

Sumerian hyphenation: an attempt at codification

Transliterated Sumerian is written with the syllable signs making up individual words hyphenated together. Decisions about the hyphenation of Sumerian words thus relate to the concept of a ‘word’ in Sumerian (see J.A. Black, ‘Sumerian lexical categories’, in *Zeitschrift für Assyriologie* 92 (2002), 60–77). In Sumerian studies, uncertainty about word boundary has led to maximalist (e.g. lugal-ki-en-gi-ki-uri-ke₄, nitaḥ-kalag-ga, me-sikil-bi) and minimalist (e.g. lugal ki-en-gi ki-uri-ke₄, nitaḥ kalag-ga, me sikil-bi) uses of hyphens in transliteration. There is so far no standard in the field. Accordingly ETCSL has established and implemented the procedures described below. This is an interim statement (29th February 2004), and inevitably will be subject to revision.

The usual criteria for identification of word boundary do not exist, e.g. pause, stress. Also words are frequently written on clay tablets with no space between them. Clitics are frequent in Sumerian, and it is not always clear if something is a word in its own right or only a clitic. Clitics are attached to words or clauses, and in all probability carried no stress of their own. A whole sequence of clitics can be attached to e.g. the last word of a clause. They may often have been phonologically integrated with their hosts; the writing system certainly writes most of such clitics as if they were in sandhi with the previous word (or clitic). Finally, in Sumerian a large number of compound nouns exists.

Sometimes the judgment whether something is a lexicalised compound or an *ad hoc* phrase is based on frequency of occurrence. Sometimes the judgment is informally based on (Akkadian, European language) translation value. Ideally the judgment would be based on syntactic grounds, e.g. can the compound itself be modified by an adjective? At all events, it must be pragmatic, since language is irregular, but decisions are needed. Sometimes this will help with lemmatisation, by disambiguating homographs. Sometimes the decision is taken on grounds of convenience. Multiword verbs exist in which the lemma is composed of two words (which, in a finite verbal form, are separated by proclitics).

- Hyphenation is used to join all written syllables of a single word together, and:
- to join clitics or affixes to the preceding base (or, with verbal proclitics, the following base)
- to join all the elements of proper names e.g. kar-zid-da ‘Right-hand Quay’, kur-^{gīs}erin-kud ‘Mountain Where Cedars Are Cut’, e₂-barag-dur₂-ḡar-ra ‘House Where Daises Are Established’, amaš-kug-ga ‘Holy Sheepfold’, where it is clear that they are proper names and not e.g. metaphorical descriptions
- to join ‘compound’ nouns (i.e. believed by us to be lexicalised), in various categories:
 - (i) co-ordinate compounds, e.g. maš₂-anše ‘wild animals’, gud-ab₂ ‘cattle’, gud-udu ‘cattle and sheep’, u₂-šim ‘vegetation’; NOT an ki ‘heaven and earth’

(ii) right-headed compounds of two nouns (very few in number), e.g. an-edin ‘high (part of the) desert’, saġ-men ‘(a type of) crown’

(iii) left-headed compounds (the largest category):

a) where a noun is followed and modified by:

- another noun, e.g. dumu-nitaġ ‘male offspring’, lu₂-erim₂ ‘enemy’, kur-ur₂ ‘heart of the mountains’, an-ur₂ ‘horizon’, an-pa ‘zenith’, an-šaġ₄ ‘interior of heaven’, mul-an ‘constellation’, eš₂-gana₂ ‘(field-)measuring line’, bad₃-si ‘crenellation’, i₃-udu ‘sheep fat’, ġiš-nu₂(-d) ‘bed’, gu₂-erim₂ ‘enemy’; dumu-gir₁₅ ‘free man’, ur-gir₁₅ ‘dog’, eme-gir₁₅ ‘Sumerian (language)’.

- another noun in the genitive, e.g. gi-dub-ba ‘stylus’ (lit. reed of tablet), lu₂-inim-ma ‘witness’, e₂-nu₂-da / a₂-nu₂-da ‘bedroom’, maš₂-ud₅-da ‘buck goat’.

- an adjective, e.g. eš₃-maġ ‘major shrine’, sipad-tur ‘junior shepherd’, lu₂-tur ‘child’, saġ-giġ₂ ‘black-headed (person)’ (= human being), e₂-gal ‘palace’, kug-babbar ‘silver’, inim-sig ‘slander’, šaġ₄-gal ‘hunger’ (cf. šaġ₄ gal ‘big heart’), kiġ₂-sig ‘morning meal’ (cf. kiġ₂-nim ‘late afternoon meal’); NOT sukkal maġ ‘chief minister’, NOT lu₂ zid ‘just person’.

b) participles in -a, e.g. kaš-de₂-a ‘wedding offering, party’, di-til-la ‘completed judgment’; but note: a₂ aġ₂-ġa₂ ‘command’ (see below).

c) nominals derived from finite verbal forms, e.g. ki-in-dar ‘crevice’, ki-in-du ‘earth’.

Included here are compounds where the second element is a verbal noun but in which the first element is the head.

There are also a number of compounds representing agents (professions) and instruments (tools), e.g. dub-sar ‘scribe’, igi-nu-du₈ ‘(type of) slave’, gaba-ri ‘rival’, kišib-la₂ ‘seal-bearer (also: hand, paw)’, gana₂-ur₃ ‘harrow’, which may be thought of as expressions in which the head (lu₂ or ġiš) is missing, for instance (lu₂-)dub-sar ‘(person) writing tablets = scribe’.

There is a group of (mostly) places and buildings with ki- (+N, or +V, or +Adj): ki-tuš ‘seat’, ki-gub ‘station’, ki-ġar ‘socle’, ki-gal ‘underworld; socle, shrine’, ki-bal ‘rebel land’, ki-erim₂ ‘enemy territory’, ki-gul-la ‘destitute person’, ki-sikil ‘maiden’, ki-še-er ‘profit’, ki-lul-la ‘murder’, ki-nu₂ ‘sleeping place’, ki-a-naġ ‘offering place’, ki-us₂ ‘foundations’, ki-šu-peš₍₅₎ ‘cult place’, ki-buru₁₄ ‘harvest area’, (lu₂) ki-inim-ma ‘witness’, ki-maġ ‘tomb’, ki-ulutim₂ ‘matrix’;

or with e₂- (+N, or +Adj): e₂-gal ‘palace’, e₂-maš₂ ‘animal stall’, e₂-dub-ba-a ‘scribal school’, e₂-gi₄-a ‘bride’, e₂-nu₂-da (a₂-nu₂-da) ‘bedroom’.

There is a group of people with lu_2- (+N, or +V, or +Adj): lu_2 -kur₂ ‘enemy’, lu_2 -kar(-ra) ‘fugitive’, lu_2 -lul(-la) ‘liar’, lu_2 -saḡ ‘guard’, lu_2 -tur ‘child’, lu_2 -erim₂ ‘wicked, enemy’, lu_2 -sa-gaz ‘robber’, NOT lu_2 zid ‘just person’.

Also with gu_2- : gu_2 -en ‘lord’, gu_2 -bar ‘nape of the neck’, gu_2 -erim₂ ‘wicked, enemy’, gu_2 -gal ‘pre-eminent’, gu_2 -gid₂, gu_2 -ḡiri₃ ‘breach’, gu_2 -kiḡ₂, gu_2 -nida, gu_2 -saḡ ‘?’.

There is a group of compounds with -saḡ: ud₅-saḡ ‘lead goat’, ab₂-saḡ ‘lead cow’, kun-saḡ ‘cultic building, stairway’, dumu-saḡ ‘firstborn son’, e₂-saḡ ‘granary’, dub-saḡ ‘superior, better’, us₂-saḡ ‘(?)’, pu₂-saḡ ‘hollow, pit’. But NOT kaš saḡ, i₃ saḡ, ga saḡ ‘best beer, butter, milk’, u₃-tud saḡ ‘principal offspring’.

There are also some hypercompounds of the above types, e.g. a₂-šu-ḡiri₃ ‘limbs’, an-ub-da ‘regions of the heavens’, lu_2 -ki-sikil ‘maiden’, lu_2 -eš₂-gid₂-da ‘surveyor’, gi-izi-la₂ ‘torch’, dub-sar-maḥ ‘senior scribe’, dub-sar-tur ‘junior scribe’, gaba-šu-ḡar ‘rival’, niḡ₂-ur₂-limmu₂ ‘quadrupeds’, šag₄-gada-la₂ ‘a type of priest’, eme₂-ga-la₂ ‘wetnurse’, er₂-šag₄-ne-ša₄ ~ ‘lamentation prayer’, NOT e₂ niḡ₂-gur₁₁-ra ‘treasure house’.

(iv) for nouns derived with nam- or nu-: nam-lugal ‘kinship’, nam-dub-sar ‘scribal skills’, nu-kiri₆ ‘gardener’, nu-banda₃ ‘overseer’.

(v) for nouns formed from finite verbal forms: u₃-na-dug₄ ‘letter’, ga-an-tuš ‘tenant’.

(vi) There is a group of compounded adjectival forms (a list of at least 37 which can be considered to be lexicalised, for syntactic or semantic reasons), compounded with the verbs -zu ‘knowing’, -tuku ‘having’, -du/di ‘doing’ and -ḡal₂ ‘being/having’. These are hyphenated. They will be lemmatised as single lemmata.

kug-zu	‘wise’
gal-zu	‘skilful’ (cf. gal-zu niḡ ₂ -nam-ma ‘wise in everything’)
gal-an-zu	‘clever’
usu-tuku	‘powerful’
a ₂ -tuku	‘powerful’
ni ₂ -tuku	‘reverent’
ḡiš-tuku	‘attentive’ (cf. sipad ḡiš-tuku-ni-ir ‘to his attentive shepherd’)
lipiš-tuku	‘furious’
gu ₂ -tuku	‘perfect, splendid’
ḥul-du	‘wicked(-acting)’
erim ₂ -du	‘hostile(-acting)’
zid-du	‘righteous(-acting)’
maḥ-di	‘illustrious’
gal-di	‘important, impressive’

a ₂ -bad-ḡal ₂	‘protective’;
a ₂ -ḡal ₂	‘strong’, also <i>neg.</i> a ₂ -nu-ḡal ₂ (cf. a ₂ -ḡal ₂ kalam-ma-me-en ‘you are the strong one of the Land’)
a ₂ -nun-ḡal ₂	‘impressively strong’ (cf. a ₂ -nun-ḡal ₂ e ₂ -kur-ra ‘most powerful in E-kur’)
an-ta-ḡal ₂	‘exalted’
dug ₃ -ḡal ₂	‘well-disposed (?)’
gaba-ḡal ₂	‘forceful’
(gu ₂ -)erim ₂ -ḡal ₂	‘wicked’ (freq. also in headless use, parallel to (gu ₂ -)erim ₂ ‘wicked, enemy’)
ḥul-ḡal ₂	‘evil’, <i>neg.</i> ḥul-nu-ḡal ₂ (varies with ḥul ‘evil’)
igi-ḡal ₂	‘wise’, also <i>neg.</i> igi-nu-ḡal ₂ (cf. also igi-ḡal ₂ -tuku; also used participially: igi ḡal ₂ ‘taking account of’; and cf. the noun igi-ḡal ₂ ‘insight, wisdom’)
inim-ḡal ₂	‘eloquent’
kug-ḡal ₂	‘holy’ (but note kug-ḡal ₂ = gu ₂ -gal/ḡal ₂ ‘canal inspector’)(varies with kug ‘holy’)
la-la-ḡal ₂	‘charming’
me-te-ḡal ₂	‘seemly’
mud-ḡal ₂	‘creative (?)’
ne ₃ -ḡal ₂	‘mighty’ (written with ne ₃ (PIRIĜ), also ne-ḡal ₂)
nir-ḡal ₂	‘lordly, authoritative’, also <i>neg.</i> nir-nu-ḡal ₂ (cf. dumu nir-ḡal ₂ an-na-ke ₄ ‘lordly son of An’)
ni ₂ -ḡal ₂	‘awe-inspiring’
peš ₁₀ -ḡal ₂	‘mighty (?)’ (also written peš-ḡal ₂)(cf. peš ₁₀ -ḡal ₂ e ₂ -kur-ra ‘mighty one of the E-kur’)
saḡ-ḡal ₂	‘safe, strong’ (in the expression: ki saḡ-ḡal ₂)
saḡ-ki-ḡal ₂	‘obstinate’
saḡ-kug-ḡal ₂	‘proud’ (var. saḡ-gu ₂ -ḡal ₂ /gal; also saḡ-gu ₂ -tuku)
su ₃ -ud-ḡal ₂	‘extended, lengthy (of time); patient-hearted’ (varies with su ₃ -ud ‘long’)
zi-ḡal ₂	‘living, alive’ (cf. id ₂ zi-ḡal ₂ kalam-ma ‘life-bringing canal of the Land’)

Expressions with niḡ₂:

The following are assumed to be lexicalised compounds with niḡ₂, on the basis mainly of frequency:

niḡ₂-a-erim₂ by analogy with niḡ₂-a₂-zig₃ and niḡ₂-erim₂ ‘violence’

niḡ₂-a₂-zig₃ ‘violence’

niḡ₂-du₇ ‘that which is fitting, proper, necessary’ (usually co-occurs with vb. pa ed₂ (ak))(n.b. same writing as niḡ₂-ul, see below)

niḡ₂-gen₆(-na) ‘justice’

niḡ₂-gig(-ga) ‘taboo’

niḡ₂-si-sa₂ ‘rectitude’, ḡustice’

niḡ₂-erim₂ ‘wickedness’

niḡ₂-hul ‘evil (abstract noun)’

niḡ₂-me-ḡar ‘(ominous) silence’

niḡ₂-ba ‘gift’

niġ₂-de₂-a ‘wedding gift’
 niġ₂-gur₁₁ ‘treasure, riches’
 niġ₂-kud ‘revenue, income, produce’
 niġ₂-mi₂-us₂-sa₂ ‘wedding gift’
 niġ₂-šam₂-ma ‘purchase’,
 niġ₂-tuku ‘rich’
 niġ₂-u₂-rum ‘acquisitions, property’
 niġ₂-ul ‘distant time in the past or future’ (n.b. same writing as niġ₂-du₇)
 niġ₂-nam(-ma) ‘everything’
 niġ₂-ar₃-ra ‘a type of flour’
 niġ₂-gu₇ ‘food’
 niġ₂-na ‘incense burner’
 niġ₂-silaġ ‘dough’
 niġ₂-TUG₂, see 2.1.7:1168 anše niġ₂-TUG₂-bi il₂-a-da [*hapax*; probably refers to an
 object (‘pack-saddles’)]
 niġ₂-bun₂-na ‘turtle’
 niġ₂-ur₂-limmu₂ / niġ₂-ur₂-4 ‘quadruped’
 niġ₂-zi-ġal₂ ‘living creatures; wild animals’

various textiles:

^{tu}g₂niġ₂-barag ‘quilt, coverlet’
^{tu}g₂niġ₂-dara₂ ‘loincloth’
^{tu}g₂niġ₂-la₂ ‘binding, bandage’
^{tu}g₂niġ₂-lam₂ ‘(a ceremonial garment)’.

N.b. some hypercompounds of the above exist:
 nam-niġ₂-erim₂, nam-niġ₂-si-sa₂, both in 1.3.1

Other expressions with niġ₂ are not hyphenated.

On the other hand, hyphenation is *not* used for:

- **any form (finite, non-finite or nominal)** of a recognised compound (multi-word) verb not covered by the above. E.g. a₂ aġ₂-ġa₂, inim bal-e, inim sig, su-lum mar; di kud ‘making judgements’ (but di-kud ‘judge’) (see below) etc.
- *ad hoc* verbal phrases or nominal phrases e.g. lu₂ tur ‘small person’ (cf. lu₂-tur ‘child’), ħar-ra-an zu ‘knowing the roads’, sun₄ la₂ ‘wearing a beard’ cf. sun₄ za-gin₃ la₂ ‘wearing a dark beard’
- demonstratives such as ne(-en), re(-en) (contrast -bi and -e)
- the various expressions with ud ‘day’:
 ud zal-la/zal-le-da ‘(during) the course of the day’

ud te-en-še₃/ta ‘to/from evening’

ud šu₂-uš ‘sunset’

ud sa₉(-a) ‘midday’ cf. ġi₆ sa₉(-a) ‘midnight’, ġi₆ u₃-na(-k) ‘dead of night’

and with ud = ‘time’:

ud da-ri₂ ‘(for) all time’, ud ul-li₂-a-še₃/ta ‘(in/from) those (ancient) days’ ud re-a ‘(in) those (distant) days’.

The syntactic status of some of these is a little uncertain. Note, however, ud-sakar ‘(new) moon’.

• expressions with na-me ‘any; (with negative verb) no’:

niġ₂ na-me ‘anything’

lu₂ na-me ‘anyone’

diġir na-me ‘any god’

ud na-me ‘any time, ever’

ki na-me ‘anywhere’.

There will be a relatively small list of agreed exceptions not covered by the above categories, e.g. ama-ar-gi₄ ‘freedom’ (where the second element is a verbal noun), giri₁₇-zal ‘joy’ (where the second element may or may not be a verbal noun).

Note also:

nam-tar (n.) ‘fate’, but nam tar (vb. phrase) ‘who determines destinies’

šu-luḥ (n.) ‘a specific ritual’, but šu luḥ (vb. phrase) ‘who washes the hands; washing the hands’

di-kud (n.) ‘judge’, but di kud (vb. phrase) ‘who judges; act of judging’

ki-ġar (n.) ‘socle; substitute’, but ki ġar ‘who founds’

igi-ġal₂ ‘wisdom’, ‘wise’, but igi ġal₂ (vb. phrase) ‘who attends to, takes account of’

ġiš-ḥur ‘plan, design’ (ġiš-ḥur kal-kal ‘precious plans’; cf. ġiš...ḥur ‘to design’)

There are some problem cases to watch out for, e.g. does the combination noun + unaffixed non-finite verbal form derive from a multi-word (‘compound’) verb or not?

Extra areas

Note the following:

šita₂ saġ 50 ‘50-headed mace’, gud si dili ‘one-horned bull’ (but saġ-dili ‘bachelor’), ab₂ igi gun₂ ‘bright-eyed cows’, niġ₂-ur₂-limmu₂ ‘four-legged things’, uġ₃ saġ-gig₂ ‘black-headed people’

nam gal ki aġ₂-da ‘from great love’, nam gal ki aġ₂ ^ddumu-zid-da-ke₄ ‘from great love of Dumuzid’, nam gal ḥul₂-la-da ‘in great joy’, nam gal₄-la-na-še₃ ‘because of her genitals’.