

Vorwort

Die Arbeit ist aus meiner Habilitationsschrift hervorgegangen, die ich am 31.12.2021 an der Neuphilologischen Fakultät der Universität Heidelberg eingereicht habe. Der Ausgangspunkt für die Arbeit ist meine langjährige Tätigkeit als Redaktorin und als IT-Projektleiterin am *Dictionnaire étymologique de l'ancien français* – DEAF: Über 20 Jahren lang redigierte ich dort Artikel zum Alt- und bisweilen auch zum Mittelfranzösischen, und ich habe auch meine Edition eines mittelfranzösischen medizinischen Fachtextes (GuiChaulMT) als Schülerin der im DEAF durch Frankwalt Möhren etablierten Schule erarbeitet. Mein erstes «digitales Vorhaben» begann 1999 mit der Entwicklung der digitalen Version der Bibliographie (DEAFBiblél); es folgte kurz darauf die eines digitalen Erfassungssystems für die 1,5 Mill. Zettel des DEAF-Zettelkastens und ab 2006 dann die einer technischen Infrastruktur für die neue, genuin digitale Artikelredaktion des Wörterbuchs: das DEAF-Redaktionssystem mit Datenbanken, Lemmatisierungswerkzeug, Import- und Exportroutinen und zahlreichen Benutzerschnittstellen. Ab 2014 leitete ich die Entwicklung einer Schnittstelle von Korpusdatenbank und Wörterbuchdatenbank (DocLing-DEAF) mit Eingriffen auf der Ebene der Datenformate und der elektronischen Redaktionssysteme von DocLing und DEAF. Das Redaktionssystem des DEAF und meine Erfahrungen in der IT habe ich 2014 auch mit in die technische Leitung der neuen digitalen Version des *Dictionnaire de l'ancien gascon électronique* – DAGél getragen. Für dieses Projekt habe ich die Anpassung des Redaktionssystems an die Bedürfnisse der Redaktion des DAGél durchgeführt, und ab 2015 war dieses für DAGél im Einsatz.

Aus meinen Erfahrungen an der Schnittstelle von inhaltlichen – meine Arbeitsfelder sind v.a. Texteditionen, historische lexikalische Semantik und natürlich die historische Lexikologie und -graphie – und technischen Blickwinkeln entstand das Bedürfnis, die philologische Arbeit mit den Möglichkeiten der Informatik oder Computerlinguistik so zu verbinden, dass ein Mehrwert für die Wissenschaft generiert werden kann. Aus meinen Digitalisierungs- und Retrodigitalisierungsprojekten formte sich ein Ganzes mit einer übergeordneten Zielsetzung: alles erarbeitete Material unserer Disziplinen auf eine Weise zugänglich zu machen, die die bisher etablierten Möglichkeiten übertreffen möge. So kam ich zu den Technologien des Semantic Web und zu *Linked Open Data*. Es wurde schnell klar, dass ich eine Vielzahl von Themen im Rahmen meiner Arbeit würde anschneiden müssen: das sprachliche Zeichen mit Ausdrucks- und Inhaltsseite, die Lemmatisierung alt- und mittelfranzösischer Schreibformen, die Bedeutung der Bedeutung, Textphilologie, Korpuslinguistik und -lexikographie etc. und das Ganze immer auch aus dem Blickwinkel der Digital Humanities, das heißt, die Konvertierung der Daten in Standardformate des Semantic Web gemäß dem Paradigma von *Linked Open Data*, die semantische Annotierung der Daten mithilfe von Ontologien etc. Die vorliegende

Arbeit ist leider nicht der Ort, die einzelnen Themen in der Ausführlichkeit und Tiefe zu behandeln, die ihnen angemessen wäre. Da sich aber die Arbeit an eine heterogene Leserschaft richtet, die nicht notwendigerweise Kenntnisse zur Datenverarbeitung («Digital») oder zum Inhaltlichen («Humanities») besitzt, bemühe ich mich hier so weit in die Themen einzuführen, wie es für das Verständnis einer Arbeit der «Digital Humanities» nötig ist.

Daten, die zu umfangreich für die Arbeit sind, stelle ich auf GitHub zur Verfügung. Dazu zählen der Appendix mit weiterführenden Materialien zu Teilen III und IV der Arbeit sowie umfangreicher Code und Graphikdateien. Cf. <https://github.com/SabineTittel/Historische-lexikalische-Semantik-und-Linked-Data>.

Die Arbeiten an den beiden Langzeitprojekten der Heidelberger Akademie der Wissenschaften DEAF und DAG, die einen großen Raum der vorliegenden Arbeit einnehmen, wurden in Heidelberg im Dezember 2021 eingestellt. Die Artikel des DEAF^{él} mit DEAF^{plus} und DEAF^{pré} sind zur Zeit der Drucklegung der Arbeit noch erreichbar auf <https://deaf.hadw-bw.de/> und werden im Frühjahr 2024 auf die Seiten der Universitätsbibliothek Heidelberg zur Langzeitvorhaltung umziehen: <https://deaf.ub.uni-heidelberg.de>. DEAFBibl^{él} wird seit 2023 auf dem Webserver des Projekts *Wissensnetze in der mittelalterlichen Romania* (ALMA, <https://www.hadw-bw.de/alma>) vorgehalten, gepflegt und regelmäßig durch Frankwalt Möhren aktualisiert, cf. <https://alma.hadw-bw.de/deafbibil/>.

DAG^{él} war, inklusive aller Suchfunktionen, bis Dezember 2020 auf dem Server <https://dag-server.adw.uni-heidelberg.de> erreichbar. Im Januar 2021 zog die Onlinepublikation des DAG unter der Ägide seines Direktors, Martin Glessgen, auf den Server <https://dag.linguistik.uzh.ch> der Universität Zürich. Der DAG wird aufgehen in dem neuen, pan-galloromanischen Projekt *Lexique étymologique de la Galloromania médiévale* (LEGaMe, Direktion Martin Glessgen). Da die Konzeption der Onlinepräsentation, der URI-Struktur, der Suchfunktionen, der Einbindung einer neuen, pan-romanischen Bibliographie (inklusive DEAFBibl^{él}) und alles weiter damit Zusammenhängende aber noch nicht abgeschlossen sind, halten wir zum jetzigen Stand für URIs und Abbildungen an der Referenz «DAG^{él} auf <https://dag.linguistik.uzh.ch/>» fest [Stand 08/2023]. Leider ist es allerdings so, dass die Artikel des DAG^{él} dort nicht frei zugänglich, sondern nur über die Anmeldung mit einem Nutzernamen sichtbar sind. Ich füge daher Bildschirmfotos für die Illustration des DAG-Materials an den wichtigen Stellen ein.

Im Folgenden möchte ich Worte des Dankes an alle die Menschen richten, die zum Gelingen der Arbeit beigetragen haben: Zunächst möchte ich mich sehr herzlich bei Prof. Dr. Anette Frank (Computerlinguistik, Heidelberg), Prof. Dr. Sybille Große (Romanistik, ib.) und Prof. Dr. Vahram Atayan (Institut für Übersetzen und Dolmetschen, ib.) für die Betreuung meiner Habilitationsschrift bedanken. Ihre Expertise hat wesentlich zum Gelingen der vorliegenden Arbeit beigetragen, die an

der Schnittstelle zweier Disziplinen, der Linguistik und der Computerlinguistik, zu verorten ist. Die Interdisziplinarität ist die notwendige Bedingung für diese Arbeit. Zugleich manifestieren sich in der Schnittstelle die Schwierigkeiten, die eine interdisziplinäre Arbeit mit sich bringt: die Diskurse, die Übereinkünfte, die Schreib- und Zitierstile der Disziplinen und andere Dinge unterscheiden sich. Den wissenschaftlichen Standards beider Seiten gleichermaßen zu entsprechen, ist nicht trivial. Die Probleme, die dabei entstanden, konnten aber mithilfe von Frau Frank, Frau Große und Herrn Atayan überwunden werden.

Ein herzlicher Dank geht an Prof. Dr. Martin-Dietrich Glessgen, Universität Zürich, und Prof. Dr. Wolfgang Raible, Emeritus der Universität Freiburg, die die Initiazündung dafür gaben, meine Erfahrungen in Form der vorliegenden Arbeit aufzuarbeiten. Ebenfalls einen großen Dank möchte ich den Herausgeberinnen und Herausgebern der Beihefte zur Zeitschrift für romanische Philologie, Prof. Dr. Éva Buchi, Prof. Dr. Claudia Polzin-Haumann, Prof. Dr. Elton Prifti und Seniorprof. Dr. Dres. h.c. Wolfgang Schweickard, für die Aufnahme meiner Arbeit in diese schöne Reihe der Romanistik aussprechen. Beim Verlag De Gruyter konnte ich an die langjährige, kompetente Zusammenarbeit mit Dr. Ulrike Krauss, Dr. Gabrielle Cornefert und Team anknüpfen: vielen Dank auch dafür.

In meine Danksagung möchte ich meinen Arbeitgeber, die Heidelberger Akademie der Wissenschaften, einschließen. Er ermöglichte meine Teilnahme am *2nd Summer Datathon on Linguistic Linked Open Data* (SD-LLOD 2017) im Juni 2017 in Cercedilla, Spanien. Im Rahmen des Datathon konnte ich das Projekt «GuiChaulMTél» beginnen und an der Transformation der Daten des DEAF in *Linked Open Data* mit OntoLex-Lemon arbeiten. Ich bedanke mich zudem sehr herzlich bei Dr. Helena Bermúdez-Sabel (Universidad Nacional de Educación a Distancia, Madrid) und Prof. Dr. Christian Chiarcos (Universität Frankfurt), ohne deren Mitarbeit das Ergebnis dieser beiden Projekte sicher ein anderes gewesen wäre.¹ Eine weitere, interdisziplinär äußerst fruchtbare Kooperation, für die ich sehr dankbar bin, konnte ich mit Frances Gillis-Webber (University of Cape Town, Südafrika) und mit Prof. Dr. C. Maria Keet (University of Cape Town, Südafrika) installieren.

Meine Arbeit wäre ohne die fachliche und persönliche Umgebung des DEAF so nicht entstanden: Ich schulde die allergrößte Dankbarkeit meinen Kolleginnen und Kollegen, insbesondere Dr. Stephen Dörr, Prof. Dr. Frankwalt Möhren und Prof. Dr.

¹ Im Rahmen des Datathon haben weiterhin Yifat Ben-Moshe (*K Dictionaries*, Tel Aviv), Mariana Curado Malta (Polytechnic of Porto, Portugal), Frances Gillis-Webber (University of Cape Town, Südafrika) und Maxim Ionov (Goethe Universität Frankfurt) zum Erfolg der beiden Projekte beigetragen: vielen Dank!

Thomas Städtler, denen ich von Herzen verbunden bin.² Alle drei waren stets zu konstruktiven Diskussionen bereit, die nicht immer zu deckungsgleichen Meinungen führten und somit umso fruchtbarer waren. Frankwalt Möhren legte mir unermüdlich kleine Zettelchen mit Ideen und Hinweisen auf den Schreibtisch und war mit seiner Erfahrung und seinem Wissen eine Unterstützung, deren Wert unschätzbar ist. Ich danke schließlich auch Marcus Husar sehr, der mir bei allen Problemen mit den (DEAF-)Daten immer sehr hilfreich zur Seite gestanden hat.

Ein ganz persönliches Dankeschön möchte ich meinen Eltern sowie Anne Schmid-Wiedersheim dafür sagen, dass sie mich in der Kinderbetreuung in der Zeit der Habilitation so unterstützt haben. Und nicht zuletzt danke ich Dr. Conny Kühne von ganzem Herzen. Er ist mir nicht nur in der Informatik ein wichtiger Diskussionspartner und strenger Lehrer, sondern er ist auch mein Fokuspunkt in allen Bereichen des Lebens.

Abschließend ein Hinweis bezüglich der Diskussion darüber, wie man Gender in der Sprache ausdrücken kann: Die Diskussion hat auf die Inhalte der vorliegenden Arbeit so geringen Einfluss, dass ich mir erlaube Abstand zu nehmen vom Gebrauch des Gendersternchens oder ähnlichen Markierungen. Dies geschieht aus Gründen der Vereinfachung des Leseprozesses und des Fokussierens auf den Inhalt. Sowohl beim Gebrauch eines generischen Maskulins als auch eines generischen Feminins beziehe ich alle momentan bekannten und im gesellschaftlichen Diskurs thematisierten Geschlechter ein.

Publizierte Vorarbeiten

Die vorliegende Arbeit war in ihrer ersten Konzeption in Teilen kumulativ angelegt und integriert demzufolge Inhalte von mir – allein oder im Team – veröffentlichter Ausführungen. Diese sind vor oder zeitgleich mit der vorliegenden Arbeit entstanden oder waren Folgepublikationen der Arbeit. Alle Veröffentlichungen sind von mir in Gänze überarbeitet; manches ist in deutscher Übersetzung integriert, manches nur in Auszügen in dieser Arbeit erhalten (z.B. als Strukturvorgabe eines Kapitels) oder in Auszügen in eine Publikation eingeflossen, vieles in dieser Arbeit enorm erweitert. Im Folgenden lege ich im Detail dar, welche Seiten welcher Publikation mit welchen der vorliegenden Seiten in Verbindung stehen, um eine größtmögliche Transparenz zu schaffen: Tittel (2009) ist in Kap. 11.3.6 integriert. Tittel (2012b, 299–303) diente als Idee für Kap. 12.4.1, ist dabei vollständig überarbeitet und in beträchtlichem Maße erweitert; Tittel (2018b) ist in die Kap. 12.4.4 und 12.4.5 integriert, Tittel (2012a, 181–188) in Kap. 13. Tittel (2015b, 129–132) ist in

² Und ich habe mich nachgerade bemüht, die «Kärrnerarbeit» sinnvoll unterzubringen.

Kap. 13.2.3 aufgegangen und Tittel (2015b, 133–145) in Kap. 13.2.4 und 13.4 Schließlich sind aus Kap. 14.4 Tittel (2018a), aus Kap. 17.5 Tittel/Bermúdez-Sabel/Chiarcos (2018) und aus Kap. 17.1.1 Tittel/Chiarcos (2018) hervorgegangen. Inhalte aus Gillis-Webber/Tittel (2019), Tittel/Gillis-Webber (2019), Gillis-Webber/Tittel/Keet (2019) und Gillis-Webber/Tittel (2020) sind in Kap. 6.1.10, Kap. 8 und in Kap. 16.1 eingeflossen. Die Inhalte der Publikation Tittel/Gillis-Webber/Nannini (2020) habe ich in Kap. 16.3 (S. 393 bis 413) verarbeitet. Und zuletzt ging Tittel (2023), entstanden aus Kap. 18, noch während der Drucklegung der Arbeit in den Druck.

Typographische Konventionen

<i>kursiv</i>	zitierte Wortform/Kontext/Wort mit besonderer Betonung; auch: mathematische Gleichung
<i>verbatim</i>	Code/Quelltext
KAPITÄLCHEN	Etymon
MAJUSKEL	Titelwort eines Wörterbucheintrags
<	ist entwickelt aus
>	entwickelt sich zu
«...»	Wort in spezieller Verwendung/Zitat
‘...’	Bedeutungsdefinition
「...」	außersprachliches Bedeutungskonzept
*	Asterisk (zur Markierung einer in den Quellen nicht in dieser Schreibform dokumentierten Lexie)

Verzeichnis der Abkürzungen

a.st.	<i>ancien style</i>	AnS	Archiv für das Studium der neueren Sprachen und Literaturen
abfrq.	<i>ancien bas francique</i>		
afr.	<i>ancien français</i>	ArchGir	<i>Archives historiques du Département de la Gironde</i>
agasc.	<i>ancien gascon</i>	ar.	arabisch
agask.	altgaskognisch	art.	<i>artésien</i>
agn.	<i>anglo-normand</i>	bourb.	<i>bourbonnais</i>
ahd.	althochdeutsch	bourg.	<i>bourguignon</i>
anfrk.	altniederfränkisch	champ.	<i>champenois</i>
ang.	<i>angevin</i>	c.o.i.	<i>complément d'objet indirect</i>
Anm.	Anmerkung	c.o.d.	<i>complément d'objet direct</i>

CSS	<i>Cascading Style Sheets</i>	intr.	intransitiv/ <i>intransitif</i>
DB	Datenbank	IRI	<i>Internationalized Resource Identifier</i>
DH	Digital Humanities		
Dok.	Dokument	Jh.	Jahrhundert
Dr.	Drittel	judéofr.	<i>judéofrançais</i>
dt.	deutsch	Kap.	Kapitel
DWS	<i>dictionary writing system</i>	kat.	katalanisch
engl.	englisch	korr.	korrigieren
esp.	spanisch	KOS	Knowledge Organisation System
evtl.	eventuell		
f.	und folgende Seite, auch: feminin/ <i>fémimin</i>	KWIC	<i>key word in context</i> = Permutiertes Register
ff.	und folgende Seiten	l.	lies
flandr.	<i>français de la Flandre française</i>	L.	(Klassifikation nach) Linné
		<i>lemon</i>	LEXicon Model for ONtologies
Fn.	Fußnote	LD	<i>Linked Data</i>
francoit.	<i>franco-italien</i>	liég.	<i>liégois</i>
franz.	französisch	LLOD	<i>Linguistic Linked Open Data</i>
frc.	<i>francique</i>	loc. adv.	<i>locution adverbiale</i>
frcomt.	<i>franc-comtois</i>	loc. verb.	<i>locution verbale</i>
frm.	<i>français moderne</i>	LOD	<i>Linked Open Data</i>
gasc.	gascon	lorr.	<i>lorrain</i>
gask.	gaskognisch	LRL	Lexikon der romanistischen Linguistik
ggf.	gegebenenfalls	lt.	<i>latin/lateinisch</i>
gr.	griechisch	lyonn.	<i>lyonnais</i>
GUI	<i>Graphical User Interface</i>	m.	<i>maskulin/masculin</i>
H.	Hälfte	mérid.	<i>méridional</i>
HAdW	Heidelberger Akademie der Wissenschaften	mfr.	<i>moyen français</i>
hain.	<i>hennuyer</i>	Mfr.	Mittelfranzösisch
hbret.	<i>haut-breton</i>	mhd.	mittelhochdeutsch
hrsg.	herausgegeben	mlt.	<i>moyen latin/mittellateinisch</i>
Hs.	Handschrift	ms.	<i>manuscrit</i>
Hss.	Handschriften	NLP	<i>Natural Language Processing</i>
HTML	<i>Hypertext Markup Language</i>		
ib.	<i>ibidem</i>	n.st.	<i>nouveau style</i>
id.	<i>idem</i>	norm.	<i>normand</i>
i.d.R.	in der Regel	o.D.	ohne Datumsangabe
insg.	insgesamt	o.O.	ohne Ortsangabe
interj.	<i>interjection</i>		

occ.	<i>occitan</i>	SPARQL	<i>SPARQL Protocol And RDF Query Language</i>
okz.	<i>okzitanisch</i>	surselv.	<i>surselvisch</i>
orl.	<i>orléanais</i>		
OWL	<i>Web Ontology Language</i>	TEI	<i>Text Encoding Initiative</i>
POS	<i>part of speech</i>	tour.	<i>tourangeau</i>
pic.	<i>picard</i>	tr.	<i>transitiv/transitif</i>
pl.	<i>pluriel</i>	UI	<i>User Interface</i>
poit.	<i>poitevin</i>	URI	<i>Uniform Resource Identifier</i>
qch.	<i>quelque chose</i>	URL	<i>Uniform Resource Locator</i>
qn	<i>quelqu'un</i>	V.	<i>Viertel</i>
R	<i>Romania</i>	vs.	<i>versus</i>
RDF	<i>Resource Description Framework</i>	VRo	<i>Vox romanica</i>
	<i>work</i>	var.	<i>variante</i>
RDFA	<i>RDF in attributes</i>	W3C	<i>World Wide Web Consortium</i>
RDFS	<i>Resource Description Framework Schema</i>	wahrsch.	<i>wahrscheinlich</i>
Rez.	<i>Rezension</i>	wall.	<i>wallon</i>
RLiR	<i>Revue de linguistique romane</i>	WWW	<i>World Wide Web</i>
saint.	<i>saintongeais</i>	XHTML	<i>eXtensible HyperText Markup Language</i>
sept.	<i>septentrional</i>	XML	<i>eXtensible Markup Language</i>
SKOS	<i>Simple Knowledge Organization System</i>	XSLT	<i>Extensible Stylesheet Language</i>
spez.	<i>speziell</i>	ZrP	Zeitschrift für romanische Philologie

Bibliographischer Hinweis: Sigel des DEAF

Für die Verweise auf Primärquellen des Alt- und Mittelfranzösischen, auf Sekundärliteratur und auf Wörterbücher verwende ich überall, wo es möglich ist, die Sigel des *Dictionnaire étymologique de l'ancien français* – DEAF (Baldinger et al. 1971–2020). Ein Beispiel: Meine Dissertation, *Die Anathomie in der Grande Chirurgie des Gui de Chauliac: Wort- und sachgeschichtliche Untersuchungen und Edition*, Tübingen (Niemeyer [Beihefte zur Zeitschrift für romanische Philologie 328], 2004, besitzt das Sigel GuiChaulMT. Die Sigel können mithilfe der Bibliographie des Wörterbuchs (*Complément bibliographique* – DEAFBibl, Möhren ⁵2021) aufgelöst werden, die als elektronische Fassung DEAFBibl⁶, Möhren 2002, auf <https://alma.hadw-bw.de/deafbibl/> frei zu konsultieren ist. In das Literaturverzeichnis dieser Arbeit nehme ich die von DEAFBibl⁶ versiegelten Publikationen nicht auf, um den Umfang des Verzeichnisses nicht unnötig zu vergrößern. Auch

für die Bezeichnungen der Handschriftensignaturen der verschiedenen Bibliotheken übernehme ich die Konventionen von DEAFBibl. Zum Beispiel bezeichne ich die Handschrift «Bibliothèque nationale de France Fonds Français 837» als «BN fr. 837».

Für das Browsen in Onlineeditionen und -wörterbüchern verwende ich die zum Zeitpunkt der Internetrecherche jeweils aktuelle Version des Browsers Google Chrome für Mac OS Ventura 13.4.1.

Verweise auf und Zitate aus dem DEAF

Wenn ich aus Artikeln des DEAF zitiere, um Beispiele für Sachverhalte zu geben oder Aussagen zu stützen, dann ziehe ich prinzipiell Artikel aus meiner eigenen Feder heran. Wo dies nicht der Fall ist, mache ich dies explizit kenntlich durch die Angabe des Redaktors des entsprechenden Artikels (z.B. «je (Personalpronomen), DEAF J 229,26 [Städtler]»). Wenn ich von IT-Projekten spreche, die ich für DEAF, DAG (*Dictionnaire onomasiologique de l'ancien gascon*, Baldinger et al. 1975–2021) und DocLing (*Documents linguistiques galloromans*, DocLing) durchgeführt habe, so sind dies grundsätzlich konzeptionelle und praktische Eigenleistungen und keine DEAF-, DAG- oder DocLing-Teamleistungen (selbstverständlich immer unter Einbezug der Bedürfnisse, Meinungen, Vorlieben der Redaktorinnen und Redaktoren des DEAF, des DAG und DocLings). Die Programmierung (DEAF-Redaktionssystem, DAG-Redaktionssystem, Integration DocLing-DEAF) wurde größtenteils geleistet durch Dr. Conny Kühne und Marcus Husar.