

## Heidelberger Akademie der Wissenschaften

### DAG – Dictionnaire de l'ancien gascon

*Der DAG ist ein Korpuswörterbuch der gaskognischen Urkundensprache des 12. bis 15. Jahrhunderts, das in der Tradition der romanistischen Grundlagenforschung im deutschsprachigen Raum steht. Es erfasst den Wortschatz der gesamten volkssprachlichen Schriftlichkeit im Sprachraum der Gascogne.*

*Sowohl empirisch wie methodisch öffnet das Projekt neue Perspektiven in verschiedener Hinsicht. Zunächst schafft der DAG ein historisch-etymologisches Wörterbuch für die einzige romanische Varietät mit Sprachstatus, für die es das bisher nicht gibt. Damit wird es auch erstmals möglich, das Verhältnis zwischen Gaskognisch und Okzitanisch und allgemein die Stellung des Gaskognischen in der Romania auf einer tragfähigen Grundlage zu beschreiben. Der im DAG verwirklichte onomasiologische Ansatz bietet dafür eine notwendige Grundlage und erlaubt einen typologischen Blick auf den Wortschatz.*

*Da alle mittelalterlichen Quellen aus der Gascogne der Gebrauchsschriftlichkeit angehören, liefert der DAG darüber hinaus ein ideales Untersuchungsmaterial zum Verständnis der mittelalterlichen Lebenswelt, wie sie sich in der Sprache der Urkunden spiegelt. Hier geht es namentlich um die Bereiche von Gewerbe, Landwirtschaft und Privatrecht. Die Gebrauchsschriftlichkeit umfasst damit die in der romanistischen Mediävistik mit Abstand am wenigsten untersuchten Textgattungen, die aufgrund ihrer lexikalischen Vielfalt ein entsprechendes innovatives Potential bergen. Der DAG ist daher sowohl für die Regionalforschung wie für die Mediävistik insgesamt von Interesse.*

*Methodisch bildet darüber hinaus das für den DAG entwickelte elektronische Redaktionssystem ein Modell, das auch für andere historische Wörterbücher eine völlig neue Bearbeitungsgrundlage liefern kann.*

*Die Vorarbeiten für den DAG begannen unter Kurt Baldinger 1955; die Publikation in Papierform setzte 1975 ein, die Digitalisierung wurde unter der Leitung von Martin Glessgen 2014 begonnen und soll ab 2018 zu einer frei zugänglichen elektronischen Version führen.*

#### Organisationsstruktur:

Das DAG hat eine komplexe, aber klar hierarchisierte Organisationsstruktur. Alle strukturellen und richtungsweisenden Entscheidungen liegen bei der projektbegleitenden Kommission der Heidelberger Akademie. Die einmal im Jahr unter der Leitung von Prof. W. Raible tagende Kommission begleitet damit die Forschungsstelle qualitativ und berät die Leitung des Wörterbuchs, parallel dazu unterstützt von einem internationalen Beratungsgremium. Der aktuelle Forschungsstellenleiter ist 2014 aus dieser Kommission hervorgegangen.

Die Redaktion der Druckversion liegt in Händen von Dr. Nicoline Winkler und zweier weiterer Mitarbeiterinnen.

Kommissionsvorsitzender: Prof. Dr. W. Raible

Mitglieder der Kommission: Prof. Dr. Ingo Appenzeller [math.-nat. Klasse], Dr. Dr. habil. Eva Buchi [Nancy], Prof. Dr. Jean-Pierre Chambon [Paris IV-Sorbonne], Dr. Dr. habil. Jean-Paul Chauveau [Nancy], Prof. Dr. Thomas T. Field [Baltimore], Prof. Dr. Martin Gleßgen [Zürich], Prof. Dr. Frank-Rutger Hausmann [Freiburg], Prof. Dr. Drs. mult. h.c. Max Pfister [Saarbrücken], Dr. Dr. habil. Gilles Roques [Nancy], Prof. Dr. Maria Selig [Regensburg], Prof. Dr. Achim Stein [Stuttgart].

Wissenschaftliche Berater: Prof. Dr. J.-P. Chambon, Dr. habil. J.-P. Chauveau, Prof. Dr. Th. Field

Forschungsstellenleiter: Prof. Dr. M. Gleßgen

Redaktion: T. Shabafrouz, M.A., Dr. N. Winkler (stellvertretende Forschungsstellenleiterin)

Mitarbeiterin: Dr. P. Burckhardt

Arbeitsstelle:

Romanisches Seminar der Universität Heidelberg

Seminarstr. 3 [Zimmer 116]

69117 Heidelberg

Tel. 06221-543156 / E-mail: daodag@urz.uni-heidelberg.de

Internetpräsenz für die Druckversion: <[www.rzuser.uni-heidelberg.de/~dc4/neu/pmain.php](http://www.rzuser.uni-heidelberg.de/~dc4/neu/pmain.php)> (mit Zugang zu Bibliographien, Publikationsliste, onomasiologischem Übersicht der behandelten Begriffsfelder, Artikelverzeichnis mit Autorenverweis)

Elektronische Ressourcen für die Digitalisierung: s. Anhang 1.

### Projektbeschreibung:

Das Gaskognische, das in der Europäischen Charta der Regional- oder Minderheitensprachen als gefährdete Sprache geführt wird, kann auf eine signifikante mittelalterliche Quellenlage zurückblicken, die es nicht zuletzt erlaubt, das Verhältnis von Gaskognisch und Okzitanisch abzuwägen.

Im Bereich der Phonetik sind die charakteristischen Züge des Gaskognischen schon um 600 erkennbar, und sie treten auch in den frühesten schriftlichen Zeugnissen ab dem 11. Jahrhundert zutage. In Morphologie und Syntax dagegen überwiegen in der gesamten Galloromania die typologischen Ähnlichkeiten, trotz erkennbarer Unterschiede in den verschiedenen Sprachregionen. Das Lexikon nimmt eine mittlere Position zwischen Phonetik und Syntax ein, bei der pangalloromanische Lexeme, gemeinsame okzitanisch-gaskognische Formen und gaskognische Individualentwicklungen sich die Waage halten.

An der weitgehenden Autonomie der gaskognischen Dialektgruppe kann daher kein Zweifel bestehen, doch gibt es außer dem DAG bislang kein Wörterbuch, das die altgaskognische Sprache gesondert oder kontrastiv erfasst. Das einzige umfassende lexikographische Werk ist ein Wörterbuch zum Teilgebiet des Béarnesischen und der Mundart von Bayonne (Vastin Lespy/Paul Raymond, *Dictionnaire béarnais ancien & moderne*) von 1887 (Neudruck 1997). Der im Entstehen begriffene *Dictionnaire de l'occitan médiéval* (DOM, München) erfasst praktisch keine gaskognischen Einträge, und das *Französische Etymologische Wörterbuch* (FEW) gibt zwar einen vergleichsweise guten Überblick über die modernen gaskognischen Dialektformen, verfügt aber über sehr wenige mittelalterliche Einträge. Mit der Schaffung des *Altgaskognischen Wörterbuchs* in den 50er Jahren des letzten Jahrhunderts sollte diese Forschungslücke gefüllt werden. Kurt Baldinger hatte sich dabei zum Ziel gesetzt, durch eine Untersuchung der Lexik des galloromanischen Südwestens die sprachgeschichtlichen Erkenntnisse mit Blick auf die Sonderstellung der Gascogne im südfranzösischen Sprachraum zu fördern.

Dem DAG liegt aufgrund der eigentümlichen gaskognischen Quellenlage ein Korpus auf relativ einheitlicher Basis zugrunde, das mit wenigen Ausnahmen nur aus Textgattungen der Gebrauchsschriftlichkeit besteht. Die juristisch-administrativen Wortfelder sind daher sehr präsent, aber auch – wie stets in den mittelalterlichen Urkunden – sämtliche Bereiche des ländlichen Lebens (Landwirtschaft, Natur- und Kulturlandschaft, Nutztiere, Utensilien, Maße) und auch des Gewerbes (einschließlich des Schiffbaus). Die geographische Situation der Gascogne zwischen Pyrenäen und Atlantikküste bringt dabei eine erstaunliche Vielfalt auch in der Binnendifferenzierung der Region zutage.

Von der diasystematischen Markiertheit her ist der juristisch-administrative Wortschatz natürlich latinisierend und auch okzitanisch geprägt – Toulouse liegt unmittelbar am Rand des gaskognischen Dialektraums. In allen anderen semantischen Bereichen treten die Regionalität und sogar die Dialektalisierung sehr deutlich zutage, und die enge Verbindung mit der gesprochenen Sprache ist hier unübersehbar.

Eine traditionelle Besonderheit des DAG ist der onomasiologische Zugriff auf die Wortdaten. Die Makrostruktur des gedruckten Wörterbuchs folgt der Gliederung nach Begriffsgruppen von R. Hallig und W. von Wartburg (*Begriffssystem als Grundlage für die Lexikographie*, 1952, <sup>2</sup>1963). Auch die Redaktion der digitalen Version erfolgt nach dem Begriffssystem, was eine höhere Urteilssicherheit bei den Definitionen garantiert und auch den einzig tragfähigen Ansatz für einen typologischen Vergleich mit den anderen galloromanischen Varietäten bietet. Natürlich erlaubt die elektronische Darstellung auch einen unmittelbaren semasiologischen Zugriff nach dem Alphabet; aber die onomasiologische Gliederung bleibt strukturbildend für jeden einzelnen Wortartikel.

Im Zentrum des Wörterbuchs steht damit der *Begriff*, eingebettet in seinen onomasiologischen Zusammenhang. Durch die Zusammenstellung bedeutungsverwandter Wörter arbeitet er die semantisch-formale Struktur der lexikalischen Ausdrucksmittel der betrachteten Sprache heraus. Die sachgemäße Definition liefert zunächst ein Hilfsmittel zum Verständnis von Texten, in der systemhaften Zusammenstellung eröffnet sie aber auch neue Perspektiven für die Sprachgeschichtsschreibung. Der onomasiologische Zugriff führt auch alle diatopischen Differenzierungen des Gaskognischen an einem Ort zusammen, während das rein alphabetische Prinzip sie getrennt behandeln würde. Auch sachgeschichtliche Erträge, wie z.B. die Häufung speziell gaskognischer Bezeichnungen in bestimmten Wortfeldern, werden durch eine onomasiologische Sicht erkennbar. Über rein wortgeschichtliche Fragen hinaus wird so der Forschung durch die Onomasiologie eine historische Dimension erschlossen und ein Beitrag zur Regionalgeschichte des Mittelalters erbracht.

Methodisch beruht die Redaktion der Einzelartikel auf der systematischen Einbeziehung der Satzkontexte, in denen die jeweiligen Lexeme auftreten. Wegweisend war der Grundgedanke Baldingers, auch jene gaskognischen Wortformen aufzunehmen, die in lateinischen (oder später französischen) Kontexten stehen bzw. morphologisch latinisiert (oder französisiert) wurden; ebenfalls berücksichtigt wurden delexikale Namensformen in Orts- und Personennamen, ggf. ihrerseits latinisiert. Bei den Ortsnamen ist allerdings zwischen Formen mit und ohne Artikel zu trennen (erstere sind sicher nach ca. 700 zu datieren und daher für die romanische Zeit unmittelbar relevant). Schließlich berücksichtigt das DAG besonders auch die Syntagmatik und die Phraseologie sowie alle metasprachlichen Elemente wie Paraphrasen.

## Projektgeschichte: I. Phase (DAG/DAO, Kurt Baldinger, 1955-2006)

Als das Forschungsvorhaben Mitte des 20. Jahrhunderts von Kurt Baldinger (1919-2007) begründet wurde, war die lexikographische Erfassung für das Okzitanische insgesamt und für das Gaskognische im Besonderen sehr lückenhaft. An diesem, im *Französischen Etymologischen Wörterbuch* unmittelbar augenfälligen Defizit setzte K. Baldinger, damals Lehrstuhlinhaber an der Berliner Humboldt-Universität, an, als er 1955 auf dem *1<sup>er</sup> Congrès international de Langue et Littérature du Midi de la France* in Avignon erste Ergebnisse zu einem geplanten Wörterbuch des Altgaskognischen vorstellte. Darin sollten – und das war eines der Alleinstellungsmerkmale seines Vorhabens – alle von den Notaren verwendeten Idiome der Region von den Anfängen ihrer Schriftlichkeit bis ca. 1600 berücksichtigt werden: das Mittellatein – vor allem im 12. und auch 13. Jahrhundert –, die eigentliche altgaskognische Scripta, und – besonders ab 1450 – das Französische. Die in unterschiedlichen sprachlichen Kontexten und in unterschiedlicher sprachlicher Kodierung erkennbaren gaskognischen Lexeme sollte dann zusammen mit den oft nur in abgelegenen Publikationen zugänglichen Kontexten in das Wörterbuch eingehen. Dieser letztlich sehr moderne und sachgemäße Grundgedanke ist auch in der aktuellen digitalen Version weiterhin gültig.

Nach Kurt Baldingers Berufung nach Heidelberg im Jahre 1957 übernahm die Heidelberger Akademie der Wissenschaften die Betreuung des Altgaskognischen Wörterbuchs. Die aus den Berliner Arbeiten gewonnenen und schon relativ umfangreichen Materialien konnten aufgrund eines Abkommens zwischen der Deutschen und der Heidelberger Akademie der Wissenschaften am 1. Januar 1963 nach Heidelberg transferiert werden.

Erste redaktionelle Arbeiten zeigten bald, dass die Einbettung der gaskognischen Belege in einen gesamtokzitanischen Rahmen sich als wertvoll erwies. Dies führte zur frühzeitigen Erweiterung des Projekts auf das parallel zum DAG angelegte gesamtokzitanische Wörterbuch DAO (*Dictionnaire onomasiologique de l'ancien occitan*); die Belegstellen zum DAO wurden in einem Supplément (DAOSuppl) publiziert, was allerdings eine relativ starke Redundanz zwischen den drei Publikationsreihen (DAG, DAO, DAOSuppl) implizierte. Inhaltlich wurde mit DAO/DAG erstmals ein mittelalterliches romanisches Sprachgebiet in seiner umfassenden Mehrsprachigkeit abgebildet, wohlgermerkt unter Einbeziehung der namenkundlichen Materialien.

Es sei angemerkt, dass der DAG/DAO nicht als primär etymologisches Wörterbuch angelegt war, sondern zur etymologischen Einordnung zunächst einmal auf das *FEW* von Walther von Wartburg verweist. Im Regelfall genügt dies, um den gaskognischen Wortschatz in den gesamt-(gallo)romanischen Zusammenhang zu stellen.

Mit der Redaktion wurde begonnen, als mit Dr. Inge Popelar eine von Prof. Baldingers früheren Berliner Schülerinnen in die Bundesrepublik kam und zunächst zwei Jahre mit Hilfe der DFG gefördert, dann als feste Redaktorin, finanziert durch die Heidelberger Akademie und die Bundesländer-Kommission, angestellt wurde. Dem Projekt steht seit seinem Anbeginn nur eine Redaktorenstelle zur Verfügung. Nach der Amtszeit von I. Popelar trat 1992 Dr. Bernhard Henschel aus Ostberlin ihre Nachfolge an. Nach seinem Ruhestand 1996 löste ihn Dr. Nicole Winkler geb. Hörsch ab. 1999 übernahm Tiana Shabafrouz M.A. die Hälfte der Redaktorenstelle von N. Winkler.

Die ersten Faszikel DAO und DAG wurden 1975 publiziert. Seit 1990 erfolgt die regelmäßige Publikation eines Faszikels pro Jahr. Im Projektverlauf erschienen:

- 10 Faszikel DAO, 10 Faszikel DAOSuppl und 11 Faszikel DAG,
- ein *Complément bibliographique* zum DAO,
- ein alphabetischer Indexband.

Der erreichte Stand der Arbeiten von DAO/DAG umfasst im *Begriffssystem* von Hallig/Wartburg den Teil A (Universum) mit 1499 Artikeln zu den Themen Gestirne und Meteorologie, Erdkonfiguration und Bodenschätze, Pflanzen und Säugetiere.

## Projektgeschichte: II. Phase (DAG\_1300, Jean-Pierre Chambon, 2006-2014)

2006 wurde auf Rückfragen der Wissenschaftlichen Kommission in Anbetracht der anberaumten Projektlaufzeit das Gesamtkonzept und seine Realisierung bis 2020 neu überdacht. Als Grundproblem stellte sich eine Diskrepanz zwischen den schwachen dem Projekt zur Verfügung stehenden Ressourcen (eine einzige Redaktionsstelle) und dem Umfang des Forschungsvorhabens heraus (die onomasiologische Aufarbeitung des gesamten verfügbaren Wortschatzes des Altokzitanischen und Altgaskognischen in allen Textsorten, unter Berücksichtigung der mittel-lateinischen Quellen und der regionalfranzösischen Parallelförmigkeiten von den Anfängen bis ca. 1600). Konkret wurden in den ca. 50 Jahren der ersten Projektphase 22 der ca. 170 angelegten Zettelkästen redigiert und in Faszikeln publiziert. Selbst wenn der Publikationsbeginn 1975 als Stichdatum und damit 30 Jahre für diese Wörterbuchstrecke angesetzt würden, wäre für die verbleibenden 150 Kästen mit rund 200 Jahren Redaktionszeit zu rechnen gewesen. Bei einer Weiterarbeit nach der bisherigen Konzeption wäre daher bis 2020 nur ein minimaler Teilbereich des Gesamtvorhabens umgesetzt worden.

Angestrebt wurde nun eine Lösung, bei der das Projekt so dimensioniert würde, dass bei vergleichbarer wissenschaftlicher Qualität im vorgegebenen Zeitraum ein abgeschlossenes Wörterbuch entstehen könnte – unter Wahrung seiner wissenschaftlichen Kohärenz. Für die neue Konzeption legte Jean-Pierre Chambon, Ordinarius für *Langue et littérature d'oc* an der Universität Paris IV – Sorbonne, im Frühjahr 2006 eine Projektskizze vor, die seine langjährige Erfahrung als Leiter des *Französischen Etymologischen Wörterbuchs* zum Tragen brachte. Als Mitbetreuer konnten die Kommissionsmitglieder Jean-Paul Chauveau und Thomas Field gewonnen werden.

Im Zentrum der Redimensionierung des Projekts stand die Konzentration auf den Kern des Unternehmens, die Erforschung des Gaskognischen. Dies ergab sich zunächst aus der Natur der Materialsammlung, die für das Gaskognische wertvolle Primärmaterialien liefert, während sie für das Okzitanische vor allem sekundäre und tertiäre Quellen auswertet. Auch verfügt das Projekt ganz konkret durch seine ursprüngliche Ausrichtung auf die gaskognischen Quellentexte und deren kontinuierliche Aufstockung über eine der bestbestückten, wenn nicht die bestbestückte Bibliothek zum Altgaskognischen in Europa. Hinzu kommt, dass die Forschungslücke für den altgaskognischen Wortschatz weiterhin das deutlich größere Desiderat bildete. Wie Thomas Field es auf dem *VIII<sup>e</sup> Congrès de l'Association Internationale des Études Occitanes (AIEO)* in Bordeaux (12.–17.9.2005) formulierte: «Il est incontestablement vrai que l'ancien gascon est le grand inconnu de la diachronie galloromane».

Aus dieser Neusicht und Neubewertung in- und externer Faktoren entwickelte sich das neue Modell DAG, das ab Faszikel 12 (erschienen 2009) umgesetzt wurde. Im Gegensatz zu der vielfältigen Zielsetzung des DAO/DAG und seiner bisherigen chronologischen Breite widmet es sich der Erfassung des frühen Altgaskognischen und wertet damit diejenigen Daten aus dem vorhandenen Textkorpus aus, die aufgrund der langjährigen Recherchen in gaskognischem Quellenmaterial von Beginn an im Zentrum des Interesses standen und den größten wissenschaftlichen Ertrag in Aussicht stellten.

Weitere Einschränkungen (chronologische Restriktion auf die Zeit bis 1300, Ausschluss von Onomastika, Verzicht auf die Behandlung der Mehrsprachigkeit) führten zusammen mit der Reduktion auf jeweils einen einzigen Belegkontext pro Lexem zu einem kompakten DAG, der durch den Wegfall der komplexen Architektur des vorherigen dreiteiligen Systems übersichtlicher und leichter handhabbar wurde.

Das *Begriffssystem* von Hallig-Wartburg wurde als Gliederungsprinzip beibehalten, allerdings nun mit einer Beschränkung auf den Menschen und seine soziale Umgebung. Die Redaktion verfolgt ebenfalls weiterhin die sorgfältige Erfassung der Lexeme, deren Auftreten jeweils einzeln in den Texteditionen überprüft, im Kontext definiert sowie datiert und lokalisiert werden. Der Anteil an lexikographischen Ersthebungen bleibt erwartungsgemäß hoch.

Dem Projekt wurde zum 1.12.2008 von der Heidelberger Akademie eine zusätzliche Teilzeitstelle (25%) für projekt- und redaktionsbegleitende Arbeiten eingeräumt. Die Stelle wurde mit Dr. Petra Burckhardt besetzt. Sie widmet sich der Verwaltung der Bibliothek sowie den bibliographischen Basisarbeiten, die den Wörterbuchnutzern den Zugriff auf die versiegelten Quelleninformationen sichern. Zusätzlich zu der online verfügbaren Siglenliste des neuen DAG hat Frau Burckhardt eine Anschlussbibliographie zu den in DAG 1–11 benutzten gaskognischen Quellen erstellt. Sie beinhaltet die bibliographische Information zu den nach 1300 verfassten Quelltexten, sowie zu den regionalfranzösischen und mittellateinischen Quellen des Südwestens.

Auf der Grundlage der neuen Konzeption sind zum DAG\_1300 bis zum Juni 2016 sieben Faszikel erschienen (Faszikel 12, 2009 – Faszikel 18, 2015); in Vorbereitung ist das Faszikel DAG 19, das Ende 2016 erscheinen wird. Diese acht Faszikel behandeln die Begriffsfelder:

*A IV Les animaux:* Les oiseaux – Les animaux marins – Les poissons – Les amphibiens – Les insectes

*B I L'homme, être physique:* Le sexe – Le corps humain – Les organes et leurs fonctions – Les organes de la parole – Les sens et les sensations physiques – Les mouvements et les positions – Le sommeil et l'état éveillé – La santé et la maladie – La vie humaine – Les besoins de l'être humain

*B II Domaines intellectuel et affectif:* Généralités: l'âme, l'intellect – La perception – La conscience – La mémoire – La pensée – Les sentiments et émotions – La volonté – La morale

*B III L'homme, être social a) La vie de société:* La famille, la descendance – le mariage

### Projektgeschichte: III. Phase (DAG\_1500 bzw. DAGél, Martin Glessgen, 2014-2020)

2011 wurde von der wissenschaftlichen Kommission die Leitung des DAG Prof. Martin Glessgen, Ordinarius für romanische Philologie in Zürich, übertragen. Die von der Bund-Länder-Kommission nach der Evaluierung 2013 nahegelegte Digitalisierung des DAG führte seit 2014 zu einer Doppelausrichtung des Projekts, einerseits der Fortführung der Printversion des Wörterbuchs, andererseits der elektronischen Erfassung und Aufbereitung der umfassenderen gaskognischen Materialien des Heidelberger Zettelkastens bis 1500 (DAGél).

Die Print-Publikation soll künftig von Nicoline Winkler allein im Rahmen einer 50%-Stelle weitergeführt werden. Dies wird zu einer Reduktion von zu erwartenden sechs auf vier oder maximal fünf Faszikel bis 2020 führen. Die ersten 22 Kästen der verzettelten Materialien ergaben damit 11 Faszikel (1975-2006, A I – A III im *Begriffssystem*), die weiteren 38 Kästen 8 Faszikel (2006-2016, A IV – B II i), die nächsten 35 Kästen sind auf 4 Faszikel angelegt (2016-2020, B II b k – B IV b 3). Dies bedingt eine Konzentration auf Schlüsselkonzepte, die allerdings ohnehin angesichts der fortschreitenden Zeit erforderlich geworden wäre. Die Teile ab B IV b 4 (insgesamt 57 Kästen waren schon zuvor nicht mehr in der Planung der Printversion vorgesehen. Angestrebt in den Faszikeln 20-22 ist die Publikation von Schlüsselkonzepten in den Bereichen:

– La société: Famille: parenté, baptême, décès, l'éducation

Langue: langage, communication de la pensée, la langue écrite

Rapports de société: la vie de société, l'étiquette, l'entraide, la conversation

– L'homme au travail: Travail et l'organisation du travail

Monde professionnel: métiers et professions (secteur agricole/élevage, horticulture; secteur artisanal, industriel; travaux de ménage, serviteurs, domestiques)

– Le commerce / La propriété: Commerce et finance

Propriété: possession, les formes de la propriété, héritage, cession, saisie

Für die elektronische Version des DAG haben Martin Glessgen und Dr. Sabine Tittel vom DEAF (*Dictionnaire Étymologique de l'Ancien Français*) seit 2014 ein umfassendes Konzept entworfen, das sich im Wesentlichen an das Redaktionssystem und den Workflow des DEAF anlehnt und die dortigen Erfahrungen und Programmierungen nutzt (informatische Umsetzung durch Dr. Conny Kühne und Markus Husar, unter der Leitung von S. Tittel). Begonnen wurde parallel dazu mit der Sichtung der verzettelten Materialien und ihrer elektronischen Erfassung sowie, ebenfalls parallel, mit einer vollständigen Überarbeitung der Bibliographie und der verwendeten Sigel (Petra Burckhardt). Seit April 2016 ist die gesamte erforderliche Architektur verfügbar, mit Ausnahme eines für das Jahresende geplanten Printexports der fertigen Wortartikel. Durch ein Abkommen mit dem Verlag De Gruyter im Jahre 2012 konnte der Kommissionspräsident Prof. Wolfgang Raible erreichen, dass die Rechte an einer digitalen Verwertung des DAG/DAO ganz bei der Heidelberger Akademie liegen.

Die Sichtung der Zettelkästen ergab, dass die ersten 22 Kästen durch die Publikation des DAG/DAO vollständig ausgewertet wurden. Für die 93 Kästen, die in der Logik des DAG\_1300 redigiert wurden bzw. werden, verbleiben jedoch zahlreiche wertvolle Materialien bislang unausgewertet. Die Digitalisierung umfasst daher alle 153 Kästen ab der Kastennummer 23. Sie folgt der ursprünglichen Konzeption Baldingers in der Einbeziehung der gaskognischen Lexeme in mittellateinischem oder französischem Kontext sowie in Eigennamen; sie verzichtet dagegen wie Chambon auf alle Wörter, die außerhalb der Gascogne niedergeschrieben wurden, weil hierbei nur wenige Primärquellen ausgewertet wurden und weil das große Forschungsdesiderat für das Altgaskognische weiterhin besteht. Auch chronologisch wurde ein Mittelweg gewählt: Die Erfassung reicht nun bis 1500. Damit müssen etwa 40% aller Zettel (ges. ca. 130.000 Zettel) erfasst und ca. 20.000 Lexeme redigiert werden.

Die radikale Verkürzung der geplanten Redaktionszeit von einstmalen 200 auf letztlich 5 Jahre erfordert den Verzicht auf eine (systematische) Prüfung der erfassten Formen und Kontexte in den zugrundeliegenden Editionen oder gar Manuskripten. Dafür ist in einer weiteren Phase die Einbeziehung der von insbesondere von Thomas Field aufbereiteten elektronischen Textkorpora vorgesehen. Die erforderliche Programmierung wird im Rahmen des DEAF bereits vorbereitet.

Die Finanzierung dieses Projekts erfolgt durch zusätzliche Mittel, die von der Akademie auf Zeit zur Verfügung gestellt wurden: Zusätzliche Stellenprozente für die Projektpartnerin Sabine Tittel und die neue Redaktorin Petra Burckhardt; Einbeziehung der derzeit noch von Tiana Shabafrouz besetzten Stelle in den Workflow der Digitalisierung (aktuell: Korrektur der in die Datenbank eingegebenen Zettel); Werkverträge für studentische Hilfskräfte in Heidelberg und Zürich zur Erfassung der Zettel, für verschiedene Redaktoren (aktuell in der Einarbeitungsphase) und für die unentbehrlichen Programmierarbeiten. Das finanzielle Volumen der verschiedenen Werkverträge bleibt dabei unter den Kosten für eine volle Stelle.

Unter der Voraussetzung, dass diese Mittel weiter zur Verfügung stehen – insbesondere auch für die nötigen informatischen Erweiterungs- und Modernisierungsarbeiten –, ist in den verbleibenden viereinhalb Jahren der Projektlaufzeit mit einer zumindest sehr weitgehenden Erfüllung der Redaktionsstrecke zu rechnen. Eine bestimmte Verlängerung wäre sicherlich hilfreich; sie würde nicht nur die Reduktion der verfügbaren Arbeitskraft für die Druckversion auffangen, sondern vor allem eine sinnvolle und perspektivenreiche Ausweitung ermöglichen (systematische Einbindung von Textkorpora, Verknüpfung mit dem altokzitanischen Wörterbuch DOM, Ergänzung jüngerer Materialien, Vorbereitung einer erweiterten Druckversion).

Zur näheren Beschreibung des Digitalisierungsprojekts s. die beiliegenden Materialien:

- Anhang 1: Projekt DigiDag – Allgemeiner Überblick
- Anhang 2: DigiDag – Kästenverwaltung
- Anhang 3: Handreichung zur Eingabe der Zettel des DAG
- Anhang 4: Directives pour la rédaction d'articles pour le DAG
- Anhang 5a/b: Artikel *austor* (Druckversion DAG\_1300 + elektronische Version DAGél)

Arbeits- und Zeitplan für die kommenden Jahre im Überblick (2017-2020)

<b>Zielvereinbarung</b>		
	<b>Arbeitsplan</b>	<b>Veröffentlichungsplan</b>
<b>2016</b>	Redaktionelle Fertigstellung DAG 19 Überarbeitung der Bibliographie (Datenbank) Fortlaufende Erfassung des Zettelkastens Fortlaufende Redaktion der elektronischen Artikel Programmierung des PDF-Exports der elektronischen Artikel	DAG 19 – La société
<b>2017</b>	Redaktionelle Vorbereitung DAG 20 Fortlaufende Erfassung des Zettelkastens Fortlaufende Redaktion der elektronischen Artikel Neuprogrammierung der bisher von der Programmbibliothek Databinder verwalteten Elemente Fortlaufende Programmierung zur Vernüpfung des Wörterbuchs mit Textdatenbanken	
<b>2018</b>	Redaktionelle Vorbereitung DAG 21 Fortlaufende Erfassung des Zettelkastens Fortlaufende Redaktion der elektronischen Artikel Fortlaufende Programmierung zur Stabilisierung und Ausweitung des Datenbanksystems	DAG 20 – La société; L'homme au travail Online-Stellung des elektronischen Wörterbuchs
<b>2019</b>	Redaktionelle Vorbereitung DAG 22 Fortlaufende Erfassung des Zettelkastens Fortlaufende Redaktion der elektronischen Artikel Fortlaufende Programmierung (Schnittstellen)	DAG 21 – L'homme au travail
<b>2020</b>	Erfassung der ersten 11 veröffentlichten Faszikel Fortlaufende Redaktion der elektronischen Artikel Fortlaufende Programmierung (Schnittstellen)	DAG 22 – Le commerce / La propriété

Heidelberg, den 25. Juli 2016



Prof. Dr. Martin Gleßgen

gez.

Prof. Dr. Wolfgang Raible



# **Projekt DigiDAG – Allgemeiner Überblick**

Stand: 14.07.2016 MG, ST

# 1 Online-Repertorien

Alle Domänen sind mit einem Sicherheitszertifikat versehen und passwortgeschützt.

(1) System Datenerfassung **Christine** <<https://christine.adw.uni-heidelberg.de>>: Erfassung und Verwaltung der Primärmaterialien (Zettelkästen, Bibliographie und onomasiologisches System)

- "Zettelkästen" = erlaubt die Digitalisierung der **Zettelkästen**, s.u. *4 Digitalisierung der Kästen: Workflow*.
- "Sigel" = Datenbankversion der **Bibliographie**, die die (alte) Druckversion ersetzt. Erlaubt das Erfassen und Editieren der bibliographischen Einträge. Informationen werden semantisch ausgezeichnet. Strukturiert die Erfassung der Zettel ("Zettelkästen"): Nur Sigel, die in DAGBiblél erfasst und als "gascognisch" markiert sind, können in der elektronischen Fassung eines Zettels eingegeben werden; nur Sigel, die in DAGBiblél eine Datierung haben, die älter oder gleich 1500 ist, können auf einem Zettel erfasst werden; s.u. *3 Bibliographie*.
- "Onomasiologische Kategorien" = Datenbankversion des *Begriffssystems* von Hallig/Wartburg, auf die bei der Artikelredaktion zurückgegriffen werden kann. Strukturiert die Erfassung der Zettel ("Zettelkästen"): Nur Kategorien, die erfasst sind, können in der elektronischen Fassung eines Zettels eingegeben werden.

(2) Redaktionssystem **Marie** <<https://marie.adw.uni-heidelberg.de>>: Kern des Datenbanksystems. Erlaubt die Redaktion von Wortartikeln ausgehend von den erfassten lemmatisierten Zetteln (Etymologie, Formeninventar, Bedeutungen und Syntagmen, Kontexte, Einbettung in das onomasiologische System), s.u. *8 Redaktion der Artikel*

(3) Homepage der Online-Publikation **DAGél** <<https://dag.adw.uni-heidelberg.de>> : Zeigt die redigierten Artikel in Wörterbuchform und gibt Zugriff auf das DAG [Vorläufig noch keine online freigeschalteten Artikel; Artikelansicht als registrierter und angemeldeter Benutzer möglich]

- interne Unterseite <<https://dag.adw.uni-heidelberg.de/administration>> = erlaubt den XML-/ XHTML-Export und die Onlinestellung von redigierten Artikeln.
- die Suche auf [../administration](https://dag.adw.uni-heidelberg.de/administration) funktioniert mit dem Joker % ("a%" entspricht dem regulären Ausdruck "a.\*", also "a" am Anfang eines beliebigen Wortes); die Suche auf <<https://dag.adw.uni-heidelberg.de>> verwendet "." für genau ein Zeichen und ".\*" für beliebig viele Zeichen.

(4) Redaktionshilfe **Ordner DigiDAG auf GoogleDrive**

<<https://drive.google.com/drive/folders/0B-aan2IlenkpaW5BUEJMY0FiaU0>>: Diverse Materialien und Tabellen

- "DAG-Elementeliste": Elemente der Datenbankstruktur und Eigenschaften der verschiedenen Felder.
- "DigiDAG-Kästenverwaltung": Eintrag aller Kästen mit Angabe der jeweiligen Bearbeiter und des Bearbeitungsstandes, s.u. *5 Digitalisierung der Kästen* und *6 Übersicht über die Kästen*.
- "Handreichung zur Eingabe der Zettel des DAG": detaillierte Beschreibung, was bei der Zetteleingabe zu beachten ist, s.u. *5 Digitalisierung der Kästen: Workflow*.
- "DigiDAG-Sigelkorrekturliste" sowie "unklare Sigel": enthält Sigel, die beim Sortieren Unklarheiten aufgeworfen haben; darin findet sich dann die Angabe, ob das jeweilige Sigel eingegeben werden muss oder nicht (bzw. der Zettel, auf dem das Sigel steht).
- "FAQ Zetteleingabe": enthält Fragen zur Zetteleingabe, Informationen zu speziellen Sigeln, eine Liste, wo Blancs und Festausschlüsse eingegeben werden müssen, etc.
- "Handreichung Redaktion von DAG-Artikeln" und "directives pour la rédaction d'articles pour le DAG": deutsche und französische Version eines Benutzerhandbuchs für die Redaktion (und die Leser), in Arbeit.

## 2 Datenbankaspekte

Sämtliche Datenbankaspekte liegen in der Verantwortung von Sabine Tittel.

Hardware:

- a) Server: Verwendung des alten DEAF-Servers mit neuer Festplatte und neuem Netzteil. Der Server steht im Hiwi-Raum des DEAF.
- b) Arbeitsplätze für Hiwis in Heidelberg: Die Arbeitsstelle wird im Raum der DEAF-Hiwis sein; die Arbeitsplätze werden geteilt. Bildschirme sind verwendbar, aber z.T. alt.

Software:

- a) 2 MySQL-Datenbanken (Maria-DB); NoSQL-Datenbank (Mongo-DB)
- b) Zwei Browser-unabhängige Benutzeroberflächen, die beide auf dem Open Source Framework Wicket basieren; dazu Hibernate; Databinder; Jetty; Programmiersprache: Java.
  - 1. Datenerfassung für Zettel, Bibliographie und onomasiologische Kategorien; Daten werden in einer MySQL-Datenbank gespeichert und von dort in die zweite MySQL-DB importiert.
  - 2. Artikelredaktionssystem (adaptierte Version des DEAF-Redaktionssystems); arbeitet auf der zweiten MySQL-Datenbank.
- c) XHTML-Export für Onlinepublikation und Langzeitsicherung der Daten.
- d) XML-Export und Export in eine NoSQL-DB für eine maximale Performanz der Suchfunktionen.

NB Die Programmarchitektur des DAGél (und damit auch des DEAFél) ist im Prinzip stabil, verwendet jedoch mit dem Databinder eine inzwischen überholte Software, die eine Aktualisierung der übrigen Programme verhindert. Diese Programmbibliothek müsste daher im kommenden Jahr durch eine proprietäre Programmierung ersetzt werden, was einen bestimmten informatischen Arbeitsaufwand bedeutet.

### 3 Bibliographie

= elektronische Version der Bibliographie: DAGBiblél

- Erfassung und Bearbeitung auf *Christine*, unter "Sigel"
- ersetzt die (alte) Druckversion
- erlaubt das Erfassen und Editieren der bibliographischen Einträge. Informationen werden semantisch ausgezeichnet. Bedingt und strukturiert die Erfassung der Zettel.

Bsp.

- LRougeAuchD
- lokalisiert (Auch/Gers) und datiert (13<sup>e</sup>–15<sup>e</sup> s., cp. ca 1450) [ein mittleres numerisches Datum wird für die automatische chronologische Ordnung angegeben, hier 1400]
- vollständiger Titel: *Livre rouge du Chapitre Métropolitain de Sainte-Marie d'Auch*, p. pour la Société historique de Gascogne par J. Duffour, Archives historiques de la Gascogne 2<sup>e</sup> série, fascicule XI<sup>me</sup> et XII<sup>me</sup>, Paris/Auch (Champion/Léonce Cocharaux) 1907–1908.
- Verweis auf interne Sigel: *AccordAuch*, *CoutAuch*, *CoutBassoues 1295*, *CoutBassoues 1325*, *OrdBlasfD*, *TransactAuch*.

z.B. CoutAuch = «Coutumes d'Auch de l'an 1301», dans *LRougeAuchD* 206–240.

- Hinweis auf älteres Sigel: LRougeAuch verweist auf LRougeAuchD

Übersicht Zettelkästen Sigel Onomasiologische Kategorien Benutzerverwaltung Abmelden

### Sigel bearbeiten

Sigelvorgänger Sigelvarianten

Name	LRougeAuchD		
Lokalisierung	Auch [Gers]	Sprache	agasc./mlt.
Datierung	13 <sup>e</sup> –15 <sup>e</sup> s. [cop. de ca. 14	num. Datierung	1400
Einzeldatum		num. Einzeldatum	
Text	<pre>1 &lt;i&gt;Livre rouge du Chapitre Métropolitain de Sainte-Marie d'Auch&lt;/i&gt;, p. pour la Société historique de Gascogne par J. Duffour, Archives historiques de la Gascogne 2<sup>e</sup> série, fascicule XI<sup>me</sup> et XII<sup>me</sup>, Paris/Auch (Champion/Léonce Cocharaux) 1907–1908&lt;br/&gt;[includ: &lt;siglum&gt;AccordAuch&lt;/siglum&gt;, &lt;siglum&gt;CoutAuch&lt;/siglum&gt;, &lt;siglum&gt;CoutBassoues 1295&lt;/siglum&gt;, &lt;siglum&gt;CoutBassoues 1325&lt;/siglum&gt;, &lt;siglum&gt;OrdBlasfD&lt;/siglum&gt;, &lt;siglum&gt;TransactAuch&lt;/siglum&gt;]</pre>		
Literaturtyp	Primärliteratur		
<input checked="" type="checkbox"/> Altgaskognisch <input type="checkbox"/> Altokzitanisch	Standort		
<input type="checkbox"/> Eingabe nicht erforderlich	DAG		
Abbrechen	Zwischenspeichern	Speichern	Löschen

b i sc sup

↑ ç ë ï œ ÿ « » ‹ › „ “ , ‘ - → ...

– die Daten der Bibliographie sind in die Zettelerfassung integriert, bei der somit nur gültige Sigel erfasst werden können, d.h. ausschließlich Sigel, die in der Bibliographie als gascongnisch markiert sind und die nicht jünger als 1500 datiert sind. Der Versuch, andere Sigel zu erfassen, scheitert am Speichern

NB Beim DEAF gibt es diese Prüffunktion nicht: Sigel müssen dort mit einer Datei mit händisch erfassten Korrekturen (Stephen Dörr) abgeglichen werden.

– die Überarbeitung ist nahezu abgeschlossen und hat aufgrund der Inkonsistenzen der Druckversion (bes. Klärung überholter bzw. redundanter oder teilredunder Sigel) und der Notwendigkeit einer präzisen Strukturierung (Titel, Lokalisierung, Datierung von Ms. und Kopie, numerisches Datum) ca. 18 Monate à 10 H / Woche in Anspruch genommen, im Schnitt 30 min je Titel (Petra Burckhardt).

– die Bibliographie wird laufend aktualisiert, wenn neue überholte Sigel bei der Zettelerfassung auftreten.

#### Inhaltliche Frage:

– die Bibliographie wurde im wesentlichen vor ca. 15 Jahren abgeschlossen; neuere Publikationen wurden nur punktuell eingearbeitet; die Frage, ob neuere Publikationen berücksichtigt werden können, müsste eingehend geprüft werden, was aber im Laufe des aktuellen Projekts eine relevante Verzögerung bewirken würde, auch wenn die Quellenpublikationen für das Gascognische nicht sehr zahlreich sind. Von grösserem Ertrag wäre an dieser Stelle die Einbindung der von Thomas Field digitalisierten Textkorpora (s.u.).

## 4 Sortierung der Zettel

Die Sortierung erfolgt durch Hilfskräfte in Heidelberg und in Zürich.

### *Auswahlkriterien für die Sortierung*

#### Grundprinzip:

Aufgenommen werden:

- (1) ALLE Formen aus der Gascogne bis 1500
  - (2) für diese Formen relevante Verweisfichen aus der Lexikographie
- Sonst nichts.

#### Ad (1):

- es werden im Prinzip alle Formen aufgenommen, die IN der Gascogne bzw. im Béarn niedergeschrieben wurden, zunächst einmal egal in welcher sprachlichen Markierung bzw. in welchem sprachl. Kontext; d.h.
  - > gascognische Formen in gascognischem Kontext (*lo segrement deus **abiroers***)
  - > gascognische oder augenscheinlich latinisierte gascognische Formen in lateinischem Kontext (*Res quoque et mercandise que **affretabuntur** in navibus que...*)
  - > ggfalls. gascognische, französische oder englische Formen in französischem oder englischem Kontext, aber in der Gascogne niedergeschrieben (*depuis qu'ils sont entrés ... gecter le **last** de leursdicts navires ...*)
- das gilt alles auch für Ortsnamen oder Personennamen: auch die werden aufgenommen, wenn sie in der Gascogne geschrieben wurden
- der *Terminus ad quem* ist 1500

NB Bei der Erfassung in der Datenbank werden die entsprechenden Markierungen vergeben (gasc., lat., fr., NL gasc., NL lat., NP etc.)

NB Bei der Redaktion werden KEINE lateinischen Wörter behandelt, die entweder schon vor 700 belegt sind oder die als "latin savant" Teil des eigentlichen Mittellateins sind. Auch keine französischen Wörter, die keine Interferenz mit dem Gascognischen aufweisen. Diese Unterscheidung erfolgt aber erst auf der redaktionellen Ebene, weil die erfassenden HiWis nicht ständig die lateinischen, mittellateinischen und französischen Wörterbücher wälzen können.

#### NICHT aufgenommen werden:

- alle späteren Formen
- alle Formen aus anderen Gegenden (okzitanisch, französisch, englisch bzw. latein außerhalb der Gascogne), also auch nicht Toulouse, Limoges, Nantes etc.

#### Ad (2):

- Verweise auf die großen Wörterbücher werden aufgenommen (FEW, Levy, Raynouard, NED); ebenso auf andere Quellen (Jal, Nicot etc.)
- allerdings NUR, wenn eine Primärform für das entsprechende Lexem aus der Gascogne vor 1500 im Zettelkasten vorhanden ist!

- außerdem geht es hier nur um den Verweis; die ggfalls. auf den Zetteln aus den Wörterbüchern etc. abgeschriebenen Texte werden NICHT mit abgeschrieben (Ausnahme: FEW)
- Doubletten werden hier nur einmal erfasst (oft gibt es mehrere Verweiszettel für eine Quelle)

#### Prozedere:

- man sollte versuchen, die kleinen gefalteten Kartentaschen, auf denen ja ein onomasiologisches Sigel steht, zu behalten und zu verwenden;
- die als relevant eingestuften Formen werden alphabetisch geordnet, um den Überblick zu bewahren und die Zusammenführung zusammengehöriger Formen überhaupt zu ermöglichen;
- die in einem Kasten vorliegenden alphabetisch zusammengehörigen Formen werden dabei natürlich an EINEM Ort zusammengeführt (ggfalls. mit der Kartentasche);
- die jeweils zusammengehörigen Formen werden in einem zweiten Schritt chronologisch geordnet, was die Digitalisierung beschleunigt (und befriedigender gestaltet), auch weil die chronologische Ordnung oft auch eine sprachliche ist.

Die NICHT aufgenommenen Formen könnten dann ggfalls. in einer zweiten Phase nach der jeweiligen Entstehungsregion (und ggfalls. innerhalb derselben alphabetisch nach den Formen) geordnet werden:

- gascognisch bzw. Formen aus der Gascogne nach 1500 / béarnesisch bzw. Formen aus dem Béarn nach 1500
- okzitanisch bzw. Formen aus den okzitanophonen Regionen (ohne zeitliche Begrenzung)
- französisch außerhalb der Gascogne (ohne zeitliche Begrenzung)
- Verweisfichen auf Wörterbücher (ohne Primärbelege aus der Gascogne vor 1500)

#### Bemerkungen zu den Wörterbuchzetteln:

- alle Zettel werden wie gesagt alphabetisch (und in einem zweiten Schritt auch chronologisch) geordnet;
- dabei sollten unbedingt Primärzettel ( $\leq 1500$ ) und Wörterbuchzettel (mit Informationen zu denselben) getrennt sortiert werden;
- faktisch kann man eigentlich erst, wenn alle Primärzettel sortiert sind, erkennen, welche Wörterbuchzettel wirklich relevant sind (Wörterbuchzettel, die KEIN Äquivalent unter den Primärzetteln haben, werden ja weggelassen); das wiederum bedeutet, dass man prinzipiell bei der ersten Sortierung alle Wörterbuchzettel beibehalten muss, die potentiell relevant sein könnten, und hinterher eine Reihe dieser Zettel wieder ausscheiden muss; anders gesagt: Bei der ersten Ordnung behält man alle Wörterbuchzettel, die nicht eindeutig NACH 1500 oder nicht eindeutig NICHT-GASCOGNISCH sind.

#### Zur Lemmatisierung

Alle Formen, die man abtippt, werden lemmatisiert.

Grundregel: man wählt dazu eine Form, die im Zettelkasten belegt ist, typischerweise die Form, die am häufigsten oder relativ häufig auftritt (bei Verben im Infinitiv, bei Substantiven im Singular, bei Adjektiven mask.sg.).

Wichtig ist beim Erfassen der Formen, dass immer dasselbe Lemma gewählt wird. Die genaue Form des Lemmas ist an dieser Stelle nicht besonders wichtig, denn die gewählte Form kann

bei der Redaktion der Artikel (oder noch danach) leicht geändert werden. Das Ganze ist für die erfassenden Hilfskräfte also sehr fehlertolerant.

Bei der Redaktion wird man dann idealerweise diejenige Form wählen, die der modernen gasc. Form am ähnlichsten ist (wobei die Pyrenäen-Varietäten von den Gironde-Varietäten deutlich abweichen, was ebenfalls erst im Rahmen der Redaktion zu klären ist).

### Wörter anderer Sprachen

Die Anweisung, dass die Wörter aller anderen Sprachen mit erfasst werden, wenn die Quelle in der Gascogne liegt, kann zu Widersprüchen führen. Es gibt z.B. Zettel mit Altspanisch, datiert 1311 und versehen mit einem Sigel „Bal. 3, p.518“. Das Sigel „Bal“ wird aufgenommen, und daher müssen die Hilfskräfte nach den geltenden Grundregeln, den Zettel aufnehmen. Gleicher Fall bei einer spanischen Form mit dem Sigel „Bull Soc Pau“.

Das Sigel hat also Priorität (d.h.: Sigel wird erfasst => Zettel wird erfasst), und das Löschen der ggfalls. aufgenommenen spanischen (oder klassisch lateinischen) Zettel erfolgt sehr rasch und einfach bei der Artikelredaktion.

### Varia

– Es kann Abweichungen zwischen der Hallig-Wartburg-Nummer auf einem Tütchen und auf einem Zettel und auf einem Kasten geben; es gilt im Prinzip die Form auf dem Zettel, aber wenn ein Lemma als solches erkannt wurde, sollte immer dieselbe Nummer verwendet werden (die bei der Redaktion korrigiert werden kann).

– Alle Zettel, auf denen ein Wort abgehakt ist (meist mit Bleistift), sind bereits bearbeitet worden und in die Redaktion der ersten 10 Faszikel des DAG\_1300 geflossen.

[NB Die wenigen Zettel, die ein „K“ oben tragen, sind von den Redaktorinnen mit einer älteren Maske (Hilfskraft Fink) eingegeben worden; z.Z. 257 Datensätze]

– Die im DAG\_1300 bereits publizierten Zettel werden dennoch neu elektronisch erfasst; eine etwaige Korrektur erfolgt dann händisch im Vergleich mit dem DAG im Fortgang der Redaktion (das ist am wenigsten zeitraubend und erleichtert die Redaktion).



## 5 Digitalisierung der Kästen: Workflow

(1) Ordnung der Kästen: Trennung der verschiedenen Kategorien (gasc. vor 1500, nach 1500, okz./frz., jeweils einschließlich der Wörterbuchfichen, s.o. 4 *Sortierung der Zettel*); für das Gasc. vor 1500: Ordnung nach Lemmata (und nach Chronologie)

(2) Digitale Erfassung der Zettel vor 1500 in der Eingabemaske (*Christine*, "Zettelkästen"):

(2.1) Der "Überblick" in *Christine* zeigt, welche onomasiologische Strecke in einem Kasten behandelt wird, die Kastennummer (des DAG-Büros), wer ihn bearbeitet, in welchem Bearbeitungszustand sich der Kasten befindet [bis hin zum Export in das Redaktionssystem *Marie*] und wie viele Zettel für einen Kasten erfasst sind:

Übersicht **Zettelkästen** Sigel Onomasiologische Kategorien Benutzerverwaltung Abmelden

### Zettelkästen

Filtern Zettelkasten erstellen Stapelverarbeitung

Zeige 25 bis 36 von 38 Anfang Zurück 1 2 **3** 4 Weiter Ende

Name	Kastennummer	Eingabe	Eingabe-Kontrolle	Inhalte-Kontrolle	Besitzer	Anzahl Zettel			
B IV b 7 (-2)	115	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Seraina Montigel	1654	Öffnen	Bearbeiten	Löschen
B IV b 7 (-3)	116	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Seraina Montigel	1351	Öffnen	Bearbeiten	Löschen
B IV b 7 (-5)	117	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Seraina Montigel	1454	Öffnen	Bearbeiten	Löschen
B IV b 7 (4)	113	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Sarah Schramowski	488	Öffnen	Bearbeiten	Löschen
B IV g	136	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Theresa Meerwarth	55	Öffnen	Bearbeiten	Löschen
B IV i 2 cc - i 2 ee	138	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Kathrin Holey	545	Öffnen	Bearbeiten	Löschen
B IV i 3 - B IV i 4 aa	139	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sarah Schramowski	548	Öffnen	Bearbeiten	Löschen
B IV i 4 bb	140	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sarah Schramowski	891	Öffnen	Bearbeiten	Löschen
B IV i 4 ee	142	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Fabrice Bernissan	420	Öffnen	Bearbeiten	Löschen
C I a	145	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Leif Imhof	822	Öffnen	Bearbeiten	Löschen
C I b 3	147	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Sarah Schramowski	109	Öffnen	Bearbeiten	Löschen
C I e-f	159	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Sascha Schultz	1368	Öffnen	Bearbeiten	Löschen

(2.2) Erfassung der einzelnen Zettel in der Eingabemaske (wiederum Überblick):

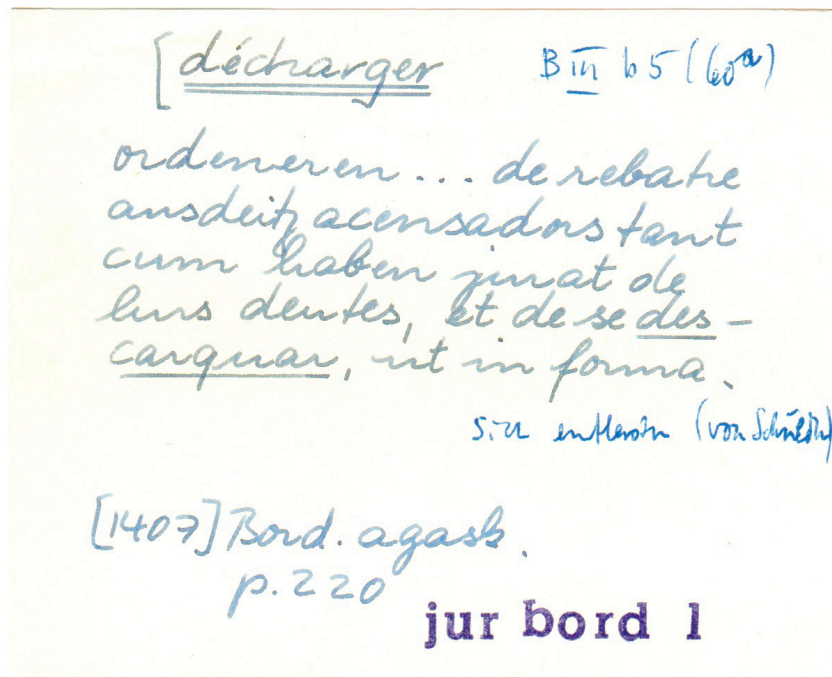
### Zettel aus B IV b 7

Filtern Zettel erstellen Stapelverarbeitung Korrektur

Zeige 1 bis 12 von 1135 Anfang Zurück **1** 2 3 4 5 6 7 8 9 10 Weiter Ende

Wort	Lemma	Sigel	Stelle	Kommentar			
acaptamentum	acaptamentum	DonFronsac	ArchGir 13,41		Bearbeiten	Korrigieren	Löschen
acapte	acapta	RecFeodD	ArchGir 5,268[197]		Bearbeiten	Korrigieren	Löschen
acapte	acapta	RecFeodD	ArchGir 5,263[185]		Bearbeiten	Korrigieren	Löschen
acapte	acapta	RecFeodD	ArchGir 5,264[189]		Bearbeiten	Korrigieren	Löschen
acapte	acapte	MillGloss	ArchGir 45,230a		Bearbeiten	Korrigieren	Löschen
acapte	acapte	ArchGir	31,460		Bearbeiten	Korrigieren	Löschen
acapte	acapte	CoutPouy-Carréjelart	ArchGir 17,27[34]		Bearbeiten	Korrigieren	Löschen
acapte	acapte	ArchGir	15,245		Bearbeiten	Korrigieren	Löschen
acapte	acapte	ArchGir	29,381		Bearbeiten	Korrigieren	Löschen
acapte	acapte	FEW	2,318a		Bearbeiten	Korrigieren	Löschen

(2.3) Einfügung der Angaben auf den Zetteln in eine Maske:



wird erfasst als:

## Zettel bearbeiten

Sigel	<input type="text" value="JurBord 1"/>	Datierung	<input type="text" value="1407"/>	<input type="text" value="1407"/>
Stelle	<input type="text" value="p.220"/>		alphanumerisch	numerisch
Wort	<input type="text" value="descarquar"/>	Einzeldatum	<input type="text"/>	<input type="text"/>
Wortart	<input type="text"/>		alphanumerisch	numerisch
Lemma	<input type="text" value="descarquar"/>	Definition	<input type="text" value="décharger"/>	
Kategorie	<input type="text" value="B III b 5"/>	enzyklopädische Definition	<input type="text"/>	
Begriff	<input type="text"/>	Kontext	<input type="text" value="ordeneren...de rebatre ansdeitz acensadors tant cum haben jurat de lurs dentes, et de se descarquar, ut in"/>	
Lokalisierung	<input type="text" value="Bordeaux"/>	Rest	<input type="text" value="Sig entle. (von Schm.)"/>	
Sprache	<input type="text" value="agasc."/>	Kommentar	<input (60\")]"="" :="" ;="" catégorie="" illisible="" rest\"="" type="text" value="Lecture : \"/>	

[der nicht erkannte Passus der Bedeutung heißt vermutlich: "sich entledigen (von Schulden)"]

(2.4) Korrektur der Zettel (neuerliche Durchsicht per "Stapelverarbeitung", gesamthaft oder in Stichproben, je nach Sicherheit der erfassenden studentischen Kraft).

(2.5) Durchsicht der Fragezeichen im „Kommentar“ durch eine Redaktorin.

(3) Nach Erfassung aller relevanten Zettel und nach Korrektur des Kastens: Import des gesamten Kastens in die Datenbank, auf der das Redaktionssystem *Marie* arbeitet.

## 6 Übersicht über die Kästen: Zahlen und Zeitplan

### Gesamtüberblick

Die Materialsammlung des DAG umfaßt 175 Kästen zu je 1.500-2.250 Zetteln. 22 davon sind bereits vollständig in Faszikeln aufbereitet (DAO/DAG). Der Kasten 148 enthält keine relevanten Daten. D.h. 152 Kästen sind zu erfassen.

### Zahl der relevanten Zettel für die Erfassung

- 1 Kasten = 35 cm; 1 cm = ca. 65 Zettel; d.h. ein ganz voller Kasten enthält ca. 2.250 Zettel (viele Kästen sind nicht ganz voll)
- Zahl der relevanten Zettel in den bis Juni 2016 digital erfassten 29 Kästen:
  - = sehr variabel: zwischen 109 und 1655 Zettel
  - = selten weniger als 360: drei von 29 Kästen
  - = 360-900: 15 von 29 Kästen
  - = 1100-1655: 11 von 29 Kästen
- gesamt enthalten die 29 Kästen genau 24.099 digital erfasste (d.h. als relevant eingestufte) Zettel
- pro Kasten entspricht dies 831 Zetteln im Schnitt, was für 152 Kästen ca. 126.000 Zettel ergibt; eine Varianz nach oben ist wahrscheinlich; aufgrund einer Reihe sehr dicht bestückter Kästen ist mit insgesamt bis zu 140.000 Zetteln zu rechnen (diese Zahl wird zur Sicherheit in der Folge bei der Zeitplanung angesetzt). Zum Vergleich: Der DEAF enthält 1,5 Mio. Zettel, also mehr als zehnmal so viel.

### Zahl der zu redigierenden Wortartikel

N.B. Terminologie: Es werden unterschieden *Lemmata*, *Wortformen*, *Bedeutungen* und *Lexeme*; *Lexeme* sollen verstanden sein als die Kombination einer *Wortform* und einer *Bedeutung*; eine polyseme *Wortform* vereint daher mehrere *Lexeme*, je nach der Zahl der an sie gebundenen *Bedeutungen*. *Lemmata* entsprechen den auf die Grundform zurückgeführten *Wortformen* (Infinitiv, mask.sg.) unter Neutralisierung der jeweiligen graphematischen Varianz; ein *Lemma* vereint also die *Wortformen*-Varianz einer lexikalischen Einheit und im Fall von Polysemie mehrere *Lexeme*.

- 8 Kästen sind in Marie integriert, 4.630 Zettel, die 1.321 Wortformen ergeben (also ein unterdurchschnittlicher Schnitt von knapp 600 Zetteln / Kasten). Die Formen werden bei der Erfassung von den Hilfskräften provisorisch lemmatisiert; bei der endgültigen Lemmatisierung durch die Redaktoren werden erfahrungsgemäß mehrere provisorische Lemmata in ein endgültiges zusammengeführt, was zu einer Reduktion der Formen führt; es ist also von ca. 1.100 zu redigierenden Lemmata auszugehen ist.
  - bei 1.100 Wortformen, ausgehend von 4.630 Zetteln, ergäben 140.000 Zettel rein rechnerisch 33.200 Wortformen. Da zahlreiche Polysemien vorliegen, geht es hierbei nicht um die Zahl der Formen sondern eher um die der Lexeme (Verbindung Form-Bedeutung).
- Ganz grob kann man beim augenblicklichen Stand für das DAGél ca. 20.000 Lemmata mit insgesamt ca. 60.000 Bedeutungen oder Syntagmen annehmen.

## Überlegungen zum Zeitplan

- theoretisch anzunehmender Arbeitsaufwand für die Erfassung:
  - = Ordnen und Lemmatisieren eines Kastens: ca. 10 H
  - = Digitalisierung: 60-80 Sek / Zettel; Prüfung und Korrektur: 20-30 Sek / Zettel
  - = für die verbleibenden ca. 115 zu ordnenden Kästen und ca. 125.000 zu digitalisierenden Zettel :  $1.100 + 3.300 = \text{ca. } 4.500$  Stunden, bei 8 H / Woche (= das realistische Arbeitspensum einer studentischen Hilfskraft) also ca. 10 Mannjahre (wohlgemerkt je 8 Wochenstunden), bei 5 in diesem Umfang arbeitenden Kräften (= 40 H / Woche bzw. 180 H / Monat) also ca. 2 bis 2,5 Jahre.
- Stand der Arbeiten bei der digitalen Erfassung:
  - = August 2015: ca. 80 Kästen primär geordnet in Heidelberg (ohne Lemmatisierung), 4 Kästen digitalisiert
  - = Juni 2016: = 88 Kästen primär geordnet (ohne Lemmatisierung), 29 Kästen digitalisiert
  - = d.h. in 10 Monaten von Sept. 2015 bis Juni 2016: 25 Kästen lemmatisiert und digitalisiert
  - = d.h. 2,5 Kästen / Monat; bei diesem Rhythmus müssten in den kommenden 48 Monaten also ca. 120 Kästen erfasst werden, was ganz knapp ausreicht.
  - = der konkrete Stand der Arbeiten lässt also vermuten, dass die angenommenen 180 H / Monat ein absolutes Minimum für Erfassung in Prüfung im geplanten Zeitrahmen darstellen.
- theoretisch anzunehmender Arbeitsaufwand für die Artikelredaktion:
  - = ca. 10-20 Minuten, also vermutlich dieselbe Zeit wie die Erfassung der Zettel für eine einfache Umsetzung, das Doppelte für eine ein wenig ausgearbeitete Redaktion.
  - = wenn man 30 min für 20.000 Wortformen ansetzt, kommt man auf ca. 10.000 Stunden; d.h. bei 10 H / Woche also ca. 20 Mannjahre. Es bräuchte also in etwa 5 Redaktoren im Umfang von wenigstens 10 H über 4 Jahre hinweg (aktuell in Einarbeitung: Petra, Jan, Fabrice, Seraina). [Anders gerechnet entspricht das dem Volumen einer vollen Stelle über fünf Jahre hinweg.]
  - = hinzu kommt die Endredaktion (Korrekturen, Ergänzungen usw.).
  - = nach den Beobachtungen bei der Erfassung wäre auch hier davon auszugehen, dass die angenommenen 220 H / Monat eher eine untere Grenze darstellen.

## 7 Überblick über die Kästen: Planung für die vier Serien

### Série I : 22 Kästen

von der Druckversion ausgehend retrodigitalisieren

### Série II : 39 Kästen

bleibt vorläufig unbearbeitet, aber zur Vorbereitung dennoch:

- (1) Prüfung der fünf redigierten Kästen zwischen 42 und 47 = max. 100 H
- (2) Digitalisierung von 41 und von den noch offenen max. drei Kästen zwischen 43 und 47 sowie noch 48 (ggfalls. später 38, 39) = max. 200 H
- (3) dann – bei späterer Gelegenheit – Redaktion des so entstandenen Blocks 40-48 (B I j - B II g 10)

B I j = La vie humaine en général

B I k = Les besoins de l'être humain (alimentation, vie sexuelle, vêtements, coiffure)

B II ... g = L'âme et l'intellect (perception, pensée, vérité, sentiments)

NB Martin Glessgen redigiert zum Test und zu Vergleichszwecken (DAG / DAGél) die Kästen 23 + 24 (A IV b = Les oiseaux)

### Série III: 54 Kästen

bleibt vorläufig ebenfalls unbearbeitet, aber zur Vorbereitung dennoch:

- (1) 84, 87, 90, 104 sind digitalisiert > in Marie importieren (84 + 104 in Heidelberg prüfen, dann alle vier an Tiana Shabafrouz zur Endkontrolle)
- (2) bei Gelegenheit Redaktion des so entstandenen Blocks 84-90 (B III b 5 + 6)

B III b 5 Le commerce, la finance (84-89)

B III b 6 La propriété (90)

### Série IV (111-169) = 60 Kästen

- (1) Tiana Shabafrouz: 115 korrigieren, dann wären 113-117 in Marie = 6085 Zettel
- (2) fünf Kästen in Prüfung: unter Hochdruck fertigmachen
- (3) Redaktion ab jetzt in diesem Teil

Reihenfolge der Bearbeitung:

(1) Reihenfolge für den Fortgang der **Digitalisierung** in Heidelberg (August-Dezember):

zunächst (bis 1. Sept. 2016):

- in Heidelberg alle in Arbeit befindlichen Prüfungen fertigmachen und in Marie integrieren
- 112 (Kasten zur Digitalisierung bei Sarah, bereits lemmatisiert)
- alle in Heidelberg in der Digitalisierung befindlichen Kästen fertigmachen, prüfen und in Marie integrieren
- erst danach: neue Kästen digitalisieren, in folgender Reihenfolge:

2016	Heidelberg	141-144	(Série IV)
	Zürich	86/88	(Série III)
	Zürich	145-156	(Série IV)
2017		41, 44, 45, 46'	(Série II)
		118-137	(Série IV)
		157-171	(Série IV)

(2) Reihenfolge für die **Redaktion**: (= Jan, Petra [hat schon einiges in B IV b 7 redigiert], Seraina, Fabrice [hat schon viel in B IV b 7 redigiert und muss korrigieren], Martin)

(2.1) 113 - 117 (ergänzt um 112, sobald Sarah fertig ist)

B IV b 7 : Le gouvernement, l'administration

B IV b 4-6 (112): Les classes sociales, la constitution, les partis

(2.2) [wenn das fertig ist, bevor der folgende Block aus IV digitalisiert ist: Kästen 84-90]

(2.3) 136 - 156

B IV g (La guerre), h (Les Belles Lettres #Achtung: Kasten 137 fehlt!#), i (Les croyances, la religion, l'Église, les fêtes)

C 1 a-d (L'existence, les sens, les mesures)

(4) [der Block aus II: 41-48]

## 8 Redaktion der Artikel

Der Workflow folgt der Struktur der Datenbank *Marie*:

- Auswahl der behandelten Lemmata und Präzisierung der Form des jeweiligen Lemmas
- Sichtung der einzelnen erfassten Zettel unter etwaiger Korrektur (bes. zum Kontext)
- eigentliche Redaktion mit Etymologisierung, Ordnung der graphematischen Varianten und Semantisierung der Belege (Bedeutungen, Syntagmen)

S. für einen detaillierten Überblick die beiliegenden *Directives pour la rédaction d'articles*

Hier eine kurze Synthese:

### *Avant la rédaction*

*Importation des fichiers de Christine à Marie:*

Button "Import von Paketen": met en œuvre l'importation des fichiers en question et synchronise les fiches avec les lemmes et les catégories onomasiologiques correspondantes.

### *Préparation de la rédaction*

- organiser d'abord les mots selon les catégories pour choisir les lemmes à traiter [= nous – traiterons désormais les mots dans un ordre onomasiologique; idéalement à partir de B IV (donc B IV, puis C)]
- relever pour une catégorie onomasiologique donnée les lemmes (= à la main)
- puis traiter les lemmes à partir de l'ordre alphabétique (pour réunir les fiches d'un mot sous différentes catégories onomasiologiques)

### *Rédaction*

- "Langartikel" "bearbeiten"

#### *1. Lemme*

- si nécessaire: changer les lemmes et/ou les sous-lemmes: "Lemma bearbeiten"
- principe général: un lemme correspond à une base protoromane (= latine écrite ou latine parlée avant *ca* 700); tout ce qui en est dérivé devient un sous-lemme

#### *2. Préparation des attestations*

- = transformation d'une fiche numérisée ("Zettel") en une attestation comme base du dictionnaire ("Beleg")
- traiter *Stelle et contexte*

### 3. Renseignement des différents champs interprétatifs

– traiter *titelblock* et *etymon* (= d'après le FEW)

– *Graphienblock / formes*

> *ordre des graphies* : purement chronologique, à l'intérieur des trois catégories

"en contexte lat.: agasc. "

"en contexte gasc.: agasc."

"en contexte fr.: agasc."

> *présentation des catégories*

EX *austor* : *en contexte latin: agasc. lt. asturio .....*

EX *affuament* (jetzt unter HL *fiu*) : *en contexte gasc.: abéarn. affuament*

puis, sous la forme balisée : *en contexte gasc.: abéarn. <cited-word index="f" language="abéarn.">affuament</cited-word>*

> cocher la petite case après chaque attestation

> indiquer la plus ancienne graphie en cliquant sur «Datum ältester Beleg (Pos. 0)». Il faut éventuellement corriger la datation (effacer des espaces, etc.). Il ne faut pas changer le *numdat*.

> on indique les graphies, même si nous n'avons qu'une seule attestation

– *définition*

> *définition*: d'abord la définition, puis l'équivalence en fr. mod.

> indiquer la "Onomasiologische Kategorie"

> dater le sens

> cocher la petite case après chaque attestation pour qu'elle soit exportée

> nom propres (ex *esparvers* CartGimC) : catégoriser comme "Unterbedeutung"



## 9 DAG / DAO – Druckversion

Die Druckversion verfolgt je nach Epoche verschiedene Konzepte:

- (1) DAG(DAO) : Konzept Baldinger = Vollausswertung des Fichiers (okz./gasc.) [Kästen 1-22]
- (2) DAG\_1300 : Konzept Chambon = Auswertung des Fichiers für Gasc. bis 1300 (in diesen Grenzen vollständig) = ca. 15 % der Zettel werden verwendet
- (3) DAG\_1300 sélectif : Konzept Winkler = Bearbeitung ausgewählter Lexeme für Gasc. bis 1300 = ca. 5 % der Zettel werden verwendet (seit 2016)

Nach Abschluss der elektronischen Version wäre an ein viertes Konzept zu denken:

- (4) DAG\_1500 : Konzept Glessgen = Ausdruck der elektronischen Version in onomasio-logischer Ordnung: Gasc. bis 1500 mit ausgewählten Kontexten = ca. 50 % des Fichiers werden verwendet (nicht: Okz., Frz. und die entsprechenden Verweis- und Lexikonfichen); Schwäche: weniger sorgfältige Bearbeitung der Details (Formen, Definitionen); Stärke: deutlich umfassendere Nomenklatur, sehr deutlich umfassendere Beleggrundlage; Umfang: bei ca. 20.000 Lexemen ca. 8000 Seiten bei der aktuellen Darstellungsform; Verdichtung durch Verzicht auf den Grossteil der Kontexte auf 3.000 Seiten (= 4 Bände) wäre sinnvoll.

NB Die bereits publizierten Bände des DAG\_1300 und in einem zweiten Schritt auch des DAG(DAO) werden gesamthaft in die Datenbank aufgenommen werden müssen; das bringt einen bestimmten Mehraufwand, liefert aber andererseits gut aufbereitete Daten, die strukturbildend wirken und die zugleich ergänzt werden können (bes. für DAG\_1300); Hyperlinks auf die gescannten Seiten sind vorgesehen.

Kastenbez.                      Kastennr. 1 nach Hause wann zurück sortiert                      lematisiert    digitalisiert kontrolliert    Zettel                      Kommentar

**Kasten 1 - 22 vollständig publiziert als DAG/DAO: muss ausgehend von der Druckfassung retrodigitalisiert werden**

**Kasten 23 - 57 publiziert als DAG 1300 (Fasz. 1-18): muss gesamthaft erfasst werden und unter Einbeziehung von DAG 1300 redigiert werden**

A IV -	23	27.3.2016	02.04.2016	12.12.2014	29.03.2016	22.06.2016	22.06.2016	807
A IV b 2 - b 5	24			12.12.2014	09.06.2015	28.05.15		132
A IV b 5 - b 8	25			13.01.2015				
A IV c - d 4	26			13.01.2015				
A IV e - i 2 (?)	27			15.12.2014				
Addenda zu C				17.12.2014				
B I a - c	29			23.02.2015				
B I d - e	30	01.02.2016		15.03.2015	15.04.2015			
B I f 1 - f 5 (redigiert bis 1300)	31	05.12.2014	04.05.2015	04.05.2015				
B I g 1	32			18.12.2014				
B I g 2	34			23.02.2015				
B I g 2	35							
B I h - i 1	36			07.01.2015				
B I i 2 aa	37			07.01.2015	arah 07.03.15	arah 08.04.15	Ben 2.12.15	232 von Fabrice
B I i 2 bb	38			20.02.2015				
B I i 2 cc	39			22.12.2014				
B I j	40			29.12.2014	01.04.2015	08.07.2015		1493
B I k 1 aa	41			17.02.2016				
B I k 1 bb - cc 4 (redigiert bis 1300)	42	05.12.2014	04.05.2015	15.08.2015	15.08.2015	ha 15.08.2015	Ben, 28.12.15	523
B I k 1 cc 5 - cc 8	43	04.06.15	23.06.15	24.02.2015	23.06.15	04.09.15	22.06.2016	450
B I k 1 dd - k3 bb 1	44			17.02.2015				
B I k 3 bb 2 - ff	45			27.02.2015				

B I k 3 gg - hh	46	30.07. 2015	05.07.2016	29.09.2015	20.10.2015	06.07.2016		461
B II a - f 2	46'			21.02.2015				
B II f 3 - f 4 bb	46''	04.06.15	23.06.15	20.02.2015	23.06.15	15.11.15	in Arbeit Kathr	1423
B II g 1 - g 2 hh	46'''			07.04.2015	15.04.2015	20.10.2015		409
B II g 2 ii - g 4 bb	46''''			24.02.2015	29.04.2016	27.05.2016		670
B II g 4 cc - g 9	47			21.01.2015	07.05.2015	22.05.15	05.05.2016	513
B II g 10	48			09.02.2015				
B II h 1 - h 2 bb	49			12.07.2016				
B II h 2 cc - h 2 ee	50			16.02.2015				
B II h 2 - ff 1	51			26.02.2015				
B II h 2 ff 2 - ff 4	52			16.02.2015				
B II h 3 - h 3 aa 5	53	01.02.2016		15.04.2015	14.05.2015		in Arbeit	
B II h 3 aa - h 3 bb	54			22.02.2015				
B II h 3 cc	55			16.02.2015				
B II h 3 dd - i 2	56			14.01.2015				
B II i 2	57			16.02.2015				

**Kasten 58 - 111 (sehr) partiell für DAG 1300 zur Redaktion vorbereitet: muss gesamthaft erfasst und redigiert werden, ggfalls. unter Einbeziehung der parallelen Redaktion für die Druckfassung**

B II i 2 - i 3	58			19.02.2015				
B III a 1 aa 1 - aa 3	59	01.02.2016		14.05.2015				
B III a 1 aa 1 aa 5	60		18.02.2016					
B III a 1 aa 6 - a 2 ε	61		18.02.2016					
B III a 2	63		18.02.2016					
B III a 2 aa 2	62		18.02.2016					
B III a 2 bb	64			01.03.2015				
B III a 3 aa	65							
B III a 3 bb - cc	66							

B III a 3 dd - ff 4	67	27.02.2015				
B III b 2 aa 1 (1)	68					
B III b 2 aa 1 (2)	69					
B III b 2 aa 2 - aa 4	70					
B III b 2 bb - cc 2	71					
B III b 2 cc 3 - cc 6	72					
B III b 2 dd - ee	73					
B III b 2 ff	74					
B III b 2 gg - hh	75					
B III b 3 aa 1	76					
B III b 3 aa 2	77					
B III b 3 aa 3	78					
B III b 3 bb (1)	79					
B III b 3 bb (2)	80					
B III b 3 bb (3)	81					
B III b 3 bb (4)	82					
B III b 3 bb (5)	83					
B III b 4 ??	84					
B III b 5 (1)	85	29.06.2015	20.07.2015	30.07.2015	arbeit kerstin	363
B III b 5 (2)	86	03.11.2015	18.02.2016	30.11.2015	30.11.2015	760
B III b 5 (3)	87		17.02.2016	10.03.2016	10.03.2016	1123
B III b 5-1		05.02.2015	01.10.2015	01.10.2015		Kasten 87 ?
B III b 5 (4)	88	24.02.2015				
B III b 5 (nach 1300	89	24.02.2015	12.07.2016			1304
B III b 6 (1)	90	18.11.2015	27.12.2015	27.12.2015	07.05.2016	1389
B III b 6 (2)	91					
B III b 6 (3)	92					
B III b 6 (4)	93					
B III b 6 (5)	94					
B III b 7 aa	95					
B III b 7 bb	96					
B III b 7 cc - dd	97					

<i>B III b 7 dd</i>	98						
<i>B III b 7 ee</i>	99						
<i>B III b 7 ff</i>	100						
<i>B III b 7 gg - ll</i>	101						
<i>B III b 8 - b 8 aa</i>	102						
<i>B III b 8 aa - bb</i>	103						
<b>B III b 8 cc - b 9 aa</b>	<b>104</b>		<b>29.04.2015</b>	<b>05.12.2014</b>	<b>Benjamin, 29.4.2015</b>		<b>732</b>
<i>B III b 8 ee - b 9 aa</i>	105						loose Yettel
<i>B IV a 1 aa - bb</i>	106						
<i>B IV a 2 aa</i>	107						
<i>B IV a 2 aa - bb 6</i>	108						
<i>B IV b 1 - b2</i>	109						
<i>B IV b 2</i>	110						
<b>B IV b 3</b>	<b>111</b>	<b>05.02.2015</b>	<b>01.10.2015</b>	<b>01.10.2015</b>	<b>05.05.2016</b>		<b>in arbeit</b>

**112-169 nicht im Plan der Printversion: müssen voll erfasst und redigiert werden**

<b>B IV b 4 - b 6</b>	<b>112</b>	<b>03.08.2015</b>		<b>14.09.2015</b>	<b>14.09.2015</b>	<b>n Arbeit Sarah</b>	
<i>B IV b 7 (-4)</i>	113	02.04.2015	03.06.2015	03.06.2015	22.09.2015	15, von Sarah	488
<i>B IV b 7</i>	114	24.02.2015		10.05.2015	10.05.2015	02.06.2015	1135
<b>B IV b 7 (-2)</b>	<b>115</b>	<b>24.02.2015</b>		<b>30.04.2015</b>	<b>30.04.2015</b>	<b>02.07.2015</b>	<b>1655</b>
<i>B IV b 7 (-3)</i>	116	24.02.2015	20.12.2015	20.04.2015	20.04.2015	05.09.2015	1354
<i>B IV b 7 (-5)</i>	117	24.02.2015	20.12.2015	15.06.2015	15.06.2015	14.11.2015	1454
<i>B IV b 7 (6)</i>	118						
<i>B IV b 8 - c 1</i>	119						
<i>B IV c 1</i>	120						
<i>B IV c 2 (1)</i>	121						
<i>B IV c 2 (2)</i>	122						
<i>B IV c 3 (1)</i>	123						
<i>B IV c 3 (2)</i>	124						
<i>B IV c 3 (3)</i>	125						
<i>B IV c 3 (4)</i>	126						
<i>B IV c 3 (5)</i>	127						

B IV c 3 (6)	128						
B IV c 3 - c 4	129						
B IV c 4 (1)	130						
B IV c 4 (2)	131						
B IV d - e	132						
B IV f 1 - f 2	133	01.10.2015	15.01.2016	15.01.2016			
B IV f 2 cc - 2 ee	134	01.10.2015	15.01.2016	15.01.2016			
B IV f 2 - f 3	135						
B IV g	136	01.10.2015	15.01.2016	15.01.2016			in arbeit
B IV h - i 2 bb	137						
B IV i 2 cc - i 2 ee	138	29.06.2015	02.05.2016	20.07.2015	30.07.2015	15.05.2016	545
B IV i 3 - i 4 aa	139	02.04.2015	03.06.2015	03.06.2015		15.03.2016	in Arbeit Kathr 548
B IV i 4 bb	140	02.04.2015	03.06.2015	03.06.2015	22.06.2016	16, von Sarah	891
B IV i 4 cc - dd	141	05.02.2015	31.03.2015	31.03.2015	31.03.2015		
B IV i 4 ee	142	05.02.2015	03.06.2015	31.03.2015	31.03.2015		420
B IV i 4 ff	143	05.02.2015	31.03.2015	31.03.2015			
B IV i 4 gg	144	05.02.2015	31.03.2015	31.03.2015			
C I a	145	18.02.2016		04.12.2014			in arbeit
C I b - C I b 2	146	18.02.2017		04.12.2014			
C I b 3	147			04.12.2014	25.05.2016	26.05.2016	109
C I b 4 aa	148			15.12.2014	22.04.2015	g: kein agasc.	
C I b 4 bb - C I c 1	149	18.02.2016		04.12.2014			
C I c 2 - C I c 3	150	18.02.2017		04.12.2014			
C I c 3 - C I c 4	151			04.12.2014			
C I c 4 Forts. 1	152	18.02.2016		04.12.2014			
C I d 4	153	18.02.2016		04.12.2014			
C I d - Fortsetzung	154	18.02.2016		27.10.2015	27.10.2015		
C I d Forts. 2	155	18.02.2016		16.12.2014			
C I d Forts. 3	156	18.02.2016		04.12.2014			
C I e	157			04.12.2014			
C I e Forts.	158			08.12.2014			
C I e-f	159			04.12.2014	04.12.2014	08.06.2015 27.03.2016	1368
C I f (30a)	160			04.12.2014			

C I f (30b-x)	161								
C I f (30x)	162			04.12.2014					
C I f/g	163			04.12.2014					
C I g (1)	164								
C I g (2)	165	28.12.2015	05.01.2016	04.12.2014	Dez. 2015	23.02.2016	11.07. 2016		1266
	166								
C I h	167	25.06.2016	06.07.2016	04.12.2014					in arbeit
C I i - C II	168			31.05.2015					
C I i Forts.	169?			04.12.2014					

*Nachträge (1)*

*Nachträge (2)*

Kasten ist schon in Marie importiert

Kasten befindet sich im Hiwiraum

Kasten ist fertig (Ampel komplett) und befindet sich wieder im DAG-Büro für Bearbeitung durch Tiana

Kastennummer 57-111: sind in Publikationsstrecke des DAG bis 2020 und sind sehr partiell "vorredigiert"

sind publiziert bis 1300: Kästen 23-56

Priorität des Erfassens

# Handreichung zur Eingabe der Zettel des DAG

Die Handreichung folgt der Reihenfolge der Eingabefelder der Eingabemaske des DAG.

## Inhaltsverzeichnis

### [Inhaltsverzeichnis](#)

[Allgemeines](#)

[Sigel](#)

[Welches Sigel kommt in dieses Feld, wenn 2 Sigel vorhanden sind?](#)

[Wichtiges](#)

[ArchGir 11 Gloss](#)

[Stelle](#)

[Wichtiges](#)

[Wort](#)

[Wortart](#)

[Lemma](#)

[Wichtiges](#)

[Kategorie](#)

[Problem](#)

[Begriff](#)

[Lokalisierung](#)

[Sprache](#)

[Datierung](#)

[Wichtiges](#)

[Numerische Datierung](#)

[Wichtiges](#)

[Einzeldatum](#)

[Definition](#)

[Achtung](#)



[Enzyklopädische Definition](#)

[Kontext](#)

[Wichtiges](#)

[Rest](#)

[Achtung](#)

[Kommentar](#)

[Kontrolle der Zetteleingabe durch eine Kollegin / einen Kollegen](#)

## Allgemeines

Alle Hiwis haben zusätzlich zu dem Zugang zur DAG Seite auch einen Zugang zu den Dokumenten des GoogleDrive-Ordners erhalten. Auf dieser Seite sollte unter „Kästenverwaltung“ jeder Kasten, den die Hiwis zu Hause haben, und der aktuelle Stand der Arbeit eingetragen werden (sortiert, lemmatisiert und digitalisiert). Im GoogleDrive-Ordner unter „Sigelkorrekturliste“ gibt es zudem eine sehr nützliche Tabelle mit Sigeln, die nicht in der Datenbank des DAG gespeichert sind und zu denen jeweils steht, ob sie aufgenommen werden müssen oder nicht. In dieser Liste, stehen auch die richtige Schreibweise zu Sigeln, die auf den Zetteln falsch geschrieben wurden, und unter dieser Schreibweise in der Datenbank des DAG nicht auffindbar sind (wie z.B <CoutCatBouz> statt richtig <CoutCastBouz>).

Grundsätzlich wird alles von den Zetteln abgeschrieben, ausser UB-Signaturen der Bibliothek Heidelberg.

## Sigel

Als Sigel wird das Primärsigel aufgenommen.

Welches Sigel kommt in dieses Feld, wenn 2 Sigel vorhanden sind?

Bibliographie anschauen, es gibt verschiedene Fälle:

Fall 1) ein Sigel ist in einem anderen enthalten (z.B. das Sigel CartFrons: aus der Bibliographie ist ersichtlich, dass CartFrons IN ArchGir 38,1903,1–35 integriert ist). In diesem Fall wird <CartFrons> in das [Sigel](#)-Feld aufgenommen, <ArchGir 38> und die Seitenangabe in den [Rest](#).

Fall 2) es steht nichts zum Sigel in der Bibliographie (aber das Sigel ist vorhanden): In diesem Fall ist das eine Sigel meistens ein sehr häufig vorkommendes wie z.B ArchGir, Monl, CoutGascT usw. Dies sind meist grosse Werke, in denen viele Sigel drin enthalten sind. Dieses häufigere Sigel kommt in den [Rest](#), das andere Sigel in das [Sigel](#)-Feld (da es sich vermutlich um das Primärsigel handelt). Die Seitenangabe kommt zum Sigel im [Rest](#).

Fall 3) ein Sigel ist nicht in der Bibliographie vorhanden: In dem Fall wird dieses Sigel in den [Rest](#) geschrieben, das bekannte in das [Sigel](#)-Feld und die Seitenzahl in das Feld [Stelle](#).

## Wichtiges

Wenn sich das Eingabefenster rosa färbt, wird der Zettel nicht aufgenommen, da das Sigel dann keinen agasc. Text beschreibt.

**Memento:** Bitte prüft nach der vollständigen Eingabe eines Kastens nochmal nach, ob alle Sigel korrekt sind, die Petra in der Zwischenzeit vervollständigt bzw. geklärt hat. Diese Sigel befinden sich in der Datei auf dem GoogleDrive "FAQ Zetteleingabe".

Die Prüfung erfolgt ganz einfach und schnell wie folgt: Zettelkasten zur Bearbeitung öffnen, alle Zettel nach "Sigel" sortieren, indem man auf "Sigel" klickt (evtl. zweimal, um die alphabetische Reihenfolge von Z-A auf A-Z umzustellen). Dann kann man die Sigel, die Petra korrigiert hat, aufsuchen und prüfen, ob alles korrekt ist.

Nicht vergessen: Man muss auch die alten, nun korrigierten Namen der Sigel suchen, um diese gegebenenfalls zu korrigieren.

## ArchGir 11 Gloss

Bei diesen Zetteln muss die Datierung nachgesehen werden, bevor man entscheiden kann, ob der Zettel aufgenommen werden muss oder nicht. Das kann man selbst machen, da alle ArchGir-Bände auf Gallica digitalisiert zugänglich sind:

<http://gallica.bnf.fr/ark:/12148/cb32701447v/date.r=Archives+historiques+du+d%C3%A9partement+de+la+Gironde.langDE>

Dort muss man entweder in ArchGir 11 den korrekten Verweis auf das Wort suchen.

<http://gallica.bnf.fr/ark:/12148/bpt6k341104.image.langDE.r=Archives%20historiques%20du%20d%C3%A9partement%20de%20la%20Gironde>

und dann in dem entsprechenden Band + Seite das Wort suchen. Dort steht am Seitenrand das Datum. Wenn auf dem Zettel gleich der Verweis auf Band + Seite steht, kann man gleich diesem folgen.

## Stelle

Als Stelle wird auch die genaue Angabe zum Band usw. angegeben: z. B. Sigel: <ArchGir>, Stelle: <5, p.120> (kein Abstand zwischen dem <p.> und der Zahl!). Sigel:< FEW>, Stelle: <5,311a> (auch kein Abstand!).

**Memento:** bei eindeutigen Stellenverweisen sollte das <p.> auch gleich gelöscht werden. => ArchGir 5,120.

## Wichtiges

Bei Stellenangaben der Wörterbücher steht: Band,Seite. Ohne Abstand zwischen Komma und Seitenangabe. Bsp: <FEW 10,50b>.

## Wort

Das Wort genauso übernehmen, wie es auf dem Zettel steht. Nur EIN WORT in dieses Feld schreiben! Wenn mehr als ein Wort unterstrichen ist, kann dieser Ausdruck in das Feld [Enzyklopädische Definition](#) geschrieben werden. Dann handelt es sich meist um eine bestimmte Verwendung, z.B. eine Kollokation. Dazu gibt es oft auch eine eigene Definition, die zur ex. Kollokation in die [Enzyklopädische Definition](#) geschrieben wird.

Wenn mehrere Wörter im Kontext unterstrichen sind, die offensichtlich nicht (als Wendung o.ä.) zusammengehören, dann legt man zwei Zettel mit je einem der unterstrichenen Wörter an. Wenn nur eines dieser Wörter klar zu der im Kasten enthaltenen onomasiologischen Kategorie gehört und das andere Wort klar einer – unbekanntem – Kategorie angehört, dann kann man,

- a) wenn die Bedeutung des Wortes klar ist, die entsprechende Kategorie raussuchen und eintragen, oder

- b) wenn die Bedeutung unklar ist, die Kategorie des Kastens übernehmen und muss dann in das Kommentarfeld schreiben „Kategorie unklar“.

Die Grossschreibung darf (und zugleich muss) nur bei Toponymen und Anthroponymen in dieses Feld übernommen werden.

## Wortart

<Substantif>, <adjectif> usw. Dieses Feld nur ausfüllen, wenn die Wortart auf dem Zettel steht, ansonsten leer lassen.

Bei den Zetteln des Wörterbuchs Raynouard (Rn) steht hinter Substantiven <(m)> oder <(f)>, hinter den Verben <(v)> (jeweils ohne Punkt). Diese Information wird in das [Wortart](#)-Feld übernommen. Es wird, auch wenn es sich klar um ein Substantiv oder ein Verb handelt, nichts anderes in das Feld geschrieben, als das, was auf dem Zettel steht. Ausnahme: Bei m, f, oder v, etc. wird ein Punkt hinzugefügt. => m., f., v., etc.

## Lemma

Als Lemma werden die Verben im Infinitiv (Grundform), die Substantive im Singular, die Adjektive in ihrer maskulinen Form im Singular angegeben.

Bei lateinischen bzw. latinisierten Wörtern wird das gaskognische Lemma angegeben, wenn bereits ein solches vorhanden ist. (Keine lateinischen Lemmata „erfinden“.)

Oft befindet sich ein unterstrichenes Wort oben auf dem Zettel: Wenn es identisch ist mit dem Wort im Kontext bzw. eine davon abgeleitete Form ist (Singular eines Substantivs im Plural/Infinitiv eines Verbs), kann es als Lemma benutzt werden. Oft ist es jedoch eine neufranzösische Entsprechung des agask. Wortes, die nicht in Anführungs- und Schlusszeichen steht. In dem Fall kann sie – wenn es keine Definition auf dem Zettel gibt, in das Feld [Definition](#) übernommen werden.

## Wichtiges

Wenn mehrere Zettel zum gleichen Wort auftauchen unter verschiedenen Schreibweisen, wird als Lemma für alle diese Zettel eine EINZIGE Form gewählt, die am ehesten standardisiert bzw. neutralisiert wirkt oder am öftesten vorkommt.

Das Lemma ist nicht unbedingt endgültig, es wird noch überarbeitet. Es ist also nicht schlimm, wenn das Lemma nicht korrekt ist. Hauptsache, es werden die zusammengehörigen Formen (mit verschiedenen Graphien) unter NUR EINEM Lemma erfasst.

Genau wie beim [Wort](#)-Feld, wird die Grossschreibung nur bei Toponymen und Anthronymen übernommen bzw. dort eingeführt.

## Kategorie

Dies ist die onomasiologische Kategorie aus HW (Hallig-Wartburg).

Im Normalfall steht auf dem Zettel die gleiche Kategorie wie auf dem Kasten, und wenn keine Kategorie draufsteht, wird die Kategorie des Kastens angegeben. Wenn eine andere Kategorie als die des Kastens angegeben ist, muss diese im Prinzip übernommen werden (s.u. "Problem").

Oft steht auf den Zetteln rechts oben eine (oft mit rotem Kugelschreiber geschriebene) Kategorie, die von HW abweicht (und die man insofern auch nicht in unserer Datenbank finden kann). Diese Sonder-Kategorien stammen von einem ehemaligen Mitarbeiter, der sie „erfunden“ hat. Sie werden immer von einer Art Kürzel abgeschlossen, das oft einer Zahl mit einem hochgestellten Buchstaben ähnelt. Diese Sonder-Kategorien sind zu ignorieren.

## Problem

Wenn eine Kategorie vorkommt, die von derjenigen auf dem Kasten abweicht: In diesem Fall sollte die Liste der onomasiologischen Kategorien konsultiert werden, die in der Datenbank einzusehen ist (über das Formular „(Onomasiologische) Kategorien“). Inhaltlich sollte entschieden werden können, welche semantische Kategorie plausibler erscheint und ins Feld [Kategorie](#) übernommen wird. Im Zweifelsfall sollte auf jeden Fall die grobe statt der präzisen Kategorie aufgenommen werden.

## Begriff

Dieses Feld muss nicht ausgefüllt werden. Das wird anhand der in der Datenbank erfassten onomasiologischen Kategorien vom System automatisch ausgefüllt.

# Lokalisierung

Als Lokalisierung wird der Ort angegeben, der auf dem Zettel steht, egal ob der Ort in der Gascogne liegt oder nicht (jedenfalls dann, wenn das Sigel überhaupt aufgenommen werden muss).

Abgekürzte Formen werden ausgeschrieben, wenn die Bedeutung klar ist: Bord. → Bordeaux

In einigen Fällen steht ein Ort und zusätzlich <betr. Bord.> (oder irgendein anderer Ort). Bsp: <London, betr. Bord.>. In diesem Fall wird London in den Rest geschrieben und Bordeaux als Lokalisierung verwendet. Es wird also, wenn möglich, der Ort angegeben, der in der Gascogne liegt oder ihr nahe liegt.

Die Abkürzung <o. O.> bedeutet höchstwahrscheinlich 'ohne Ortsangabe'. Wenn trotzdem ein Ort dabeisteht, wird das <o. O.> zum Ort ins [Lokalisierungs](#)-Feld geschrieben. Wenn kein Ort dabeisteht, wird das <o. O.> in den [Rest](#) geschrieben.

# Sprache

Wenn eine Angabe zur Sprache auf dem Zettel steht, wird sie übernommen, wenn nichts steht, kann das Feld leer bleiben. Wenn man die Sprache des Kontextes jedoch mit einiger Klarheit erkennt (lateinisch, gaskognisch oder französisch), sollte man die entsprechende Sprache in das Feld schreiben.

# Datierung

Alphanumerisch: in diesem Feld wird die Datierung genauso übernommen, wie sie auf dem Zettel steht, d.h. mit allfälligen zusätzlichen Angaben wie <ca.> oder Fragezeichen; z.B. 1221-1223, 14e s., 1215? usw.

Wenn zwei Sigel vorkommen und zwei Daten (z. B. ACoutBordB ca.1400, OrdImpôts 1350): Als Primärsigel wird OrdImpôts aufgenommen, als [Datierung](#) ca.1400, als [Einzeldatum](#) 1350 und ACoutBordB als Sigel in den [Rest](#). Das Einzeldatum ist meistens später als die Datierung. Es kann aber, wie in diesem Fall auch früher sein. Das liegt daran, dass ein Buch, das primär

den Text mit dem Sigel „X“ und der Datierung „Y“ enthält, auch ältere (oder jüngere) Texte mit eigenen Sigeln enthalten kann.

## Wichtiges

- 1) Was tun, wenn zu einem Einzeltext innerhalb eines „Sammelsigels“ wie ACoutBord o.ä. kein Einzeldatum auf dem Zettel steht?
  - a) Man kann in der Bibliographie nachsehen, ob da ein Datum zum Text steht.  
Beispiel: JurAgenM ist auf den Zetteln ohne Datum, aber die Bibliographie nennt „1345-1355“. Das Datum kann man dann aufnehmen (numerische Angabe dann „1350“)
  - b) Wenn man nicht herausfinden kann, dann wird das Einzeldatum leer gelassen.  
Zusätzlich muss dann in den Kommentar „Datierung“ geschrieben werden. Beim Prozess der inhaltlichen Korrektur werden diese fehlenden Datierungen nachgetragen werden (von der Redaktion des DAG).
- 2) <n.st.> bedeutet 'nouveau style' und bedeutet, dass der Jahreswechsel zu einem anderen Zeitpunkt als vom 31.12 auf den 01.01 erfolgt und die Datumsangabe umgerechnet wurde (wenn der Jahreswechsel auf das Osterfest fällt, also später im Jahr, wird ein Datum im Februar 1234 nach Aussage der Urkunde zu 1235 n.st., weil für unsere Zeitrechnung dann schon 1235 angebrochen ist. <a.st.> bedeutet dementsprechend 'ancien style', interessiert uns aber letztlich weniger. Jedenfalls: <n.st.> bzw. und vor allem <a.st.> müssen in das alphanumerische Feld mit aufgenommen werden. Bsp: <1249 n.st.>

## Numerische Datierung

In diesem Feld darf nur eine einzige Zahl als Datierung stehen. Es gibt eine Tabelle, auf der steht, wie die alphanumerischen Datierungen in numerische umgewandelt werden können. (Bsp.: 13e s. = 1250).

## Wichtiges

Wenn im alphanumerischen Feld eine Datierung einer gewissen Zeitspanne steht (Bsp.: 1122-1124), wird die Datierung für das numerische Feld gemittelt (in diesem Bsp. 1123).

Die Abkürzung <o. D.> bedeutet höchstwahrscheinlich ‘ohne Datierung’. Wenn trotzdem ein Datum dabeisteht, wird das <o. D.> zum Datum ins [Datierungs](#)-Feld geschrieben. Wenn kein Datum dabeisteht, wird das <o. D.> in den [Rest](#) geschrieben.

## Einzeldatum

In diesem Feld werden zusätzliche Datierungen, z. B. von Kopien des Primärwerks angegeben. Oder, wie oben im Beispiel, von Einzeltexten mit abweichender Datierung.

## Definition

Die Definition steht meistens zwischen Anführungs- und Schlusszeichen. Eine zweite Möglichkeit ist, dass die Definition mit einer anderen Farbe dazugeschrieben wurde und auch das Wörterbuch angegeben wird, wo diese Definition herkommt. Das Wörterbuch wird in den [Rest](#) geschrieben.

Eine Hilfe zum Entscheiden, ob die Definition die korrekte ist: Wenn man die Definition im Kontext anstelle des betreffenden Wortes einsetzen kann, sodass die Bedeutung und Satzstruktur des Kontextes erhalten bleibt, dann ist die Definition korrekt.

Schreibweise: die Anführungs- und Schlusszeichen weglassen. Bei mehreren Definitionen werden die einzelnen mit <;> abgetrennt

## Achtung

Oft steht bei der Definition <LvP> /<L-R> (neues Sigel <LR>)/<LR> oder <Gloss >. Dieses Sigel kann nicht in das Feld [Definition](#) übernommen werden, also wird es in den [Rest](#) geschrieben.

## Enzyklopädische Definition

In diesem Feld werden verschiedene Dinge aufgenommen:

- 1) Ausdrücke oder Redewendungen, in denen das unterstrichene Wort vorkommt, mitsamt ihrer Definition, wenn es diese gibt.



- 2) Erklärungen zum betreffenden Wort, die die Dimension einer Definition übersteigen und auch keine Übersetzungen des Kontextes darstellen.

## Kontext

Der Kontext wird so abgeschrieben, wie er auf dem Zettel steht, mit <...> <-> usw.

## Wichtiges

- 1) Bei Problemen mit der Leserlichkeit, kann der Betreffende unleserliche Buchstabe im Kontext durch <X> ersetzt werden. Wenn z.B. beim Wort <cens> der dritte Buchstabe unleserlich ist, wird der betreffende Buchstabe durch <X> ersetzt: <cens/ceXs>. Zusätzlich sollte auch ein Kommentar zum betreffenden Wort gemacht werden (im Feld [Kommentar](#)).
- 2) Grundsätzlich sollte sehr genau auf die Unterscheidung zwischen <n> und <u> geachtet werden.
- 3) Wenn im Kontext statt drei Punkten (für usw.) nur zwei bzw. mehr als drei stehen oder es keinen Abstand zwischen <...> und dem darauffolgenden Wort hat, wird die Graphie an die Konvention des Wörterbuchs DEAF angepasst, d.h. <... Wort... Wort, ...>.

## Rest

Im Rest werden zusätzliche Sigel angegeben (und u. U. die Stelle dazu) und alle Angaben, die in keines der anderen Felder gehören.

In dieses Feld werden auch die Übersetzungen des Kontextes auf Französisch eingetragen.

Bei mehreren, nicht zusammengehörenden Angaben, werden die einzelnen mit <;> abgetrennt.

## Achtung

Wenn zwei Sigel vorhanden sind, wird, wie oben beschrieben, eines in den [Rest](#) geschrieben. Dabei muss darauf geachtet werden, dass das Sigel in der definitiven Variante der Bibliographie übernommen wird. (z. B. <ACoutBord>: dazu steht in der Bibliographie: neues

Sigel <ACoutBordB>). Es muss also immer das neue Sigel aufgenommen werden, wenn eines vorhanden ist, ansonsten kann die Schreibweise des Zettels beibehalten werden.

## Kommentar

Im Kommentar-Feld können die EintipperInnen Kommentare zum Zettel notieren. Die Kommentare sind an die Redaktoren gerichtet. Es gibt einige Kürzel, die es den Redaktoren erleichtern, das Problem oder die Unklarheiten auf dem Zettel zu erkennen:

Lecture: d. h. ein Wort oder der Kontext ist unleserlich oder man ist sich nicht sicher, wie ein Wort/Buchstabe lautet. (Wort/Stelle angeben, die nicht leserlich ist)

Compréhension: unverständlich

Général: alle möglichen Kommentare, Fragen usw.

Wenn möglich sollten die Kommentare auf Französisch geschrieben werden.

## Kontrolle der Zetteleingabe durch eine Kollegin / einen Kollegen

Nach Abschluss der Digitalisierung eines Kastens muss der Kasten mit einem Häkchen im Feld „Eingabe“ markiert werden, dass der Kasten fertig ist. Dieses Kästchen findet sich auf der Seite mit der Kastentabelle. Wenn man seinen eigenen Kasten fertig bearbeitet hat, nimmt man sich einen fremden Kasten zur Gegenkontrolle vor. In der Datei „DigiDAG-Kästenverwaltung“ auf dem GoogleDrive schreibt man seinen Namen zum entsprechenden Kasten ins Feld „kontrolliert“.

Bei der Kontrolle muss man Zettel für Zettel überprüfen, ob alle Angaben in den richtigen Feldern stehen und ob Tippfehler gemacht wurden. Hier die Zeit nehmen, die nötig ist!

Wenn die Kontrolle fertig ist, dann markiert man das mit einem Häkchen im Feld „Eingabe-K“ (= -Kontrolle) auf der Seite mit der Kastentabelle.

(Das Kästchen „Inhalt-K“ (= -Kontrolle) muss die Redaktion des DAG anklicken, wenn sie den Inhalt (auf verschiedene Dinge hin) kontrolliert hat.)

Stand 30.09.2015, Seraina Montigel, gegengelesen von Sabine Tittel und Martin Glessgen;  
Überarbeitet von Kerstin Kerber in Absprache mit Sabine Tittel

## ***Directives pour la rédaction d'articles pour le DAG***

### **Table des matières**

La rédaction d'articles pour le Dictionnaire d'Ancien Gascon (DAG) avec le système de rédaction « Marie » ....	2
1 Informations générales .....	2
1.1 Les caractères spéciaux .....	2
2 Sous <b>Datenpflege</b> (sur la page d'accueil de « Marie ») .....	2
2.1 <b>Zettel</b> (fiche) .....	2
2.2 <b>Sprachen</b> (langues) .....	2
2.3 <b>Autres Langues</b> .....	2
2.4 <b>Wortarten</b> (catégories grammaticales) (pos = part of speech) .....	3
2.5 <b>Tertiärliteraturverwaltung</b> .....	3
2.6 <b>Import von Paketen</b> (l'importation de paquets) .....	3
2.7 <b>Import von Kategorien</b> (l'importation de catégories) .....	3
2.8 <b>XML-Schema-Verwaltung</b> (l'administration du schéma XML) .....	3
2.9 <b>Verwaltung der XML-Elemente</b> (l'administration des éléments XML) .....	3
3 La rédaction d'articles .....	3
3.1 Informations générales .....	3
3.2 Choisir une catégorie .....	4
3.3 le rédacteur et le correcteur de l'article .....	4
3.4 <b>Changer le lemme</b> .....	4
3.4.1 Le choix des lemmes .....	4
3.5 La vérification des fiches et la transformation en références .....	5
3.5.1 La correction des contextes .....	5
3.5.2 Les conventions pour l'entrée .....	6
3.6 <b>Titelblock</b> (le bloc du titre) .....	6
3.7 <b>Etymologie</b> (l'étymologie) .....	7
3.8 <b>Graphien</b> (les graphies) .....	7
3.9 <b>Bedeutung</b> (la définition) .....	8
3.9.1 Changer le niveau de la définition (HL → UL / UL → HL) .....	9
3.9.2 La catégorie onomasiologique .....	9
3.9.3 La définition .....	9
3.9.4 Le contexte visible / invisible .....	9
3.9.5 La datation des références .....	10
3.10 Le balisage .....	10
3.10.1 Informations générales .....	10
3.10.2 Les différentes balises - leur signification et leur utilisation .....	10
4 Exporter les articles et les regarder après .....	13

## La rédaction d'articles pour le *Dictionnaire de l'ancien gascon (DAG)* avec le système de rédaction « Marie »

### 1 Informations générales

En cas de problèmes (techniques ou autres) le mieux serait que le rédacteur s'adresse directement à Sabine Tittel. Si le logiciel produit un rapport d'erreurs, celui-ci devrait être transmis tel quel pour faciliter le travail à l'informaticien qui doit résoudre le problème.

#### 1.1 Les caractères spéciaux

Il existe une série de caractères spéciaux qu'on peut créer à l'aide de différentes combinaisons de touches.

Le cadre avec les caractères spéciaux apparaît dès qu'on appuie en même temps sur les trois touches **touche de commande (Ctrl)** et **shift** et **y**. On laisse appuyé les deux premières touches et appuie sur la touche **y** autant de fois que nécessaire pour que le caractère spécial apparaisse. À ce moment-là, il faut lâcher les trois touches et le système rajoute le caractère spécial souhaité.

Il existe d'autres combinaisons de touches (**touche de commande (Ctrl)** + **shift** + **1-6**), avec lesquelles le rédacteur peut rajouter directement le caractère spécial en question :

<b>touche de commande (Ctrl) + shift</b>	+ 1	correspond à	«
	+ 2	correspond à	»
	+ 3	correspond à	“
	+ 4	correspond à	”
	+ 5	correspond à	‘
	+ 6	correspond à	’

### 2 Sous **Datenpflege** (sur la page d'accueil de « Marie »)

#### 2.1 **Zettel** (fiche)

Il est possible de chercher sous **Zettel** une fiche dans l'ensemble de fiches déjà importés sur « Marie ».

#### 2.2 **Sprachen** (langues)

Sous **Sprachen** il y a une liste de toutes les langues et tous les dialectes possible. Cette liste peut être élargie s'il est nécessaire.

**sans localisation**: si le rédacteur a trois références picardes (pic.), les références (A, B und C) et une référence sans indication de langue ou dialecte (D), le rédacteur doit l'écrire de la manière suivante : pic. A, B, C, D sans loc. Il utilise donc **sans localisation** pour accentuer que l'élément munie avec cette information ne possède pas la même propriété que les éléments précédents. Dans notre exemple sans localisation indique que l'élément D n'est pas pic.

#### 2.3 **Autres Langues**

Dans **Autres Langues** il y a une liste de toutes les langues qui peuvent apparaître comme adstrates, superstrates ou comme étymons. La liste est très longue et peut ultérieurement être élargie si nécessaire.

## 2.4 Wortarten (catégories grammaticales) (pos = part of speech)

Dans **Wortarten** on trouve toutes les catégories grammaticales possibles lors de la rédaction d'un article.

## 2.5 Tertiärliteraturverwaltung

### 2.6 Import von Paketen (l'importation de paquets)

Pour pouvoir rédiger des articles avec le système « Marie » on doit importer les données du système « Christine » où les fiches ont été digitalisées et lémmatisées auparavant.

On peut seulement importer les caisses qui sont digitalisées complètement et qui ont été corrigées par quelqu'un (ce qui signifie que les trois cadres dans « Christine » sont cochés).

Lors de l'importation d'un paquet de données, le rédacteur clique sur **Import von Paketen**. Ensuite apparaissent les paquets prêts pour l'importation. On peut choisir un paquet et l'importer en cliquant sur **importieren**.

### 2.7 Import von Kategorien (l'importation de catégories)

Si le système onomasiologique de Hallig/Wartburg semble trop imprécis lors de la numérisation (sur « Christine ») on peut introduire de nouvelles catégories onomasiologiques et désigner les fiches avec ces nouvelles catégories onomasiologiques. Afin que les fiches désignées par une telle catégorie puissent être importées dans « Marie », le rédacteur doit importer cette nouvelle catégorie onomasiologique dans « Marie ». Sinon, le système ne reconnaît pas la catégorie et ne comprend pas ce qu'il doit faire avec les fiches correspondants.

Après l'importation de la nouvelle catégorie, le rédacteur doit cliquer sur **Alle Kategorien** et **synchronisieren**. Ainsi, il synchronise toutes les catégories y compris la nouvelle catégorie.

### 2.8 XML-Schema-Verwaltung (l'administration du schéma XML)

### 2.9 Verwaltung der XML-Elemente (l'administration des éléments XML)

Sous **Verwaltung der XML-Elemente** apparaît une liste de tous les balises qui peuvent être utilisées lors de la rédaction d'un article. On peut élargir la liste en cas de besoin.

## 3 La rédaction d'articles

### 3.1 Informations générales

Lors de la rédaction du **DEAF** pré on ne consulte pas de sources primaires pour le contrôle. Lors de la rédaction du **DEAF** on le fait en revanche. Dans ce dernier cas, on contrôle toutes les dates et les vérifie avant de les publier. La démarche lors de la rédaction d'articles pour le **DAG** est plus proche de celle du **DEAF** pré, mais avec néanmoins une part de contrôle: on contrôle en effet seulement le contexte et la définition dans les sources primaires qui paraissent étranges au rédacteur.

Les articles présentent la structure suivante :

- tête de l'article

- étymologie

- bloc des variantes graphiques (attestations structurées, avec leur datation)

- bloc des sens (attestations avec définition et datation)

## 3.2 Choisir une catégorie

Suivant le cas, on peut ordonner tous les lemmes importés dans « Marie » selon l'alphabet ou selon la catégorie onomasiologique, en choisissant alors **Artikelauswahl alphabetisch** ou **Artikelauswahl nach Kategorie**.

Avant la rédaction, le rédacteur choisit une catégorie onomasiologique (cf. le document avec l'ordre chronologique des catégories à traiter; actuellement, on travaille sur B IV, puis sur C). C'est plus facile de travailler différents lemmes appartenant à une même catégorie onomasiologique. De cette manière on peut mieux se familiariser avec la terminologie et on a une meilleure vue d'ensemble sur la catégorie.

Le rédacteur note à la main tous les lemmes d'une catégorie et les recherche ensuite de nouveau dans l'ordre alphabétique avant de commencer la rédaction des articles correspondants. De cette façon, on peut retrouver toutes les formes qui ont été placées dans différentes catégories onomasiologiques ou qui réunir des lemmes qui ont été séparés par erreur lors de la saisie des fiches.

Dès qu'on clique sur **Artikelauswahl nach Kategorie** apparaissent les lemmes ordonnés d'après la catégorie. En cliquant sur le cadre **Hauptlemma** on peut chercher un mot particulier ou une catégorie.

**Caractères de remplacements** qui peuvent être utilisés lors de la recherche :

- **ac%** : le système filtre tous les mots qui commencent par <ac> et continuent avec n'importe quel nombre de lettres → *accepte, acapte, acaptement* etc.
- **ac\_pte**: à la place où se trouve le caractère de remplacement peut se trouver n'importe quelle lettre → *accepte, acapte* etc.

## 3.3 le rédacteur et le correcteur de l'article

Dès que le rédacteur a choisi une catégorie et est prêt pour la rédaction du premier article du premier lemme dans l'ordre alphabétique, il écrit les premiers trois lettres de son nom de famille dans le cadre **Autor** à droite. On fait cela avant de commencer avec la rédaction pour que personne ne commence à rédiger le même article.

Dès que l'article est rédigé, le rédacteur coche la case **redigiert**. Après la correction de l'article le correcteur écrit aussi les trois premières lettres de son nom de famille dans le cadre **Korrektor** et clique sur le cadre **korrigiert**. On ne peut pas utiliser d'autres abréviations. On doit donc écrire « Gle » et pas, par exemple, « MGle » (première lettre du prénom et les trois premières lettres du nom).

Important : si l'on écrit quelque chose dans un cadre quelconque, on doit sauvegarder cette étape de travail avant de continuer la rédaction !

## 3.4 Changer le lemme

Si rien n'a été rédigé, on peut changer les lemmes et les sous-lemmes pour les rendre identiques (dans n'importe quelle direction). Si un premier lemme a été rédigé, un deuxième (que l'on veut unir au premier) non, il faut changer le deuxième sous le nom du premier (quitte à rechanger le nom après): c'est toujours la cible qui reste inchangée. Si les deux mots ont déjà été rédigés, il faut comparer auparavant les deux articles définitifs, parce que les informations du lemme changé seront perdues (sauf les références!).

### 3.4.1 Le choix des lemmes

Si un HL [= Hauptlemma = lemme principal] \**administrar* n'est attesté qu'une seule fois et sous la forme concrète *amenistrar*, on doit changer le lemme en *amenistrar* (= on évite des HL non attestés). Si autant *administrar* qu'*amenistrar* sont attestés, les formes héréditaires sont favorisés: ce sont elles

qui forment les lemmes / HL. Les formes savantes sont alors placées sous ces HL en tant que sous-lemmes / UL [Unterlemma] (puisque l'on considère qu'ils sont étymologiquement secondaires par rapport aux mots héréditaires).

EX: *affiuament* est un UL pour le HL *affiuar*, *affiusament* un UL pour le HL *\*affiusar* (cf. FEW: «apr. *afeuzar* (besonders gasc.)»). Dans le cas concret, le tout est placé sous un lemme HL *\*fiu* < *\*fehu*.

Principe général: un lemme correspond à une base protoromane (= latin écrit ou latin parlé avant *ca* 700); tout ce qui en est dérivé (= après *ca* 700) devient un sous-lemme.

### 3.5 La vérification des fiches et la transformation en références

Dès que le rédacteur a choisi un lemme qu'il va rédiger, il clique sur **Zettelüberprüfung**. On doit corriger toutes les fiches avant de les changer en références en choisissant **Beleg hinzufügen**. Pour les étapes de travail suivantes on ne tient compte que des références. Pour cette raison, le contrôle des fiches est très important. À cette occasion, le rédacteur doit contrôler tous les cadres et, le cas échéant, changer la catégorie si elle est mal choisie. Si la catégorie grammaticale manque le rédacteur doit la choisir de la liste. C'est aussi possible qu'il manque la localisation ou la langue. Pour obtenir ces informations, le rédacteur peut consulter la bibliographie. Pendant cette étape de travail, le rédacteur peut écrire des commentaires invisibles (**blinde Kommentare**) pour chaque fiche, comme p. ex. des remarques par rapport à la graphie ou à la langue. Ces commentaires facilitent ensuite la rédaction du **Titelblock** ou du **Graphienblock** où il ne doit plus regarder chaque fiche en consultant ses commentaires invisibles. Il est recommandable d'écrire des commentaires courts parce que sinon on ne voit pas tout le commentaire lors de la rédaction du **Titelblock** ou du **Graphienblock**.

Important: tous les cadres marqués avec une étoile rouge doivent être remplis.

Pour la conversion de la datation en une datation numérique, le rédacteur peut consulter le document **Datumsangaben**. Exemple : datation : 13<sup>e</sup> s. → datation numérique : 1250

Important : Si la datation est 1200-1350, on doit calculer la moyenne pour la datation numérique. Dans l'exemple, la datation numérique serait donc 1275. La datation peut très souvent être précisée en consultant les livres dont on cite un petit contexte. La grande partie de ces livres se trouvent à Heidelberg et le rédacteur peut les consulter lui-même. S'il n'a pas la possibilité de les consulter à Heidelberg, il peut élaborer une liste des fiches dont les datations doivent encore être précisées. Ensuite un collaborateur à Heidelberg peut les vérifier.

#### 3.5.1 La correction des contextes

EX, [avantatge :]

sur la fiche : *alqual Assalhit, no filh, premier nat, eu laissi avantatges tot quant que eu iei.*

on corrige : *al\_qual Assalhit, [m]o filh, premier nat, eu laissi avantatges tot quant que eu iei ...*

#### **explication:**

(1) Les fiches ne sont normalement relativement fiables, mais pas totalement : ils peuvent contenir des fautes produites par la personne qui a recopié la fiche, par la personne qui l'a écrite, par la personne qui a fait l'édition de texte ou même par la personne qui a copié le texte au Moyen Âge. On a donc quatre niveaux sur lesquels peuvent s'être produits des fautes que nous n'avons pas le temps de démêler.

(2) Le contexte sert à mieux comprendre le sens du mot et le savoir encyclopédique qu'il véhicule. Sa représentation dans le DAG devrait ne pas être diplomatique, mais critique ('interprétative'), d'après les critères appliqués dans les DocLing : une ponctuation moderne, un emploi des majuscules



moderne, une coupure de mots moderne (marquée comme telle) et aussi des formes de mots émendées (marquées entre crochets).

Dans le cas concret nous avons donc "al-qual" et "[m]o" ainsi qu'une espace avant les trois petits points.

### 3.5.2 Les conventions pour l'entrée

On indique l'endroit de la manière suivante : tome, virgule, pas d'espace, page  
EX «ArchGir 32, p.87» devient «ArchGir 32,87».

Si l'on écrit <p.87>, le système ajoute automatiquement une espace entre <p.> et <87> mais cette fonction ne marche pas si le rédacteur écrit lui-même l'espace. On ne devrait donc pas écrire des espaces manuellement.

indication de	exemple	explication
<b>tome et page</b>	ArchGir 3,23 non pas ArchGir t.3 p.23	sans <i>t.</i> et <i>p.</i> sans blanc entre le tome et la page
sources ne comprenant qu'un volume	ACoutBordB 31 non pas ACoutBord p.31	même chose, sans indication du tome
<b>colonnes a et b</b>	CoutMontsaunèsM 34a[28]	distinguer les deux colonnes a et b qui se trouvent sur la même page
<b>lignes, paragraphes ou chapitres</b>	Mill, ArchGir 45,10[1] et non pas Mill p. 10, l. 1	entre [ ]
<b>paragraphes avec des indications de lignes</b>	CartBigRC 74[57;4] et non pas CartBigRC p. 74, §57, l. 4	mettre tous les deux entre [ ] avec un point virgule après le numéro de paragraphe
numéro du <b>folio</b>	Mill, ArchGir 45,122[f<sup>o</sup>29;10]	sur la même page on a plusieurs lignes 10 contenues dans des folios différents
une <b>certaine période</b>	1300–1345 ou Mill, ArchGir 45,115–151	utiliser un long tiret pour signaler une certaine période
<b>notes</b>	Luc 39 n2 non pas Luc p. 39, note 3	

Il serait préférable pour l'indication des sources dotées de sigles, contenues dans un ouvrage tel que ArchGir, Monl, RecGascL, etc., d'intégrer l'ouvrage en question juste avant les autres données sous **Stelle**:

EX

Sigel: RecFeodD

Stelle: ArchGir 5,268[197]

### 3.6 Titelblock (le bloc du titre)

Sous **Titelblock** on peut indiquer différentes informations : **Lemma ist erschlossen** signifie que le lemme est reconstruit d'après les évolutions phonétiques régulières, mais que ce lemme n'est pas attesté sous cette forme.

Le rédacteur peut ignorer le cadre en haut **Bitte wählen** où sont énumérées les différentes sortes d'articles.

Dans la plupart des cas, la langue est l'ancien gascon (agasc.). Pour cette raison, cette langue est automatiquement rajoutée par le système comme choix de base. Si la langue n'est pas agasc., le rédacteur peut choisir la langue dans la liste ou, le cas échéant, choisir une deuxième langue.

important : S'il s'agit de deux langues, il est important de choisir deux langues sur le même niveau, p. ex. abéarn. et aland. ensemble, mais pas agasc. et abéarn., parce que la région où est parlé abéarn. se trouve dans la région plus grande dans laquelle on parle agasc. (agasc. est un hypéronyme de abéarn.)

Le cadre Wortart doit être rempli aussi s'il n'y a pas d'indication de la catégorie grammaticale sur la fiche. Le rédacteur devrait connaître les différentes désignations et devrait lui-même évaluer de quel catégorie grammaticale qu'il s'agit.

### 3.7 Etymologie (l'étymologie)

L'étymologie n'est renseignée que pour le HL, pas pour les UL!

– pour les UL on peut/doit en revanche ajouter, tout comme pour les HL les commentaires (cf. infra)

– le commentaire (autant dans HL que dans UL) indique:

= si le lexème/dérivé est bien attesté dans FEW ou non (enfin, on indique s'il n'est pas bien attesté; s'il est bien attesté, on ne dit rien)

= concrètement, on indique "forme / sens / syntagme manque dans FEW pour le gasc. / l'agasc.", "première attestation par rapport à FEW", év. "documentation faible dans FEW"

= quelle est sa diffusion géolinguistique dans la Gallo-/Ibéroromania

– forme du commentaire:

EX

[le dérivé s. f. manque dans FEW (seulement apr. <cited-word index="a" language="apr. ">avinhat</cited-word> adj. "planté de vignes" LvP)]

– donc: on marque les formes citées du FEW qui ne se trouvent pas dans notre corpus par une balise.

Le DAG renvoie pour l'étymologie au FEW. On le cite de la manière suivante : **FEW tome, page a/b**. Le rédacteur doit chercher l'étymon dans le FEW. Dès qu'il l'a trouvé et connaît l'endroit dans ce dernier, il peut cliquer sur **Initialwert** et le système rajoute automatiquement les balises XML dans le cadre. Ensuite le rédacteur doit écrire les informations au bon endroit entre les balises XML.

Bsp: <evidence><siglum>FEW< / siglum> <text-reference>2,318a< / text-reference></evidence> lt. <cited-word index="e" language="lt.">captare< / cited-word>

Les balises entourent toujours l'information, c'est-à dire devant le sigle « FEW » il y a toujours la balise <siglum> et après « FEW » la balise < / siglum>. Le deuxième élément de la balise commence toujours avec une barre oblique. En cliquant sur le cadre **validieren**, le système contrôle si les balises sont écrites correctement. Si le rédacteur a cliqué sur **Initialwert**, les balises sont très probablement correctes mais ils peuvent contenir des fautes si le rédacteur a écrit lui-même les balises à l'aide des fonctions au-dessous du cadre.

À propos des significations des différentes balises, voir le chapitre 3.10.2 qui les explique.

Dès que les informations se trouvent au bon endroit, le rédacteur devrait sauvegarder l'étape de travail. Ensuite apparaît à droite la présentation de l'indication étymologique comme elle serait représentée sans les éléments XML dans le dictionnaire. La représentation est colorée ce qui signifie qu'elle contient des balises.

### 3.8 Graphien (les graphies)

Le rédacteur doit toujours écrire le mot dans la bonne forme dans laquelle la graphie va être notée, c'est-à dire l'infinitif d'un verbe, le singulier d'un substantif, le singulier masculin d'un adjectif. Si l'on a seulement une ou deux attestations, cela semble exagéré. Mais dès qu'on a beaucoup de

références avec des différentes variantes graphiques, on doit ranger les formes pour qu'on puisse créer effectivement et correctement les variantes graphiques. Le système trie les entrées dans le bloc des graphies par ordre alphabétique d'après la forme des mots (et ensuite par la datation des sigles). Si l'on a toujours traité correctement les formes, tout est trié correctement. Sinon, on perd la vue d'ensemble et ne sait plus quelle référence doit être mise à côté de quelle graphie.

Dès que le rédacteur a contrôlé les fiches et les transformés en références, les différentes graphies apparaissent à gauche dans le cadre au-dessous du cadre **Graphien**.

Les graphies doivent être ordonnées à droite d'après leur langue et la langue du contexte, p. ex. « agasc. en contexte lt. ».

Important : l'ordre des catégories est toujours la suivante :

1. agasc. en contexte lat.
2. agasc. en contexte gasc.
3. agasc. en contexte fr.

À droite de chaque graphie apparaissent deux flèches rouges dirigées en haut et en bas. À l'aide de ces flèches, le rédacteur peut changer l'ordre des graphies.

S'il y a différentes graphies d'un mot dans différentes catégories sémantiques, on devrait ajouter une ligne vide entre les deux graphies. La balise pour le changement de la ligne est **br**. On doit cliquer deux fois sur **br** pour obtenir un changement de la ligne et une ligne vide entre les deux graphies.

Pour sauvegarder les références qui apparaissent à gauche à droite comme 'graphies', le rédacteur peut cliquer sur **Gr** ou tirer la référence à la main du côté gauche au côté droit dans la bonne catégorie graphique.

Si le rédacteur écrit manuellement la graphie en cliquant sur **Neue Graphie** il doit écrire le mot au bon endroit entre les balises.

Important: toujours sauvegarder l'étape de travail!

Dès que le rédacteur clique sur une graphie (à droite), un cadre apparaît au-dessous du cadre des différentes graphies dans lequel doit être indiquée la datation des graphies. On doit marquer la plus ancienne graphie comme telle. On peut le faire en cliquant sur la plus ancienne graphie et après sur **Datum ältester Beleg (Pos 0)**. En cliquant sur **Sortieren**, le système ordonne les graphies d'après leur datation.

### **3.9 Bedeutung (la définition)**

Le principe est le même sous **Bedeutungen** comme sous **Graphien** : les références à gauche doivent être déplacées à droite d'après leur sens en les tirant sur la Hauptbedeutung ou Unterbedeutung appropriée ou en cliquant sur la petite flèche (>) qui transporte la référence dans la Hauptbedeutung ou la Unterbedeutung marquée à droite. Si toutes les références ont la même signification, le rédacteur peut transporter toutes les références en même temps de gauche à droite en cliquant sur **Alle** et en tirant tous les références à droite.

Les références sont divisées en sources primaires, secondaires et tertiaires. Des sources secondaires sont des analyses de références primaires et les sources tertiaires sont des références à des dictionnaires. Quelques glossaires renvoient à une source primaire et pourraient donc en principe être traités comme des sources primaires.

Le rédacteur n'a en principe plus besoin de la colonne à gauche après avoir transporté les références à droite. Il peut cliquer sur **Linke Spalte ausblenden** et masquer ainsi la colonne gauche.

Le rédacteur doit créer un arbre des définitions à droite. Il doit donc décider quelles sont les Hauptbedeutungen (définitions primaires) et quelles les Unterbedeutungen (définitions secondaires) et ensuite sauvegarder les différentes références en cliquant sur **Neue Hauptbedeutung** ou **Neue Unterbedeutung** (sous **XML speichern**). On doit écrire la définition entre deux balises. **Alle ausklappen** signifie qu'on voit toutes les sous-catégories, **Alle einklappen** signifie qu'on voit que les Hauptbedeutungen.

S'il n'y a pas de définition du mot sur le fiche, le rédacteur doit lui-même chercher une définition avec le concours du *TLFi*, du *Petit Robert* ou du *DEAFplus*.

En principe, tous les noms propres (p. ex. *esparvers* CartGimC) doivent être catégorisés comme Unterbedeutungen.

### 3.9.1 Changer le niveau de la définition (HL → UL / UL → HL)

Si le rédacteur veut transformer une Hauptbedeutung en une Unterbedeutung, il doit tirer la Hauptbedeutung avec le curseur sur la Hauptbedeutung dont il veut créer une Unterbedeutung et sa Hauptbedeutung devient l'Unterbedeutung (elle devient rouge) de cette dernière.

S'il veut en revanche transformer une Unterbedeutung en Hauptbedeutung, il peut cliquer sur le cadre **Wurzel anzeigen** qui crée une nouvelle Hauptbedeutung qui est vide. Ensuite il peut tirer la Unterbedeutung sur cette vide Hauptbedeutung ce qui la transforme en Hauptbedeutung.

### 3.9.2 La catégorie onomasiologique

Le rédacteur doit toujours contrôler la définition du mot dans la liste des abréviations onomasiologiques de Hallig/Wartburg avant d'indiquer la catégorie sous **Onomasiologische Kategorie**. Le système rajoute de nouveau lui-même les balises et la signification de l'abréviation indiquée par le rédacteur.

### 3.9.3 La définition

On indique la définition dans une logique aristotélicienne, c'est-à-dire en passant d'une définition plus générale à une définition plus spécifique.

EX: le faucon sera défini "oiseau, oiseau de proie, ... fr. *faucon*" (la dernière entrée est en fait une traduction en français moderne).

Important : en indiquant la définition, le rédacteur doit toujours penser au fait que les mots en question se réfèrent à une époque médiévale, où les références quotidiennes et culturelles étaient profondément différentes des références actuelles. La définition doit toujours se rapporter à l'époque durant laquelle le mot a été écrit.

### 3.9.4 Le contexte visible / invisible

Dès que l'on a cliqué sur une définition apparaît au-dessous une référence correspondante. À droite de ce cadre, le rédacteur peut indiquer si le contexte doit être visible automatiquement dans l'article en cliquant sur le petit cadre.

Important : le rédacteur peut marquer les cadres de toutes les références pour que tous les contextes soient visibles.

### 3.9.5 La datation des références

Comme pour les graphies, le rédacteur peut changer l'ordre des références à l'aide des flèches rouges ou, si la référence devrait se placer dix lignes au-dessous, en cliquant sur le nombre et en indiquant ensuite le nombre de l'endroit où la référence devrait être placée. La fonction **Sortieren** ordonne les références par ordre chronologique mais pas d'après la page. Cela peut éventuellement être important quand une référence apparaît sur deux pages différentes.

On doit indiquer une datation de chaque mot, c'est-à-dire la date de la première et de la dernière attestation. Le rédacteur peut le faire en cliquant sur **Datierung**, sur **Neu**, et ensuite sur **Datierung speichern**. Si l'on n'a que deux attestations d'un mot, une de 1100 et une de 1500, on doit indiquer deux datations isolées où l'on indique sous **Anfang Dat.** sous **Ende Dat.** la même date. On indique la date de manière numérique sous **ANumdate** et **ENumdate** (pour la conversion de datations en datations numériques, voir le document **Datumsangaben**).

Les mots doivent être attestés au moins une fois par siècle pour pouvoir indiquer la datation « de XX jusqu'à XX ». Dans ce cas le rédacteur indique sous **Anfang Dat.** la date de la première attestation et sous **Ende Dat.** la date de la dernière attestation, toujours en indiquant la datation numérique correspondante.

Si les datations des attestations comportent des sautes de plus de cent ans, on doit créer des datations isolées comme décrit en haut.

Si le rédacteur écrit la date (p. ex. 1480) que dans **Anfang Dat.** et **ANumdate** et pas dans **Ende Dat.** et **ENumdate**, la datation définitive de ce mot dans l'article serait « à partir de 1480 » sans limitation jusqu'à quand le mot est attesté.

Important : en indiquant la datation on ne doit pas rajouter une espace entre le siècle et le nombre et pas non plus entre le *e* (ième) et le *s* (siècle).

Exemple : 18es.

À la hauteur du titre **Datierung** se trouve à droite en bleu **Anzeigen**. En cliquant sur ce cadre, le rédacteur voit le résultat de ce qu'il a écrit et peut contrôler s'il a fait des fautes.

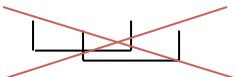
## 3.10 Le balisage

### 3.10.1 Informations générales

Dans le chapitre sur l'étymologie nous avons déjà dit qu'une balise se compose toujours de deux parties et que le l'élément qui porte le caractère de la balise est écrit entre ces deux éléments. Une série de balises peut être construite de la manière suivante :

1.  l'un après l'autre

2.  complexe

~~~~ mais PAS croisé

### 3.10.2 Les différentes balises - leur signification et leur utilisation

- **siglum**

En indiquant l'étymologie, on balise le sigle avec la balise **siglum**. Le rédacteur doit décider si le sigle sera visible ou invisible dans l'article. En cliquant sur **visible** et ensuite sur **true**, le sigle sera visible dans l'article, en cliquant sur **false**, il sera invisible.

- **text-reference**

On balise la référence au texte par la balise **text-reference**.

- **evidence**

On balise ensemble le sigle et l'endroit comme **evidence**.

- **cat-onomas**

On balise la catégorie onomasiologique par **cat-onomas**.

- **quotation**

Le contexte après le sigle et la référence textuelle est balisé comme **quotation**, toujours dans la balise **évidence**. Un texte qui est balisé comme **quotation** est toujours un texte courant sans sigle et sans endroit qui n'est pas écrit entre [].

Important : **XML** est spécifique pour le contenu, c'est-à-dire on marque p. ex. une citation avec XML. **HTML** se réfère en revanche à la mise en page et on marque p. ex. un mot en italique avec HTML. Il existe des balises qui indiquent plutôt des éléments XML et des balises qui influencent la mise en page comme p. ex. la balise **citation** qui met le passage balisé en italique.

- **citation**

Le texte balisé comme **citation** apparaît en italique dans l'article. Un texte balisé comme **citation** et **quotation** apparaît aussi en italique.

- **cited-letters**

Exemple : si on veut attirer l'attention sur le fait que la lettre f d'un mot provient de la lettre grecque φ, on balise les deux lettres comme **cited-letters**. On utilise aussi cette balise pour les terminaisons qu'on veut regarder d'un peu plus près.

- **cited-word**

On balise un mot qu'on regarde de plus près comme **cited-word**. Dès qu'un mot est balisé comme **cited-word** apparaît un nouveau cadre où le rédacteur peut choisir différentes choses.

Tout en haut apparaît **visible**. Le rédacteur peut le laisser ainsi. Au-dessous apparaît le cadre **Index** où sont indiquées différentes abréviations :

- o** n'apparaît pas dans un index ; le mot est écrit en italique
- e** il s'agit d'un étymon
- f** il s'agit d'un mot gascon (l'abréviation devrait donc être g mais elle est f parce que le système est repris du DEAF)
- a** l'abréviation pour autres langues

- **comment**

Si le rédacteur veut commenter un mot, p. ex. à l'intérieur d'un contexte, il doit **le** baliser comme **comment** et aussi le **mot** qu'il commente comme **cited-word**.

p. ex. *dorretche* (**commentaire du rédacteur** **dorretche**) cum...

- **definition**

Un mot balisé comme **definition** apparaît dans le dictionnaire entre guillemets.

- **designation**

p. ex. l'épée de Roland : **excalibur**

Une **designation** est une désignation (ici **excalibur**) qui s'est fixée bien qu'elle était au début un nom propre.

- **collocation**

Dans le bloc des définitions on balise une désignation fixe avec le tag **collocation**, p. ex. deux mots qui sont toujours utilisés ensemble. Le terme collocation est assez général et le rédacteur ne peut presque pas faire une faute en balisant des mots comme **collocation**. Mais la balise n'est pas si spécifique comme les balises **locution**, **compound**, **terminology** etc.

- **locution**

Une **locution** est similaire d'une **collocation**. Un exemple d'une **locution** verbale est « faire joie ».

- **compound**

Un **compound** est une expression lexicalisé, comme p. ex. « pomme de terre ».

- **terminology**

On doit baliser la terminologie spécifique comme **terminology**. Dès que le rédacteur choisit cette balise, il peut choisir la branche à laquelle appartient le mot.

- **usage**

La balise **usage** a différents emplois. Elle peut être utilisée pour des moyens de style ou des expressions très répandues.

- **footnote**

Le rédacteur doit baliser les notes de bas de page comme **footnote**.

- **proverb**

On balise les proverbes comme **proverb**. On indique les proverbes toujours comme Unterbedeutung dans l'arbre des définitions.

- **manuscript**

On utilise la balise **manuscript** pour marquer des manuscrits comme tels.

- **pos**

**pos** est l'abréviation de « part of speech », donc pour la catégorie grammaticale. Avec cette balise on indique la catégorie grammaticale.

- **remark**

Le rédacteur balise tous ses commentaires avec **remark**. S'il ne fait qu'un seul commentaire, le rédacteur ne doit pas ajouter un chiffre mais s'il en fait plusieurs, il doit les numéroter.

- **remark-reference**

Exemple : si à un endroit est écrit „[voir la remarque no. 2]“ on balise „no.2“ comme **remark-reference**. Par cette voie, le système dirige le lecteur directement à l'endroit où se trouve cette remarque.

- **link**

On balise tous les liens avec la balise **link**.

- **br**

Le tag **br** provoque un changement de ligne.

#### **4 Exporter les articles et les regarder après**

Dès qu'un article est rédigé, le rédacteur doit le marquer comme fini (dans le cadre à côté des trois premières lettres de son nom). Dès qu'il a marqué ce cadre, il peut aller chercher son article sur la page internet <https://dag.adw.uni-heidelber.de/administration> où il peut l'exporter en cliquant sur **exportieren**.

Sur la page <https://dag.adw.uni-heidelberg.de>, tous les collaborateurs peuvent regarder les articles rédigés bien qu'ils ne soient pas encore mis en ligne publiquement.



DAG\_1300, article *austor* :

1509 **autour:**

1. ***austor*** m. 1279 LAgC 10[10]; [1288] CoutMontsaunèsM 44a[40]; 1295 CoutFezB, ArchGir 27,390[36], ***astor*** 1279 LAgC 13[47].

FEW ACCEPTOR 24,72a: Gers *astouí*, béarn. *estour*; *astour*; *austouí*; LR sub *austour*: abéarn. *austor* ca. 1400 ForsB; [s.d.] ArchBPyr; cf. acat. *astor* Ramon Llull: DoctrPuer, CoromCat 1,464a.

◇ Rem. V. sub épervier ci-dessus.

1 *austor*

1279 Agen: En B.W. de la Mota, per lor e per lors parcerers devo I *austor* d'acapte o 30 s. de morl. per lo tenh e per l'afar de Godbes e per las apartenensas.

DAGél, articles *austor*, *austoron* :

rédação: Martin Glessgen

**AUSTOR** m.

[FEW 24,72a lt. \*ACCEPTOR.

Étymon proto-roman: \*AUCEPTOR.

REM.: Les documents latins en Gascogne connaissent par ailleurs fréquemment la forme lt. class. *accipiter*, var. *ancipiter*, cf. le contexte de l'attestation de agasc. lt. *austorio* (1286) et ce celle pour *austoron* dans (CoutFez 1295).]

### ***Inventaire des formes***

- en contexte latin: agasc. lt. *asturio* AnglGuyB Gloss 320;341 ArchGir 3,25[97], agasc. lt. *austorem* ArchGir 16, 270, agasc. lt. *austorio* ArchGir 1, 362 [47]<sup>(1)</sup>, agasc. lt. *austuro* AnglGuyB Gloss 320;341 ArchGir 3 28[109]; ArchGir 5, p. 278, agasc. lt. *austurre* AnglGuyB Gloss 320;341 ArchGir 31[118], agasc. lt. *osterio* ArchGir 5, p. 323;
- en contexte gascon: agasc. *austor* LAgC p. 10; CoutFezB 1295 ArchGir 27,390[36]; ForGén p. 34; CoutDax ArchGir 37,133[615]; CoutMontsaunèsM p. 44, agasc. *austur* AnglGuyB Gloss, agasc. *astou(r)* CoutMontsaunèsM p. 44<sup>(2)</sup>.

### ***Sens et syntagmes***

- ◆ “oiseau de chasse, autour (*accipiter gentilis*)” / A IV b 3: Les oiseaux de proie

[REM. ENCYCL.: (1) L'autour sert dans de nombreux contextes pour payer un cens (1268 ArchGir 1), parfois un cens emphythéotique (1279 LAgC) ou une redevance féodale (15<sup>e</sup> s. CoutDax); sa valeur est identifiée par l'équivalence avec 30 sous de Morlaas (1279 LAgC); cf. aussi les contextes du syntagme *austuro sauro*.

(2) L'autour sauvage est soumis à l'autorité seigneuriale (cf. par ex. la défense de sortir des jeunes autours ou des œufs du nid, 1295 CoutFezB; ca. 1400 ForGén).]

13<sup>e</sup> s. – ca. 1500:

13<sup>e</sup> s. AnglGuyB Gloss

1279 LAgC p. 10 [*En Guiscard de Berencs ..., per lors parcerers devo I austor d'acapte o 30 s. de morl. per lo tenh e per l'afar...*]

1286 ArchGir 1, 362 [47] [*unum accipitrem seu austorio domino*]  
 1295 CoutFezB 1295 ArchGir 27,390[36] [*que degun no sia ausart de prene austors ne espabers de legue ne deu loc on faran lor niser e lors hoeus.*]  
 ca. 1400 ForGén p. 34 [*que ... no pani œus d'austor ni d'esperver*]  
 15<sup>e</sup> s. CoutDax ArchGir 37,133[615] [*sera degudement per pagar l'omenadge cum es spade o lance o austor o autre cause; e asso segont la costume*]  
 ca. 1500 CoutMontsaunèsM p. 44 [*D'ayso empero so exceptatz esparviers et austors et falcos ... toutes betz, esparbes, astous et faucous*]  
 LR p. 10

◆ *\*austor sauro* “jeune autour avant la première mue” / A IV b 3: Les oiseaux de proie

1204 – 1274:

1204 ArchGir 16, p. 270 [*reddendo inde annuatim ... unum austorem saurum, pro omni servitio*]  
 13<sup>e</sup> s. AnglGuyB Gloss 320;341 ArchGir 3 28[109] [*austuro spocle [= sauro?]*]; Gloss 320;341 ArchGir 3,25[97] [*asturio sauro*]; Gloss 320;341 ArchGir 31[118] [*austurre sauro*]  
 1274 ArchGir 5, p. 278 [... *cum uno austuro sauro censuali solvendum die Beate-Marie ...*]; 5, p. 323 [... *cum uno osterio sauro vel mutato sicarium ipsius domini, ...*]

● agasc. lat. **\*austoron** m.

[REM.: Le dérivé en lt. -ONE n'est pas documenté dans la lexicographie galloromane. Le contexte des CoutFezB semble toutefois indiquer qu'il ne s'agit pas – à l'instar du type aocc. *auzelon*, gasc. *auzerou* – d'un “jeune autour avant la première mue” (puisque'il s'agit d'un oiseau capable de faire un nid et de pondre des œufs), mais plutôt d'une espèce de rapace bien déterminée, ressemblant à l'autour et à l'épervier. L'aigle de Bonelli que nous proposons comme identification possible vit exclusivement sur la Péninsule ibérique, d'où les oiseaux de chasse ont pu être amenés en Gascogne. Le suffixe indiquerait donc une approximation.]

◆ “oiseau de chasse de la famille des accipitridae, év. aigle de Bonelli (*Hieraaetus fasciatus*) qui est répandu sur toute la Péninsule ibérique et qui ressemble par son plumage et son cri à l'autour, légèrement plus petit” / A IV b 3: Les oiseaux de proie

1276; 1295:

1276 CoutFezB p. 14, Monl 6 [*nullus sit ausus capere austur[i]ones [ed. austurcones] seu ancipitres de legua(m) ubi nidum consueverunt facere aut ova eorum*];  
 1295 CoutFez 1295 Monl Monl 6,44<sup>(3)</sup> [*quod nullus sit ausus capere austurones seu ancipitres seu sparberios*]

<sup>(1)</sup> La forme agasc. latinisée glose dans ce passage la forme lat. méd. *accipiter*.

<sup>(2)</sup> Le passage introduit une forme plus latinisante (*austors*) à côté d'une forme plus orale (*astous*).

<sup>(3)</sup> L'énumération des trois rapaces laisse supposer trois espèces différentes qui sont nommées dans l'ordre de leur taille (l'épervier est plus petit que l'autour, celui-ci plus petit que l'«austoron»).