Title

Old Literary Tibetan scogs (CT sogs) "among others": Etymology, constructions, and idiomatisation

Permalink

Journal

Author

Bialek, Joanna

Publication Date

2022

DOI

Copyright Information

Peer reviewed

Himalayan Linguistics

Old Literary Tibetan scogs *(CT* sogs*) "among others": Etymology, constructions, and idiomatisation*

Joanna Bialek

Humboldt University, Berlin

ABSTRACT

The paper is the first attempt at reconstructing the word-family of *scogs* "among others" and detecting its manifold uses in Old Literary Tibetan (OLT) texts. The morpheme is traced back to a v4-stem of the verb √stsog (CT v1 *gsog*), lit. "to cause to assemble", itself derived from the verb root √tsog by means of the causative prefix *s-*. After discussing its probable cognates and demonstrating historical links between them, I examine ten types of construction attested in the OLT corpus which contain the morpheme: finite clause, participial adverbial clause (I & II), post-head relative clause, R-dislocation, relative clause extraposition, pre-head relative clause, off-subject nominalisation, and idiomatic phrase (I & II). The paper sketches the lexicalisation path taken by the morpheme from a finite verb to an idiomatic phrase.

KEYWORDS

Old Literary Tibetan, syntax, verb morphology, etymology

Old Literary Tibetan scogs *(CT* sogs*)* *"among others": Etymology, constructions, and idiomatisation*

Joanna Bialek
Humboldt University, Berlin

1 Introduction

Old Literary Tibetan (OLT) *scogs* (CT *sogs*) bridges distinct historical periods and varieties of Tibetic languages to the present day with its clearly recognisable semantics. Its modern meaning "et cetera, such as" (Gs: 1135c; see also CDTD: 8867) is attested as early as in the entry for the year 710/11 of the *Old Tibetan Annals* (OTA). And maybe it was exactly this seemingly transparent semantics that made the morpheme uninteresting to modern scholars so that no study has been devoted to it. However, when viewing from a closer perspective one discovers a wide variety of constructions, in which *scogs* is involved, especially in OLT but also in CT. Actually, I'm not aware of any other lexical morpheme in OLT that would participate in as many as ten distinct syntactic constructions – so many have been identified in the corpus under scrutiny. This situation is certainly a side-effect to the lexicalisation process *scogs* underwent: from a fully lexical verb to an idiomatic phrase. To the extent allowed by the preserved corpus of OLT, this paper is intended to sketch the lexicalisation path taken by the morpheme whose starting-point must be located in Proto-Tibetic (PT).

The paper is divided into four parts. In section *Variant forms* (2.), I discuss spelling variants of *scogs* and separate them from their homonyms. Part 3. *Etymology* clarifies the origins of the lexeme by identifying its cognates and sketching historical processes that link them to each other. Part 4. *Variety of constructions* examines the types of construction with *scogs* that have been identified in the corpus. Finally, *Distribution and chronology of constructions* (5.) provides statistical data for distinct constructions and seeks to align them chronologically. Unless absolutely necessary, each construction is illustrated in the paper with only two or three examples but their overall frequency in the corpus is provided in this final section.

The study is based on the corpus of 226 texts available in the OTDO database (status as of 09.10.2022) which have yielded 345 examples of the morpheme scattered through 106 texts (see Appendix A). Whenever possible passages cited in the paper have been checked with the texts on Gallica or IDP.[1]

2 Variant forms

In OLT, the prevailing form of the morpheme is *scogs* (Table 1).[2] The form *sogs*, identical with CT *sogs*,[3] is scarcely attested; OTDO lists only three occurrences. Two texts, PT 149 and Or.15000/256, consequently use the form *bsogs*. The latter variant occurs side by side with *bscogs* and *scogs* in PT 981. In all its occurrences, the morpheme *scogs* can be identified with the morpheme under investigation. The same concerns the variant *bscogs*, whereas *bsogs* and *sogs* possess homonyms.[4] Since it is sometimes difficult to differentiate between the superscribed *r-* and *s-* in OLT manuscripts, it is possible that the variant *rcogs* shall in fact be read as *scogs*. Only PT 239 and PT 981 use the form *rcogs* but they do so side by side with at least two other forms.[5] The variant *gsog* in Or.15000/315: r1 is isolated and *scags* in Or.8212/187: 76 is a clear misspelling.[6] Table 1 contains all the variants identified in the OTDO database:

scogs	*bscogs*	*bsogs*	*rcogs*	*sogs*	*gsog*	*scags*
328	18	6	6	3	1	1

Table 1. OLT orthographical variants of *scogs*

Table 2 lists the texts that contain variants different than *scogs*:

	scogs	*bscogs*	*rcogs*	*bsogs*	*sogs*	*gsog*	*scags*
PT 37		10					
PT 126		1					
PT 149				3			
PT 239	9	1	1				
PT 981	2	1	1	1			
PT 1043					1		
PT 1051		1					
PT 1060					1		
PT 1092		1					
ITJ 737-2	2	2					
Or.8212/187	16						1

---

in Bialek (2020b). In examples (8), (13), and (15) the variant spelling *lascogs* has been broken down to *las scogs* for the sake of interlinear morpheme translation; a colon < : > in front of *scogs* marks the division.

[2] *schogs* in PT 1166: 7 on OTDO is a misreading for the correct *scogs* of the text.

[3] Although *sogs* is the standard spelling in CT, queried for *scogs* BCRD provides over one thousand examples from Buddhist literature.

[4] However, it is not clear whether *bscogs* in ITJ 738: 1v4 should be identified with our *scogs*.

[5] According to de Jong (1989: 3), "Manuscript B (= ITJ 737-2 – JB) agrees almost entirely with manuscript E (= PT 981 – JB)". This fact explains the use of *scogs* and *bscogs* in both texts. The concurrence of three variants in PT 981 and PT 239 suggests that the latter manuscript might have been composed in the same circles or even written down by the same scribe. Dotson and Helman-Ważny's comparative study (2016: 144ff.) neither includes PT 239 nor considers the variant forms of *scogs* as a potential identification Merkmal.

[6] OTDO gives *scags* also in PT 1283: 367 but the manuscript clearly reads *scogs*.

Or.15000/256				2		
Or.15000/315						1
Or.15000/426v	1	1				
Ɣbis 2					1	
Ldan 1			1			
Ldan 2			2			
Ldan 3			1			

Table 2. Distribution of orthographical variants across texts

The variant *sogs* postdates the sound change *sc-* [sts-] > *s-* [s-]. More relevant for the following discussion are the variants with the prefix *b-*, especially *bscogs*. The texts in which the latter occurs were either composed in Buddhists circles (PT 37, PT 149, PT 239), translated from other languages (PT 981, ITJ 737-2), contain divinations (PT 1051),[7] or were written by non-native Tibetan speakers (PT 1092, Or.15000/256, Or.15000/426v). None of them can be attributed with any probability to a native speaker of a Tibetic language. Neither can the texts be reasonably assumed to be copies of older originals, in which case the form *bscogs* could have been an archaism. In general, words in *bsc-* are rare (altogether 31 occurrences including *bscogs*) and found in texts of non-Tibetan provenance. Whatever motivated their composers to replace *sc-* by *bsc-*, this seems to have been a case of hyper-correction rather than a true archaism. Since the oldest datable attestations of *scogs*, like those in the OTA or in the Bsam and Źol inscriptions, unanimously show the onset *sc-* I accept *scogs* as the original form of the morpheme.

3 Etymology

The closest identified cognates of *scogs* are diagrammed in Figure 1. The primary object of this study, the lexeme *scogs*, is enclosed in a text-box; dashed arrows show inflection, solid ones derivation.

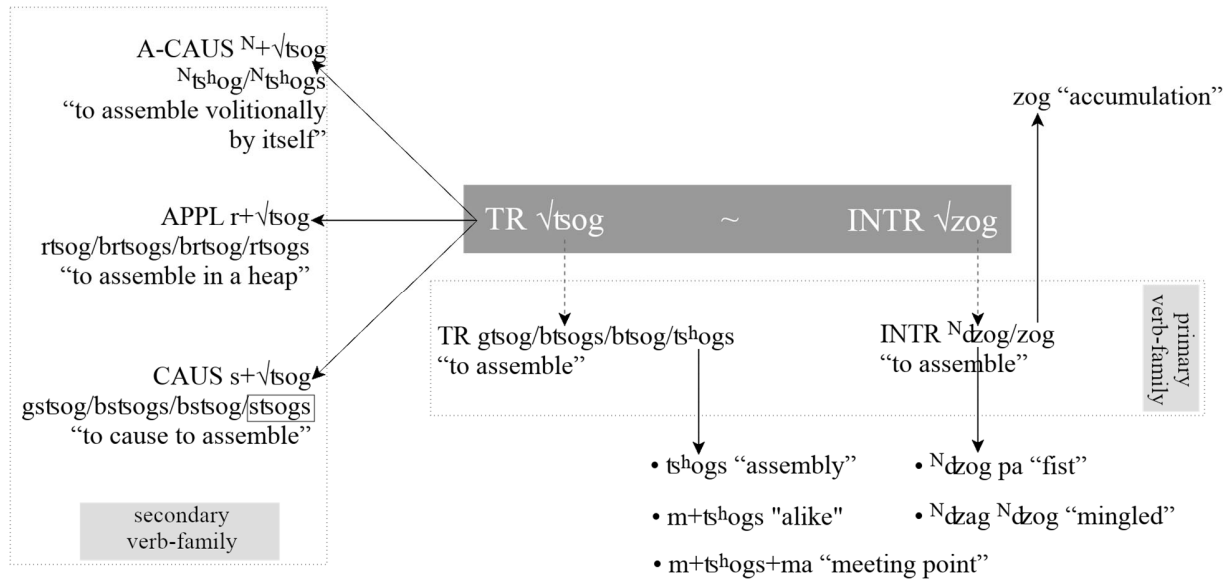[7] Divination texts are notorious for their non-standard orthography.

Figure 1. Word-family of *scogs*

The data in Figure 1 are presented with their etymological forms and meanings, which I shall confront with lexemes actually attested in OLT. But prior to that a few words are due on the core elements of the word-family: the verb roots √tsog and √zog. These have been reconstructed on the basis of the identified OLT cognates and are supported by the documented alternation voiceless transitive vs voiced intransitive in onsets of inherited verbs in PT.[8] Due to Schiefner's law that operated in one of the ancestors languages (most probably Proto-Bodish), PT did not contrast voiced affricates and voiced fricatives in syllable-initial position: *dzog > PB *zog > PT *zog (Bialek 2021a: xiiif.).

Coblin (1986: 108f.) reconstructed PT roots *tshog and *dzog[9] and connected them to Ch. 曹 *cáo* and 遭 *zāo*, reconstructing the latter as PTH *dzagw → OC *dzagw → *dzâu* "to come together; a crowd" and PTH *tsagw → OC tsagw → *tsâu* "to meet with", respectively. More recent reconstructions by Schuessler (2007) and Baxter/Sagart (2014) exhibit no final consonant:

		Schuessler	Baxter/Sagart
曹 *cáo*		LH dzou	MC dzaw
		OCM *dzû	OC *N-tsˤu
遭 *zāo*		LH tsou	MC tsaw
		OCM *tsû	OC *tsˤu

Table 3. Reconstructions of Ch. 曹 *cáo* and 遭 *zāo*

---

[8] See most recently Bialek (2020a: 267), Jacques (2021), Bialek (2021a: xiiif.).

[9] His list of LT cognates includes: *ɣjog/bcogs/bcog* "to heap together, mix up together"; *ɣchogs/chogs* "to assemble, meet; *chogs* "a crown, assembly"; *sogs*, OLT *scogs/bscogs* "to accumulate, gather together" (ibid., p. 109). *ɣjog* was related to *ɣchogs* already by Csoma de Kőrös (1834: 223a).

Thus, it is doubtful whether the said Chinese lexemes are historically related to √tsog and √zog. But Schuessler (2007: 325) accepted LT lexemes collected by Coblin as cognate with Ch. 聚 *jù*, stating "Affiliation with *jù* is likely in light of etymological parallels in WT (my LT – JB) […] and *zú* 卒 'group' (→ *zāo* 遭)." However, among the lexemes listed s.v. 聚 *jù* (ibid.) only 族 *zú* "clan, clansmen" is reconstructed with a final velar plosive:

	Schuessler	Baxter/Sagart
族 *zú*	LH dzok	MC dzuwk
	OCM *dzôk	OC *[dz]ˤok

Table 4. Reconstruction of Ch. 族 *zú*

Since for none of the Chinese verbs listed by Schuessler a final velar plosive has been reconstructed I deem the link to PT √tsog and √zog dubious for now. On the other hand, a survey of STEDT has not yielded any potential cognates of PT √tsog and √zog so that the final -*g* may be considered an innovation or an outcome of clipping with a thus far unidentified morpheme.

Now I turn to the historically attested lexemes of Figure 1. The analysis follows the division of OLT verbs into primary and secondary verb-families (see Bialek 2020a: 266ff.).

3.1   Primary verb-family

The primary verb-family consists of two verbs which are not attested lexicographically with the given conjugations. Instead, Csoma de Kőrös (1834: 223a), on whose authority the entry in Jäschke is based (J: 467a), provides the following conjugation: v1 *ɣjog* v2 *bcogs* v3 *bcog* "to gather, heap together". Conjugations with a voiceless root consonant in v2 and v3 but a voiced one in v1 slot are virtually unknown.[10] Two factors might have contributed to the confusion in the presumed conjugation of *ɣjog*: 1. ambitransitive semantics (cf. Eng. *to assemble* or *to gather*); and 2. replacement of the original verbs by derivatives of the secondary word-family (see below). Despite the uncertainty, I reconstruct two distinct verbs:[11]

> TR   v1 *gcog v2 *bcogs v3 *bcog v4 *chogs* (VII) "to collect"
> INTR   v1 *ɣjog v2 *zog (III) "to come together"

Of these only one form is attested in OLT: *chogs*. Ex. (1) illustrates its use together with the noun *chogs* "assembly" derived from v4 *chogs* by conversion:[12]

---

[10] The only parallel example is CT *ɣjur* (v2 *bčur*(*d*), v3 *bčur/gčur*, v4 *čurd/ɣjurd/čhur*; Hill 2010: 98b) "to struggle against, to resist" (J: 178b) cognate with *gčur* (v1 *gčur/bčur/ɣjur/ɣčhur*, v2 *bčur*(*d*), v3 *gčur/bčur*, v4 *čhur*(*d*), *čurd*; Hill 2010: 80a) "to be squeezed" (CDTD.V: 345). There exists a great conundrum in the sources not only concerning the inflected forms but also the semantics. I assume that conjugations of at least two verbs have been mingled and the original v3 'got lost' or rather was replaced for we most probably find it glossed as a distinct verb *gźur*.

[11] Sigla in round brackets provide the paradigms according to Bialek (2020a: 284, 286).

[12] For derivation by conversion from v4-stems, see Bialek (2020a: 302f.).

(1) *myī   čhig=la   chogs   dguɣ   chogs=na //   spuṅ+sad+zu+ce=ɣo //* (PT 1287: 317)
     man   INDF=ALL assembly   nine(ABS)   collect:SPASS=INESS   Spuṅ+sad+zu+ce(ABS)=FNL

   If all (lit. nine) accumulations are collected in one man, it is Spuṅ-sad Zu-ce.

Two further lexemes can be considered derivations from *chogs*:

   mchogs "alike"
   mchogs ma "fontanelle" < *"meeting point" (?)

By virtue of their morphology, both lexemes must have been derived from *chogs*. Since their derivation from the noun "assemble" seems improbable, I presume that they have been derived directly from the v4-stem *chogs*. The uncertain semantics of the prefix *m-* hinders the reconstruction of their etymological meanings but the logic seems to have been:

   mchogs – things or people that have been assembled share some feature, under which aspect they can be conceived of as resembling each other
   mchogs ma – the place where the bones of the skull meet

I consider *mchogs ma* to be derived directly from *chogs* and not from *mchogs*. Although lexicographical sources most usually provide the variant form *mchog ma*, the word-internal *-s* is confirmed by Chik [ʦoqsma] "fontanel" (CDTD: 6941); *-s* underwent elision at the syllable boundary between two consonants: *-s > Ø / -g_ₒm-*.[13]
   A survey of CT canonical texts has yielded multiple examples of *gcog*, *bcogs* and *bcog*, all of which turned out to be either misspellings or homophones of other lexemes, most commonly CT *bcags* (v1 *ɣchog*) "to cudgel" and *bcog pa* "dirty".
   zog is attested dialectally as "goods, commodities; merchandise; thing; cattle, livestock" (CDTD: 7406). It is a collective term denoting valuables that came in one's possession. The noun was derived from v2 by conversion with the etymological meaning *"what has come together".[14] v1 *ɣʝog* is attested in CT, cf.:

(2) *de=las   ɣʝog+paɣi   bdud+rciɣi   thig+leɣu   chogs*
     DEM=DEL assemble+NR:GEN   nectar:GEN   drop:DIM   multitude(ABS)
     (Śūnyatāsamādhivajra, *Dpal he ru kaɣi man ṅag raṅ byin gyis brlab paɣi rim pa źes bya ba*, D 1262, *rgyud ɣgrel*, ña 335v6)
     multitude of little drops of the nectar that assemble from that[15]

---

[13] Both forms *mchogs ma* and *mchog ma* are attested in written sources (see BDRC). A similar alternation between written forms and dialectal reflexes can be observed, for example, in: CT *phyag ma* "broom" (J: 348a, s.v. *phyag*) ~ Bal [pʰjaxma], Kar [pʰjaꭓsma], Tsha [pʰjaqsma], Wan [pʰjaksma] (CDTD: 5226); CT *chag ma* "sieve" (J: 444a, s.v. *chag pa*) ~ Bal [ʦʰaxma], Kar [ʦʰaꭓsma], Tsha [ʦʰaqma], Chik [ʦʰaqsma] "id." (CDTD: 6739). Simplification of word-internal consonant clusters has been well documented for compounds (see Bialek 2018a.1: 209ff.) which is assumed to be basically the same process.

[14] As against Bialek (2021a: xvi), I now consider *zog* to have been derived from the intransitive verb. The remaining arguments are still valid, including *zoṅ* as a secondary form of *zog*.

[15] Instances of *ɣʝogs* in CT turned out to be misspellings or misreadings of *ɣʝegs* "to ascend" (J: 466b).

Two further lexemes have been derived from v1 *ɣjog*:

> *ɣjog pa* "khu chur" (BYD: 461a) < *"the assembling one" as a collective term for all the fingers placed together in the palm[16]
>
> *ɣjag ɣjog* "mixed, mingled" (J: 463b) is a reduplication according to the well-established pattern, in which the second syllable retains the vowel of the original lexeme, whereas the first syllable changes it into *a* (Uray 1954: 220)[17]

This primary verb-family has to be reconstructed from scratch because, in those word-families that include both primary and secondary verb-families, the former have largely been replaced by the latter. In particular, the verb, on whose root the secondary verb-family is based, in our case √tsog, in fact ceases to exist.[18] From the conjugation *gcog/*bcogs/*bcog/*chogs* only the last form has survived, albeit in very restricted usage; the vast majority of *chogs* in OLT can be identified as nouns.

3.2 Secondary verb-family

Roots of secondary verb-families are derived from one of the roots of the primary verb-family by means of the prefixes:

- autocausative *ɣ-* [N-]
- applicative *r-* [r-]
- causative *s-* [s-][19]

In the case of the primary roots √tsog and √zog, it was the former, transitive root that functioned as the base for the derivation:[20]

- N+√tsog *"to assemble volitionally by itself" > OLT *ɣchog/ɣchogs* (INTR II) "to assemble"
- r+√tsog *"to assemble in a heap, to heap up" > OLT *rcog/*brcogs/*brcog/*rcogs* (TR VII) "to mingle"
- s+√tsog *"to cause to assemble" > OLT *gsog/*bsogs/*bsog/*scogs* (TR VII) "to assemble"

ɣchog and *gsog* have been linked to *ɣjog* of the primary verb-family already by Csoma de Kőrös (1834: 223a) and Coblin (1986: 108f.). The following discussion casts more light on the verbs of the secondary verb-family.

---

[16] Das added the meaning "to fold the fist" (D: 1056b) which I was not able to confirm in other sources. Uray (1954: 185) related *ɣjog pa* "fist" to CT *ɣjog* glossed as "to heap together". According to this interpretation, a fist were something that heaps things together – for me a less appealing image.

[17] *ɣjag ɣjog* was related to the verb *ɣjog* by Uray (1954: 225). The word together with its variants seems to be known only in the spoken language (see CDTD: 7001).

[18] As demonstrated in Bialek (2020a: 282), from the pair √du ~ √tu only inflected forms of the latter have remained in the language. The former disappeared after *ɣdu* and *sdud* had been derived from it.

[19] Bialek (2020a: 323).

[20] An asterisk preceding the form indicates that it is not attested in OLT.

3.2.1 ^{N}+√ʦog

Contrary to the data provided in later lexicographical sources,[21] *γchog* was the original v1-stem of the verb ^{N}+√ʦog:

(3)   *bcan+po=γi spur=la=nī   γchog* / (PT 1287: 46)
      bcan+po=GEN   body.HON=ALL=FOC assemble(NPFV)

   [One] assembles around *bcan po*'s body.

Likewise the CT compound *γchog čhas* "goods" (< **γchog paγi čhas* "things that come together") confirms the lack of *-s*; BDRC provides not even one instance of *!γchogs čhas*. Two occurrences are reported from the Lha-sa canon (see BCRD), which is notorious for its misspellings, but the comparison with Sde-dge (via RKTS) again reveales only *γchog čhas*.

3.2.2 *r*+√ʦog

The relation of *r*+√ʦog to the remaining lexemes of the word family has escaped the attention of scholars for it is not attested as a verb in CT, where it has been replaced by the form *bcog* as in *bcog pa* "unclean, dirty, nasty, vile" (J: 435b).[22] In OLT, only v3 *brcog* seems to be attested:

(4)   *brcog=gis   ma   γthuṅ=śig* (ITJ 737-1: 173; *apud* OTDO)
      DPASS:pollute=ERG NEG   NPFV:drink=IMP

   Because [the fluid] has been polluted, do not drink it!

In the relative clause *nal g.yam rcog paγi myi dag* (PT 126: 93), lit. "humans who mingle/pollute *nal g.yam*" (?), *rcog* seems to be the v1-stem but the passage remains unintelligible to me. Besides, one encounters *rcog* with the variant spelling *brcog* as a noun in the following phrases:

   (*b*)*rcog ṅan* (PT 1046B: 18; PT 1047 passim; ITJ 740: 148; *apud* OTDO) "bad pollution"
   rcog daṅ thab čhe ba (PT 1055: 47–8; *apud* OTDO) "pollution and turmoil that are great"[23]
   brcog sdig (ITJ 737-1: 371; *apud* OTDO) "pollution and sin"

Since derivation by conversion from v1-stems is not documented in OLT, I assume that the original form of the noun was *brcog* and it was derived from v3 *brcog*.[24] The translation "pollution" is only tentative as the lexeme must have denoted rather a source of pollution, an entity from which pollution emanates because it has been contaminated. The semantic change from the reconstructed ***"to assemble in a heap, to heap up" to "to pollute" could have been obtained via the intermediary "to mingle": gathering things together in heaps results in mingling them and so contaminating the set with objects that are considered less pure. In context of OLT, the applicative *r-* is understood as a derivational morpheme that allows promoting a non-core argument like oblique or even adjunct to

---

[21] For an overview of the forms, see Hill (2010: 241). Jäschke has *γchogs/chogs* (J: 460b).
[22] Dialectal data confirm the onset *rc-* (see CDTD: 6683 & 6684) as against CT *bc-*.
[23] The translation of *thab* is tentative. I connect it to CT *thab thob* "=thom thom" (J: 229b), ultimately probably related to *γthab* "to fight".
[24] For deverbal derivation from v3-stems, see Bialek (2020a: 297).

direct object. The most common semantic roles of the promoted elements cross-linguistically are those of benefactive, instrument, and locative, but others are likewise attested (https://wals.info/chapter/109; 04.03.2022). In r+√tsog it was the locative adjunct that has been promoted:

*"X$_S$ collects (s+√tsog) hay$_O$ <u>in sheaves</u>" ~ "X$_S$ heaps up (r+√tsog) <u>sheaves$_O$</u>"

Since applicative *r-* ceased to be productive in PT, OLT verbs in *r-* attest to a high level of lexicalisation.[25] However, the sense of confusion or mingling has already surfaced in the discussion, namely in *yjag yjog*.

3.2.3 *s+√tsog*

scogs, the primary object of this paper, was originally a verb (see section 4.1 below) and in OLT it still exhibits verbal behaviour in participial constructions, relative clauses, or nominalisations. Another inflected form of the same verb is attested as *bsogs*:

(5)   (a) *dags+rī   dkar+po=laɣ* ~~las~~ /   *pho+gśen thod+/+dkar   brgya   bsogs=te* [...]
            sunny_slope   white=ALL   male_*gśen* white_turban   hundred(ABS)   PFV:gather=GER

        (b) *srībs+rī   nag+mo=laɣ /   mo+gśen   źu+bub /   brgya   bsdus* / [...] (PT 1285)
            shady_slope   black=ALL   female_*gśen*   headgear   hundred(ABS)   PFV:gather

        On the white sunny slope of the mountain, [one] gathered one hundred male *gśen* with white turbans. [...] On the black shady slope of the mountain, [one] gathered one hundred female *gśen* with a headgear. [...]

(5) establishes *bsogs* (a) and *bsdus* (b) as near-synonyms and indicates that *bsogs* should be analysed as an active v2-stem. The clauses "[One] gathered male/female *gśen*" are repeated in the text fourteen times and, apart from (5a), they always involve the verb *bsdus*. However, another ritual text again has *bsogs* in the same context:

(6)   (a) *bdags+{ri}* /   (2r48) *dkar+po=la /   po+gśen   thod+kar   brgya   bsogs=te* // [...]
            sunny_slope   white=ALL   male_*gśen* white_turban   hundred(ABS)   PFV:gather=GER

        (b) *[srīb]s+[ri]   nag+po=la /   (2r49) ~~ma~~ mo+gśen   źu+[bub] /   brgya*
            shady_slope   black=ALL   female_*gśen*   headgear   hundred(ABS)

        [b]sogs=te / [...]   (ITJ 734; *apud* OTDO)
        PFV:gather=GER   white=ALL

        On the white sunny slope of the mountain, [one] assembled one hundred male *gśen* with white turbans. [...] On the black shady slope of the mountain, [one] gathered one hundred female *gśen* with a headgear. [...]

---

[25] I will discuss applicative *r-* more thoroughly in a forthcoming publication (see Bialek In Preparation). Another pair of cognate verbs with applicative *r-* and causative *s-* is OLT *brnan* "to add onto" and *bsnan* "to add".

The analogous structures of (a) and (b) in (6) allow us to identify *sogs* with *bsogs* in this context (as can be seen in Table 1, *bsogs* and *sogs* are also attested as orthographical variants of *scogs*). According to the OTDO database, in OLT this verb is attested altogether fifteen times as *bsogs* and once as *sogs* (ex. (6)). Unfortunately, all its occurrences come either from ritual texts on the verge of unintelligibility (PT 1060 passim; PT 1285: r66; ITJ 734: 2r48; ITJ 738: 1v4; ITJ 739: 16v6) or from texts translated from other languages (PT 986: 82; PT 1283: 147). The juxtaposition with *bsdus* in (5) confirms that *bsogs* was a v2-stem of a transitive verb. On the other hand, its semantics suggests a close paradigmatic link with *scogs*. On these grounds, I reconstruct the conjugation of s+√tsog as:

PT	v1 *gstsog	v2 *bstsogs	v3 *bstsog	v4 *stsogs
OLT	v1 *gsog	v2 *bsogs*	v3 *bsog	v4 *scogs*

Since all the examples of v2 *bsogs* come from presumably late texts the form either postdates the sound change [sts-] > [s-] or the change was facilitated in non-initial position and first occurred after prefixes. The latter hypothesis is supported by the statistical counts on the OTDO data put together in Table 5:

sc-	*bsc-*	*gsc-*
919	32	17

Table 5. The cluster *sc-* in OLT

Thus, *scogs* and *bsogs* represent two forms of one conjugation with the verb root s+√tsog. The seeming differences between their argument structures (S$_{ABS}$ Q$_{DEL}$ *scogs* vs [S$_{ERG}$] O$_{ABS}$ *bsogs*; see section 4.1) can be explained by two factors: 1. the voice distinction stative passive *scogs* vs active *bsogs*;[26] and 2. restricted number of their occurrences as verbs, the majority of which, in the case of *bsogs*, come from parallel passages in ritual texts that do not provide enough comparative material.

Since in OLT forms in *r-* are etymologically distinct from those in *s-* Zeisler's hypothesis that *rcogs* might be the original form of *scogs* (2016: 470) must be rejected. Modern dialects have preserved the distinction between reflexes of OLT *rc-* [rts-] and *sc-* [sts-]; the former has left reflexes with [ts] whereas the latter has yielded [s] (see CDTD). The only dialect that seems to have merged them is Mdzo with, e.g., [tsal] ~ [tsar] for OLT *scal* (CDTD.V: 998) and [tsa] for OLT *rca* (CDTD: 6601). WDro [sōk] (CDTD: 8867) for OLT *scogs* but [tsōkpā] (CDTD: 6683) for OLT *rcog pa* proves unanimously the etymological *s-* in onset.

After its idiomatisation, *scogs* separated from the verb that, after the sound change [sts-] > [s-], continued as (*g*)*sog* "to gather, heap up, hoard up; to assemble" (J: 579a) into the classical times. Concerning the semantic change from the verb form "was/has been assembled" to the idiom "among others", compare the etymology of Eng. *among*: "O[ld] E[nglish] *ongemang* (from *on* 'in' + *gemang* 'assemblage, mingling')" (OED: 43a). Regarding the reconstructed relationship between the OLT *rcog* and *scogs*, it is also worth mentioning the etymology of Eng. *to mingle*: "origin M[iddle] E[nglish]: frequentative of obs[olete] *meng* 'mix or blend' (rel[ated] to *among*)" (OED: 910a).

---

[26] For the voice opposition, see Bialek (2020a: 288ff.).

3.3 bsogs *vs* bsags

There exists a great conundrum concerning the inflectional forms of (*g*)*sog* in CT; Jäschke listed forms with the root vowel *o* together with those whose root vowel was certainly *a* (J: 579a, s.v. *sog pa*), whereas Hill (2010: 304) and CDTD.V (1335 & 1336) separated them into two lemmata. The fact remains that the verbs are synonyms and seem to have the same inflectional forms in CT but for v2 and v3 where the one has *o* and the other *a*. Compare (6) with the following passage:

(7) *rgyal+po čuɣu+buɣ+ɣwaṅ=gis / dmag+myi maṅ+po bsags=te /* (PT 986: 30)
     king   Čuɣu+buɣ+ɣwaṅ=ERG   soldier   many(ABS) PFV:gather=GER

King Čuɣu-buɣ-ɣwaṅ gathered many soldiers.

The preceding section has demonstrated that all verbal cognates of *scogs* have the vowel *o* throughout their conjugations. Without going into detail that would require a separate in-depth study, I assume that the verb with the root vowel *a* is etymologically distinct from √tsog and √zog and should most probably be linked to CT *ɣjag*, *ɣchag*, *chag pa* etc. The latter set may go back to the verb roots √tsag (cf. v3 *bscag* in PT 1283 passim) and √zag. The meaning "to assemble" attested for *gsog/bsags/bsag/sog*(*s*) in CT derived from "to save, to put aside" as in sifting.[27]

4 Variety of constructions

Based on the OTDO corpus it was possible to discern between as many as ten distinct constructions that involve the syllable *scogs* or any of its orthographical variants. These are:

- Finite clause (PRED)
- Participial adverbial clause proper (PAdvC.I)
- Participial adverbial clause idiomatic (PAdvC.II)
- Post-head relative clause (PostH-RC)
- R-dislocation (R-disl)
- Relative clause extraposition (RCE)
- Pre-head relative clause (PreH-RC)
- Off-subject nominalisation (Off-SN)
- Idiomatic (ID.I)
- Idiomatic (ID.II)

The following sections examine the constructions, looking at their specific characteristics and probable origins.

4.1 *Finite clause (PRED)*

In the following example *scogs* occurs clause-finally as the main verb in its clause:

---

[27] CDTD.V relates both word-families, √tsog/√zog and √tsag/√zag, to each other, adding LT *rceg* "to pile up" and *rcig* "to build" (993, 1019, 1335, 1336). I do not see any rationale behind this linkage.

(8) *thugs+glud* *lu+gu* *dkar+po* *gčig* / (99) *gla+sgaṅ* *bču+gsum* / *mdaɣ+rgyud*
     ransom(ABS) lamb   white   one   *gla+sgaṅ* thirteen   *mdaɣ+rgyud*

~~*čhe thaṅ*~~ {*bču+gsum*}=*las*   :*scogs*=*te* // (PT 1042)
              thirteen=DEL   gather:SPASS=GER

The ransom was assembled from one white lamb, thirteen [blades of the medicinal grass] *gla sgaṅ*, [and] thirteen [pieces of] the *mdaɣ rgyud* plant.

In (8) *scogs* is the predicate of a passive clause, whose subject is *thugs glud* and the remaining elements form the oblique argument. The passage agrees with the prototypical word order of intransitive clauses with oblique: SQV (S = subject; Q = oblique; V = verb). The subject *thugs glud* only refers to those elements explicitly listed in the oblique argument, i.e. to one white lamb, thirteen blades of the medicinal grass *gla sgaṅ*, and thirteen pieces of the *mdaɣ rgyud* plant. The logical relationship between the referents of *thugs glud* and the constituting elements of the ransom is that of equality; the subject refers to exactly these ingredients and nothing else:

$$Q = S$$
"The referents of the oblique argument are identical with the referents of the subject"

4.2 *Participial adverbial clause (PAdvC)*

A considerable group of constructions with *scogs* encompasses phrases with the HEAD constituent postposed to a participial adverbial clause ending in *scogs=ste*. Their structure can be schematically presented as:

$$[X=las \ scogs=ste]_{PAdvC} \ HEAD$$
Lit. "having been assembled from X, HEAD"

The HEAD NP is semantically more encompassing and includes in its meaning X that can consist of one or more prototypical elements of its class but does not list all possible elements of the HEAD set. The relation between X and the HEAD can therefore be described as that of strict inclusion:

$$X \subset HEAD$$
"X is strictly included in HEAD"

Syntactically, this construction can be compared to:

(9) *dkyīl+(64)+ɣkhor*=*la*   *ɣʝegs*=*te* //   *bod*=*kyīs*
     central_circle=ALL   ascend:PFV=GER   Tibetan=ERG

gcīgs   *bzuṅ*=*ño* // (ST Treaty E; *apud* OTI: 37)
edict(ABS)   PFV:seize=FNL

Having ascended the central circle, the Tibetans accepted the edict.

In (9), *bod*, the subject of the intransitive *ɣǰegs* and the transitive *bzuṅ*, is stated in the second clause and only the context allows one to understand it as coreferential with the zero marked subject of the first clause. There is no coreference marker in the first clause. Since (9) is the only thus far identified example of an analogous construction, I take it as a model for the analysis of *scogs=ste*. In accordance with this model, the postposed HEAD NP is coreferential with the underlying subject of the preceding verb. (10) is the only example of *scogs* with an agentive argument, here in adjunct function:[28]

(10) [*sku+gśe{n}*   *mǰol+bon+po=rnams* / (48)=*kyis* //   *do+ma=la*   *scogs=te*]PAdvC
       sku+gśen   *mǰol+bon+po=*PL=ERG   companion_ horse={DEL}[29]   assemble:SPASS=GER

       [*rkaṅ+ɣgros=kyī*   *rnams*]HEAD   *gśog+yugs=kyis*   *gdab* // (PT 1042)
       livestock=GEN   parts(ABS)   wing_stick=ERG   DPASS:strike

       Having been assembled by *sku gśen mǰol bon po*s from among companion horses, herded livestock was struck with a wing-stick.

The phrase *rkaṅ ɣgros kyī rnams*, being the subject of the passive *gdab*, is coreferential with the zero coded element of the first clause. I analyse *scogs=ste* as a passive participle, lit. "being/having been assembled". This verbal reading is confirmed by the inclusion of a locative adjunct in (11):

(11) *zla+goṅ=gi*   *bu+chaɣ+*(51)*+rgyud+ɣpheld=gyī*   *lag=na*   [*bran   źiṅ   ɣbrog   sog*
       Zla+goṅ=GEN descendant=GEN   hand=INESS serf   field   pasture   fallow_land

       (52) *chal=las*   *scogs=te*]PAdvC   [*dbaṅ=ňo=čog*]HEAD /   *blar*   *myi*   *bźes*
            forest=DEL   assemble:SPASS=GER   sway=DEM=PL   authority:TERM   NEG   take_back

       (53) *myi*   *dbrī*   *gźan=gyis*   *myi*   *dprog* (Źol N; *apud* OTI: 9)
            NEG   DPASS:diminish   other=ERG   NEG   DPASS:take_away

       Whatever is under one's sway, serfs, fields, pastures, fallow lands, [or] forests gathered in the hands of the descendants of Zla-goṅ, shall not be taken back to the authorities, diminished, or taken away by others.

dbaṅ ňo čog "whatever is under one's sway" is coreferential with the zero coded subject of *scogs* that is specified as an action of assembling or collecting worldly goods in one's hands. The locative adjunct (*lag=na*) confirms the still verbal character of *scogs*, on which the =*na* marker depends in this case.

The participial construction is attested in two semantic types discussed separately in the following sub-sections.

4.2.1 *Participial adverbial clause proper (PAdvC.I)*

In type I, the element X consists of one or more nouns joined in a coordinative phrase and the HEAD expresses a superordinate category for the object(s) listed as X. This is the prototypical and

---

[28] Cross-linguistically agentive passives are much less common than agentless passives (DeLancey 1981: 634). This tendency explains why *scogs* is preponderantly found without an agentive argument.

[29] As is well-known (see also section 5.2 below), *la scogs* is a later orthographic variant of *las scogs* and therefore, in all respective cases, *la* is glossed as delative marker but enclosed in braces: {DEL}.

more frequently encountered construction with *scogs=ste*, in which the HEAD remains coreferential with the zero coded subject of *scogs*.

In (12) the HEAD is the determinative phrase *stod pyogs gyī pho ña* "emissaries of Upper Regions" that is specified by the preceding clause in *scogs=ste*:

(12) [*ban+ɣǰag nag+po=daṅ gog=daṅ / śīg+nig=las scogste*]_{PAdvC} /   [*stod+pyogs=gyī*
      Ban+ɣǰag   black=COM   Gog=COM   Śig+nig=DEL assemble:SPASS:GER upper_region=GEN

      pho+ña]_{HEAD} *pyag*   *ɣcald* / (Or.8212/187: 20)
      emissary(ABS) hand.HON request.HON:PFV

      Having been assembled from among the Black Ban-ɣǰag, Gog, and Śig-nig peoples, emissaries of Upper Regions paid homage.

In (12), the subject of *scogs* (the HEAD element) remains the subject in the following clause. The situation changes when the subject of *scogs* receives another syntactic role in the subsequent clause so that the literal translation of *scogs=ste* is not possible any more. This seems to be an outcome of the extension of contexts in which the construction 'X=*las scogs=ste* HEAD' was used, resulting in a figurative re-interpretation of *scogs=ste* as an indicator of a specifying list "HEAD being X, among others; HEAD such as X, among others". Examples of the extended use of the construction include HEAD elements as direct object (exx. (13)–(14)), oblique (15), and postposition (16):

(13) *spyi   groṅs+ṅog=du gyur=pa=la /   [thog+maɣi   ɣon+kuṅ / tha+maɣi*
      generally   dead=TERM   (PFV)become=NR=ALL front_part:GEN trench   back_part:GEN

      gtaṅ+khuṅ=las :scogs=/(94) =te]_{PAdvC} [*sa*]_{HEAD} *lan gsum=las {lhag+pa} thur=du*
      gift_pit=DEL   assemble:SPASS=GER   place(ABS)   turn three=DEL   more   down=TERM

      rkor   myi   ruṅ=ṅo // (PT 1042)
      (NPFV)dig:TERM NEG   be_proper=FNL

      Generally, it is not proper to dig down places such as offering trench of the front part [or] gift pits of the back part, among others, more than three times for those who died.

Another example with the HEAD as direct object in its own clause includes a more complex passage, in which the specifying clause in *scogs=ste* contains a set of nominalised clauses:

(14) [*spyīr   legs=payī   bka+gros   gsol=čiṅ /   las=su*
      generally be_good=NR:GEN advice.HON give.HON=DUR work=TERM

      byas=pa=(9)=las   scogs=te]_{PAdvC} /
      do:PFV=NR=DEL   assemble:SPASS=GER

      [*dpen=pa*]_{HEAD}   *byed~byed / sñīṅ   ñe~ñeɣo* // (Źwa W; *apud* OTI: 17)
      beneficial=NR(ABS) do:NPFV~RDP / heart(ABS) be_close~RDP:FNL

      Doing what is beneficial – the given advice, that was generally good, and what [he] did as [his] works, among others, – [he] remained loyal.

In (14) the noun *dpen pa* "what is beneficial" is the direct object of *byed byed* but its content is exemplified with two nominalised clauses: **spyīr legs paɣī bka gros gsol pa* "the given advice that is generally good" and *las su byas pa* "what [he] did as [his] works". This complex construction was permitted following the weakening of *scogs* as a verb and the accompanying semantic bleaching.

(15)　　*ɣdi=rnams*　　[*g.yen　dgu=las　:scogs=te*]PAdvC //　(89) [*mṅaɣ+dbaṅ　so~so=na*]HEAD　*mčhis=pas* //
　　　　DEM=PL(ABS) division　nine=DEL　assemble:SPASS=GER　　　domain　　　　separate=INESS　exist=NR:ERG

　　　　gtad=kyaṅ　　　*mčhir*　　　*myi btub=ste* /　　*myi sman=no* // (PT 1042)
　　　　DPASS:deliver=ADD　go(NPFV):TERM　NEG PFV:be_able=GER　NEG be_beneficial=FNL

　　　Since these [animals] exist in separate domains such as the nine divisions, even if delivered, not being able to go, [they] would not be beneficial.

mṅay dbaṅ is the HEAD of the oblique NP of the verb *mčhis* but it is also the underlying subject of the participial adverbial clause *g.yen dgu lascogs te*.

(16)　　　[*mo=daṅ*　　*rmyi+ltas=las*　(32)　*scogs=ste*]PAdvC /　[*čīɣi　　phyir=yaṅ*]HEAD　*ruṅ=ste* /
　　　　　divination=COM　dream_omen=DEL　　　　assemble:SPASS=GER what:GEN because_of=ADD　be_proper=GER

　　　　myī gźig=go /　　　*myī spaṅ=ṅo* // (Skar; *apud* OTI: 23)
　　　　NEG DPASS:destroy=FNL NEG　abandon(DPASS)=FNL

　　　Whatever being the reason [for these statements], divination or (lit. and) dream-omens among others, [the shrine of the Three Jewels and the religion of the Buddha] shall not be destroyed nor abandoned.

(16) is still more intricate and demonstrates further lexicalisation of *scogs*. The subject of *scogs* must be co-referential with the indefinite pronoun *či yaṅ* "whatever" and the latter is used in the postpositional phrase *čīɣi phyir yaṅ* "for whatever reason".

4.2.2 Participial adverbial clause idiomatic (PAdvC.II)

This is the most problematic of the discussed constructions and the following analysis is deemed tentative. The overt structure of the construction resembles that of a participial adverbial clause discussed in the preceding section but with one important variation:

$$[\text{X}=las\ scogs=ste]_{\text{PAdvC}}\ \text{Y}$$
$$\text{"Y such as X"}$$

The HEAD of the participial adverbial clause is replaced by Y which indicates that this constituent cannot be construed as the subject of *scogs*. Instead, Y is identical with the subject of X which in the scheme represents clause(s) nominalised off of the subject grammatical relation (GR). In (14) we have already seen that the X element may itself consist of nominalised clauses. The following examples likewise contain lists of nominalised clauses as the X element, all of which are nominalised off of the subject.

(17)   γuṅ=gī   γog=du   [ɣ̌jaṅ=(335)=la čhab+srid   mjad=de   mywa dkar+po
       DEM=GEN bottom=TERM Ɣ̌jaṅ=ALL   dominion(ABS) make.HON=GER Mywa white(ABS)

       dpyaγ phab //   mywa nag+po   γbaṅsu   bkug=pa=la   scogste]_{PAdvC} /
       tax(ABS) (PFV)throw Mywa black(ABS) subjects:TERM PFV:summon=NR={DEL} assemble:SPASS:GER

       (336) [rgyal+po γdī=las]_Y // dbu+rmog brcan=źiṅ / čhab+srīd mthaγ
             king   DEM=DEL   helmet.HON be_mighty=DUR dominion   end(ABS)

       skyes=pa=nī   sñan+čhad gduṅ+rabs=kyis   (337) ma mjad=do // (PT 1287)
       be_born:PFV=NR(ABS)=FOC early   generation.HON=ERG   NEG make.HON=FNL

> Thereafter, none helmet was mightier and no generation formerly ruled a dominion[30] of [such] extended boarders more than this king, [who], having enforced policy against the Ɣ̌jaṅ, imposed taxes [on] the White Mywa [and] subjugated the Black Mywa and so on.

A more literal translation would be: "the king such as the one who, among others, having enforced policy against the Ɣ̌jaṅ, imposed taxes [on] the White Mywa [and] subjugated the Black Mywa". Here the X are not elements of the set termed *rgyal po* but rather characterise *the* king. This is a remarkable shift in the semantics of *scogs=ste* by which the Y constituent (*rgyal po*) is removed from the subject role of *scogs*: !"The king was assembled from the one(s) who, having enforced policy against the Ɣ̌jaṅ, imposed taxes [on] the White Mywa, [and] subjugated the Black Mywa and so on". Thus, the phrase *la scogste* has become an idiom. The same process is confirmed in (18):

(18)   [čīγi   thabs=kyīs=kyaṅ myī dprog   myi bźes=par   gnaṅ=ba=las
       what:GEN means=ERG=ADD   NEG DPASS:take_away NEG take_back=NR:TERM grant=NR=DEL

       scogste]_{PAdvC}   (46) naṃ~naṃ+źa~źar / blon   snaṅ+bzaṅ+γdus+koṅ=gi
       assemble:SPASS:GER   for_ever~RDP:TERM councillor Snaṅ+bzaṅ+γdus+koṅ=GEN

       bu+cha+γphel+rgyud / srīd   bde=źiṅ   brtan=(47)=baγī [gcigs]_Y gnaṅ=ba
       descendant(ABS)   dominion(ABS) happy=DUR firm=NR:GEN   edict(ABS) grant=NR

       γdī /   nam~nam+źa~źar gnaṅ=źīṅ myi bsgyur   myī bčos=par   (48) ṅaγi
       DEM(ABS) for_ever~RDP:TERM grant=DUR NEG DPASS:change NEG DPASS:alter=NR:TERM   1SG:GEN

       źa+sṅa=nas=kyaṅ   dbu+sñuṅ   gnaṅ // (Źwa W; *apud* OTI: 18)
       front.HON=EL=ADD   oath.HON(ABS) grant

> I personally also grant an oath so that this edict, of a happy and firm realm, granted for ever and ever to councillor Snaṅ-bzaṅ Ɣdus-koṅ's offspring ([an edict] that has been granted so that, among others, by no means would anything be removed or taken [from them]) while bestowed for ever and ever, shall not be changed nor altered.

The Y constituent is *gcigs* which is coreferential with the subject of the first *gnaṅ* but not with that of *scogs*. (18) is even more complex for *gcigs* is simultaneously the reference point of three nested constructions:

---

[30] For the polysemous *čhab srid* and its phraseology in OLT, see Bialek (2022).

- *scogs=ste*
  *čīүi thabs kyīs kyaṅ myī dprog myi bźes par gnaṅ ba las scogste gcigs "edict such as the one that has been granted so that by no means will anything be removed or taken [from them] among others"
- Determinative phrase
  *srīd bde źiṅ brtan baүī gcigs "edict [that guarantees] (lit. of) a happy and firm realm"
- Head-internal relative clause
  *naṃ naṃ ža žar blon snaṅ bzaṅ үdus koṅ gi bu cha үphel rgyud gcigs gnaṅ ba үdī "this edict that has been granted for ever and ever to councillor Snaṅ-bzaṅ Үdus-koṅ's offspring"

Finally, the entire construction headed by *gcigs* is the subject of the consecutive passive verbs *gnaṅ*, *myi bsgyur*, and *myi bčos*. Despite its compound structure there can be no doubt that *gcigs* is not coreferential with the subject of *scogs*: !"The edict has been assembled from among what has been granted so that by no means would anything be removed or taken [from them]."

The level of complexity demonstrated by (18) could only be reached in a written text but even (17) poses serious problems to the analysis. I understand the construction as idiomatic extension of the PAdvC.I construction which has resulted from broadening of the scope of elements that could form the X constituent (seen already in (13)–(16)). The semantic shift was further fuelled by the inclusion of nominalised clauses as in (14) under X. Initially, the HEAD element in 'X=*las scogs=ste* HEAD' was the subject of *scogs* only but with the introduction of off-subject nominalisations in the X position, a tighter link was established between Y and the subject of X to the disadvantage of *scogs*. The single stages can be schematically reconstructed as (NC = nominalised clause; SN = off-subject nominalisation):

X *las scogs te* HEAD	"being assembled from X, HEAD"
X_{NC} *las scogs te* HEAD	"being assembled from among [those things to which [one] did X]$_{NC}$, HEAD"
X_{SN} *las scogs te* HEAD	"being assembled from among [those who did X]$_{SN}$, HEAD"
X_{SN} *las scogs te* Y	"Y such as [the one who did X]$_{SN}$"

Be that as it may, with only three examples documented in OLT (see Appendix A), this construction plays a minor role in the repertoire of *scogs*.

4.3 *Post-head relative clause (PostH-RC)*

A nominalised clause with the predicate *scogs* can be used as a post-head relative clause:

[HEAD$_{ABS}$ [X *las scogs pa*]$_{RC}$]$_{NP}$
Lit. "HEAD, that was assembled from among X", i.e. "HEAD such as X"

The HEAD and the relative clause form together one NP that can take an argument slot in the host clause and receive a case marker: absolutive in (19) & (22), terminative in (20), or ergative in (21). The HEAD can be a collective term (exx. (19) & (21)), explicitly marked plural (22), or unmarked plural (20). The HEAD element has a broader scope of denotation which includes the referents

mentioned in the relative clause; X denotes special cases from among the elements subsumed under the HEAD. The examples collected thus far indicate that X usually does not list all the elements referred to by the HEAD. Thus, the relation is one of inclusion:

$$X \subseteq \text{HEAD}$$

"X is included in HEAD" or "X is a subset of HEAD"

In (19) four animal species are counted as *gles pa*:

(19) [*gles+pa*]_{HEAD} [*stag=daṅ* (127) *gzig / dom=daṅ dred=las ɓscogs=pa*]_{RC}
     gles+pa(ABS) tiger=COM leopard brown_bear=COM yellow_bear=DEL assemble:SPASS(ABS)

   maṅ+por mčhis=na / (PT 126)
   many:TERM come:PFV=INESS

   If *gles pa* such as tiger and leopard, brown bear and yellow bear, became abundant [...].

Although the meaning of *gles pa* remains unknown, it is clear that it must have been a noun and hypernym of the terms tiger, leopard, brown bear, and yellow bear. The interpretations given by Hill (2021: 116f.) do not convince; it seems that the term might have rather denoted big predators. This reading is confirmed by the co-occurrence of the term with *gčan pa* (PT 126: 133), certainly the same word behind *gčan gzan* "beast of prey". Thebo [le^L-bo] "boar" (see Lin 2014: 249) might contain the same stem.

(20) attests to a further lexicalisation within the PostH-RC in which the oblique argument of *scogs* has a singular referent; it counts one instance of an "evil realm". Thus, by the time of the composition of (20), the relative clause with *scogs* has undergone semantic specialisation, becoming even more idiomatic. This is confirmed by the fact that the literal translation *"evil realm, [one] assembled from hell" is logically not very appealing to say the least.

(20) (12) [*sdig+yul*]_{HEAD} [*sems+čan+dmyal+ba=las rcogs=pa*]_{RC}*r *skyes=te*
        evil_realm hell=DEL assemble:SPASS:NR:TERM be_born:PFV=GER

   sdug+bsṅal=baɣī naṅ=du rgyun=tu ɣkhor=ro / (Ldan 1; *apud* OTI: 60)
   suffering=NR:GEN inside=TERM stream=TERM circle(NPFV)=FNL

   Being reborn in an evil realm such as hell, [one] will always roam in suffering.

In (20) the entire phrase functions as an oblique argument of the predicate *skyes* and so receives terminative marking. The same idiomatisation as in (20) is also observed in (21) where the collective term *ɣbaṅs* "subjects" receives a PostH-RC that mentions only one person:

(21) *de=nas* [*yul ṅas+po=ɣi* / (188) *ɣbaṅs*]_{HEAD} // [*dbaɣs+dbyī+chab=la scogs=pa*]_{RC}*s //
     DEM=EL land Ṅas+po=GEN subjects(ABS) Dbaɣs+dbyi+chab={DEL} assemble:SPASS=NR:ERG

   bcan+po=ɣi mchan gsol // (PT 1287)
   bcan+po=GEN name.HON(ABS) give.HON

Thereafter, subjects of the Ṅas-po land such as Dbaɣs Dbyi-chab offered a name for the *bcan po*.

The same construction can be recognised in (22) although *scogs* is here truncated to one syllable devoid of the preceding *la*(*s*) and carrying no clitic. This is a side effect of the metrics since, as it seems, the passage was originally composed in verses:

(22) [[*śa*] *lhu* *bźi*]ₕₑₐD [*go* *scogs=ni*]ᵣC *myiṅ+po*[*ɣi*]
     meat portion four(ABS) *go*(ABS) assemble:SPASS=FOC brother:GEN

     sñiṅ+lan=du *khyer=čig* / [*śa* *lhu* *bźi*]ₕₑₐD
     sñiṅ+lan=TERM carry(SPASS)=IMP meat portion four(ABS)

     [*go* *scogs=*[*ni*]ᵣC (49) *sriṅ+moɣi* *dur=du* *čhug=śig* / (PT 1068)
     go(ABS) assemble:SPASS=FOC sister:GEN tomb=TERM put_in(SPASS)=IMP

     Regarding four portions of meat such as *go*, carry [them] as brother's *sñiṅ lan*! Regarding four portions of meat such as *go*, put [them] in sister's tomb!

The interpretation of *go* is uncertain. The next clause explicitly mentions *mgo braṅ smad lṅa* "head and internal organs, the five" indicating that *go* referred to a body part different than the head (*mgo*). Still, the construction with *scogs* makes the reading of *go* as denoting some part of the body the most plausible one. Maybe the etymological meaning *"upper part" was intended?

Since the construction 'HEADₐBS [X *las scogs pa*]ᵣC' forms one NP it can not only obtain case marking but also participate in any other type of construction characteristic of NPs in OLT. (23) provides an example of appositional phrases in which the said construction takes the first position and is followed by an appositive, itself a complex NP:

(23) [*lha=ɣi* *dkor*]ₕₑₐD (11) [*mduṅ* *raṅ* *ɣdebs=daṅ* / *ral+gyī* *raṅ*
     deity=GEN movable(ABS) lance self NPFV:throw=COM sword self

     gčod=daṅ / *khrab* *raṅ* *gyon=daṅ* / *phub* *raṅ* *bzur=la*
     NPFV:cut=COM armour self don=COM shield self parry={DEL}

     scogs=pa]ᵣC / [*ɣphrul=gyi* *dkor* (12) *čhed+po* *mṅaɣ=ba=ɣī* *rnams*]ₐPPV
     assemble:SPASS=NR(ABS) might=GEN movable great(ABS) possess=NR=GEN parts(ABS)

     bdag=la *scal=na* *phod=čes* *gsol=to* / (PT 1287)
     1SG.HML=ALL (PFV)give=INESS be_able=QUOT give.HON(PFV)=FNL

     "I dare [to fight against you] if [you] grant me the deity's movables such as a self-throwing lance, self-cutting sword, self-donning armour, and self-parrying shield – the great movables of transformational mights, that [you] possess." Thus [he] said.

The structure of (23) can be schematically presented as:

[HEAD [X *las scogs pa*]ᵣC]NP.1 [NP.2]ₐPPV
Lit. "HEAD, that has been assembled from among X, which is NP.2"

19

(23) illustrates a correctly formed appositional phrase in which the appositive, NP.2, refers to the HEAD of NP.1 which is itself modified with a post-head relative clause (RC).

4.4   R-dislocation (R-disl)

Constructions introduced in this section have the following structure:

[[X *las scogs pa*]_{NP.1} HEAD]_{NP.2}
"those assembled from X, the HEAD"

The entire construction forms one NP that may function as predicative (24), subject (exx. (25) & (28)), direct object (26), or indirect object (27) in the hosting clause. X *las scogs pa* invariably stands in absolutive and the case marking is added after the HEAD in accordance with the semantic role of the construction in the hosting clause. The referent of X can be singular (25), giving a prototype of the class referred to by the HEAD, or plural (exx. (24), (26)–(28)). Examples (26)–(28) demonstrate that X can consist of a coordinative phrase, listing a random number of hyponyms of the HEAD. The enumeration in X is not always exhaustive so that the logical relationship of inclusion is established between X and the HEAD:

$$X \subseteq \text{HEAD}$$
"X is included in HEAD"

(24)   [*ɣčhī+bdag=gyī   bdud=las   rcogs=pa*]_{NP} //   [*khams* (r6.2) *gsum=gī   dgra*
         lord_of_death=GEN evil_spirit=DEL assemble:SPASS=NR(ABS) realm   three=GEN   enemy

thams+čad=las]_{HEAD} *rgyal=bar*   *śog=śīg* // (PT 239)
all=DEL   be_victorious=NR:TERM   go(SPASS)=IMP

May [the deceased] be victorious over all the enemies of three realms, evil spirits of the lord of death, among others!

(25)   [*mgar=la   scogs=pa*]_{NP}   [*blon+po   kha+čhig*]_{HEAD}
         Mgar={DEL} assemble:SPASS=NR(ABS) councillor   some(ABS)

glo+ba   ɣdrīṅ=na // (PT 1287: 331)
breast(ABS)   NPFV:be_away=INESS

When some councillors, Mgar among others, were becoming disloyal [...].

(26)   (52) [*ɣbu+śīṅ+kun=daṅ   zīn+ču=daṅ   ga+ču=las   scogs=pa*]_{NP} /   [*rgyaɣī*
            Ɣbu+śiṅ+kun=COM   Zin+ču=COM   Ga+ču=DEL   assemble:SPASS=NR(ABS) Chinese:GEN

mkhar   maṅ+pho]_{HEAD} *phab=/=ste* / (Or.8212/187)
stronghold   many(ABS)   (PFV)throw=GER

[*źaṅ* Rgyal-zigs and *źaṅ* Stoṅ-rcan] conquered many strongholds of the Chinese: Ɣbu-śiṅ-kun, Zin-ču, and Ga-ču, among others.

(27)   *blon*   *snaṅ+bzaṅ+(26)+ɣdus+koṅ=gi*   *bu+cha+ɣphel+rgyud*   *dmaṅs=kyi*   *rnams=kyaṅ*
      councillor   Snaṅ+bzaṅ+ɣdus+koṅ=GEN   descendant(ABS)   commoner=GEN parts=ADD

      [*gcaṅ=(27)=daṅ*   *stoṅ=las scogs=pa*]_{NP}   [*sgor bde=baɣī*   *rnams*]_{HEAD}
      gcaṅ=COM   stoṅ=DEL assemble:SPASS=NR(ABS) *sgor*   favourable=NR:GEN parts(ABS)

      źaṅ+lon   *yī+ge+(28)+čan=gyi*   *thaṅ=du*   *gnaṅ=ba=daṅ* / (Źwa E; *apud* OTI: 21)
      aristocrat   possessing_letters=GEN   rank=TERM   grant=NR=COM

The descendants of the councillor Snaṅ-bzaṅ Ɣdus-koṅ, even those among commoners, are granted favourable *sgor*s such as *gcaṅ* and *stoṅ*, among others, in the rank of aristocrats possessing letters.

(28)   (5)   // *lha+ǰī*   *čhen+po*   *gnaṃ+lhaɣi*   *sras=las* /   [*čho+byi*   *che*
           Lha+ǰi   great(ABS)   celestial_deity:GEN   son.HON=DEL Čho+byi   *che*

      gsum=daṅ   (6)   [*mdaɣ*]+*myī ce gsuṃ=daṅ* /   *ce ce gsuṃ=daṅ*
      three=COM   Mdaɣ+myi   *ce* three=COM   Ce *ce* three=COM

      phyug+po ce gsuṃ=daṅ /   (7)   *gyīm+po ce gsuṃ=las scogs=pa*]_{NP}
      Phyug+po   *ce*   three=COM   Gyim+po   *ce*   three=DEL   assemble:SPASS=NR(ABS)

      [*spun*]_{HEAD}   *myī+ɣpha*[*gs*]+(8)+[*rigs*]=*su*   *gyes=so* (Źwa F; *apud* OTI: 67)
      kinsmen(ABS)   man+noble+lineage=TERM   divide:PFV=FNL

From the great Lha-ǰi, the son of the celestial deities, proceeded kinsmen such as three *che* of Čo-byi, three *ce* of Mdaɣ-myi, three *ce* of Ce, three *ce* of Phyug-po, and three *ce* of Gyim-po as the noble lineage.

I call this construction R-dislocation because it seems to be the closest possible approximate of rightward movement (R-movement) of a clause constituent in OLT. The examples concern the R-movement of the subject of the verb *scogs* according to the pattern:

<div align="center">

S [X *las scogs pa*]_{PostH-RC} → [X *las scogs pa*] S
"S, which was assembled from among X" → "assembled from among X, the S"

</div>

The subject is moved to the right end of the relative clause. The cataphoric reference is not established in NP.1 by a pronoun but by a morphological zero like in other languages with zero anaphora such as Japanese (see Givón 2001.2: 268). When compared with the examples of PostH-RC, the passages with R-disl attest to a more complex character of the HEAD constituent. Namely, in R-disl the HEAD is generally more complex and contains elements in attributive position, adjectives or determiners, that are absent from the HEAD-NP of PostH-RC. (28), the only exception to this observation, has a complex X. It has been argued that R-disl, as a word-order changing device, is a topic-coding construction (Givón 2001.2: 253). In OLT, however, R-disl appears to be mainly motivated by the heaviness of the constituents of a PostH-RC; if the HEAD of a PostH-RC contained postposed elements or the PostH-RC was of considerable length, the HEAD-NP was moved rightward, probably to facilitate the processing of information by placing the main constituent closer to the hosting verb.

4.5   Relative clause extraposition (RCE)

Extraposition of a constituent results in a discontinuity of the order of elements in a clause. In OLT an extraposed relative clause assumes a place in a clause not directly preposed to the NP which it modifies but separated from the latter by another NP. In (29), the NP *brag mar gyī bsam yas las scogs pa* specifies the NP *gcug lag khaṅ* but does not directly precede the latter. Instead, the locative adjunct *dbuṅ mthar* is inserted between the two NPs:

(29)   [*brag+mar=gyī*   *bsam+yas=las*   *scogs=*(14)*=pa*]_{RC} //
Brag+mar=GEN   Bsam+yas=DEL   assemble:SPASS=NR(ABS)

    dbuṅ+mthar   [*gcug+lag+khaṅ*]_{HEAD}   *brcīgs=ste* //
centre+end:TERM temple(ABS)   PFV:build=GER

    dkon+(15)+*mčhog*   *gsum=gyī* *rten*   *bcugs=pa=daṅ* // (Skar; *apud* OTI: 23)
jewel   three=GEN support(ABS) PFV:establish=NR=COM

Having built temples, Bsam-yas of Brag-mar among others, in central and border regions, [one] established the support for Three Jewels.

(30) is a much more complicated case and some doubts might be raised as to its interpretation as an RCE:

(30)   [*dkon+mčhog*   (24) *gsum=gyī* *rten*   *bcugs=pa=daṅ* /   *saṅs+rgyas=kyī* *čhos*
jewel   three=GEN support(ABS) PFV:establish=NR=COM Buddha=GEN   teaching(ABS)

    mjad=pa   (25) *myī* *gtaṅ*   *ma* *žig=par* //   *gduṅ+rabs+rgyud=kyīs* *yi+dam*
make.HON =NR   NEG DPASS:abandon NEG (PFV)ruin=NR:TERM generation=ERG   oath(ABS)

    bčaγo=(26)*=źes*   *γbyuṅ=ba=las*   *scogs=pa*]_{RC} //   *bcan+po* *yab+sras* //
make(DPASS):FNL=QUOT NPFV:occur=NR=DEL assemble:SPASS=NR(ABS) *bcan+po*   father.HON+son.HON

    rǰe+(27)+*blon*   *kun=kyīs* [*dbu+sñuṅ=daṅ* *bro*]_{HEAD} *bor=te* //   *gcigs=kyī* *yi+ge=daṅ* /
lord+councillor   all=ERG   oath.HON=COM   vow(ABS) PFV:throw=GER edict=GEN   text=COM

    (28) *rdo+rīṅs=la*   *bris=pa*   *bžīn=du=yaṅ*   *mjad=do* // (Skar; *apud* OTI: 23)
    stone_pillar=ALL   PFV:write=NR(ABS) like=TERM=ADD make.HON=FNL

The *bcan po*, the father and the son, together with all the lords and councillors, having sworn an oath (such as the one that comes: "A vow will be made by [each] generation that from now on for ever and ever the support, one has established for the Three Jewels, and the Buddha's teaching, one has been practising, will not be abandoned nor fall into ruin."), acted in accordance with what one had written in the text of the edict and on the stone pillar.

In (30), *dbu sñuṅ daṅ bro* is construed as coreferential with the subject of *scogs*; the main part of the construction is understood to be *γbyuṅ ba las scogs pa* […] *dbu sñuṅ daṅ bro*, lit. "[the words of] the oath have been collected from what comes". The X element is therefore a nominalised clause ending in *γbyuṅ ba* and whose content consists of direct speech. The direct speech is the subject of *γbyuṅ*. Since *dbu sñuṅ daṅ bro* is separated from its modifier *γbyuṅ ba las scogs pa* by the agentive subject of *bor*, I consider the construction another example of RCE.

RCE has developed from R-dislocation:

R-disl: [[X *las scogs pa*]_{NP.1} HEAD]_{NP.2}
RCE: [X *las scogs pa*]_{NP.1} Y HEAD_{NP.2}

Y is any constituent of the hosting clause that does not belong to HEAD; in (29) it is a locative adjunct, whereas in (30) an agentive subject. The phrase [X *las scogs pa*]_{NP.1} is extraposed and placed at the beginning of the hosting clause.

4.6 *Pre-head relative clause (PreH-RC)*

Pre-head relative clauses are the primary relativisation strategy in OLT. Their structure can be schematically illustrated as:

$$[… V=NR:GEN]_{RC} \ HEAD_{NP}$$
"HEAD that V"

A closer examination has revealed that in pre-head relative clauses of OLT the HEAD constituent is preponderantly relativised off of the subject GR.[31] The survey of *scogs* has yielded three potential cases of PreH-RC:

(31) [*byaṅ+čub+sems+dpaɣ   čhen+po   brgyad=daṅ / khro+bo   gñīs=las*
      bodhisattva   great   eight=COM   fierce_deity two=DEL

      scogs=paɣī]_{RC}   [*ɣkhor*]_{HEAD} (PT 16: 28r4; *apud* OTDO)
      assemble:SPASS=NR:GEN   retinue(ABS)
      a retinue that has been assembled from eight great *bodhisattva*s and two fierce deities

(32) [*ṅos   bźīɣī   mgon+po=daṅ /   phyogs+skyoṅ=daṅ /   lha+klu   sde*
      direction four:GEN protector=COM   side+guardian=COM   deity+serpentine_being class

      brgyad=las   scogs=paɣī]_{RC}   [*rten*]_{HEAD} (PT 16: 28v1; *apud* OTDO)
      eight=DEL   assemble:SPASS=NR:GEN   support(ABS)
      a support that has been assembled from protectors of four directions, guardians of the directions, and eight classes of deities and serpentine beings

The relative clauses in (31) and (32) contain numerals that stipulate the number of entities included in the referent of the HEAD constituent. As for the OLT corpus, this is unusual in constructions with *scogs*.

---

[31] Bialek (In Preparation).

(33) [*dṅo*[*s*]+*sdig*   *bśags=pa=la*   *scogs=paɣī*]ᴿᶜ
      true_misdeed(ABS)   PFV:confess=NR={DEL}   assemble:SPASS=NR:GEN

[*bsod+nams*]ʜᴇᴀᴅ (ITJ 751: 38v1; *apud* OTDO)
merit(ABS)
the merit that has been collected from confessing true misdeeds

PT 16/ITJ 751 abounds in ungrammatical constructions. Since PreH-RC with *scogs* are attested only in this text, it might be just another instance of erroneous syntax. What looks like PreH-RC might have been R-disl, in which some ignorant copyist added genitive to the HEAD constituent. This is even more probable as the text indeed contains nine examples of R-disl (see Appendix A). One must note however that PreH-RC is one of the most common constructions with *scogs/sogs* in CT. Whatever the case, the HEAD of the relative clause in (31)–(33) is relativised off of the subject GR and thus confirms that the above examples accord to the general pattern of PreH-RC in OLT. (31)–(33) can be schematically represented as:

[X *las scogs paɣi*]ᴾʳᵉᴴ⁻ᴿᶜ HEADɴᴾ
"HEAD that was assembled from X"

In the above examples the referents of the HEAD constituents seem to be equal to the referents of X, thus:

X = HEAD
"The referents of X are identical with the referents of the HEAD"

It is conspicuous that this small sample only contains examples of equality whereas strict inclusion seems to be preferred by the R-disl constructions.

4.7   *Off-subject nominalisation (Off-SN)*

scogs also occurs in constructions that are best understood as off-subject nominalisations; the clitic *=pa* added to a clause causes its nominalisation with two possible interpretations:

I. The nominalised clause expresses the subject of the verb, in (34) of an intransitive clause, in (35) of a transitive clause with a v2-stem, and in (36) of a transitive clause with a v1-stem:

(34) *glo+ba*   *riṅs=pa* (PT 1287: 95)
      breast(ABS)   be_away:PFV=NR(ABS)
      the one who was disloyal

(35) *dmag   dru+gu+yul=du*   *draṅs=pha* (ITJ 750: 255)
      army   Dru+gu+land=TERM   lead:PFV=NR(ABS)
      those who led the army to the Dru-gu land

(36) | *myi=ɣī* | *myig* | *bya+myīg* | *ltar* | *ɣdug=pa* | *ɣog=nas* | *ɣgebs=*(38)*=pa* | *gčhig* (PT 1287)
| human=GEN | eye | bird+eye(ABS) | like | COP=NR(ABS) | bottom=EL | NPFV:cover=NR | one(ABS)

one who closes his human eyes, that are like bird eyes, from below

The nominalised clauses lack the subject NP which is instead coded by the clitic *=pa* at the right end. Nominalisation off of the subject GR is the primary nominalisation strategy in OLT.

II. The nominalised clause expresses the object of the verb:

(37) | *myī* | *čhig=gīs* | (208) | *bstod=pa* (PT 1287)
| human | INDF=ERG | | PFV:praise=NR(ABS)

what a single man praised

(38) | *bcan+po=ɣi* | *sñan=du* | *ṅag+re+khyuṅ=gīs* | *gsol=paɣ* // (PT 1287: 324)
| *bcan+po*=GEN | ear.HON=TERM | Ṅag+re+khyuṅ=ERG | give.HON=NR(ABS)

what Ṅag-re-khyuṅ related to *bcan po*

Nominalisation off of the object GR is extremely rare in OLT and occurs regularly only in the introductory formula of direct speech like in (38). (37) comes from a passage adapted from the Chinese *Shǐjì* 史記 (Takeuchi 1985) whose paraphrase in the OTC displays features foreign to Tibetan syntax.

Thus, in OLT, clause nominalisation by means of *=pa* is possible only off of the subject GR and when introducing direct speech.[32] With respect to later Tibetic languages this construction has sometimes been dubbed 'headless relative clause' (Huber 2003: 12). The same type of construction may have *scogs* as its predicate:

(39) | [[*līɣi* | *qa+ma+ča=la* | *scogs=pa*]Off-SN / | *thabs*
| Khotanese:GEN | *qa+ma+ča*={DEL} | assemble:SPASS=NR(ABS) | rank

| *gser=daṅ* | [*g.yu*] | *scal=pa=la* | *scogs=pa=yaṅ*]Off-SN /
| gold=COM | turquoise(ABS) | (PFV)give.HON=NR={DEL} | assemble:SPASS=NR=ADD

| *rce+rǰe* | *zaṅs=paɣi* | *ɣog=na* | *mčhis* // (PT 1089: r24)
| prefect | copper=NR:GEN | bottom=INESS | exist

Those assembled from among the Khotanese *qa ma ča*s,[33] who were given ranks which are gold and turquoise, are below the town prefect of copper rank.

(39) contains two off-subject nominalisations which are however difficult to render as such in English since *līɣi qa ma ča la scogs pa*, lit. a "those assembled from the Khotanese *qa ma ča*s", is the direct object of *scal* whose clause is nominalised and re-used as oblique of the second *scogs*: […]

---

[32] This statement is weakened by the sole thus far identified nominalisation off of the direct object GR: *las su byas pa* (Źwa W 8) "what [one] did as [one's] works" (see ex. (14)). It might be that off-object nominalisation was soon extended to contexts other that direct speech but it seems to be a very rare construction in OLT.

[33] OLT *qa ma ča* < Khotanese Saka *āmāca* (Skt. *āmātya*; P. *amacca*; Uray 2007 [1979]: 131), the highest Khotanese official (Takeuchi 2004: 55a).

thabs gser daṅ g.yu scal pa la scogs pa yaṅ "[humans] assembled from […], who were given golden and turquoise ranks".

(40)   [*rī+dags=la   mdaɣ   rṅul=pa=las   scogs=pa*]_{Off-SN} (PT 1072: 32)
       game=ALL   arrow(ABS) aim=NR=DEL   assemble:SPASS=NR(ABS)
       Lit. [humans] assembled from among those who aimed an arrow at a game

In (39) and (40) the NPs in *scogs pa* refer to all humans that participated in the respective event; *all* Khotanese with gold and turquoise ranks were of lower position than any town prefect of copper rank and *all* those who aimed an arrow at a game were equally held responsible for their action. Apparently, with plural referents the NPs denoted equality: an exact set that consisted of all those who fulfilled the condition. In these cases, X *las scogs pa* can be paraphrased with "the group (consisting) of X".

4.8   Idiomatic scogs (ID.I & ID.II)

Already in OLT one encounters *scogs* in the idiomatic phrase "¹such as; ²and the like, and so forth; among others", the most popular construction with *scogs/sogs* in CT and apparently the only one known in modern dialects (CDTD: 8867). It is attested as early as 710/11 in the OTA (see ex. (41) below) which fact proves the considerable time depth of the lexicalisation. In this OLT construction, *scogs* continues to be used together with *la(s)* and followed by the nominal *=pa*. The only noticed exception concerns example (43) where *scogs* is the last syllable in the line and this might have triggered the fortuitous omission of the expected *=pa*.

The construction consists of only one NP whose last elements are invariably *la(s) scogs/sogs pa*. It can be schematically represented as:

[X *las scogs pa*]_{NP}
"X and the like; X, among others"

X can be a simple NP with one referent or a complex NP with multiple referents; it may even be a nominalised clause (see exx. (42), (44), & (45)). Its referents are representatives of a particular group or a set of events which themselves are not referred to anywhere else in the clause. There exists no coreference between the NP that ends in *la(s) scogs pa* and any other constituent of the clause. The idiomatic character of the phrase marks itself in the lack of the subject argument of *scogs* or any other phrase in the clause which could be taken for coreferential with the underlying subject. The whole NP functions as a normal NP and so can take case markers depending on its semantic role in the clause.

(41)   *źaṅ bcan+to+re+lhas+[byīn]=las   scogs=pas /   gñe+bo   bgyī*(177)*ste /* (ITJ 750)
       źaṅ Bcan+to+re+lhas+byin=DEL   assemble:SPASS=NR:ERG assistant(ABS) PFV:do:GER
       źaṅ Bcan-to-re Lhas-byin, among others, acted as personal assistant.

(42)   (328) */:/ bcan+po   khri+ɣdus+sroṅ*[*=gis*] *//*   [...] *phag+rgod=la   bśan=gyīs   mjad /*
              bcan+po   Khri+ɣdus+sroṅ[=ERG]   wild+boar=ALL   PFV:kill=ERG make.HON

g.yag+rgod   sg{r}og=du   bčug //
wild_yak(ABS)   fetter=TERM   PFV:put_in

(329) *stagī   rna+ba=la   bzuṅ=ba=la   scogs=paɣ //*
    tiger:GEN   ear=ALL   PFV:seize=NR={DEL}   assemble:SPASS=NR(ABS)

thugs   sgam+ba=ɣī   steṅ=du / sku+rcal   čhed+pos   bsnan=te / (PT 1287)
mind.HON   profound=GEN   top=TERM   skill.HON   great:ERG   PFV:add=GER
By means of [his] great physical skills *bcan po* Khri Ɣdus-sroṅ added [to the quality] of [his] profound mind[34] [things] such as slaughtering wild boars, fettering wild yaks, [and] seizing tigers by [their] ears.

In (42) the nominalised sentence *phag rgod la bśan gyīs mjad / g.yag rgod sg{r}og du bčug // stagī rna ba la bzuṅ ba la scogs paɣ* is the direct object of the verb *bsnan*. The construction cannot be analysed as an RCE modifying *sku rcal čhed po* because this would leave the causative verb *bsnan* without direct object. On the other hand, the ergative in *sku rcal čhed pos* dismisses this very phrase as a potential direct object, contrary to the reading proposed by Hill: "On top of his profound mind he added great energy." (2013: 176).

(43)   *ɣdre+srin=la   scogs=pa   nag+po=ɣi   phyogs=nī /   ɣphel=čiṅ*
    demon={DEL}   assemble:SPASS=NR   black=GEN   side(ABS)=FOC   increase(PFV)=DUR

    dar=te /   lo+ñes=daṅ   nad=la   scogs
    spread(PFV)=GER   bad_harvest=COM   disease:{DEL}   assemble:SPASS

(15)   *rgyun=du   byuṅ=nas //* (ITJ 737-3; *apud* OTDO)
    stream=TERM   (PFV)occur=EL

The party of the bad (lit. black) ones, demons and the like, increased and spread; bad harvest, diseases and the like occurred constantly.

    The first *scogs* in (43) belongs to R-dislocation, whereas the second one forms the idiomatic phrase under discussion. In (44) the construction forms a determinative phrase with the following NP and therefore stands in genitive:

(44)   (11) *čhos+rgyal   čhen+pos   phrin+las=su   či*   (12) *mjad=pa=daṅ /*
    *dharma*_king   great:ERG   deed.HON=TERM   what(ABS)   make.HON=NR=COM

    dbu+rmog   brcan+poɣi   byīn=(13)*=gyis /   čhab+srid   skyes=pa=las   scogs=paɣi*
    helmet.HON   mighty:GEN   splendour=ERG   dominion(ABS)   grow:PFV=NR=DEL   assemble:SPASS=NR:GEN

    skyes=pa=las   scogs=paɣi   (14) *gtam=gyi*
    grow:PFV=NR=DEL   assemble:SPASS=NR:GEN   account:GEN

    yī+ge /   źib+mo   gčīg=ni /   gud=(15)*=na   yod=do //* (Ɣphyoṅ; *apud* OTI: 13)
    text   detailed   one(ABS)=FOC   separate_place=INESS   exist=FNL

---

[34] Lit "on top of [his] profound mind".

A detailed text of an account of what the great *dharma* king did as deeds and of growing of the realm due to the splendour of [his] mighty helmet and the like exists elsewhere.

(45) illustrates the use of the construction in a postpositional phrase with the postposition *phyir*:

(45)   *ñi+ṅog*   *rgyal+(388)+po*   *gźan*   *dpyaɤ*   *phab=pa=daṅ* /   *rgyal+pran*
     under_the_sun   king   other(ABS)   tax(ABS)   (PFV)throw=NR=COM   petty_king(ABS)

     ɤbaṅs=su   *bkug=pa=la*   *scogs=pa*   *phyir* //   *čhab+srid*
     subjects=TERM   PFV:summon=NR={DEL}   assemble:SPASS=NR(ABS)   because_of   dominion(ABS)

     mthaɤ   (389) *bźir*   *bskyed* / (PT 1287)
     end   four:TERM   PFV:extend

[The *bcan po*] extended the realm in four directions in order to, among others, impose taxes on other kings under the sun and subjugate petty kings as [his] subjects.

The construction also occurs in postpositional phrases with the relator nouns *naṅ* (PT 1075: 13; *apud* OTDO), *mdun* (PT 1084: 8; *apud* OTDO), and the honorific *spyan ṅa* (for *spyan sṅa in Or.8210/S.2228: C9; *apud* OTDO).

In terms of its historical development, the idiomatisation of *las scogs pa* followed the re-interpretation of the off-subject nominalisation with plural referents. In the original 'X *las scogs pa*', lit. "[humans] assembled from among X", the generally expressed X (see exx. (39)–(40)) was replaced by concrete entities so that the entire construction has been extended to groups whose members were explicitly listed like in (46):

(46)   *dgun+ɤdun*   *draɤī*   *bye+gror*   *blon+[č]e*
     winter+council(ABS)   Dra:GEN   Bye+gro:TERM   grand_councillor

     [čuṅ]+(11)+bzaṅ=daṅ /   *ɤbal+ldoṅ+cab=daṅ*   *blon*   *maṅ+pho+rǰe=daṅ* /
     Čuṅ+bzaṅ=COM   Ɤbal+ldoṅ=cab=COM   councillor   Maṅ+pho+rǰe=COM

     źaṅ ɤbriṅ+rcan=las   *scogs=pas*   *bs[du]s[t]e* / (Or.8212/187)
     źaṅ Ɤbriṅ+rcan=DEL   assemble:SPASS=NR:ERG   PFV:gather:GER

Grand councillor [Ɤbro] Čuṅ-bzaṅ [Ɤor-maṅ], [councillor] Ɤbal [Skyes-bzaṅ] Ldoṅ-cab, councillor [Čog-ro] Maṅ-pho-rǰe [Khyi-čhuṅ], and *źaṅ* Ɤbriṅ-rcan [Khyi-bu] convened the winter council at Bye-gro of Dra.

X *daṅ* Y *daṅ* Z *las scogs pa* could be rendered as "a group of X, Y, and Z". (46) is especially instructive for we know that councils were convened by concrete individuals whose names are regularly recorded in the OTA. Accordingly, the winter council of 747/8 was convened by exactly these four individuals listed in (46). Thus, splitting up of the collective into plural of individuals was the first step towards idiomatisation:

Off-SN: COLLECTIVE *las scogs pa* "a group of COLLECTIVE", lit. "[humans] assembled from COLLECTIVE"

ID.I: X *daṅ* Y *daṅ* Z *las scogs pa* "a group of X, Y, and Z"

When the number of elements was subsequently reduced to two, their interpretation as a group became disputable and so the meaning of *las scogs pa* must have been broadened to also house these cases. Moreover, X, Y, and Z, being nominal slots, could now be replaced by any type of NP, including nominalised clauses like in (42), (44), and (45). This process added to the semantic change in the construction for now not only physical entities, like humans or objects, but also abstract events could be enumerated.

Despite the semantic change and the shift towards idiomatisation, *las scogs pa* remained morphologically transparent to a certain degree even in CT. Namely, one encounters it in OLT with the idiomatic meaning and the form *la*(*s*) *scogs te* (cf. PT 1072: 55; ITJ 740: 284; ITJ 751: 38r3–4, 39r4); the replacement of *=pa* by *=ste* indicates that the verbal origin of *scogs* is still being recognised. Due to its differing morphosyntax I separate it from ID.I and label it ID.II. Its structure can be schematically presented as:

<div align="center">

X *las scogs te*
"X and the like; X, among others"

</div>

Despite the fact that ID.II rarely occurs in the OLT corpus, I was not able to discern any semantic difference between this construction and ID.I, both are highly idiomatised and headless. But an important contrast is their function in the host clause: only ID.I can take an argument role.

4.9   Summary

This brief summary provides a systematised overview of the OLT constructions with *scogs* as identified in this paper. Table 6 lists eight main constructions together with their subtypes.

Type	Construction	Etymological meaning	Set relations
PRED	S X=*las scogs*	S has been assembled from X	X = S
PAdvC.I	X=*las scogs=ste* S [...]	having been assembled from X, the S [...]	X ⊂ S
PAdvC.II	X=*las scogs=ste* Y [...]	Y such as X	
PostH-RC	[HEAD$_{ABS}$ [X=*las scogs=pa*]$_{RC}$]$_{NP}$	HEAD, that has been assembled from X, [...]	X ⊆ HEAD
R-disl	[[X=*las scogs=pa*]$_{NP.1}$ HEAD]$_{NP.2}$	those assembled from X, the HEAD [...]	X ⊆ HEAD
RCE	[[X=*las scogs=pa*]$_{NP.1}$ Y HEAD]$_{NP.2}$	those assembled from X, the HEAD [...] Y	X ⊂ HEAD
PreH-RC	[[[...] V=NR+GEN]$_{RC}$ HEAD]$_{NP}$	HEAD that V [...]	X = HEAD
Off-SN	[X=*las scogs=pa*]$_S$	those (*=pa*) assembled from X	X = S
ID.I	X=*las scogs=pa*	X, among others	
ID.II	X=*las scogs=ste*		

<div align="center">

Table 6. OLT constructions with *scogs*

</div>

The rightmost column supplies logical relations held between referents of the two constituents of the *scogs* constructions, the X and the HEAD or S. One could hope that these may

help us to establish whether the enumeration expressed by X was complete or open-ended. The latter case would necessitate the addition of "among others, and the like, etc." to the translation. The corpus examined in the study indicates that the relation of equality characterises complete enumeration, whereas that of strict inclusion points at open-ended enumeration. Unfortunately, two very popular constructions, PostH-RC and R-disl, remain ambiguous regarding the relations between the referents and so do not facilitate the interpretation of their enumerations. Needless to say, the relations, whatever they may be, are not intrinsic to the syntactic constructions themselves but are the natural outcome of the original semantics of the verb *scogs* that codes the action of forming plurality: "to assemble, to gather, etc."

The majority of the constructions are well-attested in OLT with other predicates as well. PostH-RC, R-disl, and RCE depend on each other and could have evolved historically only in this order but, apart from the PostH-RC, the other two have not been described for any Tibetic language thus far. The well-known participial meaning of *=ste* comes otherwise to the fore in clause linkage with coreferential subject, albeit in the prototypical construction it is the first clause (ending in *=ste*) that contains the overt subject.

Our understanding of the single constructions does not yet allow determining the pragmatic motivation behind the use of the one or the other in a given context. In particular, the establishment of the differences between the PAdvC, PostH-RC, R-disl, RCE, and PreH-RC could deliver important insights into cognitive-pragmatic processes underlying the selection of each of the constructions. DeLancey (1999: 244), followed by Zeisler (2011: 157), argued that the difference between the pre- and post-head constructions is that of restrictive vs non-restrictive relative clauses. However, in terms of their original motivations, it seems that the primary function of PostH-RC was to supply cataphoric grounding for new information (Givón 2001.2. 177f.) as this construction appears to have evolved in OLT out of paratactic constructions (Bialek In Preparation). Be that as it may, all the constructions require much more dedicated research in order for us to fully appreciate their contribution to the syntacto-pragmatic complexity and richness of OLT.

5 Distribution and chronology of constructions

Now that I have identified (hopefully) all constructions based on *scogs*, the question arises as to their distributional patterns in OLT. Are certain constructions more common in some texts than in others? Do certain constructions tend to take one particular variant form of the lexeme? Can their distribution contribute to dating or classifying OT texts? The following discussion shall help answer these and perhaps also other questions.

5.1 Distribution and relative chronology

Table 7 provides statistical counts for each construction based on the OTDO database and ordered according to their frequency:

Construction	n=345
Idiomatic I	177
R-dislocation	41
Participial adverbial clause proper	34
Post-head relative clause	27
Off-subject nominalisation	11
Idiomatic II	5
Participial adverbial clause idiomatic	3
Pre-head relative clause	3
Relative clause extraposition	2
Finite clause	1
Incomplete	28
Uncertain	13

Table 7. Statistics of *scogs*-constructions in OLT

28 examples are incomplete which means that some of their elements are missing, usually due to paper damage. These as well as the thirteen uncertain cases could not be classified. The uncertain examples come from passages that are not intelligible to me at the moment. Appendix A provides data on the distribution of the constructions in concrete texts.

The interpretation of the data is hindered by the fact that most of the texts are rather short and contain only a small number of relevant examples. 60 texts attest to a single occurrence of a construction with *scogs*, whereas only twenty provide five or more such examples. Among the texts with one construction, the idiomatic use of *scogs* (ID.I) prevails with 32 occurrences, followed by PostH-RC (7), PAdvC (4), R-disl (3), and Off-SN (1). The following remarks concentrate mainly on those 46 texts that contain more than one occurrence of a *scogs*-construction.

The high participation rate of idiomatic constructions with *scogs* may indicate either the advancement of idiomatisation throughout the period of OLT or a later date of texts that display some preference for this construction. According to data in Appendix A, texts that contain constructions considered more lexicalised (such as PAdvC.II and ID.II), also contain their prototypical variants. Thus, ID.II occurs only in texts with the standard idiomatic expression (ID.I), whereas idiomatic PAdvC occurs mainly when proper PAdvC is likewise present (exception: Źwa E).

Somehow surprisingly, PAdvC appears to be in complementary distribution to PostH-RC with only one text, PT 1287, attesting to both. Table 8 illustrates their distribution.

	PAdvC.I	PAdvC.II	PostH-RC
PT 16/ITJ 751			3
PT 239			2
PT 1068			5
PT 1089			2

PT 1096			1
PT 1287	1	1	3
ITJ 740			1
ITJ 750			1
S.07133			2
Ldan 1			1
PT 37	4		
PT 981	2		
PT 986	6		
PT 1042	6		
PT 1071	2		
PT 1073	1		
PT 1111	1		
PT 1283	1		
Or.15000/256	1		
Or.8212/187	1		
Skar	1		
Żol	1		
Żwa W	2	1	

Table 8. Distribution of PAdvC and PostH-RC

PT 1287 is a heterogeneous composition, or rather a patchwork, and as expected PAdvC occurs in different parts of the text than PostH-RC:

PostH-RC	l. 11	Chapter on Dri-gum
PostH-RC	l. 188	Chapter on Gnam-ri Slon-mchan
PostH-RC	l. 193	Chapter on Gnam-ri Slon-mchan
PAdvC.I	l. 333	Eulogy of Khri ɣdus-sroṅ
PAdvC.II	l. 335	Eulogy of Khri ɣdus-sroṅ

The chronological motivation behind the pattern is supported by the observation that PostH-RC is attested in ITJ 750 for the year 705/6, whereas PAdvC in Or.8212/187 comes from the year 756/7. It seems then that PostH-RC is an older construction prevailing maybe until the early 8th c. but subsequently supplanted by PAdvC. The former remained in use most probably as an archaism for it occurs readily in ritual and religious texts (PT 16/ITJ 751, PT 239, PT 1068, ITJ 740). From the mid-8th c. onward, PAdvC seems to have been preferred in official style of the Central Tibetan inscriptions (Skar, Żol, Żwa W), in legal texts (PT 1071, PT 1073), and in translations or adaptations from foreign languages (PT 981, PT 986, PT 1283). Although PAdvC seems to be a younger construction, this fact alone does not suffice as a basis for dating the OT texts.

In the section on R-dislocation, I have argued that this construction most probably developed from PostH-RC. Now, by comparing Table 8 with Table 9 one notices that texts that have PAdvC but not PostH-RC do contain R-disl (PT 37, PT 1073, PT 1111, PT 1283, Or.8212/187).

	PAdvC.I	R-disl
PT 37	4	1
PT 1073	1	2
PT 1111	1	1
PT 1283	1	2
PT 1287	1	4
Or.8212/187	1	2
Skar	1	1
PT 16		3
PT 149		1
PT 239		4
PT 2204c		1
ITJ 151		1
ITJ 737-1		1
ITJ 737-3		2
ITJ 740		2
ITJ 751		6
Khrom		1
Lčaṅ		1
Ldan 4		1
Źwa E		1
PT 981	2	
PT 986	6	
PT 1042	6	
PT 1071	2	
Or.15000/256	1	
Źol	1	
Źwa W	2	

Table 9. Distribution of PAdvC.I and R-disl

It occurs that R-disl evolved out of PostH-RC and slowly began to replacing it.[35] Parallel to this development, PAdvC gained in importance, maybe because the displacement of PostH-RC by R-disl has left some pragmatic gap in the system which had to be closed. Figure 2 presents two chains of this parallel development.

---

[35] Four texts include both PostH-RC and R-disl (PT 16/ITJ 751, PT 239, PT 1287, ITJ 740) and therefore could be examined in a future study for syntacto-pragmatic motivations behind the two constructions.

$$\text{PRED} \longrightarrow \text{PostH-RC} \longrightarrow \text{R-disl} \longrightarrow \text{RCE}$$
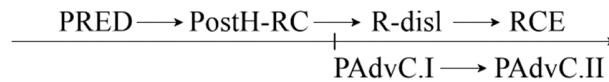$$\text{PAdvC.I} \longrightarrow \text{PAdvC.II}$$

Figure 2. Parallel development of R-disl and PadvC from PostH-RC

Furthermore, since PreH-RC with *scogs* does not surface prominently in the corpus (even though this is the primary relativisation strategy in OLT), I assume that this construction developed not independently but out of R-disl. It quickly gained in popularity to become, beside ID, the main *s(c)ogs*-construction in CT.

Off-SN and ID are headless constructions, consisting of only one but highly lexicalised element. Since Off-SN by means of the =*pa* clitic has been inherited from a parent language of PT and must have been common to many historical Trans-Himalayan languages (see DeLancey 2011: 345; Bialek 2021b: 275ff.), it must have been present in the language from its very beginning. By virtue of their shared formal characteristics, I assume that the ID-constructions derived from Off-SN. The sporadic occurrence of Off-SN in the corpus, restricted to texts related to judicial practice (PT 1071, PT 1072, PT 1089, ITJ 739), might have resulted from its general replacement by ID in the language. It has been preserved in specialised, juridical language that is frequently characterised by depersonalised expressions, for which Off-SN is best suited.

In addition to the above proposed chains of mutually dependent constructions a further chronological relation may exist between PostH-RC and ID. Namely, these two constructions have complementary distribution in Central Asian contracts:

PostH-RC   [*dpaṅ po/gñer pa*]$_{\text{HEAD}}$ [X *la scogs pa*]$_{\text{RC}}$ *ɣi dpaṅ rgyas btab pa*[36]
                         [contract] sealed with witness seals of witnesses/*gñer pa*s, who were assembled from X[37]

ID   (*dpaṅ la*) [X *la(s) scogs pa*]$_{\text{ID}}$*ɣi dpaṅ rgyas* [...]$_{\text{ERG}}$ *btab pa*[38]
                         (concerning the witnesses,) [contract] sealed with witness seals of [persons] such as X

The difference between PostH-RC and ID could be described as demotion of the HEAD from the subject position in the first to oblique in the second construction, but not all examples of ID contain the element *dpaṅ la*. Accordingly, a more accurate explanation would be the elision of the HEAD constituent from PostH-RC. This provides a second path for the idiomatisation: 1. Off-SN → ID (see above); and 2. PostH-RC → ID. The temporal precedence of PostH-RC over ID can also be inferred from the semantics of the constructions; to wit, a list of people sealing an official document must by definition be closed – only persons whose names are overtly stated are involved in the contract, nobody else. PostH-RC was better suited to fulfil this formal condition in official documents and first with its obsolescence could ID step in. Since this alternation between PostH-RC and ID has so far been observed only in Central Asian contracts, one may speculate that ID,

---

[36] Attested in PT 1086: 5–6; ITJ 914: 4–5; ITJ 1018: b10–b11; ITJ 1274: 6–7; ITJ 1374: 5–6; S.7133: a9–a10, b7–b8.

[37] Takeuchi (1995: 259) describes the function of *gñer pa* as "officials whose task was to levy or collect taxes, tributes, penalties etc."

[38] Attested in PT 1098: 11–13; PT 1101: 9–11; PT 1115: 7–8; PT 1162: 6–7; PT 1166: 6–8; PT 1297_1: 10–12; PT 1297_3: 9–11; PT 1297_6: 1.5–1.6, 2.7–2.8; PT 2124: c2–c3; ITJ 844: 7–8; ITJ 850: 9–11; ITJ 1379: r5–6; Or.15000/256: 6; Or.15000/486: r2–r4; Or.15000/530: r7–r9; Or.8212/194a: 4–5; Or.8212/194b: 3–4.

gaining in popularity, was generalised from other textual contexts and locally facilitated the ousting of PostH-RC.

In conclusion, the constructions discerned can be grouped in three sets as shown in Figure 3:
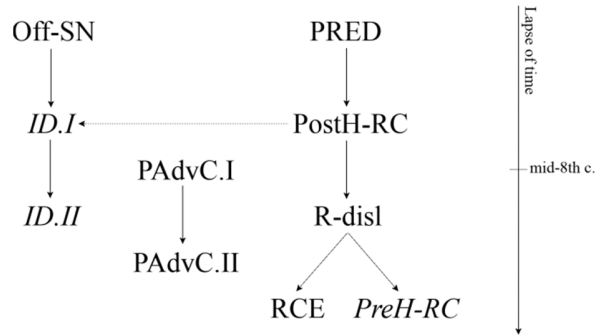
Figure 3. Relative chronology of *scogs*-constructions

The occurrence of a subsequent stage does not presuppose the disappearance of the preceding construction. The diagram only illustrates the mutual structural and semantic dependence of single constructions on each other. Italicised constructions prevail in CT, although some of the remaining ones can likewise be sporadically encountered. For instance, PAdvC.I surprisingly occurs in Mi-la Ras-pa's songs: *rma bya daṅ ɣ̌ol mo la sogs te mǰes paɣi bya* (*apud* WtS.21: 401a, s.v. *ɣ̌ol mo*), lit. "having been collected from peacocks and finches, birds that are pretty [sing]".

5.2   las scogs *vs* la scogs

Another factor that can allow us to better understand the chronology of the constructions and maybe that of the texts themselves is the form of the oblique marker. Namely, three variants are known: *las scogs*, *la scogs*, and *lascogs*.[39] In addition, *scogs* devoid of the case marker is likewise attested several times but preponderantly in metrical passages. Whereas in *la(s) scogs* and *scogs* the main syllable can be represented by various variants (*scogs*, *bscogs*, *bsogs*, *rcogs*, *sogs*, or *gsog*), the linked spelling *lascogs* only occurs if the second syllable has the preconsonantal *s-* in absolute onset. Appendix B summarises the distribution of the variants across the texts.

The distribution of the variants is almost complementary: twenty-nine texts have only *las scogs*, forty-two texts have only *la scogs*, and seven texts have both forms (see Appendix B). The co-occurrence of *las scogs* or *la scogs* with *lascogs* may be fortuitous but in any way is of no relevance for the following discussion. Instead I shall concentrate on the co-occurrence of *las scogs* with *la scogs* in a text and their potential correlation with particular *scogs*-constructions. Table 10 illustrates the distribution of the variants across the constructions in seven texts that attest to both *las scogs* and *la scogs*.

---

[39] This analysis disregards the syllable that may follow *scogs*. For an overview of OLT and CT variants of *la(s) scogs*, see Zeisler (2016: 468ff.).

	las scogs	*la scogs*	*lascogs*
PRED			PT 1042
PAdvC.I	PT 37	PT 1071	
	PT 1042		
PAdvC.II			
PostH-RC	PT 16	PT 239	
		ITJ 751	
R-disl	PT 16		
	PT 37	PT 149	
	PT 239		
	ITJ 751		
RCE			
Pre-H-RC	PT 16	ITJ 751	
Off-SN	PT 1071		
ID.I	PT 16	ITJ 737-2	
	PT 37		
	PT 149		
	PT 239		
	ITJ 751		PT 1042
	PT 1071		
ID.II		PT 1071	
		ITJ 751	
In	ITJ 737-2	ITJ 751	
Un	PT 16	PT 1071	
	PT 239		

Table 10. Distribution of *las scogs* and *la scogs*

This juxtaposition yields a true mosaic of combinations. PT 1042 has both *las scogs* and *la scogs* in PAdvC.I, PT 16/ITJ 751 and PT 239 in R-disl, PT 1071 in Off-SN, and PT 37 and PT 239 in ID.I. The variations in the use of *las scogs* and *la scogs* are difficult to account for and apparently do not correlate with the type of construction. In PT 16/ITJ 751, *la scogs* is used in the first occurrence of the construction and consequently between ll. 37v1 and 39r4. It is possible that in a draft version of the text the latter fragment was written by a different person than the rest of the text. The single *la scogs* in PT 37 must be a scribal error. The distribution of *las scogs* and *la scogs* in PostH-RC and ID of Central Asian contracts (see Table 11) suggests a higher proportion of and increase in the use of *la scogs* with ID, a construction presumably derived from PostH-RC in this genre (see the preceding section).

	las scogs	*la scogs*	*lascogs*
PostH-RC	2	2	3
ID	2	5	11

Table 11. Distribution across PostH-RC and IDs

la scogs seems to have gained the upper hand in the first half of the 9th c. and was thereafter adopted in official writing of Central Tibetan chancelleries. *Terminus ante quem* for the acceptance of *la scogs* as the standard spelling can be set in 822 – the date of the ST Treaty inscription. This, however, does not mean that the variant might not have been occurring in texts composed prior to that date owing to the influence from the spoken language. The final *-s* after a vowel was still pronounced when the text of the ST Treaty was composed (Preiswerk 2014: 66f.). Hence, *las > la* was not a consequence of the sound change -V*s* > -V but rather resulted from the simplification of word-internal consonantal cluster *-s+sc-* [-s+sts-] following the popularisation of the *scogs*-constructions, especially ID.I. Idiomatisation contributed to the perception of *las scogs* in ID as one lexical unit and so shifted the word boundary from *las#scogs* to #*las₊scogs*.

Even though these insights do not suffice to date texts with any certainty, the probability for an original *official* text with *la scogs* to have been composed after ca. 820 is relatively high. We know however that some of the texts in their actual version contain passages that must have been originally composed much earlier but underwent revision and/or copy-editing. This is certainly the case with the OTC but most probably also with the so-called *Ɣa-ža Annals* (ITJ 1368). Here the use of *la scogs* points to a date of the revision rather than to that of the composition – a distinction that must not be neglected. On the other hand, *las scogs* itself cannot be taken as a marker of antiquity since it might have continued in compositions that intentionally applied archaic style. This form is sporadically encountered even in texts unanimously dated to the CT period (Zeisler 2016: 469); a search on BCRD yields three instances in Sde-dge *Bstan ɣgyur* and BDRC adds further examples from other editions of the canon. Zeisler's speculation that *las scogs* might be a secondary form derived by consonant migration (*la scogs > lascogs > las scogs*; ibid., p. 470f.; see also Zemp (2018: 71, fn. 34)) is contradicted by the distribution of the forms; for instance, in Central Tibetan inscriptions or in the OTA *las scogs* has clearly older attestations. Moreover, in OLT verbs of gathering take the allative case marking for the entity around which one gathers like *spur la* "around the body" in (3). Thus, *"X Y la scogs'* (e.g., in (8)) would have to be interpreted as !"X was assembled around Y", which for obvious reasons is false in all the examined contexts.

6 Conclusions

The paper is the first attempt at reconstructing the word-family of *scogs* and detecting its manifold uses in OLT texts. The morpheme is traced back to the v4-stem of the verb √stsog (CT v1 *gsog*), lit. "to cause to assemble", itself derived from the verb root √tsog by means of the causative prefix *s-*. After discussing its probable cognates and demonstrating historical links between them, I have examined the constructions attested in the OLT corpus which contain the morpheme. Eight distinct constructions have been discerned: finite clause, participial adverbial clause (I & II), post-head relative clause, R-dislocation, relative clause extraposition, pre-head relative clause, off-subject nominalisation, and idiomatic phrase (I & II). Besides discussing the syntactic features of the constructions I have also considered the referential relationships between their constituents with the hope that this could help resolve the character of enumerations, complete versus open-ended. The obtained results are unfortunately not clear-cut in the case of post-head relative clauses and R-dislocation but seem to be unambiguous for the remaining constructions, at least as far as the OLT corpus is concerned. In contrast to my initial asumption when I first started collecting the data for the paper, it appears that the highly idiomatised usage of *scogs* known from CT is already well

attested in the oldest available records of Tibetic languages. This is yet another hint at the remarkable capability of the language to alterate rapidly when trying to keep pace with the dynamics of social and political changes during the eventful period of the Tibetan Empire.

ABBREVIATIONS

√	reconstructed verb root
!	historically/logically impossible form or translation
#	word boundary
=	clisis
+	morpheme boundary within a lexical word
~	reduplication
.	grammatical categories of one morpheme
:	morpheme boundary not shown in the original
_	one lexical word
{*ba*}	text corrected to *ba*
[b]	text reconstructed
b̵	text deleted
1SG	1st person singular
Ɣbis 2	Ɣbis-khog 2 inscription
Ɣphyoṅ	Ɣphyoṅ-rgyas inscription
A-CAUS	autocausative
ABS	absolutive
ADD	additive
ALL	allative
APPL	applicative
APPV	appositive
Bal	Balti
Brag A	Brag-lha-mo A inscription
Bsam	Bsam-yas inscription
CAUS	causative
Ch.	Chinese
Chik	Chiktan
COM	comitative
COP	copula
CT	Classical Tibetan
DEL	delative
DEM	demonstrative
DIM	diminutive
DPASS	dynamic passive
DUR	durative
E	east-facing inscription
EL	elative
Eng.	English

ERG	ergative
EOT	Early Old Tibetan
FNL	sentence final
FOC	focus
GEN	genitive
GER	gerund
HML	humble
HON	honorific
ID	idiomatic construction
IMP	imperative
INDF	indefinite
INESS	inessive
INTR	intransitive
ITJ	IOL Tib J
Kar	Kargil
Khri	Inscription at the tomb of Khri Lde-sroṅ-brcan
Khrom	Khrom-čhen inscription
Lčaṅ	Lčaṅ-bu inscription
Ldan 1	Ldan-ma-brag 1 inscription
Ldan 2	Ldan-ma-brag 2 inscription
Ldan 4	Ldan-ma-brag 4 inscription
LH	Later Han Chinese
Lho	Lho-brag inscription
MC	Middle Chinese
Mdzo	Mdzorganrabar
N	north-facing inscription
NEG	negation
NP	noun phrase
NPFV	imperfective
NR	nominaliser
O	object
OC	Old Chinese
OCM	Minimal Old Chinese
Off-SN	off-subject nominalisation
OLT	Old Literary Tibetan
OTA	*Old Tibetan Annals*
OTC	*Old Tibetan Chronicles*
Or.	Oriental Collections of the British Library
P.	Pāli
PAdvC	participial adverbial clause
PB	Proto-Bodish
PFV	perfective
PL	plural
PostH-RC	post-head relative clause
PreH-RC	pre-head relative clause

PT	1. Pelliot tibétain; 2. Proto-Tibetan
PTH	Proto-Trans-Himalayan
Q	oblique
QUOT	quotative
R-disl	R-dislocation
RC	relative clause
RCE	relative clause extraposition
RDP	reduplication
S	subject
Skar	Skar-čuṅ inscription
SPASS	stative passive
Skt.	Sanskrit
TERM	terminative
TR	transitive
Treaty	Sino-Tibetan Treaty inscription
Tsha	Tshangra
V	verb
v1, v2, v3, v4	verb stems
W	west-facing inscription
Wan	Wanla
WDro	Western Drokpas
LT	written Tibetan
Źol	Źol inscription
Źwa	Źwaɣi-lha-khaṅ inscription
Źwa F	Fragmentary inscription near Źwaɣi-lha-khaṅ

References

Baxter, William H.; and Sagart, Laurent. 2014. Old Chinese reconstruction, version 1.1 (20 September 2014). Accessed 05.08.2017.

BCRD = The Buddhist Canons Research Database: http://www.aibs.columbia.edu/databases/New/index.php

BDRC = Buddhist Digital Resource Center: https://www.tbrc.org/

Bialek, Joanna. 2018a. *Compounds and compounding in Old Tibetan. A corpus based approach* 2 vols. Marburg: Indica et Tibetica.

Bialek, Joanna. 2018. "The Proto-Tibetan clusters *s*L- and *s*R- and the periodisation of Old Tibetan". *Himalayan Linguistics* 17.2: 1–50. doi: https://doi.org/10.5070/H917238831.

Bialek, Joanna. 2020a. "Old Tibetan verb morphology and semantics: An attempt at a reconstruction". *Himalayan Linguistics* 19.1: 263–346. doi: https://doi.org/10.5070/H919145017.

Bialek, Joanna. 2020b. "Towards a standardisation of Tibetan transliteration for textual studies". *Revue d'Etudes Tibétaines* 56: 28–46.

Bialek, Joanna. 2021a. "Comments on Jacques' 'The directionality of the voicing alternation in Tibetan'". *Journal of the Southeast Asian Linguistics Society* 14.1: xii–xxiii.

Bialek, Joanna. 2021b. "Social roots of grammar: Old Tibetan perspective on grammaticalization of Kinterms". In: Lange, Diana; Ptáčková, Jarmila; Wettstein, Marion; and Wulff, Mareike (eds.), *Crossing Boundaries. Tibetan Studies Unlimited*, 253–288. Prague: Academia Publishing House.

Bialek, Joanna. 2022. "The concept of power in Old Literary Tibetan. A lexicological analysis of *čhab srid*." *Oriental Studies: Local and Global Perspectives* 45: 24–46.

Bialek, Joanna. (In Preparation). *Old Tibetan annals. A comprehensive text grammar*.

BYD = Rnam-rgyal Che-riṅ. 2001. *Bod yig brda rñiṅ chig mjod*. Pe cin: Kruṅ goyi bod rig pa dpe skrun khaṅ.

CDTD = Bielmeier, Roland; Haller, Felix; Häsler, Katrin; Huber, Brigitte; and Volkart, Marianne (eds.). 2013 (manuscript). *Comparative dictionary of Tibetan dialects*.

CDTD.V = Bielmeier, Roland; Häsler, Katrin; Haller, Chungda; Haller, Felix; Hein, Veronika; Huber, Brigitte; Volkart, Marianne; Preiswerk, Thomas; Ngawang Tsering; Widmer, Manuel; and Zemp, Marius. 2018. *Comparative dictionary of Tibetan dialects (CDTD). Volume 2: Verbs*. Berlin: De Gruyter Mouton.

Coblin, W. South. 1986. *A Sinologist's handlist of Sino-Tibetan lexical comparisons*. Nettetal: Steyler Verlag.

Csoma de Kőrös, Alexander; and Sangs-Rgyas Phun-Tshogs. 1834. *Essay towards a dictionary, Tibetan and English*. Calcutta: Baptist Mission Press.

D = Das, Sarat Chandra. 1902. *A Tibetan-English dictionary with Sanskrit synonyms*. Calcutta: The Bengal Secretariat Book Depot. [Reprint: Delhi: Motilal Banarsidass, 2000].

de Jong, Jan Willem. 1989. *The story of Rāma in Tibet: text and translation of the Tun-huang manuscripts*. Stuttgart: Franz Steiner Verlag Wiesbaden.

DeLancey, Scott. 1981. "An interpretation of split ergativity and related patterns". *Language* 57.3: 626–657.

DeLancey, Scott. 1999. "Relativization in Tibetan". In: Yadava, Yogendra P.; and Glover, Warren William (eds.), *Topics in Nepalese linguistics*, 231–249. Kathmandu: Royal Nepal Academy.

DeLancey, Scott. 2011. "Finite structures from clausal nominalization in Tibeto-Burman". In: Yap, Foong Ha; Grunow-Hårsta, Karen; and Wrona, Janick (eds.), *Nominalization in Asian Languages: Diachronic and typological perspectives*, 343–359. Amsterdam: John Benjamins.

Dotson, Brandon; and Helman-Ważny, Agnieszka. 2016. *Codicology, paleography, and orthography of early Tibetan documents*. Wien: Arbeitskreis für tibetische und buddhistische Studien, Universität Wien.

Gallica: http://gallica.bnf.fr

Givón, Talmy. 2001. *Syntax: An introduction* 2 vols. Amsterdam: John Benjamins Publishing.

Gs = Goldstein, Melvyn C. 2001. *The new Tibetan-English dictionary of modern Tibetan*. Berkeley: University of California Press. [Reprint: New Delhi: Munshiram Manoharlal, 2004].

Hill, Nathan W. 2010. *A lexicon of Tibetan verb stems as reported by the grammatical tradition*. München: Bayerische Akademie der Wissenschaften.

Hill, Nathan W. 2013. " 'Come as lord of the black-headed' - an Old Tibetan mythic formula". In: Cüppers, Christoph; Mayer, Robert; and Walter, Michael (eds.), *Tibet after Empire. Culture, Society and Religion between 850-1000*, 169–179. Lumbini: Lumbini International Research Institute.

Hill, Nathan W. 2021. "The envoys of Phywa to Dmu (PT 126)". *Revue d'Etudes Tibétaines* 60: 84–143.

Huber, Brigitte. 2003. "Relative clauses in Kyirong Tibetan". *Linguistics of the Tibeto-Burman Area* 26.1: 1–14.

IDP = International Dunhuang Project: http://idp.bl.uk/

J = Jäschke, Heinrich August. 1881. *A Tibetan-English dictionary*. London. [Reprint: New York: Dover Publications, 2003].

Jacques, Guillaume. 2021. "The directionality of the voicing alternation in Tibetan". *Journal of the Southeast Asian Linguistic Society* 14.1: 32–38.

Karmay, Samten G. 1998-2005. *The arrow and the spindle: Studies in history, myths, rituals and beliefs in Tibet* 2 vols. Kathmandu: Mandala Book Point.

Lin, You-Jing. 2014. "Thebo". In: Sun, Jackson T.-S. (ed.), *Phonological profiles of little-studied Tibetic varieties*, 215–267. Taipei: Academia Sinica.

OED = Stevenson, Angus, and Waite, Maurice (eds.). 2011. *Concise Oxford English Dictionary*. Oxford: Oxford University Press.

OTDO = Old Tibetan Documents Online: http://otdo.aa.tufs.ac.jp/

OTI = Iwao, Kazushi; Hill, Nathan W.; and Takeuchi, Tsuguhito. 2009. *Old Tibetan inscriptions*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Preiswerk, Thomas. 2014. "Die phonologie des alttibetischen auf grund der chinesischen Beamtennamen im chinesisch-tibetischen Abkommen von 822 n. Chr.". *Zentralasiatische Studien* 43: 7–158.

RKTS = Resources for Kanjur & Tanjur Studies: https://www.istb.univie.ac.at/kanjur/rktsneu/sub/index.php

Schuessler, Axel. 2007. *ABC etymological dictionary of Old Chinese*. Honolulu: University of Hawaii Press.

STEDT = Sino-Tibetan Etymological Dictionary and Thesaurus: http://stedt.berkeley.edu/~stedt-cgi/rootcanal.pl

Takeuchi, Tsuguhito. 1985. "A Passage from the *Shi Chi* in the *Old Tibetan Chronicle*". In: Aziz, Barbara N.; and Kapstein, Matthew (eds.), *Soundings in Tibetan civilization*, 135–146. Delhi: Manohar.

Takeuchi, Tsuguhito. 1995. *Old Tibetan contracts from central Asia*. Tokyo: Daizo Shuppan.

Takeuchi, Tsuguhito. 2004. "The Tibetan military system and its activities from Khotan to Lob-nor". In: Whitfield, Susan (ed.), *The silk road. Trade, travel, war and faith*, 50–56. Chicago: Serindia Publications.

Uray, Géza. 1954. "Duplication, gemination and triplication in Tibetan". *Acta Orientalia Academiae Scientiarum Hungaricae* 4: 177–244.

Uray, Géza. 2007. "The Old Tibetan sources of the history of Central Asia up to 751 AD: A survey". In: McKay, Alex (ed.), *The History of Tibet*, 118–141. London: RoutledgeCurzon.

WALS = The World Atlas of Language Structures: https://wals.info/

WtS = Franke, Herbert; Hartmann, Jens-Uwe; and Höllmann, Thomas O. (eds.). 2005. *Wörterbuch der tibetischen Schriftsprache*. München: Verlag der Bayerischen Akademie der Wissenschaften.

Zeisler, Bettina. 2011. "For love of the word: a new translation of Pt 1287, the Old Tibetan Chronicle, chapter I." In: Imaeda, Yoshiro; Kapstein, Matthew T.; and Takeuchi, Tsuguhito (eds.), *New studies of the Old Tibetan documents: Philology, history and religion*, 97–213. Tokyo: Tokyo University of Foreign Studies.

Zeisler, Bettina. 2016. "las.stsogs etc. – On internal cues for dating Old Tibetan documents". *Zentralasiatische Studien* 45: 467–491.

Zemp, Marius. 2018. *A grammar of Purik Tibetan*. Leiden: Brill.

Joanna Bialek
jbialek108@gmail.com

Appendix A

The table presents all occurrences of *scogs* in the OLT corpus grouped according to the type of construction in which they are attested.

	PRED	PAdvC.I	PAdvC.II	PostH-RC	R-disl	RCE	Pre-H-RC	Off-SN	ID.I	ID.II	In	Un	Total
PT 1071		2						8	24	1		5	40
Or.8212/187		1			2				14				17
PT 986		6							10				16
PT 1287		1	1	3	4				6				15
ITJ 151					1				14				15
PT 16				2	3		2		4			1	12
PT 239				2	5				3			1	11
ITJ 751				1	6		1		1	2			11
PT 37		4			1				5				10
PT 1047									6			3	9
PT 1042	1	6							1				8
PT 1283		1			2				4		1		8
ITJ 740				1	2				3	1			7
Skar		1			1	2			3				7
PT 1072								1	4	1			6
PT 1068				5									5
PT 1089				2				1	2				5
PT 1096				1					4				5
ITJ 750				1					4				5
Źwa W		2	1						2				5
PT 981		2							2				4
ITJ 737-2									1		3		4
Khri											4		4
PT 149					1				2				3
PT 1073		1			2								3
PT 1078bis									3				3
PT 1084									3				3
PT 2204c					1				1		1		3
ITJ 737-3					2				1				3
ITJ 1368											3		3
Or.15000/481									3				3
Lčań					1				2				3
Źol		1									2		3
PT 1111		1			1								2

PT 1297_2							1		1		2
PT 1297_3							2				2
PT 1297_6							2				2
ITJ 734									2		2
ITJ 737-1				1			1				2
ITJ 844							2				2
Or.15000/256	1						1				2
Or.15000/426							2				2
S.07133			2								2
Khrom				1					1		2
Ldan 2							2				2
Źwa E		1		1							2
PT 126			1								1
PT 366							1				1
PT 1043				1							1
PT 1051							1				1
PT 1060										1	1
PT 1075							1				1
PT 1086			1								1
PT 1087							1				1
PT 1088_1							1				1
PT 1092							1				1
PT 1094							1				1
PT 1095							1				1
PT 1098							1				1
PT 1101							1				1
PT 1115							1				1
PT 1162							1				1
PT 1166							1				1
PT 1285									1		1
PT 1286			1								1
PT 1290	1										1
PT 1297_1							1				1
PT 1297_4							1				1
PT 2124							1				1
PT 2125									1		1
ITJ 474	1										1
ITJ 504	1										1
ITJ 738										1	1
ITJ 739						1					1
ITJ 850							1				1
ITJ 914			1								1
ITJ 1018			1								1

ITJ 1274									1			1
ITJ 1374				1								1
ITJ 1375		1										1
ITJ 1379									1			1
Or.15000/123										1		1
Or.15000/138										1		1
Or.15000/229										1		1
Or.15000/315										1		1
Or.15000/429										1		1
Or.15000/467									1			1
Or.15000/486									1			1
Or.15000/496									1			1
Or.15000/530									1			1
Or.8212/194a									1			1
Or.8212/194b									1			1
Or.8212/1529										1		1
Or.8212/1834c									1			1
S.02228									1			1
S.12243									1			1
Ɣbis 2									1			1
Ɣphyoṅ									1			1
Brag A										1		1
Bsam									1			1
EndCell											1	1
Ldan 1				1								1
Ldan 4					1							1
Lho									1			1
Treaty										1		1
Źwa F					1							1

Appendix B

The table presents the distribution of *las scogs*, *la scogs*, *lascogs*, and *scogs* in OLT. It includes also uncertain and incomplete passages unless the very position before *scogs* is missing as in 4 cases (*scogs* represents any of the attested variants of the syllable).

	las scogs	*la scogs*	*lascogs*	*scogs*
PT 16/ITJ 751	16	7		
PT 37	9	1		
PT 149	2	1		
PT 239	7	4		
PT 1042	1	3	4	
PT 1071	2	35	3	
ITJ 737-2	2	1		
PT 126	1			
PT 981	4			
PT 1047	9			
PT 1060	1			
PT 1072	6			
PT 1073	3			
PT 1075	1			
PT 1092	1			
PT 2124	1			
ITJ 504	1			
ITJ 734	1			1
ITJ 750	5			
ITJ 914	1			
ITJ 1274	1			
Or.15000/256	2			
Or.15000/315	1			
Or.8212/187	16		1	
Or.8212/1529	1			
Ɣphyoṅ	1			
Bsam	1			
Khri	2			
Lčaṅ	3			
Ldan 1	1			
Lho	1			
Skar	7			
Źol	3			
Źwa F	1			
Źwa W	5			
Źwa E	2			

PT 366		1		
PT 986		13	3	
PT 1043		1		
PT 1051		1		
PT 1084		3		
PT 1086		1		
PT 1088_1		1		
PT 1089		5		
PT 1094		1		
PT 1095		1		
PT 1096		5		
PT 1098		1		
PT 1111		2		
PT 1115		1		
PT 1283		8		
PT 1286		1		
PT 1287		15		
PT 1297_2		2		
PT 1297_4		1		
PT 2125		1		
PT 2204c		3		
ITJ 151		12	3	
ITJ 474		1		
ITJ 737-1		2		
ITJ 737-3		3		
ITJ 850		1		
ITJ 1368		3		
ITJ 1374		1		
Or.15000/123		1		
Or.15000/426		1	1	
Or.15000/429		1		
Or.15000/467		1		
Or.15000/481		3		
Or.15000/486		1		
Or.15000/530		1		
S.02228		1		
S.12243		1		
Ɣbis 2		1		
Khrom		2		
Ldan 2		2		
Ldan 4		1		
Treaty		1		
PT 1078bis			1	1

PT 1087			1	
PT 1101			1	
PT 1162			1	
PT 1166			1	
PT 1290			1	
PT 1297_1			1	
PT 1297_3			2	
PT 1297_6			2	
ITJ 740			7	
ITJ 844			2	
ITJ 1018			1	
ITJ 1375			1	
ITJ 1379			1	
Or.15000/138			1	
Or.15000/229			1	
Or.15000/496			1	
Or.8212/194a			1	
Or.8212/194b			1	
Or.8212/1834c			1	
S.07133			2	
Brag A			1	
EndCell			1	
PT 1068				5
PT 1285				1
ITJ 738				1
ITJ 739				1