



Sumerian language

- [Context](#)
- [Phonemes](#)
- [Morphemes](#)
- [Word classes](#)
 - [Nouns](#)
 - [Verbs](#)
- [Noun phrases](#)
 - [Non-core adverbial case markers](#)
 - [Core adverbial case markers](#)
 - [Adnominal case marker](#)
- [Clauses](#)
 - [Clause structure](#)
 - [Clause type](#)
- [Abbreviations](#)

Context

Sumerian is a long-extinct language documented throughout the ancient Middle East, in particular in the south of modern Iraq. It is arguably the first language for which we have written evidence, the rival candidate being ancient Egyptian. This evidence is spread over more than 3,000 years, the first sources dating to the late fourth millennium BCE and the last to the first century AD. When Sumerian ceased to be spoken is difficult to determine; according to some estimates this took place during the early second millennium BCE. Afterwards it was an élite language, used only in royal, ritual and scholarly contexts, the first two of which often overlapped.

Sumerian is a language isolate, that is no languages related to it have so far been convincingly identified, although many of its grammatical features are attested in other living languages outside of the Indo-European family to which English belongs.

The other language for which we have extensive written evidence in the ancient Middle East is Akkadian. Being a Semitic language with modern counterparts, Akkadian is much better understood than Sumerian. Part of the evidence for both languages consists of lexical lists, some indicating how cuneiform signs were pronounced and others giving Akkadian equivalents to Sumerian words. Consequently our reconstruction of the sounds and meanings of different cuneiform signs has been much influenced by our understanding of Akkadian. The word Sumerian is itself an anglicisation of an Akkadian term, *Šumeru*, the language's users referring to it instead as *Emegir*, possibly meaning 'native tongue'.

Inevitably uncertainties remain in our reconstruction of Sumerian. A further problem in describing the language is that it varied across both space and time. The following account aims to be no more than an introduction to some of Sumerian's basic grammatical features, following the conventional hierarchy of its phonemes (smallest sound units), morphemes (smallest grammatical units), words, phrases (groups of related words without a finite verb)

and clauses (groups of related words with a finite verb).

Phonemes

Sumerian is thought to have had eight vowel sounds: short and long *a*, *e*, *i*, and *u*. Vowel length is, however, not indicated in transliteration, that is the sign-by-sign representation of the cuneiform. (Throughout this introduction, italics are used to refer to Sumerian sounds, and bold to refer to signs.)

Fifteen consonants are usually recognised in transliterating the language: *b*, *d*, *g*, *ĝ* (as in *siĝ*), *ḫ* (as in *loḫ*), *k*, *l*, *m*, *n*, *p*, *r*, *s*, *š* (as in *šip*), *t* and *z*.

Two adjacent vowels typically contract. For various reasons, including instances in which such contraction does not occur, Sumerian has been argued to have further, weak, consonants, in particular *h*, *y* and ' (a glottal stop). These are again not specified in transliteration.

In addition, an early stage of the language is thought to have had a further consonant whose identity remains uncertain. In the period of the corpus this consonant appears to have merged in some contexts with *d* and in others with *r*, or sometimes simply to have been dropped.

The other consonants found in the corpus (and in the sign list) occur in non-Sumerian names. All of these alphabetic representations need to be regarded as approximations.

Morphemes

The term morpheme is used to refer to the smallest grammatical units in a language. Most linguists distinguish between bases and affixes, an English example being *runs* in which *run* is a base and *s* is an affix. As this example indicates a base morpheme can correspond to the higher grammatical category of the word.

Many linguists also recognise a class of morphemes, referred to as clitics, which are indeterminate on a continuum between affixes and words, English examples being *'m* in *I'm* and *'s* in *the man's dog*. The term bound morpheme is then used to include both affixes and clitics. One criterion for identifying a clitic is that it is indifferent to the class of the word to which it attaches, as in, for example, *the man's dog*, *the man who was running's dog* and *the man who was shouting loudly's dog*. A further characteristic of *'s* is that it is always phrase-final, that is it always occurs at the end of a noun phrase.

Unsurprisingly the grammatical function of an affix in one language can be performed by a clitic or word in another, some of the functions of the English preposition *to*, for example, being performed by a dative affix in Latin. In Sumerian the functional equivalents to prepositions (and postpositions in some other languages) are referred to as case markers. Like English *'s* they are always phrase-final and are indifferent to the class of the word to which they are attached. They are thus also classifiable as clitics, or more precisely as enclitics, indicating that they lean backwards onto their host, the word to which they are phonologically bound.

The plural of certain nouns in Sumerian is indicated by a morpheme, called a plural marker, which also behaves in a similar way. The principal other enclitics in the language are the determiners, morphemes which modify a noun and typically also have a corresponding pronominal form (an English example being *that* which is a determiner in *that book* but a pronoun in *give me that*). In addition, as in English the verb **me** 'to be' occurs both as a word and in enclitic forms.

In the transliteration conventions followed by the corpus most enclitics are preceded by a hyphen linking them to their host. The exceptions are the demonstrative and indefinite determiners, although **bi**, which functions both as a demonstrative ('this') and as a possessive ('its, their'), is preceded by a hyphen in both functions.

Because so many bound morphemes are enclitics, rather than affixes as in a language like

Latin, morphological change in most Sumerian word classes is limited. Many word classes are morphologically invariant, and for nouns and adjectives variation is restricted to base-reduplication. Verbs are the striking exception: these can occur in highly complex affixed forms which also feature base-reduplication.

Word classes

The minor word classes in Sumerian are numbers, conjunctions, interjections, adverbs, adjectives and circumpositions (functional equivalents to such complex English prepositions as *behind* and *in front of*), as well as related sets of pronouns (personal, demonstrative, indefinite, interrogative and reflexive) and determiners (possessive, demonstrative and an indefinite). Unlike English, Sumerian has no definite or indefinite article. The primary word classes are nouns and verbs.

Nouns

Sumerian nouns can be subcategorised into two classes on the basis of gender, the distinction being between human nouns (referring to people and deities) and non-human nouns (referring to animals and inanimates). This is a semantically based distinction to which there are some socially conditioned exceptions, **saġ** 'slave', for example, sometimes being construed as a non-human noun. In addition, animals and inanimates can be personified in literary compositions and thus construed as human nouns.

This gender distinction is only morphologically apparent in most parts of the language's third person pronominal system (first and second person reference necessarily being solely human). It is also syntactically apparent in restrictions on how the case markers and the plural marker are used.

Only a noun phrase whose head (grammatically dominant word) is a human noun can contain a plural marker, non-human nouns consequently being indeterminate in terms of number. However, this plural marker appears to have an individualising force and if reference is to a group of people or deities it is omitted, that is the noun is construed as if it were non-human. This is particularly the case for nouns with a group meaning, such as **erin₂** 'troop'. Similarly non-human pronominal morphemes can be used to refer to groups of people or deities. (As far as we can judge some Sumerian signs were pronounced in the same way. The subscript numeral in the earlier example is a modern convention to associate a sound sequence with a particular sign. So, for example, **du**, a form of the verb 'to go', was written with the sign referred to as DU, while **du₃** 'to erect' was written with the sign referred to as KAK.)

While the plural marker is restricted to human nouns, base-reduplication occurs in both classes of noun, appearing to express a form of totality (*all*).

Along with verbs, nouns are the principal open word class, that is the class of word most likely to form new members. New nouns are primarily formed by compounding, such as **dub-sar** 'scribe' (literally '(someone) tablet writing').

Verbs

A particular characteristic of verbs is their ability to distinguish tense and/or aspect, that is to locate an action in time or to express its quality in some way. They can be subcategorised in terms of their semantics, syntactic requirements and regularity. In addition, they have different finite and non-finite forms.

Reconstructing the tense and/or aspect system of Sumerian verbs is difficult. Most scholars agree that the primary distinction is between a completed and an uncompleted action, but differ as to whether this reflects a distinction of tense (past versus non-past) or of aspect (completive versus incompletive). Many scholars have therefore adopted instead two terms used by Akkadian grammarians, *ḥamtu* ('quick') and *marû* ('fat') respectively. For the sake of convenience, in this introduction the primary distinction is assumed to be aspectual.

In intransitive finite verbal forms, that is those without a direct object, completive aspect is unmarked while incompletive aspect is indicated by the suffix *ed* immediately after the base. In languages like Latin, a person-number-gender (PNG) suffix is used to express in pronominal form the subject of the verb (as in *am+o* 'I love', *am+as* 'you love' etc.). The same applies to Sumerian intransitive verbs, the PNG suffix following *ed* in the case of incompletive forms.

Sumerian extends this principle to also marking the direct object in transitive verbal forms. In such verbs it is the distribution and form of these subject and direct object morphemes (consisting of prefixes to the base, suffixes after the base and some circumfixes on either side of the base) that serve to express the aspectual distinction.

The difference between finite and non-finite verbal forms is partly morphological, the latter having far fewer morphemes than the former. Among the morphemes excluded from Sumerian non-finite forms are PNG affixes, the aspectual distinction being expressed instead with an aspect suffix. Non-finite forms are more nuanced and have stronger temporal connotations than finite forms, distinguishing between completive (marked with *a* and having past reference), habitual (zero-marked (\emptyset) and having present reference), and incompletive (marked with *ed(a)* and having non-past reference). The only other affix non-finite forms can have is a prefix *nu* expressing negation. The non-finite forms function as verbal adjectives (participles) and nouns (gerunds), and in non-finite relative and adverbial clauses (for example, of purpose and time).

In terms of semantics Sumerian verbs can be divided into two classes, stative verbs that refer to persisting states or situations, such as **pel** 'to be defiled', and dynamic verbs which refer to an action or process, such as **ra** 'to beat'. In the same way as English stative verbs are excluded from progressive aspect (*I am knowing Sumerian*, for example, being an unacceptable clause), Sumerian stative verbs are excluded from incompletive aspect. Consequently only context indicates whether they have past or non-past reference. In non-finite verbal forms the distribution of stative verbs appears to be syntactic, intransitive verbs occurring in completive aspect and transitive ones in habitual aspect. All verbs can have reduplicated bases: in stative instances this expresses intensity, and in dynamic instances iterativity.

However, this semantic distinction is less of one of classes and more one of usages, some verbs occurring in both classes. For example, **pel** can also be used dynamically to mean 'to make something be defiled', that is 'to defile' something, in which meaning it can occur in incompletive aspect.

In terms of syntactic requirements Sumerian verbs can be divided into four principal classes: intransitive verbs which require no object, such as **uš₂** 'to die'; extended intransitive verbs which require a non-direct object, such as **kur₉** 'to enter' into a place; transitive verbs which require only a direct object, such as **du₃** 'to erect' something; and, finally, extended transitive verbs which require both a direct and a non-direct object, such as **ġar** 'to place' something on something.

Again, however, this is more a distinction in usages rather than classes, some verbs occurring both intransitively and transitively. For example, **uš₂** can be used transitively with the meaning 'to kill' ('to make someone die').

This ability of a verb to determine its syntactic environment is sometimes referred to as its valency, and the non-direct object as a complement, as distinct from an adjunct, a phrase that can be easily deleted from a clause (*in London* being a complement in *he lives in London* but an adjunct in *he bought a house in London*).

A further syntactic class has only one member, **me** 'to be'. While the English verb *to be* has

both a copular (linking) and a locational function, **me** has only a copular function, location being expressed instead by the verb **ĝal₂**. Sumerian **me** conjugates like an intransitive stative verb and consequently is never found in incomplete forms. It differs, however, in that it always requires what is referred to as a predicative complement, such as a noun or an adjective, which refers back to the subject of the verb (as in *he is the king* and *he is handsome*). The copular verb occurs both as a word and as an enclitic attached to its predicative complement. It also has various abbreviated forms, such as the third-person negative **nu**, which are arguably also enclitics.

In terms of regularity Sumerian verbs can be divided into four classes: regular verbs (the majority) which have the same base regardless of aspect and/or number (such as **ra** 'to beat'); suppletive verbs which have a different base depending on aspect and/or number (such as singular **tuš** 'to sit' but plural **durun**); reduplicating verbs which have a (partly) reduplicated base in incomplete aspect (such as complete and habitual **ĝar** 'to place' but incomplete **ĝa₂-ĝa₂**); and a small class of extending verbs which have a base extended with a consonant in incomplete aspect (such as complete and habitual **te** 'to approach' but incomplete **teĝ₃**).

Expansion of the class of verbs is primarily by multiword constructions in which a nominal element and a verb combine in a semantic unit. The nominal element is typically a noun, in particular a body-part, functioning as the verb's direct object. These multiword constructions can be divided into two types. In one the verb is **dug₄** 'to say' or **ak** 'to do' and the semantic load is carried by the nominal element, an example being **šu dug₄** 'to tend' (literally 'to express the hand', although the Sumerian word order is the reverse of the English). A wider range of verbs occurs in the other subclass and the semantic load is more evenly spread, an example being **ĝiš tag** 'to sacrifice' something (literally 'to touch wood' to something, again in the reverse order). The high incidence of multiword constructions in Sumerian means that it has many more non-direct objects than a language such as English.

Noun phrases

At the level of the noun phrase Sumerian is left-headed, that is the noun which is the head of the phrase occurs at its beginning. In outline the sequence is noun, modifier(s), determiner, plural marker and then case marker. However, a few adjectives, **kug** 'shining' and **gal** 'big', sometimes precede the noun. The plural marker only occurs in a phrase with a human noun as its head. And the indefinite and most demonstrative determiners do not occur with modifiers.

Like English 's the case markers are always phrase-final. And in the same way that 's is sometimes called a genitive, they are referred to with the same types of label as are used for case affixes in other languages. The case markers can be divided into three groups. All except one typically indicate the syntactic role which a phrase plays in relation to a verb in a clause, consequently being described as adverbial. Those adverbial case-markers that are functionally equivalent to English prepositions can be termed non-core and those that mark the subject and any direct object of the verb as core. The final case marker, the genitive, is adnominal only, that is it functions only to indicate a relationship between noun phrases.

Non-core adverbial case markers

The non-core adverbial case markers include the ablative (*ta* 'from'), allative (*še* 'to(wards)'), comitative (*da* 'with'), dative (*ra*, 'to/for'; restricted to phrases with a human noun as the head), directive (*e* 'in(to) contact with'; restricted to phrases with a non-human noun as the head), and locative (*a* 'in(to)'; again restricted to phrases with a non-human noun as the head). These case markers occur at the end of phrases that can be complements or adjuncts, depending on the valency of the verb in the clause. A similar but more nuanced set of morphemes is

incorporated in finite verbal forms.

Two further non-core adverbial case markers are only used to express adjuncts and have no equivalent morpheme in finite verbs, the adverbial (eš'in the manner of') and the similative (*gin* 'like').

Core adverbial case markers

The two final adverbial case markers have a more grammatical, core function. Most languages have a strategy for distinguishing the subject of a transitive verb from its direct object. In English this is done mainly by word order, although a case system still operates in pronouns (*he hates him*). This distinction made, different languages mark the subject of an intransitive verb in different ways. In English both subjects are marked in the same way (*he runs*). This is referred to as nominative-accusative alignment. However, in Sumerian noun phrases the subject of an intransitive verb is marked in the same way as the direct object of a transitive verb. This is referred to as ergative-absolutive alignment, the subject of a Sumerian transitive verb being marked by an ergative case marker *e* (morphologically the same as the directive from which it possibly derives), and the intransitive subject and transitive direct object being zero-marked with what is termed the absolutive case marker.

However, ergative-absolutive alignment applies only in noun phrases which have a noun as their head. Noun phrases with a personal pronoun as their head can be regarded as nominative-accusative in alignment because they have the same zero case marking regardless of the transitivity of the verb whose subject they are. Sumerian is thus one of many languages which have a syntactic split determined by the class of the word functioning as the head of the noun phrase.

Adnominal case marker

The genitive case marker, *ak* 'of', is adnominal only and consequently also has no equivalent morpheme in finite verbal forms. Typically a genitive noun phrase occurs embedded within another noun phrase ending with an adverbial case marker:

itid šu-numun-a-ka
itid šu-numun=ak=a
month Šu-numun=GEN=LOC
'in the month of Šu-numun'

As this example indicates there is not always a one-one correspondence in the writing system between bound morpheme and sign, the genitive *ak* here being written across two signs, and the locative *a* as part of one sign. (The abbreviations used in the morphemic analysis of the Sumerian are explained at the end of this introduction.)

However, in the same way as an English direct object can be shifted to the beginning of a clause and its original position marked by a pronoun (*snakes, I hate them*), so too can a Sumerian genitive noun phrase be front-shifted, a possessive determiner then marking its original position:

e₂-a ni₂ gal-bi
e=ak ni gal=bi=Ø
temple=GEN awesomeness great=3NHUM.POSS=ABS
'the great awesomeness of the temple'
(literally 'of the temple its great awesomeness')

In this example *a* is the reduced form of the genitive that occurs when it is not followed by a vowel.

Because the genitive is the only solely adnominal case marker its semantic field covers much more than possession, also being used, for example, to express a location:

bad₃ iri kug-ga-ka-ni
 bad iri kug=ak=ani=Ø
 wall city holy=GEN=3HUM.SG.POSS=ABS
 'her wall in the Holy City'

This example indicates another characteristic of the writing system, the *g* which precedes the genitive arguably having no phonological significance but simply being a graphic resumption of the preceding consonant.

Clauses

At the level of the clause Sumerian is right-headed, that is the verb which is the head of the clause occurs at its end, the typical sequence being subject, object and then verb. However, because a finite verbal form includes PNG affixes expressing in pronominal form the core functions of the subject and any direct object of a verb, a Sumerian clause can consist of only a finite verb.

Clause structure

In addition to the core PNG affixes, a finite verbal form can include further prefixes expressing a complement or an adjunct of a verb in pronominal form. These morphemes consist of a set of non-core PNG prefixes and a set of 'case' prefixes that are related to the non-core case markers. Just as a case marker follows a noun in a noun phrase, so the 'case' prefixes are postpositional to the PNG prefix. More than one 'case' prefix can occur in a finite verbal form, but only the first can be preceded by a PNG prefix. There are, however, some restrictions on the occurrence of the non-core PNG prefix:

ba-ra-an-e₃-en
 ba-ra-n-e-en
 MID-ABL-3HUM.SG.S-go out-1HUM.SG.DO
 he drove me out of it

In this example a further prefix, *ba* (whose functions are described in more detail below), excludes the presence of a non-human PNG prefix before the ablative *ra* (its form after a vowel, being *ta* after a consonant).

A clause of this type can be expanded to include noun phrases:

agrig-ge e₂-ta ba-ra-an-e₃-en
 agrig=e e=ta ba-ra-n-e-en
 steward=ERG house=ABL MID-ABL-3HUM.SG.S-go out-1HUM.SG.DO
 'the steward drove me out of the house'

In an expanded clause the core PNG affixes are always retained in the verb. The non-core PNG prefix and 'case' prefixes are, however, sometimes omitted.

The further prefixes possible in a finite verbal form include middle *ba*, ventive *m(u)* and a vowel-initial prefix. In outline a finite verb can consequently have the following structure: vowel, ventive, middle, non-core PNG, 'case(s)', core PNG, (reduplicated) base, aspect, core PNG.

The middle prefix *ba* is a single morpheme which is ambiguous with a sequence of two morphemes from which it arguably derives, that is the non-core non-human PNG prefix *b* ('it, them') and the dative prefix *a* ('to'). If the argument is correct, it helps to explain the absence of the non-core non-human PNG prefix in the previous examples, the etymology of the middle prefix excluding repetition of one of the prefixes from which it derives. Middle *ba* can, however, be followed by a non-core human PNG prefix.

The middle prefix is restricted to completive aspect. Its range of functions, some of which

remain unclear, is broader than the conventional term middle suggests.

With stative verbs it has an inchoative function, that is it expresses the coming into existence of a state: **ba-an-tuku** 'he married her' (literally 'he came to have her'). In this and the following examples both the third person transitive direct object and the third person intransitive subject are zero-marked in the verb, a further example of ergative-absolutive alignment in Sumerian.

With dynamic verbs *ba* is used in particular when the subject of the verb is affected by the action of the verb. Consequently it functions as an invariant reflexive indicating that the endpoint of the verb's action is the same entity as the subject of the verb: **ba-an-zuḥ** 'he stole it (for himself)'. By extension it was also used to form the equivalent to the English middle voice in which the subject of the verb no longer has an agentive role but continues to be affected by the action of the verb: **ud ba-bur₂** 'the weather improved'. Expressing such agentless or spontaneous events is often referred to as an anticausative function (although in the external world this type of event obviously does have a cause). The functional range of *ba* was extended still further to include forming the equivalent to the English passive, that is to non-spontaneous events which consequently do have an implied agent: **ba-ḥul** 'it was destroyed'.

The ventive (or cislocative) prefix *m(u)* has a more restricted range of functions. It can be regarded as orienting a verb towards the speaker or narrator and can occur in both completive and incompletive aspects.

The vowel-initial prefixes are *i* and *a(l)*. The former has no semantic function but is used with dynamic verbs before two consonants, and before an otherwise unprefixated verbal base or core PNG prefix; *a(l)* performs a similar function with stative verbs. However, *a(l)* is also used with dynamic verbs, in which occurrences it has a semantic function. In completive aspect this includes expressing a statal passive: **al-du₃** 'it is built'. In incompletive aspect its functions possibly include expressing a habitual action:

šag ₄ -ga-ni	ab-ḥul ₂ -le-en ₆
šag=ani=Ø	a-b-ḥul-en
heart=3HUM.SG.POSS=ABS	STAT-3NHUM.DO-be happy-1HUM.SG.S
'I will (habitually) make his heart happy'	

However, this still leaves many incompletive instances of *a(l)* unaccounted for.

Clause type

Clauses can be analysed in terms of their status, the basic distinction being between main clauses which can stand on their own and subordinate clauses which can't, and also in terms of their contribution to what is termed discourse function, that is, for example, whether they make a statement or express a question.

Most Sumerian clauses simply make a statement and, like the English indicative, are zero-marked. However, the same applies to closed questions, that is ones requiring only a yes or no answer, the implication being that they were signalled with a change in intonation (compare *you're going out?*). Open questions are signalled by an interrogative pronoun (such as **a-na** 'what?') or adverb (such as **me-še₃** 'where?').

More complex types of discourse function, such as commands, prohibitions and wishes, are expressed primarily by morphological changes to the finite verb. In most cases a verb-initial prefix is added. However, for second-person positive commands the imperative is used in which what are prefixes in other verbal forms are instead suffixes (compare *dites-le-moi* 'tell it to me'). This is regarded as untypical behaviour for affixes and raises some doubt about where on the continuum between clitics and affixes these bound morphemes lie. Another characteristic of the imperative is that it deletes both the singular intransitive and transitive

subject. It can therefore be regarded as a further example of nominative-accusative alignment in Sumerian.

These verb-initial prefixes themselves lie on a different type of continuum, one between signalling a change in verbal mood and connecting clauses. In some contexts the prefix *ħu*, for example, has a clear modal function:

ħu-mu-na-ab-šum₂-mu

ħu-mu-n-a-b-šum-u

MOD-VENT-3HUM.SG-DAT-3NHUM.DO-give-3HUM.SG.S

'he should give it to him'

In other contexts it combines modality with clause-connection, forming a type of conditional, and subordinate, clause (compare the English conditional subjunctive *should he give it to him, he will suffer*). Other verb-initial prefixes have a more straightforward connective function, *u*, for example, indicating that the action expressed by its verb precedes the action expressed by the verb that follows, and thus being translatable with the English subordinating conjunction *after*.

Sumerian also has the three more conventional types of subordinate clause: relative clauses which modify a noun and thus occur within the noun phrase; nominal clauses which can function as, for example, the subject of a verb; and adjunct (or adverbial) clauses which are subordinate to a main clause and have, for example, a causal or a temporal function.

In English the first two of these types can be signalled by a subordinator, *that*, and the third by a subordinating conjunction, such as *before*. The functional equivalent in Sumerian to the subordinator is a verb-final suffix *a*. Analysis of adjunct clauses is, however, less straightforward. Sumerian has very few simple subordinating conjunctions. More often a complex construction is used which begins with a noun and ends with a case marker, the suffix *a* again being bound to the verb. For example, **eġer ...-ta**, literally 'from the back that something had happened', can be translated as *after*. A less literal analysis is that in such constructions the noun has been bleached of its lexical content and combines with the case marker to form a complex subordinating conjunction.

Abbreviations

1	first person
3	third person
ABL	ablative
ABS	absolute
DO	direct object
ERG	ergative
GEN	genitive
HUM	human in gender
LOC	locative
MID	middle
MOD	modal
NHUM	non-human in gender
POSS	possessive
S	subject

SG singular

STAT stative



© Copyright 2003, 2004, 2005,
2006 The ETCSL project, Faculty
of Oriental Studies, University of
Oxford
Updated 2005-03-29 by GC

