

Handreichung für die Erstellung der Transliterationen des Basiscorpus

(1. Ausgabe: HFR, Februar 2017; letzte Änderung: 01.12.2021)

Inhaltsverzeichnis

A. Allgemeines.....	2
Überschrift des Textes	2
Statusdatei	3
Veränderungen von CTH-Nummern	4
Umgang mit indirekten Joins.....	4
Kommentare und Fußnoten.....	5
B. Formatierung.....	6
Formatierungsvorlagen.....	6
Elemente des Tafellayouts.....	9
Zeilenangaben.....	9
Textbrüche	11
Zahlen.....	12
Frage- & Ausrufezeichen; Korrekturen; Matres Lectionis; Subskribierte Zeichen	12
Klammern und unlesbare Zeichen.....	13
Sonderzeichen.....	15
Weiteres.....	15
C. Konventionen zur Transliteration.....	17
Grundsätzliches.....	17
Transliteration von Gleitlauten.....	17
E- und I-Transliterationen	17
Transliteration von stimmlosen/stimmhaften Zeichenvarianten.....	18
Einzelne Zeichen.....	18
Nomina.....	18
Verben.....	18
Determinative und Datumsangaben.....	19
Abkürzungen.....	19
Ligaturen.....	19
Sumerogramme	19
D. XML Dateiumwandlung und -verwaltung.....	20
Speicherung .xml-Dateien in /HFR BASISCORPUS XML.....	20
Dokumentation der .xml-Umwandlung und Annotation.....	20
Weiterverarbeitung der .xml-Dateien (noch zu klären).....	21
Anhang.....	22
Sumerogramme.....	22
Akkadogramme.....	28
Hethitische Lexeme.....	31

A. Allgemeines

Überschrift des Textes

- Jeder Text hat die Publikationsnummer im Format „AO:TxtPubl“ als Überschrift. Handelt es sich um eine Inventarnummer und keine Publikationsnummer, wird diese als „AO:InvNr“ formatiert. Die erste Zeile wird mit der Absatzvorlage „AO:Manuscripts“ („AO:TxtPubl“ und „AO:InvNr“ sind Zeichenvorlagen) versehen.

KBo 77.777

Bo 77777

- Wenn es sich um einen gejointen Text handelt, werden die Publikationsnummern der einzelnen Fragmente mit Eurozeichen + Ziffer in geschweiften Klammern identifizierbar gemacht, z.B.: KUB 35.6{€1} + KBo 8.3{€2}.
- Bei gejointen Texten wird als erstes das Leitfragment genannt, das nicht notwendigerweise Text {€1} sein muss. Das Leitfragment ist die Publikationsnummer (oder Inventarnummer) des ersten in der Konkordanz genannten Fragments der Tafel.
- Die Angabe der gejointen Fragmente nach dem Leitfragment sollte zuerst alle direkt joinenden Fragmente geschlossen auflisten, danach weitere Gruppen von Fragmenten, die untereinander direkt joinen, aber nur indirekt mit anderen Gruppen derselben Tafel, und zuletzt die einzelnen Fragmente, die an kein anderes Fragment der Tafel direkt anschließen. Z. B. 1 + 2 + 3 (+) 4 + 5 + 6 (+) 7: dabei stellen die Fragmente 1–3 eine Reihe von direkten Anschlüssen dar, ebenso wie auch 4–6, wobei zwischen den beiden Gruppen 1–3 und 4–6 keine direkte Anschlüsse bestehen, und 7 an kein anderes Fragment der Tafel direkt anschließt.
- Wenn durch einen neuen Join das neuangeschlossene Fragment in die Position des Leitfragments gelangt, muss diese an erster Stelle unter den Textangaben erwähnt werden, und der Dokumentenamen und die Angaben in der Statusdatei angepasst werden. Die bestehenden Fragmentnummern (mit € Zeichen) bleiben aber unverändert.
- Die CTH-Nummer wird nicht im Basiscorpus-Dokument angegeben.
- Der Text beginnt in der auf die Zeile mit Publikationsnummer folgenden Zeile. Das Dokument sollte grundsätzlich keine Leerezeilen enthalten.
- Das Dokument erhält ebenfalls die Publikations-/Inventarnummer des Leitfragments; bei gejointen Texten werden diese mit + versehen: z.B. KBo 17.35+.odt (ohne Leerzeichen vor + und ohne Klammern für die Angabe indirekter Joins). Die Angabe der indirekten Joins durch () erfolgt in der Dokumentenüberschrift.
- Bei den Autographien in der Reihe DBH soll man die Bandnummer von der Teilnummer mit „/“ in der Angabe der Publikation(en) in der Überschrift abgrenzen, z. B. DBH 46/2.19.
- In den Dateinamen wird „/“ NICHT verwendet! Stattdessen wird „_“ benutzt: 123_z; Kp 07_78; KuSa 1_1.1; DBH 46_2.15.

Statusdatei

- Konventionen und Kürzel:
 - ANN automatisch und anschließend manuell annotiert
 - F Formatierung
 - J Join
 - KF Kollation am Foto
 - KOR Korrektur
 - KOR2 zweite systematische Korrektur des ganzen Fragments/der ganzen Tafel
 - KT Kollation an der Tafel
 - MB Marburg
 - MZ Mainz
 - OF Korrektur ohne Einsicht des Fotos (d.h. /KOR OF_... statt /KOR KF_...)
 - TI (Wimis eigene Erstellung einer) Transliteration
 - Ü Übernahme (einer Transliteration) danach Quellangabe (abgekürzt)
 - Mitarbeiternamen werden durch Erstbuchstaben abgekürzt: CS, MP, SG, JL, TS, SM, AP usw.
 - # Zusätzliche Informationen (möglichst wenig benutzen; Beobachtungen lieber in einer Datei „0-Notizen_CTH ...“ im CTH Ordner im Basiscorpus festhalten)
 - /# Markierung eines formlosen Kommentars, das eventuell später informationstechnisch verwertet werden könnte
 - Leerzeichen z.B. nach @ oder & dient nur der Übersichtlichkeit, (z.B. Mac Vorschaufunktion)
 - Leerzeile nach Zeile mit & Publikations-/Inventarnr. wenn keine Transliteration vorhanden

- **Jedes** Eingreifen muss in der Statusdatei in folgender Art und Weise eingetragen werden: /
 [Arbeitsvorgang]_[Erstbuchstaben] [TT.MM.JJJJ] / [Arbeitsvorgang]_[Erstbuchstaben] [TT.MM.JJJJ]
 Zusammen von einer Person ausgeführten Arbeitsvorgänge können auch zusammen (nur mit
 Leerzeichen getrennt) angegeben werden, z.B. /KOR KF_CS 01.01.2016.
- Die Anordnung der Texte innerhalb der Statusdatei erfolgt alphabetisch nach den an erster Stelle zitierten Leitfragmenten.
- Bei **gejointen Fragmenten** sind **alle** einzelnen Joins in der Statusdatei mit den entsprechenden Sigla anzugeben. Diese Angaben müssen mit denen des Basiscorpus-Dokuments übereinstimmen. Dabei wird das „Leitfragment“ als erstes gelistet; falls die ursprüngliche Übernahme der einzelnen Joins aus verschiedenen Quellen erfolgte, müssen die Angaben bezüglich der Übernahme separat angegeben werden. Bsp.:

& KBo 24.116{€1} + KBo 34.151{€2} + KBo 17.88{€3} + KBo 20.67{€4} + KBo 40.60{€5}

/Ü MZ {€2}_CS 13.4.2016 /F_MP 15.04.2016 /Ü DBH {€5}_CS 13.4.2016 /F_MP 15.04.2016 /Ü JL
 {€4}_CS 15.04.2016 /F_MP 15.04.2016/KOR Joins tw.- zusammen_CS 27.04.2016 #{€1} keine
 Transl. #{€3} keine Transl. #{€4} tw. Transl. #Ergänze Text aus KBo 20.67JL.odt

- Im Falle eines mehrfach veröffentlichten Fragment, soll bevorzugt die ursprüngliche Publikationsnummer zitiert werden. Eine jüngere Publikation soll dann zitiert werden, wenn sie das Fragment vollständiger bzw. genauer dokumentiert.
- Joinangaben
 - einer oder mehrere neue Joins:
& KBo 71.1{€1} + KBo 71.2{€2} + KBo 71.3{€3}
/J +KBo 71.2+KBo 71.3_CS 07.05.2018
 - bei indirekten Joins
& KBo 71.1{€1} + KBo 71.2{€2} (+) KBo 71.3{€3}
/J +KBo 71.2(+KBo 71.3_CS 07.05.2018

Veränderungen von CTH-Nummern

Werden Texte einer anderen CTH-Nr. zugeordnet, sind folgende Arbeitsschritte zu vermerken:

- In der Statusdatei der alten Zuordnung in der ersten Zeile direkt nach der Textnummer: /CTH „alt“ ZU CTH „neu“_JL 4.10.2016 (z.B. /CTH 670 ZU CTH 627_JL 4.10.2016; CTH 644 ZU CTH 832_JL 4.10.2016).
- Dasselbe in der Statusdatei der neuen Zuordnung in der ersten Zeile direkt nach der Textnummer: /CTH „alt“ ZU CTH „neu“_JL 4.10.2016.
- Protokolliert im Dokument / „Organisation“ / „Veränderung_CTH_Nr.ods“; dort Angabe der Inventar-Nr., der Publikationsnr., der ehemaligen CTH.-Nr., der neuen CTH-Nr.
- Darüber hinaus soll in dem Dokument „Veränderung_CTH_Nr.ods“ durch ein „x“ vermerkt werden, ob es sich um ein Festrival oder eine andere, für uns nicht relevante CTH-Nr. handelt.
- Kann das Festrival nicht näher bestimmt werden, wird es in den Ordner „HFR Basiscorpus“ / „CTH 670“ geschoben; wird klar, dass es sich bei einem Text nicht um ein Festrival handelt, wird er in den Ordner /„HFR Basiscorpus“ / „Andere CTH-Nr.“ / „CTH „neu““ geschoben. In allen anderen Fällen wird der Text samt Vermerken in den Statusdateien direkt in /„HFR Basiscorpus“ / „xAndere CTH-Nummern“ / „CTH ...“ verschoben.

Umgang mit indirekten Joins

- Ist in der Konkordanz ein indirekter Join verzeichnet und ist das Fragment in Relation zum bekannten Teil der Tafel hinreichend lokalisierbar, werden die Fragmente gemäß der Konkordanz in einem ODT-Dokument behandelt und ggf. eine neue Joinskizze von S. Košak oder einem anderen Mitarbeiter erstellt.
- Ist der indirekte Join – auch bei vorhandener Joinskizze – unwahrscheinlich oder ausgeschlossen, werden die Fragmente getrennt behandelt und Silvin Košak um eine entsprechende Modifizierung der Joinskizze gebeten. Dabei wird in der Statusdatei direkt unter der Angabe des Leitfragments folgendes vermerkt:
 - /Join mit XXX aufgelöst_CS 3.6.2017 (bei Fragmenten, die nicht zusammen gehören können)

- /Join mit XXX aufgehoben_CS 3.6.2017 (bei Fragmenten, deren Zusammenschluss oder deren Anordnung unsicher ist)
- In den beiden letzten Fällen ist dies in die Liste / „Organisation“ / “Aufgelöste Joins” mit Angabe eines aufgelösten bzw. aufgehobenen Joins einzutragen.

Kommentare und Fußnoten

- Fussnoten sind im Basiscorpus möglichst kurz und präzise zu halten, auf korrekte sprachliche und grammatikalische Formulierungen ist zu achten. Sie beginnen mit Großbuchstaben und enden mit Punkt. Eine zuverlässige sprachliche Korrektur der Fußnoten muss ggf. durchgeführt werden.
- Im Basiscorpus selbst sollen keine zusätzlichen Vermerke stehen. Alle mit Lattenkreuzen o.ä. markierten Vermerke müssen bearbeitet und ggf. an anderer Stelle dokumentiert werden, z.B. in den Dokumenten für die Museumsarbeit für noch zu tätige Kollationen oder im CTH-eigenen Notizen-Dokument (0_Notizen-CTH ...).
- Auch in Fußnoten sind Akkadogramme, Sumerogramme, Determinative und hethitischer Text durch die betreffenden Zeichenvorlagen zu formatieren.
- Die Literaturangaben erfolgen gemäß der Hethitischen Bibliographie im Format „**Melchert H.C. 2004b**: 122“; markiert mit Stil „**AO:-LIT**“. Mehrere Literaturangaben in einer Reihe werden durch Semikolon abgegrenzt. Neue Literatur, die nicht in der Hethitischen Bibliographie verzeichnet ist, ist vom Mitarbeiter in die Liste /Spaces/HFR Basiscorpus Dateien_CTH/Organisation /HethBib_neue-Literatur.odt vollständig nach Muster der Hethitischen Bibliographie einzutragen. Diese wird weitervermittelt und so in die Bibliographie aufgenommen.
- In den Textdateien sind Fußnoten keinesfalls an die Überschrift, also entweder Publikations- oder Inventarnummer, sondern an den Text selbst ans Ende eines Wortes – nur wenn nicht anders möglich an einer Zeilenangabe (vgl. dazu unten) – zu setzen.

B. Formatierung

Formatierungsvorlagen

Absatzvorlagen:

AO>Title

AO:Subtitle

AO:Manuscripts

AO:Textline-Hur

AO:Textline-Hat

AO:Textline-Hit

AO:Textline-Luw

AO:Textline-Pal

AO:Textline-Akk

AO:Textline-Sum

AO:Text_HURinHIT

AO:Text_HATinHIT

AO:Text_LUWinHIT

AO:Text_PALinHIT

AO:Text_AKKinHIT

Zeichenvorlagen:

AO:Hittite

AO:Sumerian

AO:Akkadian

AO:Hattian

AO:Hurrian

AO:Luwian

AO:Palaic

AO:Numeral

AO:SUMGRAM

AO:AKKGRAM

AO:TxtPubl

AO:InvNr

AO:MaterLect

AO:subscr

AO:corr

AO:Determ

AO:-LineNrExpl

AO:-LIT

AO:-KolonMark

AO:-MarkupDef

AO:CTH-Nr
 AO:-Text-Gap
 AO:SIGN

Die aktuellen Zeichen- und Absatzvorlagen finden sich automatisch in den mit KOR2 umwandelten Dokumenten; überflüssige und veraltete Stile werden dort auch ggf. gelöscht bzw. ersetzt.

Falls notwendig, kann wie folgt die Vorlagen in das zu bearbeitende Dokument geladen werden:

- Ein (wie oben beschrieben) mit KOR2 umwandeltes Dokument öffnen und unter „Datei“ - „Vorlage“ - „Als Vorlage speichern ...“,
- einen Namen wählen und unter „Meine Vorlagen“ speichern,
- in der Menüleiste unter „Vorlagen“ „Vorlagen laden...“ anklicken,
- in dem erschienenen Feld „Überschreiben“ anklicken und die Handreichungsvorlage im rechten Kasten auswählen. „Ok“ klicken.

Die aktuellen Vorlagen sind in den mit der aktuellen Simtex-Version bzw. die mit KOR2 erstellten fodt-Dateien vorhanden. Bei früher umgewandelten Dateien kann die aktuelle Vorlage übernommen werden, wie oben beschrieben.

Folgendes ist bei der Markierung mit den Absatz- und Zeichenvorlagen zu beachten:

- Die Absatzvorlage ist von der Grundsprache des Textes abhängig und entsprechend zu setzen (z.B. AO:Textline-Hit für hethitische Texte).
- Die Zeichenvorlage ist „Standard“ (Default Style) für die Lexeme der Sprache, die mit der Sprache der Absatzvorlage korrespondiert (z.B. hethitisch bei der Absatzvorlage „AO:Textline-Hit“, nicht aber dann zusätzlich mit der Zeichenvorlage „AO:Hittite“) verwendet; anderenfalls ist sie manuell zu markieren (z.B. „AO:SumGRAM“ für Sumerogramme innerhalb hethitischer Texte usw.).
- Luwische, hurritische, hattische Wörter in hethitischem Kontext, die in ihrer jeweiligen Herkunftssprache flektiert werden oder die Stammform zeigen, markieren wir im Zuge der manuellen Überprüfung der Annotation als Hurritisch, Luwisch, Hattisch (z. T. ist eine automatische Markierung durch KOR2 bereits vorgesehen).
- Treten in einer Zeile verschiedensprachige Textabschnitte auf, gibt es zwei Möglichkeiten den fremdsprachigen Textabschnitt zu markieren:
 - Ist der nicht-hethitische Abschnitt syntaktisch unabhängig vom vorausgehenden hethitischen Satz (z.B. nach *verbum dicendi* wie etwa *memai* oder *halzai*), dann soll nach dem letzten hethitischen Wort die Zeile umgebrochen, und die nächste Zeile mit entsprechender Absatzvorlage der „AO:Textline-...“ Serie markiert werden: dabei soll die Zeilennummer nicht erneut angegeben werden (im Gegensatz zur früheren Konvention). Wenn auf dem fremdsprachigen Abschnitt ein neuer hethitischer Satz in der selben Zeile ansetzt, muss ebenso die Zeile umgebrochen werden, ohne Wiederholung der Zeilennummerierung. Beispiel:

Vs. I⁷ GAL-ŠU-NU ḫal-za-a-i
 an-te-eg-ga-ḫ[u-li
 [a-pé-e-ma-aš-ši kat-ta-an ar-ku-wa-an-zi]
 Vs. I⁸ te-eg-ga-ḫu-le-e te-ka-a [te-ka-ḫu-li 𐎠𐎢𐎣
 Vs. I⁹ GAL-ŠU-NU ḫal-za-a-i
 a-ak-ki-nu-n[a

- Wird der nicht-hethitische Abschnitt syntaktisch in einem hethitischen Satz eingeschlossen, z. B., wenn das Verb nach dem Abschnitt steht, dann muss der Abschnitt (ggf. mit Umbruch der Zeile, ohne Wiederholung der Zeilennummerierung) mit entsprechender Absatzvorlage aus der Serie „AO:Tex_...inHIT“ markiert werden. Beispiel:

Rs. IV 4' EGIR-ŠU-ma D IŠTAR URU ḫa-a[t-ta-ri-na
 Rs. IV 5' D ni-na-at-ta D ku-li-i[t-ta
 Rs. IV 6' DINGIR^{MES} -na ku-la-ah-e-na [
 Rs. IV 7' DINGIR^{MES} -na ku-lu-pa-te-na
 TUŠ-aš 'e'¹ -[ku-zi

- Uninterpretierte Zeichen werden mit der Zeichenvorlage „AO:SIGN“ markiert.
- Akkadische phonetische Komplemente von Sumerogrammen werden mit Bindestrich angeschlossen und als Akkadogramme formatiert: DINGIR-LIM. Die zu verwendende Schriftart ist HPMLinux Libertine O.
- Die Zeichenvorlage folgender Elemente ist aus informationstechnischer Perspektive irrelevant, wenn sie vor oder nach einem transliterierten Zeichen stehen:
 - Klammer (eckige, halb-eckige, und spitze): z. B. 'É.GAL', 'É.GAL', [É.GAL sind alle zulässige Formatierungen.
 - Fußnotenzeichen (nur am Wortende! Beachte: *innerhalb* der Fußnoten sollen Sumerogramme usw. entsprechend markiert werden)
 - Sternzeichen (bei Rasuren), z. B. zulässig sind *É.GAL*, *É.GAL* und *É.GAL*
- Innerhalb eines Ausdrucks richten sich die folgenden Elemente nach der Vorlage des Ausdrucks, e.g. in, zwischen, oder unmittelbar vor/nach einem Sumerogram oder Akkadogram. Das Prinzip ist, dass zusammenhängende Formatierung nicht unterbrochen werden darf:
 - Klammer in einem transliterierten Zeichen: É.[GA]L, aber NICHT: É.[GA]L oder É.*GAL*
 - Punkte zwischen Elementen eines Sumerograms stets mit AO:SumGRAM, auch wenn Determinative dazwischen stehen: DUMU^{MES}.LUGAL; GIŠ^D.INANNA DUMU^{MES}.LUGAL; NICHT GIŠ^D.INANNA; DUMU^{MES}.LUGAL; oder DUMU^{MES}.LUGAL.
 - Punkte zwischen Elementen eines Determinativs stets mit AO:Determ: LÚ^{MES}.ALAM.ZU,
- Einige Logogramme werden mit dem Zeichen ₁ (modifier letter shelf; U+02FD) verbunden (siehe die vollständige Liste der Logogramme in ###). In diesen Fällen wurde eine Verbindung mit Punkt oder Bindestrich aus Sicht der Transliterationskonventionen ausgeschlossen; mit ₁ wird jedoch gesichert, dass die betreffenden Logogramme als ein Wort informationstechnisch identifiziert wird. Beispiele dafür sind LÚ₁^{GIS}BANŠUR (auch LÚ^{MES}₁^{GIS}BANŠUR) sowie mehrfache heterographische Verbindungen

wie TU_7 , $ME-E$, UZU , GU_4 . Dabei soll das Zeichen `_` nicht mit Zeichenvorlage markiert sein (d.h. als Default bleiben).

Elemente des Tafellayouts

- Für die Angabe von **Paragraphen** lautet die Grundregel: Paragraphenstriche werden immer berücksichtigt, Randleisten nie; letztere dürfen aber auch nicht mit den Paragraphenstrichen verwechselt werden. D.h. im Einzelnen:
 - Generell zur Markierung von Paragraphenstrichen: `┌┌┌`
 - Paragraphenstrich vor der ersten Zeile:
 - (Vs. II) `┌┌┌`
 - (Vs. II) 1' *ma-a-an*
 - Paragraphenstrich + Ende der Kolumne mit Randleiste: `┌┌┌` (nicht `===`!)
 - Paragraphenstrich + Ende der Kolumne ohne Randleiste: `┌┌┌`
 - Doppelter Paragraphenstrich: `===`
 - Doppelter Paragraphenstrich + Randleiste: `===` (nicht `=== ┌┌┌`!)
- Das **Ende der Kolumne** wird durch eine entsprechende Bemerkung in einer separaten Zeile (Markierung: „AO:-Text-Gap“) angegeben (z.B.: **Ende Kol. IV**).
- Ist die Kolumne am Ende oder am Anfang abgebrochen, so dass sich nicht erkennen lässt, ob ein Paragraphenstrich vorhanden war, wird **nicht** automatisch ein Paragraphenstrich angegeben. Wenn er aber aus dem Kontext oder Duplikaten erschließbar ist, wird er in eckigen Klammern ergänzt.
- Ein **Wortabstand** soll in der Umschrift **nicht** imitiert werden durch Leerzeichen zwischen den Wörtern; keinesfalls dürfen Spatien **innerhalb** eines Wortes erscheinen (also nicht *i-ia-an- zi*; diese würde zu einer Analyse von *i-ia-an-* und *zi* als zwei getrennte Worte führen).
- Visuelle Wiedergabe der Texterhaltung: wenn, z. B. nur rechte Zeilenenden vorhanden sind, dann sollte der Text durch Leerzeichen (**nicht** mit Tabs) nach rechts gerückt werden.

Zeilenangaben

Die Zeilenzählung jeder Zeile wird mit der Zeichenvorlage: „^{AO:-LineNrExpl}“ markiert und folgt im Allgemeinen dem folgenden Muster (beachte die Leerzeichen!):

- Vs. II 1'
- r. Kol. 1'
- lk. Kol. 1'
- mit. Kol. (für mittlere Kolumne)
- o. Rd. (oberer Rand – zwischen Rs. und Vs. in Leserichtung)
- u. Rd. (unterer Rand – zwischen Vs. und Rs. in Leserichtung)
- lk. Rd. (linker Rand)
- r. Rd. (rechter Rand)

Für Joins gelten folgende Regeln:

- Wenn die Autographien zweier joinender Fragmente dieselbe Angabe über Kolumne **und** Zeile zeigen, wird die Zeilenangabe nur einmal angegeben:
 - {€1+2} Vs. II 3'
- Anderenfalls wird die Zeilenangabe jeweils separat angegeben, selbst wenn die Kolumnenangabe übereinstimmt:
 - {€1+2} Vs. II 1'/Vs. II 3', {€1+2} Vs. II 1'/Vs. r. Kol. 2', usw.
 - bei gejointen Fragmenten entspricht die Reihenfolge der Fragmente ihrer jeweiligen Position in der Tafel von Links nach Rechts; bei horizontal verlaufenden Brüchen, von Oben nach Unten.
- Bisläng unpublizierte Fragmente erhalten keine eigene Zeilenzählung, wenn sie sicher direkt joinen, außer wenn das unpublizierte Fragment dem publizierten vorausgeht (in letzterem müsste die Zeilenzählung nach oben korrigiert werden, was im Basiscorpus nicht gemacht wird).
- Die Zugehörigkeit einer Zeile zu einer bestimmten Kolumne infolge eines Joins erschließt sich durch die Angaben wie „Ende Vs. II“, „Rs. III bricht ab“ (für Musterbeispiele siehe unten) u. ä. am Ende der Kolumne. Die Benennung der Zeile selbst erfolgt nach der Autographie und nicht nach Joins oder Neuordnungen innerhalb einer Tafel. Erklärende Fußnoten sind durch die konsequenten Kolumnenangaben überflüssig und werden nicht gesetzt.
- Sollten dennoch Fußnoten bei den Zeilenangaben unumgänglich sein, werden diese ans Ende der Angabe gesetzt:
 - {€1+2} Vs. r. Kol. 5'/Vs. 2'¹

Für einkolumnige Tafeln folgen wir dem Muster:

- Normalerweise (mit Leerzeichen vor dem Beginn der Textzeile!): **Vs. 8**
- Wenn Vorder- und Rückseite nicht zu bestimmen sind, nur: **8**
- Wenn aber 2 Seiten erhalten und Vs./Rs. nicht angegeben sind, werden entweder die genauen Angaben der Autographie (z.B. Seite A – Seite B) verwendet, oder, wenn die Autographie gar keine Angaben macht, werden die Angaben Seite A – Seite B nach eigenem Ermessen vergeben.
- Bei fraglicher Zuordnung der Vs. bzw. Rs. (kein Leerzeichen vor dem Fragezeichen), Angabe entspricht der Autographie genau: **Vs.? 1' ; Vs.(?) 1' ; Vs. (III) 1'**

Verhältnis zu Angaben in der Autographie:

- In der Zeilenzählung halten wir uns bezüglich der Zeilenzahl an die Autographie, außer wenn diese falsch ist (übersehene Zeile usw.). In diesem Fall setzen wir die neue (korrekte) Zählung ein. Alle Zeilen, die nicht mit der Autographie übereinstimmen, bekommen ein Ausrufezeichen (z.B. **Vs. II 1'!**).
- Bezüglich der Zeilenzählung bei abgebrochenem oberem Rand der Tafel (d.h. die erste(n) Zeile(n) ist/sind nicht vorhanden) wird die Zeilenzählung mit Prime Zeichen ' (nach Unterbrechung dann doppelt, drei- und vierfach: " " " ") begonnen. Statt x+1 beginnt die Zählung mit 1'.

1 Fußnotentext.

- Zusätzliche Leerzeichen (Tabs **nicht** verwenden!) zur Markierung von abgebrochenem Text vor oder nach der eckigen Klammer müssen mit der Standardformatierung markiert sein.
- Die häufigsten Text-Gap-Markierungen lauten:
 - Vs. bricht ab
 - Ende Vs.
 - unbeschriebene Zeile
 - ... (Nummer) unbeschriebene Zeilen
 - ... (Nummer) abgebrochene Zeilen
 - Lücke unbekannter Größe (z. B. bei abgebrochenem Tafelrand)
 - Lücke unbestimmbarer Größe (z. B. bei erhaltenem Tafelrand)

Zahlen

- Zahlen im hethitischen Text werden in arabischen Ziffern mit der Zeichenvorlage „AO:Numeral“ wiedergegeben.
- Ein Numerales, das sich nicht genau bestimmen lässt, wird transliteriert als:
 - **n** (und ebenfalls als „AO:Numeral“ formatiert);
- Bestandteile unvollständig erhaltener Numeralia werden nach folgendem Muster transliteriert:
 - **[n+]1, [n+]2** usw. (und als „AO:Numeral“ formatiert);
- Zahlen bei Datumsangaben: s. unten „Determinative und Datumsangaben“;
- Bruchzahlen werden mit ihrem Unicode-Zeichen angegeben. Neben den „einfachen“ $\frac{1}{2}$ $\frac{1}{3}$ etc. können Bruchzahlen auch mit Hilfe von „FRACTION NUMERATOR ONE (U+215F)“ $\frac{1}{2}$ oder „FRACTION SLASH (U+2044)“ / zusammengesetzt werden;
- Mit Bruchzahlen zusammengesetzte Zahlen werden wie folgt gehandhabt:
 - **1 $\frac{1}{2}$** (nicht **1.5** o.ä. und beides als „AO:Numeral“ markiert mit unmarkiertem Leerzeichen dazwischen).

Frage- & Ausrufezeichen; Korrekturen; Matres Lectionis; Subskribierte Zeichen

- Die Frage- und Ausrufezeichen, die die Lesung kommentieren (!, ?, sic), erhalten die Zeichenvorlage „AO:corr“. Dabei gehören Fragezeichen in die halben Klammern (z.B. **「pu[?]」**, nicht **「pu[?]」**).
- Symbole:
 - **?**: Fraglichkeit eines Zeichens (hinter Zeichen ohne Leerzeichen, Stil „AO:corr“)
 - **(?)**: Fraglichkeit eines Strings (hinter String ohne Leerzeichen, Stil „AO:corr“)
 - **!**: Zur Angabe von abweichenden Zeichenformen sowie von Schreiberfehlern (Ausrufezeichen, Formatierung „AO:corr“) mit Angabe des tatsächlichen Zeichens in einer Fußnote. Das Fußno-

tenzeichen wird immer an das Wortende gesetzt, das Ausrufezeichen dagegen direkt hinter das betroffene Zeichen, z.B. ^{DUG}*kal-ti*².

- sic : Geprüfte Lesung gegen Autographie (hochgestellt, Stil „AO:corr“)
- Bei Zeilenangaben wird das Ausrufezeichen nur bei Auslassungen in der Autographie, die eine im Basiscorpus abweichende Zählung hervorrufen, gesetzt. Zeilenangaben erfolgen immer nach der Autographie (mit deren Frage- und Ausrufezeichen); von uns gesetzte „?“ in Bezug auf eine Spalten- oder Seitenzuordnung erfolgen nur in den Text-Gap-Angaben am Ende einer Spalte.
- Wenn die Reihenfolge zweier Zeichen vertauscht ist, werden sie in der korrekten Reihenfolge mit jeweils einem Ausrufezeichen angegeben und mit der Angabe in der Fussnote „Zeichen ... und ... vertauscht.“ versehen.
- Außerdem kennzeichnet das Ausrufezeichen ungewöhnliche Zeichenformen, die nicht in HZL geführt sind (z.B. **Ṣ̌.A**! in KUB 2.5 Vs. II 27’); dafür ist eine Fussnote nicht zwingend erforderlich:



- Bemerkungen wie **(Rasur)**, **Lücke von ca. 3 Zeilen** o.ä. werden, sofern sie in der Zeile und nicht in einer Fußnote stehen, mit „AO:-Text-Gap“ formatiert. Es gibt keine satzinitiale Großschreibung für Bemerkungen in „AO:-Text-Gap“.
- Formatierung von *kán^{an}*, *túh^u* : AN bzw. Uḫ als „AO:MaterLect“. Dieser Stil findet für alle Zusatzzeichen Verwendung, die Zeichenlesungen erklärende, also auch für AR bzw. RU in luw. *pár^{ar}-na* oder sum. *É.MÁ.URU₅^{ru}*. Diese Zusatzzeichen werden sowohl in syllabischem (*túh^{uh}*) als auch in logographischem Kontext (*É.MÁ.URU₅^{ru}*) kleingeschrieben.
- Sollte das als „AO:MaterLect“ formatierte Zeichen direkt auf ein Determinativ stoßen, ist ein Punkt dazwischen zu stellen, formatiert als „AO:Determ“, z.B.: *NAM.TÚL^{ul}.ḫ.A*
- Die Formatierung von subskribierten Zeichen wie in *wa_a* usw. erfolgt durch Zuweisung der Zeichenvorlage „AO:subscr“. Dagegen werden für Indexziffern wie in *SA₅* etc. die betreffenden Unicode-Zeichen verwendet (₀₋₉; U+2080 – U+2089). Ihre Zeichenformatierung geschieht analog zur Zeichenvorlage der Silbe. Gleiches gilt für x als Index, das nicht als „AO:subscr“ formatiert wird, sondern als Unicode-Zeichen U+2093: *tan_x*.

Klammern und unlesbare Zeichen

- Ein Zeichen gilt als nicht zerstört, wenn alle Bestandteile eindeutig vorhanden sind (vorzugsweise der Kopf eines Keils).
- **x** für nicht mehr lesbare, aber vorhandene Zeichen.
- Folgende Schreibungen werden für nicht vollständige bzw. nicht lesbare Zeichen verwendet:
 - Beginn des Zeichens im Bruch: **]x**
 - Ende des Zeichens im Bruch: **x[**
 - Beginn und Ende des Zeichens im Bruch: **]x[**

2 Text: UN.

- Zeichenraum vollständig vorhanden, Zeichen aber nicht lesbar: **x** oder **x** [oder] **x** [
- Zeichen nicht lesbar, Zusammenhang zu einem Wort nicht erkennbar:] **x x x** [
- Zeichen nicht lesbar, gehören aber zweifelsfrei zu einem Wort: **x-x-x**
- Wenn ein Wort direkt nach der schließenden Bruchklammer beginnt, muss zwischen Klammer und Wort ein Leerzeichen stehen; gleiches gilt wenn die öffnende Klammer direkt nach einem vollständig geschriebenen Wort steht:

Vs. I 3' [] **pé-ra-an** (nicht: [] **pé-ra-an**)

Vs. I 3' **pé-ra-an** [] (nicht: **pé-ra-an**[])

- Wenn ein Wortbeginn vor oder Wortende nach der Klammer unsicher ist, dann ist dies mit „(-)“ ohne Leerzeichen nach bzw. vor der Klammer zu verdeutlichen:

Vs. I 3' [] **(-)da-an**

Vs. I 3' **an-da(-)**[]

- Die Größe von Textlücken wird durch Klammern angegeben. Innerhalb der Textlücke selbst werden weder Unterstriche noch **x** gesetzt, sondern ihr Umfang allein durch Leerzeichen (**nicht** durch Tabs!) festgelegt. Zusätzliche Leerzeichen (Tabs **nicht** verwenden!) zur Markierung von abgebrochenem Text vor oder nach der eckigen Klammer müssen mit der Standardformatierung markiert sein. Ist die Breite der Lücke schätzbar, werden die Klammern geschlossen; anderenfalls bleiben sie offen:

Vs. I 3' [] **da-a-i**

Vs. I 3' **pé-ra-an** []

Vs. I 3' **da]-a-i**

Vs. I 3' **pé-ra-an** []

- Abgebrochene Zeilenenden, in denen nichts gestanden haben dürfte, können mit [] (ein einziges Leerzeichen zwischen den Klammern) markiert werden.
- Bindestriche, Punkte etc. stehen außerhalb der Klammern (z.B. **a-pé-[ez-z]i-ia-an**; [MUNUS].LUGAL; **nu-<uš>-ma-aš**).
- ʾ : Für teilweise zerstörte Zeichen.
- < > : Für Zeichen, die vom Schreiber fehlerhaft ausgelassen wurden und von uns ergänzt werden.
- << >> : Für Zeichen, die vom Schreiber fehlerhaft in den Text gesetzt wurden und von uns getilgt werden. Dabei bitte die überflüssige Bindestriche und/oder Punkte löschen, z. B. ^{LÚ<<MEŠ>>}ALAM.ZU, (nicht ^{LÚ<<MEŠ>>}ALAM.ZU,) oder **e-ku<<zi>>-zi** (nicht **e-ku-<<zi>>-zi**).
- * * : Angabe von Zeichen die über ein Rasur geschrieben wurden: ***pé-ra***-an.
- *<<...>>*: Für Zeichen, die unter der Rasur noch gelesen werden können, die aber getilgt werden sollen.
- **(Rasur)**: vorhandene Rasur, sowohl zwischen Worten oder innerhalb eines Wortes (in letzterem Fall ohne Leerzeichen davor oder danach!), ohne dass die ausradierte Zeichen noch lesbar sind, oder, dass Zeichen über der Rasur geschrieben wurden (Zeichenvorlage „AO:-Text-Gap“).

- Ist innerhalb eines Wortes ein Zeichen radiert, ohne dass es gelesen werden kann, wird dieser Umstand wie folgt wiedergegeben:
 - *pa-iz(Rasur)-zi* (Angabe **(Rasur)** mit Zeichenvorlage „AO:-Text-Gap“ formatiert)
- Wenn das Vorhandensein eines Zeichens in ergänztem Text unsicher ist, darf das Zeichen nicht in runden Klammern stehen (etwa ^D*ta-u-ri(-it)*: so NICHT!); das Zeichen ist entweder ohne Klammern zu ergänzen, oder in einer Fußnote als Ergänzungsmöglichkeit anzugeben.
- Ein Schrägstrich „/“ darf nicht für alternative Lesungen/Deutungen eines Zeichens in der Transliteration verwendet werden; es muss für eine Lesung entschieden werden, mit Angabe der Alternativlesung in einer Fußnote.

Sonderzeichen

Alle verwendeten Sonderzeichen sind in den vorgesehenen Fonts vorhanden und über die jeweiligen Tastaturlayouts zugänglich. Die für das Basiscorpus am häufigsten verwendeten Zeichen mit ihrer jeweiligen Unicode Nummer sind wie folgt:

◀	U+12471	×	U+00D7	◌	U+2086
↖	U+12039	↳	U+02FD	◌	U+2087
↗	U+1227D	×	U+2093	◌	U+2088
⟨	U+2329	₁	U+2081	◌	U+2089
⟩	U+232A	₂	U+2082	◌	U+2080
¬	U+00AC	₃	U+2083	◌	U+2032
⸀	U+2E22	₄	U+2084	◌	U+2033
⸁	U+2E23	₅	U+2085	◌	U+2034

Weiteres

- Glossenkeile:
 - seit 27.07.2016 als ◀ (U+12471) für doppelten Glossenkeil bzw. ↖ (U+12039) für einfachen Glossenkeil, und ### für den dreifachen Glossenkeil.
 - Glossenkeilzeichen werden direkt vor dem betroffenen Wort (d.h. ohne Leerzeichen dazwischen) geschrieben und erhalten dieselbe Formatierung wie das darauf folgende Wort (insbesondere Markierung als Hurritisch, Luwisch, usw.).
 - Falls Glossenkeilzeichen in einer anderen Funktion erscheinen als der des eigentlichen Glossenkeils (z.B. als Worttrenner [besonders häufig in hurritischen Texten, obwohl nicht eingeschränkt darauf], Angabe eines neuen Paragraphen, Einrückung einer Zeile usw.), werden diese vom eigentlichen Glossenkeil dadurch unterschieden, dass davor und auch danach ein Leerzeichen steht.
- Bei zusammengesetzten Zeichen (z.B. **KA×U**) ist das Sonderzeichen „×“ (U+00D7; ascii 215), nicht der Kleinbuchstabe **x** zu verwenden.

C. Konventionen zur Transliteration

Grundsätzliches

- Keine Großschreibung von Orts-, Götter- und Personennamen in Transliterationen.
- Keine Verwendung von Tabs in den Dokumenten.
- Keine Leerzeichen innerhalb eines Worts.
- Keine Leerzeilen im Dokument.

Transliteration von Gleitlauten

- Anstelle der häufig verwendeten Transliterationen von *i* bzw. *u* mit Bogen darunter oder gar von *j* und *y* ist für HFR für alle Sprachen *ia* bzw. *wa* und *IA* bzw. *WA* obligatorisch.

E- und I-Transliterationen

- Als doppel- bzw. mehrdeutig gelten die Zeichen ⟨pé/pí⟩, ⟨ke/ki⟩, ⟨ge/gi⟩, ⟨de/di⟩, ⟨ḫe/ḫi⟩ (trotz ⟨ḫé⟩), ⟨re/ri⟩, ⟨le/li⟩, ⟨né/ni⟩ (trotz ⟨ne⟩), ⟨ze/zi⟩ (trotz ⟨zé⟩), ⟨ep/ip⟩, ⟨et/it⟩, ⟨ek/ik⟩, ⟨aḫ/eḫ/iḫ/uḫ⟩, ⟨er/ir⟩, ⟨em/im⟩, ⟨ez/iz⟩.
- *e* wird bei doppeldeutigen Schreibungen in der Transkription und *bound transcription* (BT) immer dann gesetzt, wenn mit hinreichender Sicherheit aufgrund von eindeutigen Graphien oder aus sprachwissenschaftlichen Gründen ein /e/ an der betreffenden Stelle im Hethitischen anzunehmen ist. Beim Zusammentreffen eines eindeutigen *e*-Zeichens und eines doppeldeutigen Zeichens setzt sich in der Regel *e* in Transkription und BT durch. Ausnahmen dazu sind *i*-Stämme sowie *-iya/e*-Verben (siehe dazu unten).
- Wenn der hethitische Schreiber *i*-haltige Zeichen benutzt, die auch *e*-haltige Alternativen haben (z.B. ⟨mi⟩ vs. ⟨me⟩, ⟨in⟩ vs. ⟨en⟩) werden diese als *i*-haltige Zeichen in die Transliteration übernommen. Entsprechendes gilt für die *e*-haltigen Zeichen. Aus ⟨en⟩ wird dann also nicht ⟨in₄⟩ und umgekehrt. Dies wird in der BT genau umgesetzt, z.B.:
 - **ḫa-an-ni-eš-šar** → **ḫa-an-né-eš-šar** → **ḫanneššar** (weil ⟨ni/né⟩ doppeldeutig ist), aber
 - **dam-me-iš-ḫa-an-ta-an** → **dammeišḫantan** (weil beide Zeichen eindeutig sind)
- Wenn eine *e*-haltige Lesung aufgrund von eindeutigen Graphien oder aus sprachwissenschaftlichen Gründen gefordert ist, werden die KVK-Zeichen, die im HZL keinen *e*-Wert haben, mit *e*-Vokalismus transliteriert. Es ist also die *i*-Lesung durch eine *e*-Lesung mit derselben Indexzahl zu ersetzen.

Transliteration von stimmlosen/stimmhaften Zeichenvarianten

- Die Transliteration von Zeichen mit stimmloser und stimmhafter Variante (z.B. ⟨ad/t⟩, ⟨tab/p⟩, ⟨bi/pí⟩) erfolgt für alle Sprachen (außer Akkadisch) nach hethitologischer Konvention, d.h. in Kontaktstellung zu einem folgenden KV-Zeichen richtet sich die Umschrift nach diesem (z.B. **ad-da-aš**, **at-ta-aš**), in den anderen Fällen erfolgt eine Umschrift mit der stimmlosen Variante (z.B. **tap-ri**, **la-ap-la-na**).

Einzelne Zeichen

- Wir lesen ⟨qa⟩ als *ka*.

Nomina

- Der Nom. Pl. der *i*-Stämme geht auf /-ies/ aus (z.B. **ka-ru-ú-i-li-eš** mit *li* und BT **karuilies**).
- Eigennamen (Personen, Orte, Götter, Topographie) werden in der Transliteration kleingeschrieben (z.B. = ^{HUR.SAG}**ta-ba-la** und BT ^{HUR.SAG}**Tabala**).

Verben

- „*sk*-Verben“ haben entweder *-e-* oder *-a-* im Stammauslaut vor der Endung (Ausnahme Imp. 2. Sg auf *-ški*). Dementsprechend sind die *e/i*-Zeichen – wo möglich – mit *e* zu lesen. Das gilt auch für die KVK-Zeichen, die im HZL keinen *e*-Wert haben, also z.B. *kit*₉, das als *ket*₉ zu lesen ist.
- iye/a*-Verben verlieren ihr *-ie-* in der Transliteration nicht, also z.B. **ḫu-u-la-li-ez-zi** (nicht **ḫu-u-la-le-ez-zi**). Das gilt ebenso *iya/iye-* „machen“ und *tiya/tiye-* „hintreten“.
- a/-ai-* Verben (= *-a/-ae-* Verben) der *mi*-Konjugation weisen verschiedentlich den Diphthong *-ae-* im Stammauslaut auf, wiedergegeben durch *-Ca-e-* oder *-Ca-eC-* (z.B. **ḫa-at-ra-(a)eš**, **i-ga-e-e/it-ta**, [**za-a**] **m-mu-ra-a-e-e/iz-zi**) neben ambiguum *-Ca-iC-*. Wir richten uns **nicht** nach der Default-Lesung der verwendeten Zeichen (*-iz-* oder *-it-*), sondern nach der vermuteten phonologischen Wirklichkeit, lesen also **-ez-** oder **-et-** (z.B. in **ḫa-an-da-ez-zi** oder **ir-ḫa-a-et**).
- Der Ausgang der 3.Sg. Prät. Ind. Akt. der *ḫi*-Verben auf *-ai/-i(ya)-* und *-a/-ai/-e-* lautet dagegen /-ais/ (z.B. **ḫalzaiš**).
- pai/pa-* „gehen“ behält grundsätzlich *i*, also **pa-iz-zi**.
- Für „kommen“ lautet der Stamm *uwa/ue-*, d.h. die 3. Sg. Prs. **ú-ez-zi**.
- Bei den Verben auf *-ešš-* finden sich manchmal Schreibungen mit ⟨iš⟩, die auch so mit *i* transliteriert werden.
- Der Ausgang der 3. Pl. Prt. Ind. Akt. lautet /-er/. Dementsprechend sind die Zeichen – wo möglich – mit *e* zu lesen. Das gilt auch für die KVK-Zeichen, die im HZL keinen *e*-Wert haben (s.o.). Dort ist anstelle der *i*-Lesung eine *e*-Lesung unter Beibehaltung der ggf. vorhanden Indexzahl einzusetzen, z.B.

per statt *pir* in *e-ep-per*. Dabei ist zu beachten, dass es *mi-* und *hi-*Verben mit einem Pluralstamm auf *-iya-* gibt, die dementsprechend eine 3. Pl. Prt. auf *-ier/* aufweisen.

Determinative und Datumsangaben

- Das Determinativ für Götternamen ist ^D (Großbuchstabe, nicht Kapitalchen, und nicht ^d).
- Falls ein Göttername mit DINGIR.MEŠ determiniert ist, entsprechend der Konvention zu ^{D.MEŠ} abkürzen, z. B. ^{D.MEŠ}*dam-na-ša-ra-aš*.
- Männer- und Frauennamen werden mit ^m bzw. ^f determiniert.
- Wenn auf ein Determinativ für Personennamen direkt ein weiteres Determinativ folgt, sind diese mit einem Punkt zu verbinden, z.B.: ^{m.D}30.SUM-*ia*
- Datumsangaben: UD 24^{KAM} (mit Leerzeichen).
- Gleiches gilt für Jahres- oder Tafelangaben: MU 3^{KAM} und auch DUB 2^{KAM}. Beachte allerdings die Kombination mit einem akkadischen Komplement DUB.3-*PU*.
- Eine deutlich erkennbare Schreibung des Pluraldeterminativs als ^{ME.ES} wird in der Transliteration konsequent von ^{MES} unterschieden.
- Die Pluraldeterminative ^{DIDLL.HI.A} werden durch Punkte getrennt (gegen HZL sub Nr. 1) und als Determinativ markiert: URU^{DIDLL.HI.A}.
- Das sumerische Suffix .TA.ÀM „je“ für Distributivzahlen wird als Sumerogramm gekennzeichnet: 4.TA.ÀM

Abkürzungen

- Abgekürzte hethitische Worte erhalten keine besondere Markierung (z.B. *pé-an*, nicht *pé.-an!*).
- Abgekürzte Akkadogramme erhalten keine besondere Markierung (z.B. *PA, LI*).

Ligaturen

- Als Ligaturen geschriebene Zeichenfolgen werden mit einem Pluszeichen statt mit Bindestrich verbunden (z. B. *kat+ta, A+NA*).

Sumerogramme

- Sumerogramme für „Gefäß“ + „Inhalt“ sind stets ohne Punkt zu schreiben (z.B. GAL.KAŠ lies GAL KAŠ).
- Prinzip: Nomen + Adj. Enthält Spatium dazwischen, z. B. GIŠ.^DINANNA GAL/TUR
- Wenn in einem Logogramm zwischen Sumerogramm und Determinativ kein Leerzeichen sondern ein Punkt steht (sogenannte ‚Binnenderterminative‘), soll der Punkt als Sumerogramm formatiert sein, z. B. GIŠ.^DINANNA oder DUMU^{MEŠ}.É.GAL. Gleiches gilt für ‚Binnenkomplemente‘ wie AN.DAḤ^{ah}.ŠUM.

- Feste Verbindungen, e.g. LÚ^{MES} _ GÍŠ BANŠUR als solche mit dem Zeichen _ als Bindeglied gekennzeichnet (Logogrammtabellen Referenzliste in Appendix)

Für die vollständige verbindliche und aktuelle Lesungen aller Logogramme, auch der Orts- und Götternamen, siehe die jeweilige Tabellen der Logogramme, derzeit in der Dropbox: /Annotation/TABELLENaktuell/...

D. XML Dateiumwandlung und -verwaltung

Speicherung .xml-Dateien in /HFR BASISCORPUS XML

Für die endgültige Umwandlung der Basiscorpus .odt-Dateien ins .xml-Format zwecks Annotation und Weiterverarbeitung, wird folgendermaßen vorgegangen:

- Bevor überhaupt die Umwandlung ins .xml und die anschließende Annotation beginnen, müssen die .odt-Dokumente der jeweiligen CTH Nummer auf Versionsprüge in TeamDrive geprüft werden. Dies geschieht am schnellsten und bequemsten in TeamDrive, mit der Ansicht der Datei-Versionen rechts eingeschaltet, während die einzelnen Dateien links durch geklickt werden. Versionsprüge sind auf Anheb erkennbar durch die nebeneinander stehenden Versionsnummern mit einem Pfeil dazwischen, z. B. 5 ← 3 (d. h. Version 3 abgespeichert als Version 5; Version 4 wurde nicht weiter berücksichtigt).
- .odt-Dateien aus „/Spaces/HFR Basiscorpus Dateien_CTH/CTH 000“ ins .xml-Format umwandeln
- .xml-Dateien speichern unter „/Spaces/HFR BASISCORPUS XML/CTH 000_xml“
- Den entsprechenden Ordner in „/Spaces/OUTPUT_XMLbasis/CTH 000“ (übergangsweise für die automatisch annotierten aber noch nicht validierten xml-Dateien vorgesehen) löschen, sobald die xml-Dateien unter „/Spaces/HFR BASISCORPUS XML/CTH 000_xml“ gespeichert wurden.
- Die einzelnen Dateien behalten ihre bisherige Benennung bis auf die Erweiterung „.xml“ (anstelle von .odt)
- Die jeweilige Statusdatei nach „/Spaces/HFR BASISCORPUS XML/CTH 000_xml“ umziehen
- Umbenennung des jeweiligen Ordners in „/Spaces/HFR Basiscorpus Dateien_CTH/CTH 000“ als „/CTH 000_depr“

*Damit sind die .odt-Dateien **depreziert**, und bleiben zwecks Archivierung in „/Spaces/HFR Basiscorpus Dateien_CTH“ gespeichert. Der Vorteil bei TeamDrive ist die Archivierung des gesamten Bearbeitungsprozesses (Dateiversionen) sämtlicher .odt-Dokumente.*

Für HFR wird nur nach Rücksprache mit den Annotatoren Weiterarbeit/Korrektur an den .odt-Dateien ausgeführt, deren anschließende (Neu)Annotation aber in das .xml-Dokument übernommen wird. Zukünftig wird dann die Weiterarbeit an den .xml-Dokumenten mit einem .xml-Editor direkt vorgenommen.

Dokumentation der .xml-Umwandlung und Annotation

Der Annotationsvorgang an den einzelnen Texten muss in die Statusdatei eingetragen werden. *Die Statusdateien werden zunächst als .odt-Textdokumente weitergepflegt. (Ob sie zukünftig auch in ein .xml-Format umgewandelt werden, ist noch zu klären.)*

- Bei Erstellung des automatisch annotierten .xml-Formats, in der Datei „/Spaces/HFR Basiscorpus Dateien_CTH/Organisation/Übersicht aller Festrutuale.ods“ in der Spalte „XML und automatisch annotiert“ (ein einziger Vorgang): Erstbuchstaben des verantwortlichen Mitarbeiters nach Umwandlung der jeweiligen CTH Nr.

- Bei Fertigstellung der manuell-geprüften Annotation für **einzelne Dateien**, in der **Statusdatei** jeweils bei den Texten dokumentieren: /ANN_[NAME] [DATUM]
- Wenn die manuelle **Annotation für die komplette CTH-Nr. fertig** ist, Erstbuchstaben in der Spalte „manuell annotiert“ in „Übersicht aller Festrituale.ods“ eintragen

Weiterverarbeitung der .xml-Dateien (noch zu klären)

Insbesondere in Hinsicht auf Druckpublikation bedarf es einer Umwandlung der .xml-Dateien in einem mit Dokumentenverarbeitungsprogrammen kompatiblen Format.

Anhang

Sumerogramme

-*ĀM* => .ĀM

-*TA.ĀM* => .TA.ĀM

-*TA.ĀM* => .TA.ĀM

.KAM => ^{KAM}

"UDU.SÍG+MUNUS" => UDU.SÍG+MUNUS

^{(M)EŠ} *lies* ^{MES}, (Zeichen EŠ als Pluraldeterminativ)

10-*ub* => 10-*up*

10-*ub-bi* => 10-*up-pí*

10-*up-pé-na* => 10-*up-pí-na*

7.7.BI => IMIN.IMIN.BI

ÁB.NIGA => ÁB NIGA

ÁGA => ÁG

ĀM => .ĀM

AM.SI ZU₉ => AM.SI_ZU₉

AMAR.ÁB ŠE => AMAR.ÁB NIGA

AMAR.MAḤ ŠE => AMAR.MAḤ NIGA

Á^{MUŠEN} => TI₈^{MUŠEN}

AN.TAḤ.ŠUM => AN.DAḤ.ŠUM

ANŠE MUNUS => ANŠE.MUNUS

ANŠE MUNUS.AL.LÁ => ANŠE.MUNUS.AL.LÁ

ANŠE NÍTA => DÛR

ANŠE.KUR.RA MUNUS => ANŠE.KUR.RA.MUNUS

ANŠE.KUR.RA MUNUS.AL.LÁ => ANŠE.KUR.RA.MUNUS.AL.LÁ

ANŠE.KUR.RA NÍTA => ANŠE.KUR.RA.NÍTA

ANŠE.NUN.NA MUNUS.AL.LÁ => ANŠE.NUN.NA.MUNUS.AL.LÁ

ANŠE.NUN.NA NÍTA => ANŠE.NUN.NA.NÍTA

BANSUR => BANŠUR

BANŠUR.GAL => BANŠUR GAL

BUGIN GÍD.DA => BUGIN.GÍD.DA

BUGIN TUR => BUGIN.TUR

BÚGIN TUR => BÚGIN.TUR

BUNIN => BUGIN

BÚNIN => BÚGIN

^DGIŠ.INANNA GAL/TUR

^D10 => ^D10

^D30 (Zahl 30 als Name des Mondgottes: als „AO:SumGRAM“ formatieren!)

DAG-*TI* => DAG-*tí*

DAM.KI.NA => *DAM-KI-NA*

^DÉ.A => ^DÉ-*A*

^DEN.ZU *lies* ^DZUEN

DI.KUD => DI.KU₅

^DIB => ^DURAŠ

DÍLI => DÍLIM

DILIM => DÍLIM

^DIM => ^DIŠKUR

DINGIR GE₆

^DKAL. => ^DKAL (*mit Spatium danach*)
^DLAMA => ^DKAL
^DLAMA. => ^DKAL (*mit Spatium danach*)
^DMAḤ => DINGIR.MAḤ
^DU => ^D₁₀
^DUTU-*gi/ki* => ^DUTU-*ge/ke* (*hurr. Šimig/ke*)
DUG SAGI => DUG.SAGI
DUG SAGIA => DUG.SAGIA
DUG.DÍLIM.GAL => ^{DUG}DÍLIM.GAL
DÛG.GA => DU₁₀.GA
DUG.KA.GAG => ^{DUG}KA.GAG
DUG.KA.GAG.A => ^{DUG}KA.GAG.A
DUG.KAŠ => DUG KAŠ
^{DUG}KAŠ => DUG KAŠ
^{DUG}LIŠ.GAL *lies* ^{DUG}DÍLIM.GAL
^{DUG}TU₇ => ^{DUG}ŪTUL
DUMU É.GAL => DUMU.É.GAL
DUMU-*TIM* => TUR-*TIM*
DUMU.E.GAL => DUMU.É.GAL
DUMU.LÚ.ULÛ => DUMU.LÚ.U₁₉.LU
DUMU.MUNUS.LUGAL => DUMU.MUNUS LUGAL
DUMU^{MEŠ} É.GAL => DUMU^{MEŠ}.É.GAL
DÚR-*aš* => TUŠ-*aš*
DURU₅ => DUR₅
É GIŠ.ÛR.RA => É.GIŠ.ÛR.RA
É NA₄ => É.NA₄
É-*ir* => É-*er*
E.DÉ => SIMUG
É.GEŠTIN => É GEŠTIN
É.ḤI.UŠ.SA => É.DU₁₀.ÛS.SA
É.ŠÁ => É.ŠÀ
É.TU₇ => É TU₇
^ÉDU₁₀ => É.DU₁₀
EGIR KASKAL => EGIR.KASKAL
É^{MEŠ} NA₄ => É^{MEŠ}.NA₄
EN.ÉRIN => EN ÉRIN
EN.ZU => ZUEN
ERIN => EREN
ERÍN => ÉRIN
ÉRIN.MEŠ => ÉRIN^{MEŠ}
EZEN => EZEN₄
EZEN×ŠE => EZEN₄
EZEN₄ => EZEN₄
EZEN₄.ITU => EZEN₄ ITU
EZEN×ŠE => EZEN₄
GA.KALA.GA => GA KALA.GA
GA.KALAG.GA => GA KALA.GA
GA.KU₇ => GA KU₇
GAD => GADA
GAL.GEŠTIN => GAL GEŠTIN

GAL.KAŠ => GAL KAŠ
 GALÁ => GALA
 GAŠAN.LÍL => GAŠAN LÍL
 GESTIN => GEŠTIN
 GEŠTIN ḪÁD.DU.A => GEŠTIN.ḪÁD.DU.A
 GEŠTIN-*ia-na-aš* => *wi₅-ia-na-aš*
 GEŠTIN.KU₇ => GEŠTIN KU₇
 GEŠTIN.UD.DU.A => GEŠTIN.ḪÁD.DU.A
 GEŠTUG => GEŠTU
 GIS => GIŠ
 GIS => GIŠ
 GIŠ^DINANNA => GIŠ^DINANNA
 GIŠ INANNA => GIŠ.INANNA
 GIŠ.ÉRIN.ZI.BA.NA => GIŠ.ÉRIN.ZI.BA.NA
 GIŠ^{ERIN} *lies* GIŠ^{EREN}
 GIŠ^{ḪUR} => GIŠ.ḪUR
 GIŠ^{IGI.GAG} *lies* GIŠ^{ŠUKUR}
 GIŠ^{NÁ} => GIŠ.NÁ
 GIŠ^{PA} => GIŠ^{GIDRU}
 GIŠ^{ŠÀ.A.TAR} *lies* GIŠ^{TIBULA}
 GIŠ^{SAR} => GIŠ^{KIRI₆}
 GIŠ^{ŠAR} => GIŠ^{KIRI₆}
 GU.ZA GÍD.DA => GU.ZA.GÍD.DA
 GU₄ => GU₄
 GU₄ => GU₄
 GU₄ ŠE => GU₄ NIGA
 GU₄.ÁB => GU₄ÁB
 GU₄.AMAR => GU₄AMAR
 GU₄.MAḪ ŠE => GU₄.MAḪ NIGA
 GU₄.NIGA => GU₄ NIGA
 GU₄.ŠE => GU₄ NIGA
 GUB-*as* => GUB-*aš*
 GUD => GU₄
 GUD.ÁB => GU₄ÁB
 GUD.UDU => GU₄ UDU
 GÚDU => GUDU₁₂
 GUŠKIN => KÛ.SI₂₂
 ḪI.ÚŠ => DU₁₀.ÚŠ
 ḪI.ÚŠ => DU₁₀.ÚŠ
 ḪUR.ŠAG => ḪUR.SAG
 Ì.DÛG.GA => Ì.DU₁₀.GA
 Ì.DUG.GA => Ì.DU₁₀.GA
 IGI.GAG => ŠUKUR
 ILLAT => KASKAL.KUR
 IM.ME => GUDU₁₂
 INANNA.GAL => INANNA GAL
 INANNA.TUR => INANNA TUR
 ÌR => ARAD
 KA.DU => GÛ.DU
 KA.DÛ => KA.GAG

KA×DU => ZU,
 KAB.KA.GA => KAB.DU₁₁.GA
 KAR-*aš* => *kar-aš*
 KASKAL.GAL => KASKAL GAL
 KA×DU => ZU,
 KI.KAL.BAD => KARAŠ
 KILAL => KILĀ
 KI.NE => GUNNI
 KÛ.GI => KÛ.SI₂₂
 KÛ.SI₂₂ => KÛ.SI₂₂
 KU₇ => KU₇
 KUBABBAR => KÛ.BABBAR
 KUŠ.GU₄ => KUŠ GU₄
 KUŠ.UDU => KUŠ UDU
 KUŠ.UZ₆ => KUŠ UZ₆
 KUŠ₇ KÛ.SI₂₂ => KUŠ₇.KÛ.SI₂₂
 LAMMA => KAL
 LAMMA. => KAL (*mit Spatium danach*)
 LIŠ.GAL => DĪLIM.GAL
 LÚ^D₁₀ => LÚ₁₀^D
 LÚ^DĪŠKUR => LÚ₁₀^DĪŠKUR
 LÚ Ē.GAL => LÚ₁₀Ē.GAL
 LÚ GĪR => LÚ₁₀GĪR
 LÚ^{GIŠ}BANŠUR => LÚ₁₀^{GIŠ}BANŠUR
 LÚ^{GIŠ}GIDRU => LÚ₁₀^{GIŠ}GIDRU
 LÚ^{GIŠ}ŠUKUR => LÚ₁₀^{GIŠ}ŠUKUR
 LÚ^{GIŠ}ŠUKUR KÛ.SI₂₂ => LÚ₁₀^{GIŠ}ŠUKUR.KÛ.SI₂₂
 LÚ^{GIŠ}TUKUL => LÚ₁₀^{GIŠ}TUKUL
 LÚ^{GIŠ}ZA.LAM.GAR => LÚ₁₀^{GIŠ}ZA.LAM.GAR
 LÚ NIM.LĀL => LÚ₁₀NIM.LĀL
 LÚ ŠUKUR => LÚ₁₀ŠUKUR
 LÚ ŠUKUR KÛ.SI₂₂ => LÚ₁₀ŠUKUR.KÛ.SI₂₂
 LÚ ZA.LAM.GAR => LÚ₁₀ZA.LAM.GAR
 LÚ.Ē.GAL => LÚ₁₀Ē.GAL
 LÚ^{MES}Ē.GAL => LÚ₁₀^{MES}Ē.GAL
 LÚ^E.DĒ(.A) *lies* LÚ^{SIMUG}(.A)
 LÚ^EĒ.GAL => LÚ₁₀^EĒ.GAL
 LUGAL.GAL => LUGAL GAL
 LÚ^{GALĀ} *lies* LÚ^{GALA}
 LÚ^{IM.ME} *lies* LÚ^{GUDU}₁₂
 LÚ^{MES} => LÚ₁₀^{MES}
 LÚ^{MES}₁₀ => LÚ₁₀^{MES}₁₀
 LÚ^{MES}₁₀ĪŠKUR => LÚ₁₀^{MES}₁₀ĪŠKUR
 LÚ^{MES}₁₀Ē.GAL => LÚ₁₀^{MES}₁₀Ē.GAL
 LÚ^{MES}₁₀GĪR => LÚ₁₀^{MES}₁₀GĪR
 LÚ^{MES}₁₀^{GIŠ}BANŠUR => LÚ₁₀^{MES}₁₀^{GIŠ}BANŠUR
 LÚ^{MES}₁₀^{GIŠ}GIDRU => LÚ₁₀^{MES}₁₀^{GIŠ}GIDRU
 LÚ^{MES}₁₀^{GIŠ}ŠUKUR => LÚ₁₀^{MES}₁₀^{GIŠ}ŠUKUR
 LÚ^{MES}₁₀^{GIŠ}TUKUL => LÚ₁₀^{MES}₁₀^{GIŠ}TUKUL
 LÚ^{MES}₁₀^{GIŠ}ZA.LAM.GAR => LÚ₁₀^{MES}₁₀^{GIŠ}ZA.LAM.GAR

LÚ^{MEŠ} NIM.LÀL => LÚ^{MEŠ} NIM.LÀL
 LÚ^{MEŠ} ŠUKUR => LÚ^{MEŠ} ŠUKUR
 LÚ^{MEŠ} ŠUKUR.KÛ.SI₂₂ => LÚ^{MEŠ} ŠUKUR.KÛ.SI₂₂
 LÚ^{MEŠ} ZA.LAM.GAR => LÚ^{MEŠ} ZA.LAM.GAR
 LÚ^{MEŠ} É.GAL => LÚ^{MEŠ} É.GAL
 LÚ^{MU} => LÚ^{MU} MUḪALDIM
 LÚ^{QA.ŠU.DU₈(A)} lies LÚ^{SAGI(A)}
 LÚ^{SILÀ.ŠU.DU₈/DUḪ} lies LÚ^{SAGI}
 LÚ^{SIPÁ(UDU)} lies LÚ^{SIPA(UDU)}
 LÚ^{ZABAR.DIB} lies LÚ^{ZABAR.DAB}
 MÁ.TUR => MÁ TUR
 MAŠ.SÌLA => MAŠ.SÌL
 MÁŠDA => MAŠ.EN.GAG
ME (100) ist akkadographisch als ME zu transliterieren (für die Wiedergabe von Zahlen mit ME s. Unten).
 ME.EŠ und MEŠ sind zu unterscheiden
 MEŠKÁ UR.GI₇ => MEŠKÁ.UR.GI₇
 MEŠKÁ.GAL UR.GI₇ => MEŠKÁ.GAL.UR.GI₇
 MU.KAM => MU^{KAM}
 MÚ.KIRI₆ => MÚ.SAR
 MUḪALIM => MUḪALDIM
 MUNU₈ => DIM₄
 MUNUS AMA.DINGIR => MUNUS AMA.DINGIR
 MUNUS GIŠ_{PAN} => MUNUS GIŠ_{PAN}
 MUNUS NA₄ARA₅ => MUNUS NA₄ARA₅
 MUNUS.AMA DINGIR => MUNUS AMA.DINGIR
 MUNUS.AMA.DINGIR => MUNUS AMA.DINGIR
 MUNUS AMA DINGIR => MUNUS AMA.DINGIR
 MUNUS^{MEŠ} GIŠ_{PAN} => MUNUS^{MEŠ} GIŠ_{PAN}
 MUNUS^{MEŠ} NA₄ARA₅ => MUNUS^{MEŠ} NA₄ARA₅
 MURUB₄ => MÚRU
 MUSEN => MUŠEN
 NA4 => NA₄
 NAGGA => AN.NA
 NAR GAL => NAR.GAL
 NERGAL => NÈ.ERI₁₁.GAL
 NÍG.ḪAR.RA => NÍG.ÀR.RA
 NÍG.LAM => NÍG.LÁM
 NINDA Ì.E.DÉ.A => NINDA.Ì.E.DÉ.A
 NINDA KU₇ => NINDA.KU₇
 NINDA.GUR.RA => NINDA.GUR₄.RA
 NINDA.GUR4.RA => NINDA.GUR₄.RA
 NINDA.GUR₄.RA.GAL => NINDA.GUR₄.RA GAL
 NINDA.KUR₄.RA => NINDA.GUR₄.RA
 NINDA^{ÉRIN} => NINDA.ÉRIN
 NINDA^{Ì.E.DÉ.A} => NINDA.Ì.E.DÉ.A
 NINDA^{KASKAL} => NINDA.KASKAL
 NU.AL.TIL => NU AL.TIL
 NU.TIL => NU TIL
 PA^{LÚ.MEŠ} => UGULA^{LÚ.MEŠ}
 PAB => ✎
 QA.ŠU.DU₈ => SAGI

QA.ŠU.DUḪ => SAGI
 ŠÀ-BA => ŠÀ.BA
 ŠÀ.A.TAR => TIBULA
 SAG.GÉME.ÌR => SAG.GÉME.ARAD
 SAL => MUNUS
 ŠE.KIN.KUD => ŠE.KIN.KU₅
 SÍLA => SILA₄
 SILÀ.ŠU.DU₈ => SAGI
 SILÀ.ŠU.DUḪ => SAGI
 SILA₄ => SILA₄
 SIPÁ => SIPA
 SIPAD => SIPA
 ŠÌR => ŠÌR
 ŠÌR-ru => ŠÌR-RU
 ŠÌR-RU => ŠÌR-RU
 ŠÌR^{RU} => ŠÌR-RU
 SISKUR.SISKUR => SÍSKUR
 ŠU.NIGIN => ŠU.NÍGIN
 SUḪUR.LAL => SUḪUR.LÁ
 SÚR.DÛ.A => SUR₁₄.DÛ.A
 TA.ÀM => .TA.ÀM
 TE^{MEŠ} => TE^{MEŠ}

TU₇ als Bezeichnung verschiedene Suppenarten (TU₇.Ì „Fettbrühe, auch determiniert ^{UZU}TU₇ „Fleischsuppe“); aber determiniert ^{DUG} und ^{URUDU} lies ^{DUG/URUDU}ÚTUL „Topf“

TU₇ A => TU₇.A
 TU₇ A UZU => TU₇.A.UZU
 TU₇ A UZU GU₄ => TU₇.A.UZU.GU₄
 TU₇ A.UZU => TU₇.A.UZU
 TU₇ BA.BA.ZA => TU₇.BA.BA.ZA
 TU₇ ÉRIN => TU₇.ÉRIN
 TU₇ GA => TU₇.GA
 TU₇ GÚ.GAL => TU₇.GÚ.GAL
 TU₇ GÚ.GAL.GAL => TU₇.GÚ.GAL.GAL
 TU₇ GÚ.TUR => TU₇.GÚ.TUR
 TU₇ Ì => TU₇.Ì
 TU₇ KIRI₆ => TU₇.KIRI₆
 TU₇ ME-E UZU => TU₇.ME-E.UZU
 TU₇ ME-E UZU GU₄ => TU₇.ME-E.UZU.GU₄
 TU₇ ME-E UZU.GU₄ => TU₇.ME-E.UZU.GU₄
 TU₇ NÍG.ÀR.RA => TU₇.NÍG.ÀR.RA
 TU₇ SAR => TU₇.SAR
 TU₇ UZU => TU₇.UZU
 TU₇.Ì => TU₇.Ì
 TU₇.SAR => TU₇.SAR
 TU₇.UZU => TU₇.UZU
 TÚG.GÚ.È => TÚG.GÚ.È
 TÚG.GÛN => TÚG.GÛN
 TUKUL GÍD.DA => TUKUL.GÍD.DA
 TÚL => PÚ
 TÚL => PÚ

TUŠ-*as* => TUŠ-*aš*
 UD-*MI*; UD-*UM*; UD-*MA-AM*
 UD.KA.BAR => ZABAR
 UD.KAM => UD^{KAM}
 UD.SIG₅ => UD SIG₅
 UD.SIG₅.GA => UD SIG₅.GA
 UD.ZAL.LI => UD.ZAL.LE
 UD^{KAM}.SIG₅ => UD^{KAM} SIG₅
 UDU.NIGA => UDU NIGA
 UDU.NÍTA => UDU NÍTA
 UDU.ŠE => UDU NIGA
 UDU.SÍG+MUNUS (*ohne Anführungszeichen*)
 UMMEDA => ÛMMEDA
 URUDU₇TU₇ => URUDU₇ÚTUL
 URU₇GIDRU-*ti* => URU₇GIDRU-*TI*
 URU₇GIDRU-*TI* => URU₇ĤAT-*TI*
 URU₇GIDRU-*ti* => URU₇ĤAT-*TI*
 URU₇KÛ.BABBAR-*ti* => URU₇GIDRU-*TI*
 ÚTUL => TU₇
 UTÚL => ÚTUL
 ÚTUL *lies* TU₇, *wenn es sich um Speisegerichte und nicht ein Gefäß handelt*
 UZU.GU₄ => UZU GU₄
 UZU.Ì => UZU₄Ì
 UZU.ŠÀ => UZU₄ŠÀ
za-ba₄-ba₄ => ZA-BA₄-BA₄
 ZA.BA₄.BA₄ => ZA-BA₄-BA₄
 ZABAR.DIB => ZABAR.DAB
zag => ZAG
 ZÍD.DA => ZÌ.DA
 ZIZ.A => ZÍZ DUR₅

 M => m
 F => f

Akkadogramme

1 NU-TIM => 1-NU-TIM
 A-BI BI-TI => A-BI-BI-TI
 A-BU BI-TI => A-BU-BI-TI
 A-BU BI-TI-I => A-BU-BI-TI-I
 A-BU BI-TIM => A-BU-BI-TIM
 AR-ZA-AN-NU => AR-SÀ-AN-NU
 AR-ZA-NA => AR-SÀ-NA
 AR-ZA-NU => AR-SÀ-NU
 BUR-ZI-TUM => PUR-SÍ-TUM
 DAM-KE-EN-NA => DAM-KI-EN-NA
^DSÍN *lies* ^D30 (*30 als „AO:SumGram“, nicht als Numeral zu markieren!*)
^DU₄ => ^DUD
 E-A => Ê-A
 E-EN-TU₄ => E-EN-TUM

Ê-TIM => Ê-TIM
 EM-ŞU => EM-ŞÚ
 GA-AN-NU-UM => KĀ-AN-NU-UM
 GAL-AM => GAL-AM
 ȚA-AZ-ZI-IN => ȚA-AŞ-ŞĪ-IN
 ȚA-AZ-ZI-IN-NA => ȚA-AŞ-ŞĪ-IN-NA
 ȚA-AZ-ZI-IN-NE => ȚA-AŞ-ŞĪ-IN-NE
 ȚA-AZ-ZI-IN-NI => ȚA-AŞ-ŞĪ-IN-NI
 ȚA-AZ-ZI-IN-NU => ȚA-AŞ-ŞĪ-IN-NU
 ȚA-DA-A-NU => ȚA-TĀ-A-NU
 ȚA-DA-NI => ȚA-TĀ-NI
 ȚA-DA-NU => ȚA-TĀ-NU
 I NU-TIM => 1-NU-TIM
 I-DI => I-DE
 IM-ZA => EM-ŞA
 IS-TU => IŞ-TU
 IŞTU => IŞ-TU
 KA-NI-EŞ => ka-ni-eš
 KA-NI-IŞ => ka-ni-iš
 ka-pal-li => KA-BAL-LI
 KA-PAL-LI => KA-BAL-LI
 KA-YA-MA-NIM => KA-IA-MA-NIM
 KAP-RU => KAB-RU
 KU-DA-DU₄ => KU-TĀ-TUM
 KU-RI-DU => KU-RI-TŪ
 KU-UL-LU-U => QŪ-UL-LU-U
 KU-UL-LU-Ū => QŪ-UL-LU-Ū
 KU-UŞ-ŞI => KU-UŞ-ŞĪ
 KU-UZ-ZI => KU-UŞ-ŞĪ
 LA-AB-GA => LA-AB-KĀ
 LI-AĤ-E => LE-E²-E
 LI-IM ŞE-RI => LI-IM,ŞE-RI
 LI-IM ŞE-RI => LI-IM,ŞE-RI
 LŪ^{MEŠ} ZI-IT- => LŪ^{MEŠ} ZI-IT-
 LŪ¹ A-BU-BI-DU lies LŪ¹ A-BU-BI-TŪ
 LŪ¹ KUŠ₇ => LŪ¹ KUŠ₇
 LŪ¹ MU-DI- lies LŪ¹ MU-TI₄-
 LŪ¹ MU-DU- lies LŪ¹ MU-TŪ- etc.
 LŪ¹ TE-MU lies LŪ¹ TE₄-MU
 LŪ¹ ZI-IT- => LŪ¹ ZI-IT-
 ME (100) ist akkadographisch als ME zu transliterieren.
 ME-EL-KI-IT => ME-EL-QĒ-ET
 ME-EL-KI-TUM => ME-EL-QĒ-TUM
 ME-EL-QĪ-IT => ME-EL-QĒ-ET
 MU-BAR-RI-TI => MU-PĀR-RI-ŤĪ
 MU-BAR-RI-TI₄ => MU-PĀR-RI-ŤĪ
 MU-DI-ŞU => MU-TI₄-ŞU
 NA-AK-DAM-MI => NA-AK-TĀM-MI
 NA-AP-ŤA-RI
 NAM-MA-AN-TA => NAM-MA-AN-DĀ

NAM-MA-AN-TI => NAM-MA-AN-DÌ
 NAM-MA-AN-TÛ => NAM-MA-AN-DU
 NAM-MA-AN-TUM => NAM-MA-AN-DUM
 NAM-MA-TUM => NAM-MA-DUM
 NI-EŠ => NI-ĪŠ
 NI-MI-IT-TI => NĒ-MĒ-ET-TI
 NI-MI-IT-TUM => NĒ-MĒ-ET-TUM
 PUR-RU-TIM => BUR-RU-TIM
 QA-AZ-ZU => QA-AS-SŪ
 QA-ZU => QA-SŪ
 RI-IK-QÍ => RI-IQ-QÍ
 RI-KU-TI => RI-QŪ-TI
 ŠA-A-DU => SA₂₀-A-TŪ
 ŠA-A-TI => SA₂₀-A-TI
 ŠA-KĀ-RU => ŠA-GA-RU
 ŠA-PU-LU => ŠA-BU-LU
 SA₂₀-A-TU
 SA₂₀-A-TŪ
 SA_x => SA₂₀
 SA_x => SA₂₀
 SĒ-EM-MI-TA-A-TI => SĒ-EM-MI-DĀ-A-TI
 SE₂₀-ER-TUM => SE₂₀-ER-DUM
 ŠÍG ḤA-ŠAR- => ŠÍG ḤA-ŠAR-
 SILA₄ => SILA₄
 ŠU-TUM => ŠU-TUM
 ŠU-UP-PA-DU => ŠU-UP-PA-TŪ
 SUMUQAN => GÌR
 SUMUQAN => GÌR
 TE-ŠI => DI₁₂-ŠI
 TU-TI-IT-TUM => TU-DÌ-IT-TUM
 TU-U-TI-IT-TUM => TU-U-DÌ-IT-TUM
 ṬUP-PA => TUP-PA
 TUP-PI => TUP-PÍ
 ṬUP-PÍ => TUP-PÍ
 ṬUP-PU => TUP-PU
 Ú-BA-RŪ => Ú-BA-RUM
 U₄ => UD
 UM-ME-IA-NU-TIM => UM-MĪ-IA-NU-TIM
 URU ḤAT-TI und URU ḤA-AT-TI *akkadographisch, außer bei hurritischer Komplementierung*
 URU ḤA-AT-TE-n => URU ḥa-at-te-n
 URU ḤA-AT-TI-n => URU ḥa-at-ti-n
 URU ḤA-LA-AB => URU ḥa-la-ap
 URU ḤA-LA-AP => URU ḥa-la-ap
 URU ḤUR-RI => URU ḥur-ri
 UŠ-GE-EN => UŠ-KĒ-EN
 UŠ-GI-EN => UŠ-KĒ-EN
 UŠ-KI-EN => UŠ-KE-EN
 UŠ-KÍ-EN-NU => UŠ-KĒ-EN-NU
 UŠ-KI-IN-NU => UŠ-KE-EN₆-NU
 UZU QA-TAM => UZU QA-TŪ

ZA-AL-MU-TIM => ŠA-AL-MU-TIM

ZÉ-ER-TU₄ => SE₂₀-ER-DUM

ZÉ-ER-TUM => SE₂₀-ER-DUM

ZI-IA-TUM => zi-ia-tu₄

ZI-ID-DI => ZI-IT-TI₄

ZI-KU-KI => sí-QŪ-Qí

Hethitische Lexeme

-ḫi-ir => -ḫe-er

-me-it => -me-et

-ša-me-it => ša-me-et

-še-it => -še-et

-še-me-it => še-me-et

-te-it => -te-et

-ua => -wa

-ue => -we

-ya => -ia

^{NINDA} a-a-an etc.

a-ḫar-re-e-eš => a-ḫar-ri-e-eš

a-pé-ia => a-pí-ia

a-péd-da => a-pád-da

a-pí-e-da-aš => a-pé-e-da-aš

a-pí-e-da-ni => a-pé-e-da-ni

a-pí-e-el => a-pé-e-el

a-pí-e-et => a-pé-e-et

apez, apedani, apenissan

a-pí-e-ez => a-pé-e-ez

appezzi-

a-pí-e-ti-i => a-pé-e-ti-i

a-pí-ni-iš-šu-an => a-pé-ni-iš-šu-an

a-pí-ni-iš-šu-wa-an => a-pé-ni-iš-šu-wa-an

a-ru-wa-a-iz-zi => a-ru-wa-a-ez-zi

a-ru-wa-iz-zi => a-ru-wa-ez-zi

ap-pé-eš-kir => ap-pé-eš-ker

ar-kam-mi-ia-le-e-eš => ar-kam-mi-ia-li-e-eš

ar-kam-mi-ia-le-eš => ar-kam-mi-ia-li-eš

da-a-ki-tum => da-a-ki-tu₄

da-ḫi-ia-le-eš => da-ḫi-ia-li-eš

da-ki-tum => da-ki-tu₄

^D gul- => ^DGUL-

^{DINGIR} te-li-pí-nu => ^D te-li-pí-nu

^D te-eš-šu-up-pí (nicht -ub-bi)

^É pár-ku-wa => ^É pár-ku-wa

e-ip-zi => e-ep-zi

e-iz-za-az-zi => e-ez-za-az-zi

^É ḫa-al-ki => ^É ḫa-al-ki

^É tup-pa-aš => ^É tup-pa-aš

^{GADA}tanipu- (nicht kattanipu-)

ga-ri-ip- => ga-re-ep-

ga-ri-p => ga-re-p

ge-ri-e-p- => ge-re-e-p

gi-eš-ḥi => ge-eš-ḥi

gi-nu-wa-aš => ge-nu-wa-aš

gi-ri-ip- => ge-re-ep-

gi-ri-p => ge-re-p

gul-a-an-zi => GUL-a-an-zi

gul-za-a-i-ma => GUL-za-a-i-ma

gul-za-an-zi => GUL-za-an-zi

gul-za-da-na-za => GUL-za-da-na-za

ḥa-an-di-zum-ni => ḥa-an-de-zum-ni

ḥantezzi-

ḥa-an-te-ez-ze-en => ḥa-an-te-ez-zi-en

ḥa-ap-pí-na-an => ḥa-ap-pé-na-an

ḥa-aš-du-ir => ḥa-aš-du-er

ḥa-at-ta-ri-eš-na-aš => ḥa-at-ta-re-eš-na-aš

ḥa-at-tal-le-e-eš => ḥa-at-tal-li-e-eš

ḥa-at-tal-le-eš => ḥa-at-tal-li-eš

ḥa-at-ti-le-e-eš => ḥa-at-ti-li-e-eš

ḥa-at-ti-le-eš => ḥa-at-ti-li-eš

ḥa-at-ti-le-eš => ḥa-at-ti-li-eš

ḥa-at-tin-zi => ḥa-at-tan_x-zi

ḥa-az-ze-e-iš-še-it => ḥa-az-ze-e-iš-še-et

ḥa-az-zi-ik-ki-iz-zi => ḥa-az-zi-ik-ke-ez-zi

ḥa-az-zi-ki-iz-zi => ḥa-az-zi-ke-ez-zi

ḥa-az-zi-ki-u-an => ḥa-az-zi-ke-u-an

ḥa-li-en-tu- => ḥa-le-en-tu-

ḥalentu(wa-)

^Dhé-pát (nicht -bat/d)ḥa-re-eš => ḥa-ri-eš

ḥal-ke-eš => ḥal-ki-eš

ḥal-li-ia-re-e-eš => ḥal-li-ia-ri-e-eš

ḥal-li-ia-re-eš => ḥal-li-ia-ri-eš

ḥal-li-re-e-eš => ḥal-li-ri-e-eš

ḥal-li-re-e-eš => ḥal-li-ri-e-eš

ḥal-li-re-eš => ḥal-li-ri-eš

ḥal-za-a-i => ḥal-za-a-i

ḥal-zi-aḥ-ḥi => ḥal-zi-iḥ-ḥi

ḥal-zi-ša-an-zi => ḥal-ze-ša-an-zi

ḥar-ki-ir => ḥar-ke-er

ḥar-na-i-šar => ḥar-na-i^{SAR}

ḥar-ni-eš-šar => ḥar-né-eš-šar

ḥar-pa-na-al-le-e-eš => ḥar-pa-na-al-li-e-eš

hé-bat => hé-pát

hé-pát => hé-pát

hé-pát-pí-na => hé-pát-pé-na

hu-u-wa- => ḥu-u-wa-

ḥu-wa-aš-ša-na-al-le-eš => ḥu-wa-aš-ša-na-al-li-eš

i-du-re-en => i-du-ri-en

ib-ri-mu-ša => ip-ri-mu-ša
ir-ḥa-a-iz-zi => ir-ḥa-a-ez-zi
ir-ḥa-iz-zi => ir-ḥa-ez-zi
iš-kat-ta-aḥ => iš-kit₉-ta-aḥ
iš-kat-taḥ => iš-kit₉-taḥ
iš-kat-túḥ => iš-kit₉-taḥ
iš-ke₉-ta-aḥ => iš-kit₉-ta-aḥ
iš-ke₉-taḥ => iš-kit₉-taḥ
iš-ke₉-túḥ => iš-kit₉-taḥ
iš-kit₉-túḥ => iš-kit₉-taḥ
iš-pa-an-tu-uz-ze-eš-šar => iš-pa-an-tu-uz-zi-eš-šar
ka-at-re-es => ka-at-ri-es
kattera-
ka-la-ḥe-e-eš => ka-la-ḥi-e-eš
kat-re-e-eš => kat-ri-e-eš
kat-re-eš => kat-ri-eš
ki-e => ke-e
ki-e-da-an-ta => ke-e-da-an-ta
ki-e-da-aš => ke-e-da-aš
ki-e-da-ni => ke-e-da-ni
ki-e-el => ke-e-el
ki-e-et => ke-e-et
ki-e-ez => ke-e-ez
ki-e-ti => ke-e-ti
ku-wa-pé => ku-wa-pí
ku-wa-pé-ek-ki => ku-wa-pí-ik-ki
ku-wa-pé-et => ku-wa-pí-it
ku-wa-pé-et-ta => ku-wa-pí-it-ta
li-e => le-e
li-el-wa-ni => le-el-wa-ni
LÚ^Éḥé-eš-ta-a => LÚ₁^Éḥé-eš-ta-a
LÚ^Éḥé-eš-ti-i => LÚ₁^Éḥé-eš-ti-i
LÚ^Éḥi-iš-ta-a => LÚ₁^Éḥi-iš-ta-a
lu-uk-ke-ez-zi => lu-uk-ki-iz-zi
LÚ^{MES}ḥé-eš-ta-a => LÚ^{MES}₁ḥé-eš-ta-a
LÚ^{MES}ḥé-eš-ti-i => LÚ^{MES}₁ḥé-eš-ti-i
LÚ^{MES}ḥi-iš-ta-a => LÚ^{MES}₁ḥi-iš-ta-a
LÚ¹ḥé-eš-ta-a => LÚ₁ḥé-eš-ta-a
LÚ¹ḥé-eš-ti-i => LÚ₁ḥé-eš-ti-i
LÚ¹ḥi-iš-ta-a => LÚ₁ḥi-iš-ta-a
LÚ¹ma-ni-ia-aḥ-ḥi-aš_{EN} => LÚ¹ma-ni-ia-aḥ-ḥi-aš_{EN}
LÚ¹ma-ni-ia-aḥ-ḥi-aš iš-ḥa => LÚ¹ma-ni-ia-aḥ-ḥi-aš₁iš-ḥa
LÚ^{MES}₁ḥé-eš-ta-a => LÚ^{MES}₁ḥé-eš-ta-a
LÚ^{MES}₁ḥé-eš-ti-i => LÚ^{MES}₁ḥé-eš-ti-i
LÚ^{MES}₁ḥi-iš-ta-a => LÚ^{MES}₁ḥi-iš-ta-a
LÚ¹mu-ki-iš-na-aš_{EN} => LÚ¹mu-ki-iš-na-aš_{EN}
LÚ¹pa-ra-a ú-wa-an => LÚ¹pa-ra-a₁ú-wa-an
ma-al-le-e-et => ma-al-li-e-et
ma-al-le-ez-zi => ma-al-li-ez-zi
ma-ni-ia-aḥ-ḥi-eš-šar => ma-ni-ia-aḥ-ḥe-eš-šar

me-ik-ki => me-ek-ki
mu-ki-iš-ki-mi => mu-ki-iš-ke-mi
mun-an-zi => MUN-an-zi
 NINDA *a-a-an => NINDA a-a-an*
 NINDA *hur-za-zu => NINDA har-za-zu*
 NINDA *mi-in-ta => NINDA mi-in-ta*
 NINDA *mi-it-ga => NINDA mi-it-ga*
 NINDA *mi-it-ka => NINDA mi-it-ka*
 NINDA *mi-it-ta => NINDA mi-it-ta*
 NINDA *ša-ra-am-na-aš ha-a-li => NINDA ša-ra-am-na-aš ha-a-li*
 NINDA *ša-ra-am-na-aš ha-li => NINDA ša-ra-am-na-aš ha-li*
nu-us- => nu-uš-
nu-uš-sa-an => nu-uš-ša-an
nu-za-ni-in-ki => nu za-ni-in-ki
pa-ez-zi => pa-iz-zi
pa-ra-a-ap-pa-an-zi => pa-ra-a ap-pa-an-zi
pa-ra-a-pa-a-an => pa-ra-a pa-a-an
pal-wa-a-iz-zi => pal-wa-a-ez-zi
pal-wa-a-iz-zi => pal-wa-a-ez-zi
pal-wa-at-tal-le-e-eš => pal-wa-at-tal-li-e-eš
pal-wa-at-tal-le-eš => pal-wa-at-tal-li-eš
pal-wa-iz-zi => pal-wa-ez-zi
pal-wa-tal-le-e-eš => pal-wa-tal-li-e-eš
pal-wa-tal-le-eš => pal-wa-tal-li-eš
pár-ša-a-iz-zi => pár-ša-a-ez-zi
pár-ša-iz-zi => pár-ša-ez-zi
pár-si- => pár-ši-
pár-si-ia => pár-ši-ia
pár-si-ia-zi => pár-ši-ia-zi
pár-šu-ul-la-a-iz-zi => pár-šu-ul-la-a-ez-zi
pát-tal-wa-an => pít-tal-wa-an
pát-ti-ia-an-zi => pít-ti-ia-an-zi
pát-tu-lu-uš => pít-tu-lu-uš
pé-an => pé.-an
pé-an-zi => pí-an-zi
peda-
peḥḥun, peḥḥi
pé-en-zi => pí-en-zi
pé-es-kán-zi => pé-eš-kán-zi
pé-eš-ki-u-en => pé-eš-ke-u-en
pé-eš-kir => pé-eš-ker
pé-eš-kit_o-tén => pé-eš-ke_o-tén
pé-eš-si-an-zi => pé-eš-ši-an-zi
pé-es-si-ia-an-zi => pé-eš-ši-ia-an-zi
pé-eš-ši-ni-it => KAŠ-eš-ši-ni-it
pé-tén-zi => pé-tan_x-zi
pé-tin-zi => pé-tan_x-zi
pé.an => pé.-an
péd-da-an-zi => pád-da-an-zi
pí-an => pé.-an

pí-di => pé-di
pí-e-da-an-zi => pé-e-da-an-zi
pí-e-da-i => pé-e-da-i
pí-eš-ši-ia-zi => pé-eš-ši-ia-zi
pí-eš-ši-iz-zi => pé-eš-ši-ez-zi
pí-i-e-an-zi => pé-i-e-an-zi
pí-i-e-ez-zi => pé-i-e-ez-zi
pí-i-e-zi => pé-i-e-zi
pianzi (nicht peanzi)
pí-ia-zi => pé-ia-zi
pí-is-ši-ia-zi => pí-iš-ši-ia-zi
pí-iš-tin => pí-iš-tén
pí-ra-an => pé-ra-an
pí.-an => pé.-an
piš-ku-ru-nu- => pùš-ku-ru-nu-
pít-te-ia-le-e-eš => pít-te-ia-li-e-eš
pít-ti-ia-le-e-eš => pít-ti-ia-li-e-eš
pu-u-re-eš => pu-u-ri-eš
pu-un-ni-ke-eš => pu-un-ni-ki-eš
puš-ku-ru-nu- => pùš-ku-ru-nu-
qa => ka₄
š-kat-ta => š-ket₉-ta
š-kat-ta-at => š-ket₉-ta-at
š-kat-ta-ri => š-ket₉-ta-ri
š-kat-ta-ru => š-ket₉-ta-ru
š-kat-ta-ti => š-ket₉-ta-ti
š-kat-te-en => š-ket₉-te-en
š-kat-tén => š-ket₉-tén
š-ki-id-du => š-ke-ed-du
š-ki-ir => š-ke-er
š-ki-it => š-ke-et
š-ki-it-tin => š-ke-et-tén
š-ki-iz-zi => š-ke-ez-zi
š-ki-u-wa-an => š-ke-u-wa-an
š-ki-wa-an => š-ke-wa-an
š-kir => š-ker
š-kit₉-ta => š-ket₉-ta
š-kit₉-ta-at => š-ket₉-ta-at
š-kit₉-ta-ri => š-ket₉-ta-ri
š-kit₉-ta-ru => š-ket₉-ta-ru
š-kit₉-ta-ti => š-ket₉-ta-ti
š-kit₉-te-en => š-ket₉-te-en
š-kit₉-tén => š-ket₉-tén
ša-aḥ-ta-ri-le-e => ša-aḥ-ta-ri-li-e
šanezzi-
šar-me-e-eš => šar-mì-e-eš
še-ḥi-el-li => še-ḥe-el-li
šer
ši-pa-an-za-ki-iz-zi => ši-pa-an-za-ke-ez-zi
ŠIR-ki-iz-zi => ŠIR-ke-ez-zi

šu-un-ni-it => šu-un-ni-et
 šu-up-pí-iš-re-e-eš => šu-up-pí-iš-ri-e-eš
 ta-a-ḫi-ia-le-e-eš => ta-a-ḫi-ia-li-e-eš
 ta-a-ḫi-ia-le-eš => ta-a-ḫi-ia-li-eš
 ta-a-pé-še-na-an => ta-a-pí-še-na-an
 ta-an pé-da-aš => ta-an_pé-da-aš
 ta-as => ta-aš
 ta-ḫé-ia-le-eš => ta-ḫé-ia-li-eš
 ta-ḫi-ia-le-eš => ta-ḫi-ia-li-eš
 ta-ki-tum => ta-ki-tu₄
 ta-pé-ša-n => ta-pí-ša-n
 tak => ták
 ták-le-e-ez-zi => ták-li-e-ez-zi
 tar-aḫ- => tar-uḫ-
 tar-ku-um-mi-ia-iz-zi => tar-ku-um-mi-ia-ez-zi
 tar-kum-mi-ia-iz-zi => tar-kum-mi-ia-ez-zi
 teḫḫun, teḫḫi
 te-eš-šu-ub => te-eš-šu-up
 te-eš-šu-ub-bi => te-eš-šu-up-pí
 Te-li-pí-nu => te-li-pí-nu
 tuḫ => túḫ
 túḫ => túḫ
 túḫ-uḫ-ša => túḫ^{uh}-ša
 túḫuḫ => túḫ^{uh}
 ú-da-ir => ú-da-er
 ú-iz-zi => ú-ez-zi
 u-me-ni-eš => u-me-né-eš
 u-un-ni-ir => u-un-né-er
 u-un-ni-ir => u-un-né-er
 u-ur-ke-eš => u-ur-ki-eš
 um-me-it => um-me-et
^{URU}ḫa-la-ab => ^{URU}ḫa-la-ap
^{URU}ta-a-ḫar-p => ^{URU}ta-a-ḫur-p
^{URU}ta-ḫar-p => ^{URU}ta-ḫur-p
 úw => up
 wa-al-ḫi-ia-le-e-eš => wa-al-ḫi-ia-li-e-eš
 wa-al-ḫi-ia-le-eš => wa-al-ḫi-ia-li-eš
 ya- => ia-
 zí-ia-ri => zé-ia-ri
 zi-ik-ki-ez-zi => zi-ik-ke-ez-zi
 zi-ik-ki-iz-zi => zi-ik-ke-ez-zi
 zi-ik-ki-zi => zi-ik-ke-zi
 zinna-
 zi-in-ni-ez-zi => zi-in-né-ez-zi
 zi-in-ni-iz-zi => zi-in-né-ez-zi
 zi-in-ni-zi => zi-in-né-zi
 zi-in-tu-ḫe-e-eš => zi-in-tu-ḫi-e-eš
 zi-in-tu-ḫe-eš => zi-in-tu-ḫi-eš
 zi-ner => zi-nàr
 zi-nir => zi-nàr

zi-ri-ia-al-li => ze-ri-ia-al-li