird dieses Wissen bewahrt und einer breiten Öffentlichkeit zugänglich gemacht. Die Online-Präsentation des Wörterbuchs kann jederzeit erweitert, aktualisiert und verbessert werden, da die Artikel korrigiert und erneut publiziert werden können, auch wenn sie schon geschrieben waren. Ein weiterer Vorteil ist, dass jeder Artikel publiziert werden kann, sobald er fertig ist. Man muss nicht mehr auf den Abschluss eines Heftes oder Bandes warten.

Fazit

Mit dem digitalisierten Belegmaterial, dem Redaktionssystem für die Analyse des Materials sowie der Aufbereitung der Wörterbuchartikel und ihrer Basisbelege für eine Online-Präsentation ist das Bayerische Wörterbuch auf dem Weg zu einem modernen Wörterbuch und wissenschaftlichen Grundlagenwerk, das von interessierten Laien und Wissenschaftlern gleichermaßen sowohl digital als auch konventionell genutzt werden kann – unter Wahrung der tradierten hohen philologischen Qualitätsansprüche.

Abb. 4: Screenshot des Abgleichs einzelner Fragen über mehrere Fragebögen hinweg.

Abb. 5: Screenshot der experimentellen Datenaufbereitung (XML).

DIE AUTORINNEN

Dr. Edith Burkhart-Funk ist wissenschaftliche Mitarbeiterin des Bayerischen Wörterbuchs, das an der Bayerischen Akademie der Wissenschaften entsteht. Ursula Welsch ist in der IT-Abteilung der Akademie zuständig für Digital Humanities.