

himalayan linguistics

A free refereed web journal and archive devoted to the study of the
languages of the Himalayas

Himalayan Linguistics

The syntactic constructions and the semantic classifications of Tibetan verbs

Di Jiang

Chinese Academy of Social Sciences

M. L. Li

Shanghai Normal University

ABSTRACT

This paper represents a departure from traditional Tibetan grammar in terms of the classification of verbs, for it constructs verb types and relevant syntactic rules based on syntax and semantics. The paper further distinguishes twelve types of Tibetan verbs according to the numbers of different arguments and the requirements of different syntactic properties. Therefore, the classification of the syntax and semantics of verbs permits a detailed and overall reflection of the syntactic framework of all sorts of all kinds of Tibetan clause constructions, including word orders, case markers and syntactic particles. All the findings of syntactic and semantic classifications of Tibetan verbs can be applied directly to build a Tibetan grammatical information dictionary which serves as a piece of infrastructure in Tibetan natural language processing.

KEYWORDS

Tibetan; verb classification with semantics and syntax; syntactic structure; syntactic markers

This is a contribution from *Himalayan Linguistics*, Vol. 15(1): 41–54.

ISSN 1544-7502

© 2016. All rights reserved.

This Portable Document Format (PDF) file may not be altered in any way.

Tables of contents, abstracts, and submission guidelines are available at
escholarship.org/uc/himalayanlinguistics

The syntactic constructions and the semantic classifications of Tibetan verbs

Di Jiang
Chinese Academy of Social Sciences

M. L. Li
Shanghai Normal University

1 Syntactic and semantic types of Tibetan verbs

This paper discusses the classification of Tibetan verbs in support of an effort to build a Tibetan machine readable dictionary of grammatical properties in accordance with verb types and grammatical constructions of modern Tibetan.¹ Meanwhile, it is anticipated that the framework of clause pattern with verbs will be used to build a system for automatic grammar processing.

A verb in a clause is the core, governing the whole clause. The constituents it dominates comprise subject, object and others. Fundamentally speaking, the verb, the soul of clause structure and feature, builds and decides basic clause structures. Clause structures built by verbs, however, are diverse. What factor ultimately determines them?

Grammarians are aware that numerous grammatical phenomena are closely related to semantic differences. Modern grammarians, therefore, set out to channel their attention to probing the correspondence between syntactic structure and semantic structure. The different construction patterns that different verbs dictate usually are attributable to the meanings of the verbs. Verbs with different meanings yield different syntactic and semantic properties of a clause. Thus, a variety of grammatical structures is constructed. From this perspective, verbs can be classified based on their syntactic and semantic properties and the corresponding different clause patterns.

This paper distinguishes 12 types of verbs in modern Tibetan, including stative verbs, action verbs, mental verbs, verbs of perception, verbs of change, verbs of motion, *verba dicendi*, copula verbs, verbs of possession, existential verbs, interactive verbs, and causative verbs, which are with their syntactic constructions:

1. Stative verbs (STA): NP+(NP+)VP_(STA)
2. Copular verbs (COU): NP+NP+V_(COU)

¹ The grammatical system used in this paper has been represented in a series of papers, in which there are 14 types of syntactic cases, 9 types of syntactic aspects, and 4 types of evidentiality, and other grammatical categories, structural particles, and morphological markers. There is no tense category. Cf. Jiang (2005) and Jiang *Lexicon and Morphology of Tibetan Language* (In Chinese, Peking University Press, forthcoming).

3. Existential verbs (EXI): NP+NP+[LOC]+VP_(EXI)
4. Verbs of Possession (VOP): NP+[POS]+NP+V_(VOP)
5. Verbs of Change(CHA): NP+VP_(ADJ+V)+ [COP]+VP_(CHA)
6. Perceptive Verbs(PER): NP+[AG]+NP+VP_(PER)
7. Directional Verbs(DIR):
 - NP+VP_(DIR)
 - NP+VP +VP_(DIR)
 - NP+VP+[TAP]+VP_(DIR)
8. Action Verbs (ACT) :
 - NP+VP_(ACT)
 - NP+NP+VP_(ACT)
 - NP+NP+[DAT]+NP+VP_(ACT)
 - NP+[AG]+NP+NP+[FAT]+VP_(ACT)
9. Mental Verbs (PSY): NP+[AG]+NP+[OBJ]+VP_(COG)
10. *Verba dicendi* (NAR) :
 - NP+[AG]+{NP+VP}+VP_(NAR)
 - {NP+VP}+NP+[AG]+VP_(NAR)
11. Interactive Verbs (REL): NP+NP+[ITP]+VP_(REL)
12. Causative Verbs (CAV): NP+[AG]+NP+[POS]+ (NP +)VP+[CAU]+VP_(CAV)

In addition, Tibetan verbs can be classified into a number of subcategories, like receiving verbs, equivalent verbs, addressing verbs, giving verbs and the like, each of which is distinguished by argument and case marking. There are other typical structures relating to the grammar and meaning of a clause, such as comparative clauses and clauses describing the weather etc. The traditional Tibetan grammatical classification of verbs (S.J. Hu 2002), including categories such as transitive and intransitive verbs, automatic and causative verbs, volitional and nonvolitional verbs, is incapable of providing a thorough and detailed reflection of all sorts of clause structures. As a result, they can be used as an overarching classification which applies to both syntactic and semantic types.

2 Types of Tibetan verbs and a description of their grammatical patterns

This section describes verbs and their patterns, providing the basic structures and case marking patterns and illustrating these with examples.

1. Stative Verbs (STA): Stative verbs mainly signify states. In contrast to clauses with adjectival predicates (including descriptive clauses composed of copulative verbs and adjectives) stative verbs are generally used when indicating qualities. Some stative verbs are transitive verbs and others intransitives; the basic construction is NP+(NP+) VP (STA), of which the subject and object are not suffixed with case markers.

The most frequently used stative verbs are ལྷན་ (snyun) ‘to be ill’, བཟི་ (bzi) ‘to be drunk’, ཞེ་ (shi) ‘to be dead’, ལྷེ་ (skye) ‘to be born’, ལུབ་ (nub) ‘to be sunk’, ཡལ་ (yal) ‘to be gone’, etc.² There are numerous such verbs.³ For example,

- (1) ད་ ཁོ་ཚོ་བཟི་པ་རེད།
 da kho-tsho bzi-pa-red.
 Now 3pl be-drunk-REA
 ‘Now they are drunk.’⁴
- (2) ལྷོ་བཟང་ལགས་ལྷན་ཚད། འཛིན་གར་ཚོགས་གཤམ་ལེབས་བྱེ་མི་འདུག།
 blo-bzang-lags snyun tsang, vdzin-grwa tshogs gag phebs
 Lopsang was-ill since class meeting TAP come
 thub-kyi-mi-*vdug*.
 be-able-NEG-DUR
 ‘Lopsang was ill, so he cannot come to the class.’

There are various subcategories of verbs denoting states. For instance, verbs denoting ‘deficiency’ include དགོན་ (dgon) ‘to lack’, དགོས་ (dgos) ‘to be in need’, etc.; verbs denoting the meaning of ‘release’ are མཚོན་ (mtshon) ‘to express, to deliver’, མངོན་ (mngon) ‘to turn up, to emerge’, འདོན་ (vdon) ‘to take out’, འཛོན་ (vthon) ‘to arise, to be known’, ཤར་ (shar) ‘to rise, to rise up’; verbs denoting ‘delay’ are འགོར་ (vgor) ‘to postpone, to detain’, འགྱང་ (vgyang) ‘to defer, to retard’, etc. For the purposes of illustration one example suffices,

- (3) དེ་རིང་གི་ ཉི་མ་དེ་ཤར་ནས་ ཤར་སོང་། ། དམ་ཡང་ན་ ལུབ་ནས་ ཤར་སོང་།
 de-ring-gi nyi-ma de shar-nas shar-song ngam yang-na
 Today-GEN sun the east-ABL rise-PEF INT otherwise

 nub-nas shar-song
 west-ABL rise-PEF
 ‘Did today’s sun rise from the east or did it rise from the west?’

² We have compiled quite a big text-database on Tibetan language for more than twenty years, some of which have been tagged with grammatical markers. Cf. Jiang, Long and Zhang (2005), Long (2012). Most of the examples used in the paper are taken from this text-database.

³ The transliteration used in the paper are IEA-scheme, which is a pre-GB scheme in China. Cf. Jiang, D.; Long CJ. *On Characters of Tibetan Writing System : alphabetic characters, pronunciations, ISO codes, sorting orders, picture symbols, and transliterations*. Beijing: Social Sciences Academic Press, 2010. (江荻, 龙从军(著) 2010《藏文字符研究: 字母、读音、编码、字频、排序、图形、拉丁字母转写规则研究》, 社会科学文献出版社). Also in: *Transliteration of Tibetan into Latin Characters in Text*. In: Government Greenbook 2015, by Ministry of Education of China. Beijing: The Commercial Press, 2015. (藏文拉丁字母转写方案(草案)。载: 教育部语言文字信息管理司编.商务印书馆, 2015).

⁴ In the transcription of Tibetan a hyphen is used to transliterate a specific Tibetan character, called *tsheg*, which marks the boundary between two syllables. (This punctuation also occurs between words, but it is only transliterated word-internally.) The *tsheg* often corresponds to morpheme breaks, but some morpheme breaks are not marked. We follow this Tibetan studies convention here, rather than strict morphemic glossing.

2. Copula verbs (COU): Copula verbs mainly equate two noun phrases; the choice of copula is often determined by a subjective judgment or access to information. The two primary copula verbs are ཡིན་ (yin) and རེད་ (red). The clause structure required by a copula verb is NP+NP+V (COU), in which both NPs are zero-marked for the absolutive case (Huang et al. 2003).

- (4) དགོ་ཚམ་དེ་ ལོ་ ག་ཚོད་རེད།
 dge-rgan de lo ga-tshod red.
 teacher the age how-many is
 ‘How old is that teacher?’
- (5) དེ་ནི་ གུང་གོ་ཕྱོགས་ཡོངས་ནས་བསྐྱར་བཅོས་བྱེད་པའི་གནས་བབ་ཤར་རྒྱ་ཡིན་པའི་སྐབས་རེད།
 de-ni krung-go phyogs-yongs-nas bsgyur-bcos-byed-pa-vi gnas-bab
 This-PAP China comprehensive-ABL reform-NMZ-GEN situation
 shar-rgyu-yin-pa-vi rtags red.
 show-up-NMZ-GEN symbol is
 ‘This is a symbol which shows the situation of China’s comprehensive reform.’

The negative forms of ཡིན་ (yin) and རེད་ (red) are མིན་ (min) and མ་རེད་ (ma red) respectively. མིན་ (min) is a contracted form of མ་ཡིན་ (ma yin).

- (6) ཆེན་མས་ "ཚུད་ཀྱི་བུ་མིན་པ་དང་གཉིས་པ་མ་རེད་པས།"
 chen-ma-s " khyod-kyi bu min bdag-gi yin " zhes zer.
 wife-AG " 2sg-GEN son not mine-GEN is " say
 ‘The (first) wife said, “he is not your son but mine”.’
- (7) འདི་ཚེད་མངགས་ང་མདོག་ཉེས་བཅོས་པ་མ་རེད་པས།
 vdi ched-mngags nga mdog-nyes bzos-pa ma-red pas.
 this deliberately 1sg smear do-NMZ NEG- is INT
 ‘Doesn’t this deliberately embarrass me?’

3. Existential Verbs (EXI): Existential verbs signify that someone or something is in some place; the subject is in the zero-marked absolutive case and the 2nd NP in the locative case. The basic existential clause structure NP+NP+[LOC]+VP (EXI). In the Lhasa vernacular language the most common existential verbs are ཡོད་(yod), འདུག་(vdug), and ཡོད་རེད་(yovo red). There are also a small number of other existential verbs, such as གནས་ (gnas) ‘to have, to live, to exist’; བཞུགས་ (bsdad) ‘to reside’, which also call for the same type of clause construction.

- (8) ངའི་མ་ནང་ལ་ ཡོད།
 nga-vi ama nang-la yod.
 1sg-GEN mother home-LOC be
 ‘My mother is at home.’
- (9) རང་རྒྱལ་གྱི་གྲངས་ཉུང་རིགས་མང་བ་ཆེ་བ་མ། །འ་མཚམས་ལ་ ཡོད་རེད།
 rang rgyal-gyi grangs-nyung-rigs mang-ba-che-ba
 Our motherland-GEN minority-ethnics most

mthav-mtshams-la **yovo-red**.
border-areas-LOC be
'Most minorities of our motherland live in border areas.'

- (10) ཁྱིལ་མཚོ་ལ་གནས། ང་སྐམ་ས་བསྐྱེད།
khyod mtsho-la **gnas**, nga skam-sa(-r) **bsdad**.
2sg lake-LOC resides 1sg land-LOC reside
'You reside in the lake, and I live on the dry land.'

4. Verbs of Possession (VOP): In addition to existential verbs, such as ཡོད་/འདུག་ (yod/vdug) 'to get', དབང་ (dbang) 'to have, to own', there are other associated verbs denoting 'giving', like འཁྲབ་ (vthob) 'to obtain', རག་ (rag) 'to receive', འཛམས་ (vdzams) 'to equip, to possess' etc.

The clause structure for a verb of possession is NP+[POS]+NP+V(VOP), where the verb of possession requires the subject with a possessive case marker and object with a zero case marker. However, if the subject is an animate noun while the object an abstract one (most nouns), the possessive case marker of the former can be omitted. Similarly, if the subject is marked with a pause particle, again the possessive case marker may be omitted (Jiang 2003).

- (11) ངར་ བོད་ཡིག་གེ་ཚོག་མཛོད་གཞིག་ཡོད།
nga-r bod-yig-ge tshig-mdzod gcig **yod**.
1sg-POS Tibetan dictionary one have
'I have a Tibetan dictionary.'

- (12) མྱེན་ལགས་དེར་ ཉམས་ཚུང་ཞེ་དགས་ཡོད་རེད།
rgan-lags de(-r) nyams-myong zhe-drags **yovo-red**.
teacher the(-POS) experience much have
'That teacher has a lot of experience.'

- (13) ད་ལྟར་ང་རེདེ་ ཉང་འདི་ནི་ ཉང་ཡོན་བྱེ་བ་གསུམ་དང་ས་ཡུང་གུ་ཡོད།
da-ltar nga-revi tang vdi-ni tang-yon bye-ba-gsum
now 1sg-myself-GEN party this-PAP member thirty million
dang sa-yad-gu **yod**.
and nine million have
'Our party now has thirty-nine million members.'

There are other verbs classified into this sort, signifying the quasi-possessive relation, like ཟེར་ (zer) 'to name', གནས་ (gnas) 'to value', རྗེས་ (rtsis) 'to equal', used in clauses whose form is roughly the same as that of the possessive verb.

- (14) ངའི་མིང་ལ་ འོན་ལྷན་ཚེ་རིང་ཐོང་ཕྱི་ཡོད།
 nga-vi ming-la don-grub-tshe-ring zer-gyi-yod.
 1sg-GEN name-POS Dhondup Tsering name-DUR
 ‘My name is Dhondup Tsering.’

- (15) རང་རྩིས་ན་མི་གསུམ་འདུག། རང་མ་རྩིས་ ན་མི་གཉིས་ལས་ མི་འདུག།
 rang rtsis na mi gsum vdug, rang ma-rtsis na mi-gnyis
 you count if person three have you NEG-count if persons-two
 las mi vdug.
 EXC NEG have
 ‘If you are counted, there are but two persons.’

5. Verbs of Change (CHA) refer to verbs that describe the bringing about of resultative complements. Such verbs include ལྷུལ་ (sprul) ‘to change, to become’, ལྷུར་ (sgyur) ‘to turn into’, འབྱུར་ (vgyur) ‘to grow’, འབྲོ་ (vgro) ‘to transfer’, འདུལ་ (vdul) ‘to tame, domesticate’ etc. (JW Zhou et al. 2003).

The clause structure with a verb of change is NP+VP (adj/v) +[COP]+VP (CHA), in which a complement particle *ru* or *la* are compulsorily inserted between an adjective or verbal phrase denoting the result of the change in question and the verb of change, forming a resultative complement.

- (16) ང་ཚོའི་ ཞིང་ལས་ ༼ན་སྐྱེད་ ཡག་ ཅུ་ཡག་ ཅུ་ འགོ་གིས་ །
 nga-tsho-vi zhing-las-thon-skyed yag-ru-yag-ru vgro-gis
 1pl-GEN agricultural-production better-COP-better-COP become-DUR
 ‘Our agricultural production is becoming better and better.’

- (17) རང་གི་ རྩ་ཚ་ དྲག་ ཅུ་ ཕྱིན་ འདུག་གས་ །
 rang-gi na-tsha drag ru phyin-vdug gas.
 2sg-GEN illness recover COP become-RST INT
 ‘Did you recover from your illness?’

- (18) ངའི་ འཚམ་པ་ ལྷག་ ཅུ་ ལྷག་ ཅུ་ འགོ་གྱི་འདུག།
 Nga-vi vcham-pa sdug-ru-sdug-ru vgro gyi-vdug.
 1sg-GEN influenza worse-COP-worse-COP become DUR
 ‘My influenza is becoming worse and worse.’

6. Verbs of perception (PER) refer to such verbs as མཁྲིང་ (mthong) ‘to see’, ལོ་ (go) ‘to hear’, ལྟ་གོ་ (ha go) ‘to understand’, ལྟེན་ (shes) ‘to know’, དྲན་ (dran) ‘to remember’, བརྗེད་ (brjed) ‘to forget’, རོ་ལྟེན་ (ngo shes) ‘to know, be acquainted with’ etc. Perception is defined as the sense or reaction to a person’s visual, auditory, tactile and cognitive experiences. It is fundamental to such verbs that they are involuntary. The verb བརྗེད་ (brjed) ‘to forget’ etc., however, has the feature of controllability (actor can consciously conduct or not an action; the verb then is a controllable one), while the results brought about by མཁྲིང་ (mthong) ‘to see’ etc. cannot be controlled consciously by human, hence a non-controllable verb (Yuan 1993).

The clause structure for verbs of perception is: NP+[AG]+NP+VP (PER); the subject of it is followed by an agentive case marker.

- (19) ལྟན་ལགས། ད་ ངས་ གཞིན་སྙམ་པོ་ཡིང་ །
 rgan-lags, da nga-s gzhi-nas **ha-go-song**.
 teacher now 1sg-AG just understand-PEF
 ‘Teacher, I just now understand.’

- (20) བུ་མོ་འདི་འི་མིང་ངས་ བློ་བྱུང་ །
 bu-mo vdi-vi ming nga-s **dran-byung**.
 girl this-GEN name 1sg-AG remember-PEF
 ‘I remembered this girl’s name.’

A verb of perception can be followed by an object of a clause with a nominalization markers མིན་/མེད་ (min /med) ‘whether’.

- (21) ཁོང་ཚོའི་ནང་ལ་ མི་ ག་ཚོད་ ཡོད་མེད་ངས་ གསལ་པོ་ བཤམ་གྱི་མེད་ །
 khong-tsho-vi nang-la mi ga-tshod yod-med nga-s
 3.pl-GEN family-POS people how-many have-not 1sg-AG
 gsal-po shes-kyi-med.
 clearly know-NEG-DUR
 ‘I don’t know clearly whether there are many people in their family.’

7. Verbs of motion (DIR) refer to those verbs that denote movement or lack of movement. Such verbs include འགྲོ་ (vgro) ‘to go’, ཕྱིན་ (phyin) ‘to remove’, ཡོང་ (yong) ‘to come’, གྲོང་ (thad) ‘to depart’, བེབས་ (phebs) ‘to arrive’, རྒྱུགས་ (rgyugs) ‘to go’, ཤོག་ (shog) ‘to leave’, བསྐྱོད་ (bsdad) ‘to stay’, བཞག་ (bzhag) ‘to place’ etc. Such verbs serve three syntactic functions: (1) they can constitute a directional predicate clause (NP+VP), as in example 22 below; (2) they can serve as a dummy directional complement (NP+VP+VP), as in example 23 below; (3) with other verbs they can constitute a serial predicate construction, when a target particle (TAP) interposed. The latter clause structure is NP+VP+[TAP]+VP (DIR), as in example 24 below.

- (22) ལྷིང་རང་གཉིས་མགྱོགས་པོ་ རྒྱུགས་ །
 khyed-rang gnyis mgyogs-po **rgyugs**.
 you two quickly go
 ‘You two go in a hurry.’

- (23) ང་རྒྱུག་གི་ མིན་ མོན་ལ་ལོག་འགོ་ལོད་ །
 nga sgug-gi-min sngon-la log **vgro go-yod**.
 1sg wait-NEG-PRO first back go need-RST
 ‘I can’t wait, since I need to go back first.’

- (24) ཁོ་རྒྱུག་མ་བཅད་ཀླུ་ ཕྱིན་སུ་ །
 kho smyug-ma **bcad-kag phyin-su**.
 he bamboo chop-TAP go-PEF
 ‘He has gone to chop the bamboo.’

8. Action Verbs (ACT) can be subdivided into various types, including intransitive verbs, transitive verbs, ditransitive verbs and verbs that take a resultative as object. The objects of a transitive action verb may permit a patient object, indirect object (dative object), and resultative object. Whether transitive or not, the subject is always followed by an agentive case marker. Action verbs are the major verb category in Tibetan. A full discussion is not feasible here.

The basic clause construction with an intransitive action verb is NP+[AG]+VP (ACT). The basic clause construction with a transitive action verb is NP+[AG]+NP+VP (ACT).

- (25) ངས་ཕྱོད་ཚོག་ལ།
 nga-s phyin chog.
 1sg-AG go be-able
 'I can go.'

- (26) ངས་ ལམ་གཅིག་ཚོས་པ་ཡིན་ལ།
 nga-s zha-mo gcig nyos-pa-yin.
 1sg-AG hat one buy-REA
 'I bought a hat.'

The basic construction for a ditransitive clause is NP+[AG]+NP+[DAT]+NP+VP.

- (27) ཇམ་ཚེས་ངར་ ལས་ གཅིག་བརྟན་བྱུང་ལ།
 am-chi-s nga-r khab gcig brgyab-byung.
 Doctor-AG 1sg-DAT needle one hit-PEF
 'A doctor has given me an interjection.'

The basic construction for a transitive clause with a resultative object is NP+[AG]+NP+NP+[FAT]+VP(ACT).

- (28) ཇམ་ཚེས་ངར་ ལས་ གཅིག་བརྟན་བྱུང་ལ།
 khyed-rang-gyis rgya-yig bod-yig-la bsgyur-rogs-gnang.
 2sg-self-AG Chinese Tibetan-FAT translate-HON
 'Please (you) translate Chinese into Tibetan.'

9. Mental Verbs (PSY) mainly refer to verbs with a target object, psychologically emotional actions or are non-action verbs signifying psychological actions. For instance, ངོ་ཚ་ (ngo tsha) 'to be shy', དགལ་ཞེན་བྱེད་ (dgav zhen byed) 'to like', སེམས་ངན་བྱེད་ (sems ngan byed) 'to bully', སེམས་ཤོར་ (sems shor) 'to be infatuated', ལྷན་ (sun) 'to hate', བརྩི་འཇོག་བྱེད་ (brtsi vjog byed) 'to respect', རེ་སྐྱུག་བྱེད་ (re sgug byed) 'to expect', མཁྱེད་ཚུང་བྱེད་ (mthong chung byed) 'to despise', ལས་ལེན་བྱེད་ (khas len byed) 'to admit', ཐག་དོག་བྱེད་ (phrag dog byed) 'to be envious', བརྩེ་སྦྱོང་བྱེད་ (brtse skyong byed) 'to be fond of', སེམས་ཁྲལ་བྱེད་ (sems khral byed) 'to worry', བསྐྱོད་གླུ་འབྲུལ་ (bstod glu vbul) 'eulogize', ངོ་རྟོག་བྱེད་ (ngo rgol byed) 'to oppose', ཡིད་ཚེས་བྱེད་ (yid ches byed) 'to trust', ལྷུང་ལངས་ (rlung lang) 'to be angry', བྱམས་པོ་བྱེད་ (byams po byed) 'to be fond', རྟོང་སྐྱེས་བྱས་ (snying skyes byas) 'to be pitiful', དྲགས་ (dogs) 'to be suspicious' etc.

The basic clause structure for psychological verbs is NP+[AG]+NP+[OBJ]+VP (COG), namely, the subject of psychological verb clause is followed by an ergative case marker, while the target object is followed by a target case marker.

- (29) རང་གིས་ མོར་ དགའ་པོ་ཡོད་པས། །
rang (gis) mo-r dgav-po-yod pas.
2sg (AG) 3sg-OBJ love-RST-INT
'Do you love her?'

- (30) བཤམས་ ངར་ རྒྱུ་ལངས་སོང་། །
pha-ma-s nga-r rlung-langs-song.
father and mother-AG 1sg-OBJ be angry-PEF
'Father and mother are angry with me.'

10. *Verba dicendi* (NAR) include verbs of speech (to say) and verbs of thought (to talk with oneself), such as བཤད་ (bshad) 'to say', ལབ་ (lab) 'to talk, to speak', ཟེར་ (zer) 'to say, to speak, to be alleged', གསུང་ (gsung) 'to say, to talk', འདྲི་ (vdri) 'to ask', ལན་སློག་ (lan slog) 'to answer', ལྷ་ (zhu) 'to tell, to inquire', བསམ་ (bsam) 'to think, to meditate, to consider', དགོངས་ (dgongs) 'to guess, to ponder', ཤེས་ (shes) 'to realize' etc.. The syntactic feature of this sort of verbs is that they are followed by a clausal object or a nominalized verbal phrase serving as an object.

The basic clause structure for *verba dicendi* is NP+[AG]+{NP+VP}+VP (NAR), or {NP+VP}+NP+[AG]+VP (NAR), among which the clause object of verbs of speech is often suffixed by a clause particle (OCP) (Jiang 2007).

- (31) རང་སོང་ཚུ་གཙང་སྤྱད་དོ་སྤྱད་བྱེད་དགོས་ལེ་ རྒྱུ་ཉོ་ལབ་པ་ཡིན། །
nga-s khong-tsho-r gtsang-sbrar-do snang-byed-dgos-se rdzag-to lab-pa-yin.
1sg-AG 3pl-ALA hygiene pay-attention-to repeatedly tell-REA
'I repeatedly told them to pay attention to hygiene.'

- (32) ངས་གསོ་ནི་མཁའ་ ཤས་དགོས་ བསམས་བྱུང་། །
ngal-gso nyi-ma kha-shas dgos- bsams-byung.
rest day a few need consider-PEF
'I should have considered resting a few days.'

- (33) སྤྲེལ་མ་རང་གིས་ རང་ཚུ་ལྷགས་རྗེ་ཆེ་བེད་གྱི་འདུག་སེ་ལབ་རྟོགས། །
sgrol-ma rang-gis khong-tsho-r thugs-rje-che zer-gyi-vdug-se lab-rogs.
Droma 2sg-AG 2pl-DAT thanks say-DUR (OBP) speak-HON
'Droma, please give my thanks to them.'

Verba dicendi can also be followed by clause with nominalizational marker མིན་/མེད་ (min/med) 'whether'.

- (34) དེང་སང་གནམ་གཟུང་གི་ སྤྱ་སེ་ ཉོ་ལག་པོ་ ཡོད་མེད་ ལྷན་ཆ་འདྲི་རྟོགས་གནང་། །
deng-sang gnam-gru-vi spa-se nyo khag-po yod-med
recently airplane-GEN ticket buy difficult have-NEG
skad-cha vdri-rogs-gnang.
word ask-HON
'Excuse me for asking, whether it is difficult to buy the airplane ticket recently.'

11. Interactive Verbs (REL) are a special class of verb, whose meanings are relevant to two logically or conventionally corresponding or subordinating matters or events. In syntactic form, the interactive verbs usually require an interactive case marker following the correlated noun. དང་ (dang)(ITP).

The basic clause structure for interactive verbs is NP+NP+[ITP]+VP(REL). The oft-used interactive words are མཐུན་ (mthun) ‘to agree, to fit, to correspond’, ལྷན་ (stun) ‘to coincide, to adapt, to accord’, ཁ་བྲལ་ (kha bral) ‘to disrupt, to separate’, འགལ་ (vgal) ‘to violate, to contradict’, འདྲ་ (vdre) ‘to mix, to interblend’, ཟླ་ (sre) ‘to compound, to mingle’, འདྲ་ (vdra) ‘to resemble, to be alike’, ཁ་ཐུག་རྒྱུག་ (kha thug rgyag) ‘to meet, to encounter’, ལྷོ་མ་ (snyom) ‘to balance, to match’.

- (35) དང་དུང་སློབ་ཐོགས་གཞན་དག་དང་ མཐུན་སྒྲིལ་ཡག་པོ་བྱེད་དགོས་ཤིང་།
 da-dung slob-grogs gzhan-dag dang mthun-sgril yag-po-byed-dgos-red.
 also student other (ITP) unite well-do need
 ‘It is necessary to make good relations with other students as well.’
- (36) ཡིན་ནའི་བུ་མོ་དེས་ བ་མ་དང་ཁ་བྲལ་བྱས་མནའ་མར་འགོ་འདྲོད་བྱེད་གྱི་ཡོད་མ་ཤིང་།
 yin-navi bu-mo de-s pha-ma dang kha-bral-byas mnav-ma-r
 however girl this-AG parents ITP separate bride-COP
 vgro-vdod-byed-kyi-yovo-ma-red.
 become-be-willing-NEG-DUR
 ‘However, this girl was not willing to be separate from her parents and become another’s bride.’

- (37) མོ་མའི་ མ་མ་དང་ འདྲ་པོ་མིན་འདུག་།
 mo mo-vi a-ma dang vdra-po-min-vdug.
 3sg 3sg-GEN mother ITP be alike-NEG-DUR
 ‘She is not like her mother.’

12. Causative Verbs (CAV) in Tibetan include བཟུག་ (bcug) ‘to allow, to make’, བྱེད་ (byed) ‘to let, make’, བཟོ་ (bzo) ‘to ask that’.(Jiang D. 2006) The subject of a causative verb is suffixed with an agentive case marker, and the object is often a clause or a verb phrase with a causative particle (CAU) རུ་ (ru) before the causative verb. the subject of the clause is followed by a possessive case marker or an agentive case marker.

The basic clause structure for causative verbs is NP+[AG]+NP+[POS]+ (NP +) VP+[CAU]+VP (CAV).

- (38) ཁོང་གིས་གོ་སར་ ས་ཆ་དེའི་ བདག་པོ་ བྱེད་ (རུ་) བཟུག་པ་ཤིང་།
 khong-gis ge-sar sa-cha de-vi bdag-po byed-(ru) bcug-pa-red.
 3pl-AG Gesar place this-GEN master become-(CAU) let-REA
 ‘They let Gesar become the master of this place.’
- (39) རྒྱལ་བ་ཚོ་དལ་གསོ་ ཡག་རྒྱུག་རུ་ བཟུག་དགོས་།
 nad-pa-tsho ngal-gso yag-rgyag-ru bcug dgos.
 Patient-pl have a rest good-CAU let AUX
 ‘Ask the patient to have a good rest.’

The preceding discussion of Tibetan syntactic and semantic categories provides an elementary sketch of the syntactic frames of modern verbs, which can serve as a foundation for the automatic word segmentation of Tibetan texts. In practical applications, however, there may be other special verb categories, including specific clause constructions, or various clause structures with alternative word orders, forming much more complicated grammatical constructions. We hope the discussion here is a feasible starting point, and recently, some research has borne out this hope (Jiang 2011; Li et al. 2013; Kang et al. 2015).

3 Functions of Tibetan verbs in syntactic and semantic categories

This article classifies verbs into syntactic and semantic categories, in an attempt to build a machine readable dictionary of grammatical properties. In such a dictionary, each of the verbs (and other parts of speech) is provided with grammatical information distinguishing it from other verbs, noting morphology, meaning, morphological variants and syntactic/semantic category etc. Once the verb category is identified, analysis of its syntactic frame, argument valency and relevant word order becomes possible, the property of various kinds of homomorphic syntactic markers can be specified and then a comparatively accurate classification and analysis can be made (Yu 1998). We take as an example the clause ང་ཉེད་གཉིས་དང་ཁ་བྲལ་འདོད་མི་འདུག with the interactive verb ཁ་བྲལ་ (kha bral) ‘to separate’. In the clause, དང་(dang), an interactive marker, is homomorphic with a coordinating conjunction. Therefore, two different possible segmental results A or B may be produced. Furthermore, དང་(dang) and ཁ་(kha) are exactly combined into another word form: དང་ཁ་(dang kha) ‘taste(s)’, the clause will be tagged into

- A: ང་(1sg)ཉེད་གཉིས་(2dl)དང་(ITP)ཁ་བྲལ་(separate) འདོད་(AUX)མི་འདུག(NEG-RST)
 ‘I am not willing to be parted from you two’,
 or B: ང་(1sg)ཉེད་གཉིས་(2dl) དང་ཁ་(taste)བྲལ་(part)འདོད་(AUX)མི་འདུག(NEG-RST)
 ‘We two need not part tastes.’

However, since we can extract from the dictionary information on the interactive verb ཁ་བྲལ་ (kha bral), it is known that its basic grammatical frame requires the interactive marker དང་ (dang). Thus, taking advantage of an identification method based on rules, we can ascertain the validity of segmental path A.

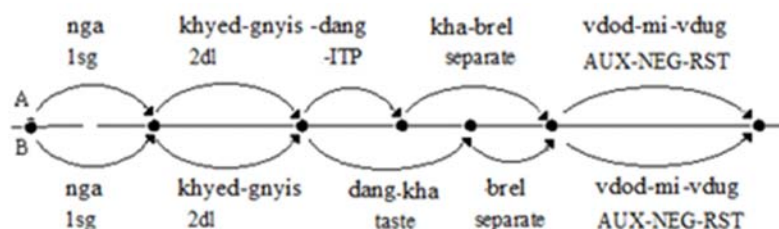


Fig 1 two different segmentation

Up till now, whether based on rules or statistics for natural language processing, a machine readable dictionary or a corpus-based knowledge representation system has always been necessary, with abundant grammatical information required in the design of word classifications and

syntactic rules available for text analyses (Jiang 2005). The coding described in this article is a preliminary step in the service of projects underway for the compilation of a Tibetan grammatical information dictionary. As regards the methods of verb classification, predecessors have conducted a great variety of studies worthy of reference (Hu 1984; Zhou et al. 2002; Chen et al. 2003; Jiang et al. 2008). The verbs in syntactic and semantic categories mainly involve syntactic facts and the construction of dictionary information. Certainly, a Tibetan machine readable dictionary should have much more information about verbs. For example, relative verbs, verbs of possession and existential verbs interact with the grammatical person of the subject, while other verbs maybe involve willingness (volitional and nonvolitional), transitivity (transitive and intransitive), causativity (causative or automatic) and commonly-matched labels of nominalization etc. In grammar, most verbs undergo morphological change (*da lta ba* basic form, *ma vongs pa* future tense, *vdas pa* past tense, *skul tshig* imperative mood), particular negative form, syllabicity (monosyllable/disyllable /multisyllable) and allomorphic form etc. It is hoped that the accumulation of information for such a dictionary will be progressively advanced in further research.

ABBREVIATIONS

ABL	ablative case	INT	interrogative word
ACT	action verb	LOC	locative case
AG	agentive case	MER	merely past aspect
ALA	allative case	NAR	narrate verb
AUX	auxiliary	NEG	negative
CAU	causative particle	NMZ	nominalization marker
CAV	causative verb	OBJ	object case
CHA	verb of change	PAP	pause particle
COG	cognition verb	PEF	perfect aspect
CON	contextual aspect	PER	perception verb
COP	complement particle	POS	possessive case
COU	copula particle	PRM	preliminary aspect
COU	copula verb	PRO	prospective aspect
DAT	dative case	REA	reality aspect
DIR	directional verb	REL	interactive verb
DUR	durative aspect	RST	resultative aspect
EXI	existential verb	STA	stative verb
FAT	factual case	TAP	target particle
GEN	genitive case	TP	interactive particle
HOP	honor particle	VOP	verb of possession
IMM	imminent aspect		

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31271337, 61132009) and the National Social Science Foundation of China (10&ZD124, 12&ZD174). Many Thanks to

Prof. Nathan W. Hill, who checked the paper with a professional point of view in Tibetan language, and the anonymous reviewers' comments.

REFERENCES

- Chen, Y. Z.; Li, B. L.; Yu, S. W. 2003. Re-study on Tibetan verb Categorization based on Case-relation and Valence Theory. In: *Language calculation and content-based text processing*. Beijing: Tsinghua University Press, Beijing. pp. 284-290. 陈玉忠, 李保利, 俞士汶 2003. 基于格关系和配价的藏语动词再分类研究. 载: 孙茂松, 陈群秀(主编): 《语言计算与基于内容的文本处理》 284—290. 北京: 清华大学出版社.
- Hu, S.J. 2002. *The Concise Dictionary of Tibetan verbs*. Chengdu: Sichuan Minzu Press. 胡书津 2002. 《藏语动词类型释要》. 成都: 四川民族出版社.
- Hu, T. 1984. An analysis of several verbal clause patterns in Tibetan Dialect. *Minzu Yuwen*. (1):1-16. 胡坦 1984. 拉萨藏语中几种动词句式的分析. 《民族语文》第 1 期.
- Huang, X.; Jiang, D. 2003. Automatic Recognition of the Subject and Object of Linking verbal Clauses in Tibetan. In: *Language computing and Content-based text processing*. Beijing: Tsinghua University Press. 167-172. 黄行, 江荻 2003. 现代藏语判定动词句主宾语的自动识别方法. 载: 孙茂松, 陈群秀主编: 《语言计算与基于内容的文本处理》 167-172. 北京: 清华大学出版社.
- Jiang D.; Long C.J.; Zhang J. C. 2005. The Verbal Entries and Their Description in a Grammatical Information-Dictionary of Contemporary Tibetan. In Robert Dale, Kam-Fai Wong, Jian Su, Oi Yee Kwong (Eds), *Natural Language Processing – IJCNLP*, 874-884. Springer.
- Jiang, D. 2003. On syntactic chunks and formal markers of Tibetan. In: *Language calculation and content-based text processing*. Beijing: Tsinghua University Press, Beijing. pp. 160-166. 江荻 2003. 现代藏语的句法组块与形式标记. 载: 孙茂松, 陈群秀(主编): 《语言计算与基于内容的文本处理》第 160-166 页. 北京: 清华大学出版社.
- Jiang, D. 2005. The Syntactic Rules and Tag Sets of Word Classes and Chunks in Modern Tibetan. In: Jiang, D., Kong, J.P. (eds.) *Advances on the Minority Language Processing of China*. China Social Document Press. 江荻 2005. 面向机器处理的现代藏语句法规则和词类、组块标注集. 载: 江荻, 孔江平(主编): 《中国民族语言工程研究新进展》 10—93 页. 北京: 社会科学文献出版社.
- Jiang, D. 2006. Identification of boundaries of object clauses in causative verb clauses in Modern Tibetan, *Journal of Chinese Language and Computing*, Vol.15(4):185-192. Singapore. 江荻 2006. 现代藏语致使动词句中宾语小句的边界识别. *Journal of Chinese Language and Computing*. Vol.15(4):185-192.
- Jiang, D. 2007. Approaches to clause objects of Narrate verbs and Their markers in Tibetan language. *Journal of Chinese information processing*. Vol.21(4):111-115. 江荻 2007. 藏语述说动词小句宾语及其标记. 《中文信息学报》第 4 期:111-115.
- Jiang, D. 2011. The tagging methods for Tibetan Irregular Verbs in texts. In: *Frontiers of Chinese Information Processing*. Beijing: Tsinghua University Press. 江荻 2011. 藏语不规则动词的信息标注方法. 载: 孙茂松, 陈群秀(主编): 《中国计算语言学研究前沿进展》. 北京: 清华大学出版社.

- Jiang, D.; Long, C. J. 2008. The types of Tibetan Objects and their markers. In: *Collections of the 100th anniversary of Professor Liu-Shengqi's birth and study on Tibetan cultures and history*. 625-658. Beijing: Press of China Tibetology. 江荻, 龙从军 2008. 藏语的宾语类型及标记分析. 《纪念柳陞祺先生百年诞辰暨藏族历史文化论集》652-658 页. 北京:中国藏学出版社.
- Kang, C. J.; Long, C. J.; Jiang, D. 2015. Tibetan names recognition research based on CRF. *Computer Engineering and Applications*. 51(3):109-111. 康才峻, 龙从军, 江荻. 2015. 基于条件随机场的藏文人名识别研究. 《计算机工程与应用》第 3 期:109-111.
- Li, L.; Long, C. J.; Jiang, D. 2013. Tibetan Functional Chunks Boundary Detection. *Journal of Chinese information processing*. Vol. 27(6): 165-168. 李琳, 龙从军, 江荻 2013. 藏语句法功能组块的边界识别. 《中文信息学报》第 6 期:165-168.
- Long, C. J. 2012. Key Issues in the Text Information Processing of Tibetan Language. *E-Science Technology and Application*. Vol.3(4)51-58. (龙从军 2012 藏语文本信息处理的几个关键问题. 《科研信息化技术与应用》, 第 3 卷第 4 期, 51-58 页.)
- Yu, S. W. (Editor) 1998. *The Grammatical Knowledge-base of Contemporary Chinese: A Complete Specification*. Beijing: Tsinghua University Press. 俞士汶等 1998. 《现代汉语语法信息词典详解》. 北京:清华大学出版社.
- Yuan, Y. L. 1993. *On imperative clauses in Mandrin*. Beijing: Beijing University Press. 袁毓林 1993. 《现代汉语祈使句研究》. 北京:北京大学出版社.
- Zhou, J. W. 2002. On Tibetan verbs. *China Tibetology* (Chinese version). 2(119-127) & 4(118-128). 周季文 2002. 论藏语动词. 《中国藏学》第 2 期:119-127, 第 4 期:118-128.
- Zhou, J. W.; Xie, H. F. 2003. *A grammar of Lhasa Tibetan language*. Beijing: Minzu Press. 周季文, 谢后芳 2003. 《藏语拉萨话语法》. 北京:民族出版社.

Di Jiang
jiangdi@cass.org.cn