Research on a Standard for POS Tagging of Contemporary Tibetan for TIP

Tsering Gya and Dbangphyug Tsering
Research Center of Tibetan Information
Qinghai Normal University
Xining, Qinghai, China 810008
zwxxzx@163.com peljorstring@yahoo.com

Abstract

The Provincial Key Laboratory of Tibetan Information Processing and Machine Translation, Qinghai Normal University has devised Part of Speech (POS) Tagging Norms for Tibetan Information Processing (TIP) based on the automatic segmentation and manual analysis of a large-scale Tibetan corpus. In order to achieve the two goals of computer automatic segmentation and tagging of the Tibetan corpus, the process of developing Part of Speech Tagging Norms was to first classify empty words and content words, and then determine the largest category. The major categories were then separated into sub-categories and different depth subclasses.

Introduction

In this new era of information processing, language processing technology is the core supporting technology. In the case of Tibetan, Tibetan raw corpora as they currently exist have not reached their full potential; the reason for this is the lack of unified norms and standards, resulting in a waste of large-scale Tibetan resources. It was our contention that the establishment of standards for segmentation and tagging of Tibetan corpora would have important significance for improving TIP. Based on the above considerations, the Provincial Key Laboratory of Tibetan Information Processing and Machine Translation at Qinghai Normal University began research into multi-level processing of Tibetan corpora in 2002.

Classification and POS marking [1],[2],[3],[4]

At the present time, ten million Tibetan words of a raw corpus (comprised of classical Tibetan literature and records, primary and secondary materials, Tibetan periodicals and newspapers, web text, contemporary Tibetan anthologies, Tibetan versions of political writings, Tibetan versions of laws and regulations, etc.) have been experimented upon, producing a segmented and tagged corpus. As a result of this experiment, in December of 2005, a draft standard for discussion "Part of Speech Tagging Norms for Tibetan Informa-

tion Processing" was put forth, followed in June of 2007 by a formal presentation entitled "Part of Speech Tagging Norms for Tibetan Information Processing (V1.0)."

This specification provides tagging codes for Tibetan parts of speech and other segmentation units for information processing, as they are applicable to TIP, as well as for use as a resource for teaching and research into the modern Tibetan language. The norms for information processing set forth in this document were in accordance with Tibetan language characteristics and requirements based on grammatical functions, with the aim of ease of computer analysis and automatic processing at the task of the division of a text into parts of speech.

POS Marking Codes

A system was devised following the international practice of using English terminology letters as tag codes, for example, using the first letter of English term "noun" — "n" — as a tag code for the Tibetan "noun," for the English term "numeral" using the third letter "m" as the tag code for the Tibetan "numeral", etc. In cases where the classification was unique for Tibetan, or the use of English[5],[6] letters to mark terms was inconvenient, the Chinese phonetic alphabet was used to generate tag codes. For example, the grammatical particles ("nāmje"; རྣམ་དབྱེ་; 格助词), use the first letter of the Chinese Phonetic equivalent "g" for the Chinese characters "格" as a tag code; similarly, for "status word" the first letter of the Chinese Phonetic equivalent "z" of the Chinese character, "状" was used as the tag code. This resulted in a standard in which part of speech was divided into seventeen major categories, twenty-one primary level categories, and sixty secondary categories.

The POS tagging scheme developed resulted in seventeen major categories, with numerous sub-divisions. The seventeen major categories were:

1. Nouns
2. Numerals
3. Quantifiers
4. Pronouns
5. Verbs
6. Adjectives
7. Adverbs
8. Interjections
9. Onomatopoeia
10. Helping words
11. Declaration words
12. Grammatical ("nāmje"; རྣམ་དབྱེ་; 格助词) particles
13. Connective particles
14. Affixes
15. Semantemes
16. Marks
17. Poetic terms

Applying the POS Marking Codes to Tibetan Corpora

The process of applying these POS codes to our Tibetan corpus was based on a simple, longest substring match to a lexicon of words with associated Part-of-Speech tags. The result of this process was a uniform application of tags to Tibetan language data. The basis for this mark-up system, with examples, can be seen in Appendix I.

Conclusion

"Part of Speech Tagging Norms for Tibetan Information Processing" was a major project in linguistics research at Qinghai Normal University. During the course of conducting research into tagging norms, it was necessary to mark large-scale Tibetan corpora with segmentation information through relying on experimental feedback and tests, and not based merely on a consideration of whether the categories were theoretically correct and feasible. In addition, consideration was also given to the normativity and stability of these categories in the field of automatic processing and machine learning, resulting in the automatic segmentation and manual analysis of large-scale Tibetan corpora described in this paper. It is our hope that through this article on research into TIP from the level of Tibetan word-processing upto the level of processing of phrases and sentences that our research will play a positive role in the future of Tibetan Information Processing.

References

[1] བསམ་གཏན། (山木旦), 郑绍功, བཀྲ་ཤིས་རབ་བརྟན། （扎喜拉旦等）. དག་ཡིག་གསར་བསྒྲིགས། /新编藏文字典. 西宁: 青海民族出版社, 1979. [Samten, Zheng Shao Gong, Tashi Rabten. *A Newly Compiled Tibetan Dictionary*. Xining: Qinghai Minority Press, 1979.]

[2] ཚེ་ཏན་ཞབས་དྲུང་། （才旦夏茸）. བོད་ཀྱི་བརྡ་སྤྲོད་བསྟན་བཅོས་སུམ་ཅུ་པ་དང་རྟགས་འཇུག་གི་རྣམ་བཞག་རྒྱ་ཆེར་བཤད་པ། / 藏文文法详解. 西宁: 青海民族出版社, 1988. [Tseten Shabdrung. *An Extensive Explanation of Tibetan Grammar*. Xining: Qinghai Minority Press, 1988.]

[3] དོར་ཞི་གདོང་དྲུག་སྙེམས་བློ། （多识洛桑图丹琼排）. །བོད་ཀྱི་བརྡ་སྤྲོད་རིག་པའི་དོན་འགྲེལ་ཕྱོགས་བསྒྲིགས།/藏语语法深义明释. 兰州: 甘肃民族出版社, 1999. [Dorzhi Dongdrug Nyemlo. *A Collection of Meaning Commentaries on Tibetan Grammar*. Lanzhou: Gansu Minority University Press, 1999.]

[4] 才让加 （ཚེ་རིང་རྒྱ།）, 吉太加 （མཆོག་ཐར་རྒྱལ།）. "基于藏语语料库的词类分类方法研究," 西北民族大学学报 (自然科学版), 2005 年 6 月第 26 卷总第 57 期, pp.39-42. [Tsering Gya, Chogthar Gyal. "Studies on a Taxonomic Approach to Part of Speech Identification in the Tibetan Corpus," *Journal of Northwest University for Nationalities. Natural Sciences*, vol. 26 no. 57 (June 2005), pp.39-42.]

[5] 俞士汶, 朱学锋, 段慧明. "大规模现代汉语标注语料库的加工规范," 中文信息学报, 2000 年第 6 期, pp.58-64. [Yu Shiwen, Zhu Xuefeng, Duan Huiming. "A Standard for Large-scale Modern Chinese Corpus Labeling and Processing," *Journal of Chinese Information Processing*, vol. 14 no. 6 (Nov. 2000), pp.58-64.]

[6] 靳光瑾, 郭曙纶, 肖航, 章云帆. "语料库加工中的规范问题—谈《信息处理用现代汉语词类标记集规范》," 语言文字应用, 2003 年 11 月第 4 期, pp.16-24. [Jin Guangjin, Guo Shu Lun, Xiao Hang, Zhang Yunfan. "On the Principle Question of Corpus Processing: An Approach to Processing Modern Chinese Punctuation and Parts of Speech," *Applied Linguistics*, no. 4 (Nov. 2003), pp.16-24.]

Appendix I

A detailed breakdown of the POS Tag categories with examples is as folows.

1. Class of Nouns (N)

1.1 Nouns (n)

ཁྱད་གཞི་བརྗོད་པའི་མིང་། The name of person or things that mainly to serve as the subject or object in a sentence.

1.1.1 General Nouns (nn)

These kinds of words are general nouns, for example, སྤྱི་ཚོགས། མི་དམངས། དུ་བ།.

1.1.2 Personal Names (nr)

མིའི་མིང་། Proper nouns for expressing people's names, for example, ཚེ་རིང་རྒྱལ། ཀརྨ་རྡོ་རྗེ། བློ་བཟང་རྒྱ་མཚོ། ཚེ་ཏན་ཞབས་དྲུང་།.

1.1.3 Gerunds (nv)

བྱ་སྤྱོད་ཀྱི་མིང་། Nouns that express a behavior or action. (Note: these words cannot be followed by numerals directly, except by ཅིག, ཞིག, or ཤིག, and must use ལན་ or ཐེངས་ to connect them to numerals.) For example, གན་རྒྱལ། སྣངས་སྣབས། གྲ་སྒྲིག དཔྱད་བརྗོད། དགག་སྒྲུབ། སློབ་སྦྱོང་། ཞིབ་འཇུག.

1.1.4 Nationalities (nm)

མི་རིགས་ཀྱི་མིང་། Proper nouns expressing the name of a nationality or tribe, for example, བོད། བོད་རིགས། བོད་པ། རྒྱ། རྒྱ་རིགས། ཧོར།.

1.1.5 Place Names (ns)

ས་མིང་། Proper nouns expressing the names of places, administrative divisions, rivers, seas, mountains, and buildings, for example, ལྷ་ས། མཚོ་སྔོན་པོ། རྨ་ཆུ། པེ་ཅིན། པོ་ཏ་ལ།.

1.1.6 Country Name (ng)

རྒྱལ་ཁབ་ཀྱི་མིང་། Proper nouns expressing the names of countries, for example, ཀྲུང་ཧྭ་མི་དམངས་སྤྱི་མཐུན་རྒྱལ་ཁབ། ཨ་མེ་རི་ཀ། འཇར་པན། རྒྱ་གར།.

1.1.7 Institutions & Teams (nt)

ལས་ཁུངས་ཀྱི་མིང་། Proper nouns for expressing the names of institutions or groups, for example, རྒྱལ་སྲིད་སྤྱི་ཁྱབ་ཁང་། རྩིས་རིག་སྡེ་ཁག སོ་རྩི་བཟོ་གྲྭ། སློབ་གསོ་ཐིན།.

1.1.8 Proper Nouns (nz)

ཆེད་བཏགས་ཀྱི་མིང་། Proper nouns for expressing professional subjects, for example, སྔགས་རིམ་ཆེན་མོ། ལིང་ཚེ། ཐིག་གཉིས་འབྲེལ་བ། གློག་ར། བར་རྟུལ། སྤུ་མཐའ།.

1.1.9 Job Titles (nx)

ཚ་ལོའི་མིང་། Proper nouns for expressing technical or professional titles, leadership positions, military and police rank, for example, སློབ་གཙོ། སྲི་དཔོན། དགེ་བཤེས། ལྷ་རམས་པ། སྟོང་དཔོན། ཀྲུའུ་རིན།.

1.1.10 Adnouns (na)

གཞི་ཆོས་བརྗོད་པའི་མིང་། Nouns formed from adjectives through the addition of the syllables པོ་ or བོ་, མ་ or མོ་, for example, དམར་པོ། གསལ་པོ། སྙན་མོ། གཙང་མ། རྡུང་པོ།.

1.1.11 Acronyms (nl)

བསྡུས་མིང་། The shortened form of a noun, for example, བཀྲིས་ལྷུན། ཞིར་ཕྱིན། ཞིར་བཟང་། སྐལ་རྒྱལ། དཔལ་རྡོར།.

1.2 Time Words (t)

དུས་ཚོད་སྟོན་པའི་མིང་། Nouns for expressing times, seasons, eras, or dynasties, for example, དབྱར་ཁ། སྔ་དྲོ། དགོང་མ། སོས་ཀ། སྟོད། ཕྱི་ཚིགས། གདོད་མ། ཆིང་རྒྱལ་རབས།.

1.3 Premise Words (s)

དུས་ཚོད་སྟོན་པའི་མིང་། Location words, such as a term plus a direction word, for example, ས་འོག ཁང་ཐོག.

1.4 Direction Words (f)

ཕྱོགས་མཚམས་སྟོན་པའི་མིང་། Words showing the relative direction of positions. The features of direction words are that they:

① cannot be modified in front by མ་ or མོ་;

② cannot be followed directly by a numeral;

③ cannot be modified in front by a degree adverb;

④ can have ཕྱོགས་ added behind them;

⑤ cannot be reduplicated.

For example, ར། ནུབ། ལྷོ་བྱང་། ཕྱི། སྟོད། སྨད། གྱེན། མས། ཡས་མས། ཕྱི་ནང་། གྱེན་ཐུར།.

2. Class of Numerals (M)

2.1 Numerals (m)

གྲངས་ཀ་བརྗོད་པའི་མིང་། Words expressing number and order.

2.1.1 Base Numbers (mj)

གྲངས་གྲངས། Words expressing size or exact number, for example, གཅིག བཞི། བརྒྱ། ཁྲི། བྱེ་བ། ཅིག ཞིག ཤིག (also the method of counting) ༡, ༢, ༣, ...

2.1.2 Ordinal Numbers (mx)

རིམ་གྲངས། Ordinal numbers that order things, for example, དང་པོ། བརྒྱད་པ། ཉི་ཤུ་པ། ཉི་ཁྲི་པ། སྟོང་ཕྲག བརྒྱ་བ་སོགས།.

2.1.3 Approximate Numbers (mg)

རགས་གྲངས། Expressing numbers of more or less, or inexactness, for example, གསུམ་བཞི། དགུ་ཙམ། སུམ་ཅུ་ལྷག གཡས་གཡོན། བརྒྱ་ཡས་མས་སོགས། འགའ། མང་ཤོས། ཉུང་ཤོས།.

2.1.4 Totals (mz)

སྤྱི་གྲངས། Words expressing a total number, formed by adding ཀ་, ག་, བོ་, or པོ་ as a suffix to the base number, for example, གཉིས་ཀ བཞི་ག བཅུ་བོ། ཉིས་བརྒྱ་བོ། ཁྲི་གཉིས་ཀ་སོགས།.

3. Class of Quantifiers (Q)

3.1 Quantifiers (q)

འཇལ་བྱེད་ཀྱི་མིང་། Words that specify people, things, or units of action.

3.1.1 Measure Words (qd)

འཇལ་གཞོར་གྱི་མིང་། For example, སར། མཁྱིད། རྒྱང་གྲགས། ཐེ། རྒྱ་མ། སྲང་། འདོམ་པ། སོར། སྟོང་ཁེ། སྨི།(meter) ཞོ།.

3.1.2 Collective Quantifiers (qj)

ཚོགས་སྤྱིའི་མིང་། For example, ཁྱུ། སྐོར། ཕོན་པོ། སྤེལ། རྡོག་པོ། ཤར་མོ། བྱུབས། པོ་ཏི། ཆ། འཁོར་ཆེན།.

3.1.3 Individual Quantifiers (qg)

ཁེར་རྐྱང་གི་མིང་། For example, ཉག་མ། ཡ། རྐང་། མིག.

3.1.4 Kind Quantifierss (qz)

རིགས་སྤྱིའི་མིང་། For example, རིགས། རྒྱུད། ཚན། རིམ་པ།.

3.1.5 Multiplying Power (ql)

བལྟབ་ཚད་ཀྱི་མིང་། For example, བལྟབ། ཆ་(ལྡེ་ཆའི་གཅིག) ལྡེ་འགྱུར། གསུམ་ལོག.

3.1.6 Degree Quantifiers (qc)

ལན་གྲངས་ཀྱི་མིང་། For example, ཐེངས། ལན། འཁོར་མ། རེབ།.

4. Class of Pronouns (R)

4.1 Pronouns (r)

རིས་བཟུང་གི་མིང་། Words that perform the function of replacing nouns and indicate repetition.

4.1.1 Personal Pronouns (rr)

གང་ཟག་རིས་བཟུང་གི་མིང་། For example, ང་། ཁོ་མོ། ཁྱེད། ཨུ་བུ་ཅག ཨུ་ཚོ། ཁྱོད་ཚོ།.

4.1.2 Demonstrative Pronouns (rz)

བརྗོད་དོན་རིས་བཟུང་གི་མིང་། For example, དེ། འདི། དེ་དག དེ་རྣམས། འདི་དག.

4.1.3 Interrogative Pronouns (ry)

དྲི་བ་སྟོན་པའི་མིང་། For example, གང་། ཅི། ཇི། སུ། དུ། ནམ། ཨེ། གྲང་། གམ། ངམ། དམ། ནམ། བམ། མམ། འམ། རམ། ལམ། སམ། ཏམ། ཅེ་ན། ཞེ་ན། ཤེ་ན།.

5. Class of Verbs (V)

5.1 Verbs (v)

བྱ་བ་བརྗོད་པའི་མིང་། Words that express actions, events, or a state of being.

5.1.1 Transitive Verbs (vt)

བྱེད་པོ་གཞན་འབྲེལ་གྱི་ལས་ཚིག Verbs that are able to take an object and an ergative (བྱེད་སྒྲ།) subject, for example, བསྐུམ། གཅོད། གདོན་པ། སྡོར་བ། བཙངས་པ། བྲིས་པ། བལྟ་བ།.

5.1.2 Intransitive Verbs (vi)

བྱེད་པོ་གཞན་དང་མི་འབྲེལ་གྱི་ལས་ཚིག Verbs that are incapable of taking an object and an ergative (བྱེད་སྒྲ།) subject, for example, འཁྱོམ། ཕལ་པ། དཀག ཐང་བ། ནོར་བ། སྡོམས། སོང་བ།.

6. Class of Adjectives (A)

6.1 Adjectives (a)

ཁྱད་ཆོས་བརྗོད་པའི་མིང་། An adjective usually follows the noun or the pronoun which it modifies. Expressing the nature, state, and characters of people or things, in a sentence primarily it act as a predicate, modifier, or complement.

6.1.1 Single Syllable Adjectives (ad)

སྒྲ་རྐྱང་མ་ཅན། Adjectives that are able to take the suffix པ། or བ། and can be modified by མི། ཤིན་ཏུ། or ཧ་ཅང་།, for example, དམར། ཆེ། བཟང་། ལེགས། མཛེས། གྲང་། དཔའ།.

6.1.2 Multi-Syllable Adjectives (as)

སྒྲ་འདུས་མ་ཅན།

6.1.2.1 Inextensible Multi-Syllable Adjectives (asb)

སྒྲ་འདུས་མའི་ཁྱད་ཆོས་སྒྲ་སྒྲིད་མི་ཆོག་པ་ཅན། Adjectives that are unable to take the suffix པ། or བ། and cannot be modified by མི། ཤིན་ཏུ། or ཧ་ཅང་།, but can be reduplicated, for example, ལྷང་ཀྱུ། མོག་མོག ལྷབ་ལྷབ། ཚད་མེད། མཐའ་ཡས། རབ་འབྱམས། དམར་དམར།.

6.1.2.2 Extensible Multi-Syllable Adjectives (ask)

སྒྲ་འདུས་མའི་ཁྱད་ཆོས་སྒྲ་སྒྲིད་ཆོག་པ་ཅན། Adjectives that are unable to take the suffix པ། or བ། and cannot be modified by མི། ཤིན་ཏུ། or ཧ་ཅང་།, occurring in the pattern where the first half lacks a vowel, and is followed by a second half that mirrors it but with a vowel, for example, རབ་རིབ། (རབ་བ་རིབ་བེ།) ལྷན་ལྷིན། (ལྷན་ན་ལྷིན་ནེ།) ཀྱག་ཀྱོག (ཀྱག་ག་ཀྱོག་གེ) .

6.2 Status Words (z)

ངང་ཚུལ་སྟོན་པའི་མིང་། Expressing the status, in a sentence primarily acting as a predicate, modifier, adverbial, or complement.

6.2.1 Shape Status Words (zx)

གཟུགས་ཀྱི་ངང་ཚུལ་སྟོན་པའི་མིང་། For example, སེ་རེ་རེ། ཐུ་ལུ་ལུ། ལམ་མེ། ཚུབས་སེ་ཚུབ།.

6.2.2 Sound Status Words (zs)

སྒྲའི་ངང་ཚུལ་སྟོན་པའི་མིང་། For example, ཤྲིར་རི་རི། ཙ་རུ་རུ། ཤྲིགས་སེ་ཤྲིག སང་ངེ་།.

7. Class of Adverbs (D)

7.1 Adverbs (d)

ཤུགས་འདོན་པའི་མིང་། An adverb can modify a verb, an adjective, indicating scope, degree etc. In the sentence it acts as adverbial adjunct.

7.1.1 Degree Adverbs (dc)

དངོས་ཚད་སྟོན་པའི་མིང་། Adverbs that only appear before an adjective, for example, ཧ་ཅང། ཆེས། ཤིན་ཏུ། ཁྱད་པར་དུ། ལྷག་པར། ངེས་པ།.

7.1.2 Frequency Adverbs (dp)

དུས་ཚད་སྟོན་པའི་མིང་། Adverbs that relate to times, procedure, or frequency, including adverbs that reduplicate, for example, འཕྲལ་དུ། མྱུར་དུ། སྔད་མར། དེ་མ་ཐག་ཏུ། སྐད་ཅིག་གིས། སྔར་ཡང་། ཡང་ནས་ཡང་།.

7.1.3 Scope Adverbs (dw)

ཁྱོན་ཚད་སྟོན་པའི་མིང་། Adverbs that relate to space, scope of meaning, or estimation, for example, མ་ལུས་པ། ཆ་ཚང་། ཁྱོད་སྡོམ། ཡོད་དོ་ཅོག ཕལ་ཆེར། ཕལ་མོ་ཆེར། ཕལ་ཆེ་བར། ཇི་བཞིན། ཇི་ལྟ་བ་བཞིན།.

7.1.4 Negative Adverbs (df)

དགག་པའི་མིང་། མི། and མ། are negative adverbs.

7.1.5 Mentality Adverbs (dx)

བསམ་ཚུལ་སྟོན་པའི་མིང་། Adverbs that modify the state of mind, for example, ཡང་དག་པར། མངོན་སུམ་དུ། དོན་དམ་པར། ནན་ཏན་དུ།.

8. Class of Interjections (E)

8.1 Interjections (e)

འབོད་ངག་གསལ་བྱེད་ཀྱི་མིང་། Words that express a sigh, a call, or a response, that could be a separate sentence, or a sentence that serves as an independent component, for example, ཀྱ་ཡེ། ཀྱེ་མ། ཨེ་མ། ཀྱོ་ཧུད། ཨ་ཙི། ཨ་ལ་ལ། ཨ་ཙ་མ། ཀོ། ཧེ།.

9. Class of Onomatopoeia (O)

9.1 Onomatopoeia (o)

སྒྲ་སྒྲུག་ཚུལ་སྟོན་པའི་མིང་། Words that are analogous in sound to the sound of something in nature, and which can not be separated into a sentence, for example, ལྷང་ལྷུང་། ཐག་ཐག དིར་དིར། བང་བང་། སང་སད། ཤིར་ཤིར།.

10. Class of Helping Words (U)

10.1 Helping Words (u)

གསལ་བྱེད་ཀྱི་མིང་།

10.1.1 Metaphor Helping Words (ub)

མཚུངས་པ་གསལ་བྱེད་ཀྱི་མིང་། Words that are connected with the objective of a metaphor, but are not separate words, cannot be reduplicated, and in general appear the end of the sentence, for example, མཚུངས། བཞིན། འདྲ།.

10.1.2 Existential Helping Words (uc)

གནས་པ་གསལ་བྱེད་ཀྱི་མིང་། Words that can apply to a single word or be appended to a past tense verb, meaning "there are." For example, ཡོད། གནས། ལྡན།.

10.1.3 Objective Helping Words (um)

དགོས་པ་གསལ་བྱེད་ཀྱི་མིང་། Generally, words that are applied to the end of a

sentence, and sometimes which cannot be used alone, for example, ཆེད། ཆེད་དུ། སླད། སླད་དུ། དོན་དུ། ཕྱིར་དུ།.

10.1.4 Determinative Helping Words (up)

དེ་ཉིད་གསལ་བྱེད་ཀྱི་མིང་། Words that are applied to the end of a sentence that express a positive or negative state, for example, མིན། མེད། ཡིན། རེད།.

10.1.5 Auxiliary Helping Words (ux)

For example, བྱེད། ཐུབ། དགོས།.

10.1.6 Tense Helping Words (us)

དུས་གསལ་བྱེད་ཀྱི་མིང་། For example, འགྱུར། ཀྱིན། གྱིན། གིན། ཚར། ཟིན། སོང་།.

10.1.7 Imperative Helping Words (uq)

For example, ཅིག་གུ། ཤོག ཅིག ཞིག ཤིག.

10.1.8 Terminal Helping Words (uz)

For example, གོ་ (ལ་སོགས་པ་བཅུ་གཅིག་དང་) ཅོ། ཞོ། ཤོ།.

10.1.9 Reason Helping Words (uy)

རྒྱུ་མཚན་གསལ་བྱེད་ཀྱི་མིང་། Words that are applied to the end of a sentence that express a reason — with or without the presence of an added ergative (བྱེད་སྒྲ་) particle, for example, སྐབས། གཞིས། རྐྱེན་གྱིས། ཕྱིར། དེ་བས། དེའི་ཕྱིར།.

11. Class of Declaration Words (y)

ཉི་སྒྲ།

12. Class of Grammatical Particles (g)

རྣམ་དབྱེ་ཞེས་པ་ནི་མིང་གི་མཐར་སྦྱར་བས་ཚིག་ཏུ་འགྲུབ་པར་བྱེད་པའི་རྐྱེན་ཏེ་སྒྲའི་ཆ་ཤས་ཉུས་པ་ཅན་ཞིག་ཡིན། Grammatical particles that are added to nouns and pronouns depending on syntax.

12.1 Ergative Particles (gx)

བྱེད་སྒྲ། The five ergative particles: གིས། ཀྱིས། གྱིས། འིས། ཡིས། Used behind a noun, pronoun, or noun phrase to form an ergative construction as the subject in a sentence that includes a transitive verb. It is also used to show a causal instrument, or the causal relationship of a sentence.

12.2 Genitive Particles (gz)

འབྲེལ་སྒྲ། There are five genitive particles: གི ཀྱི། གྱི། འི། ཡི། Used behind a noun, pronoun, or noun phrase to form a genitive construction, or in front of an-

other noun in an attributive sense to indicate the relationship of a modifier, possession, or repetition between the main words.

12.3 Locative Particles (gl)

ལ་དོན།   There are seven locative particles: སུ། རུ། ར། ལ། ཏུ། དུ། ན། Used behind a noun, pronoun, or noun phrase to form a locative construction. According to their different functions of the verb predicate, it is used to indicate either an object, adverbial, complement, or subject in a sentence.

12.4 Source Particles (gi)

འབྱུང་ཁུངས།   There are two source particles: ནས། ལས། Used behind a noun, pronoun, or noun phrase to form a source construction as an adverbial that shows the origin or source of things in a sentence.

13. Class of Connective Particles (C)

13.1 Connective Particles (c)

ཚིག་ཕྲད།

13.2 Conjunctions (cd)

དང་སྒྲ། For example, དང་།.

13.3 Word Set Decorations (cj)

For example, ཀྱང་། འང་། ཡང་།.

13.4 Remainder Words (cn)

ལྷག་བཅས་ལ་འཇུག་པའི་སྒྲ། For example, ཏེ། དེ། སྟེ། ཅིང་། ཞིང་། ཤིང་།.

14. Class of Affixes (H)

14.1 Affixes (h)

ཚིག་གྲོགས།

14.1.1 Prefixes (hq)

སྔ་རྒྱན། For example, ཨ།.

14.1.2 Infixes (hz)

བར་རྒྱན། For example, མ།, as in: རོང་མ་འབྲོག, ར་མ་ལུག, གདངས་མ་ཆར།.

14.1.3 Suffixes (hh)

ཕྱི་རྒྱན། For example, ཀ། ཁ། ག གེ ངེ་། དེ། བེ། མེ། འེ། རེ། ལེ། ནེ། སེ། པ། པོ། བ། བོ། ཅན། ལྡན། མཁན། གུ ཏུ་། བུ། ཙུ། ཅིག ཞིག བྱ། བྱེད། ཅེལ། ཚོ། ཆ། ཏ།.

15. Class of Semantemes (i)

བརླག་ཙང་གི་སྒྲ། (སྒྲ་བརླག་ཙང་ཡང་དོན་མེད་པའི་ཡི་གེ).

16. Class of Marks (l)

ཚེག་རྟགས།

16.1 Punctuation (lq)

ཚེག་རྟགས། For example, . , ; ? ! 《 》 “ — [ ] = +

16.2 Loan Words (lw)

གཡར་རྟགས། Chinese and foreign language words and symbols, for example, CCTV, A, @, %, $, &.

16.3 Vertical Lines (lz)

ཀྱེན་གྱི་ཡི་གེ For example, ༄༅།། ། །།།། ༄། ༄༅༅

17. Class of Poetic Terms (K)

མངོན་བརྗོད་ཀྱི་མིང་། For example, ནགས་ཀྱི་ཤ་ན། རྟ་ལྗང་ཅན། བཅུ་གཉིས་བདག་པོ། མགོ་སྐྱེས། ཁང་འཕྱུར་། རྟ་བདུན་དབང་པོ།.