『狄島夜話記』におけるアイヌ語について

安岡孝一 (やすおかこういち)*

1　はじめに

釋天嶺[1]『狄島夜話記』(享保 19 年 7 月 10 日) には、18 世紀初頭のアイヌ語が漢字で書かれている[2]。『狄島夜話記』は、いわゆる日本漢文で書かれており (図 1)、全体として見ると、漢字で書かれた漢文テキスト中に漢字で書かれたアイヌ語が紛れ込んでいる状態である。すこぶる読みにくい。

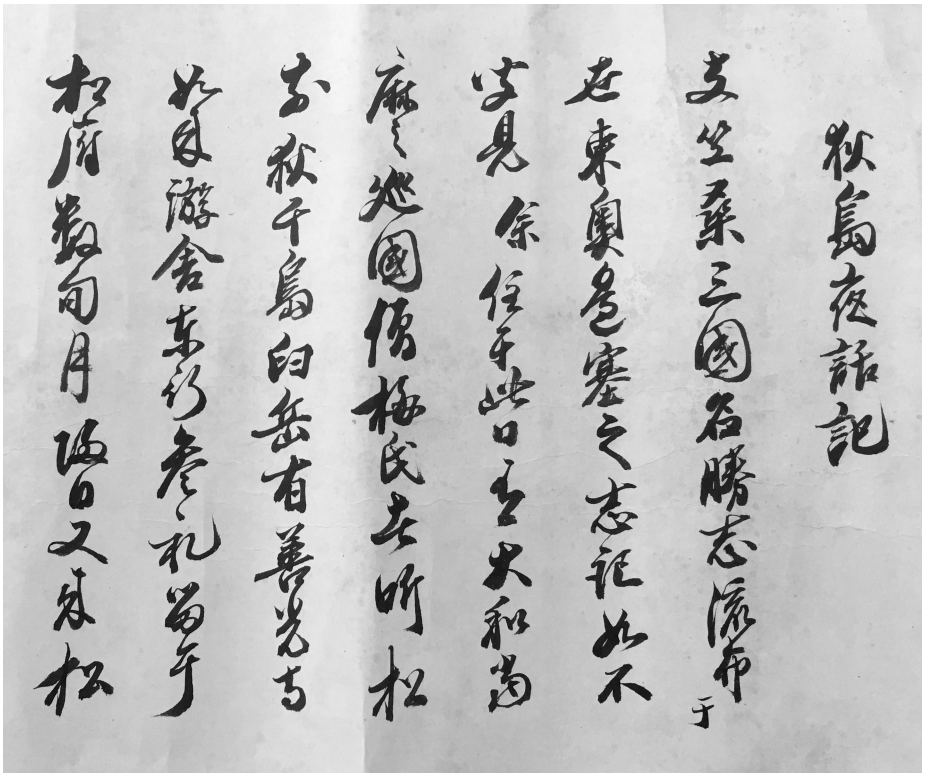


図 1: 『狄島夜話記』(東北大学附属図書館 丙C / 3-13 / 4) 冒頭部

釋天嶺『松島夜話』(松島: 東礀, 京都: 柳枝軒多左衛門, 江戸: 小川彦九郎, 元文 3 年 2 月) 巻之下 15 オ〜17 オには、『狄島夜話記』が多少表現を変えながらも収録されている。『松島夜話』には返り点や送り仮名がある (図 2) ことから、『狄島夜話記』より多少は読みやすいが、アイヌ語については割注も後注もない。

*京都大学人文科学研究所附属人文情報学創新センター

[1]松島瑞巖寺 105 世、号は性空。

[2]佐藤知己: 18 世紀前半のいくつかのアイヌ語資料について, 北海道大学文学研究院紀要, Vol.127 (2009 年 2 月), pp.29-58.

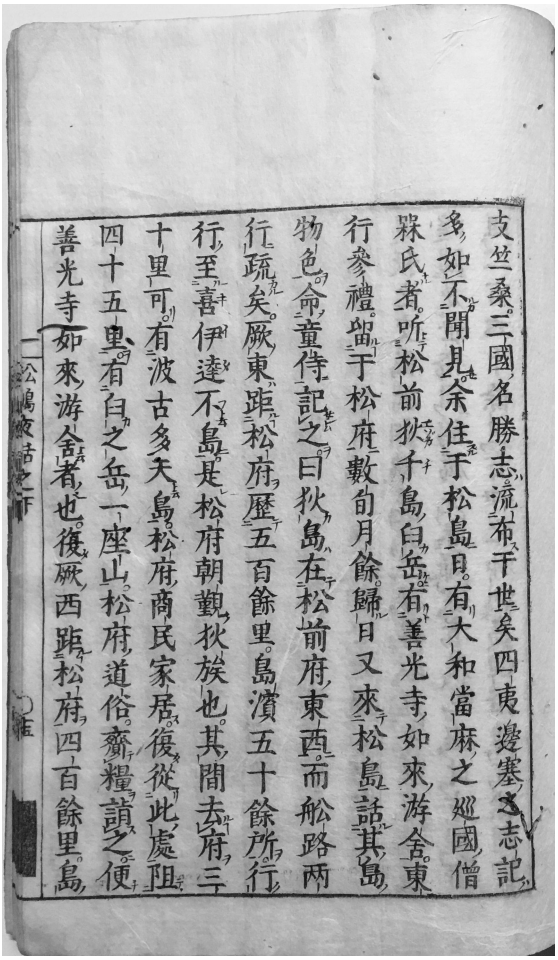


図 2: 『松島夜話』(京都大学附属図書館 1-26 / シ / 26) 巻之下 15 オ

筆者が班長を務める京都大学人文科学研究所共同研究班「古典中国語コーパスの応用研究」(班員: Christian Wittern、池田巧、李媛、劉冠偉、久保旭、守岡知彦、山崎直樹、二階堂善弘、鈴木慎吾、師茂樹、藤田一乘) では、現在、日本漢文の文法解析に精力を傾注しており、その道具立てとして、Universal Dependencies[3](以下「UD」と呼ぶ) を日本漢文に適用すべく研究を進めている。『狄島夜話記』は、日本漢文の中にアイヌ語が紛れ込む、というヤヤコシイ形式を取っており、相手にとって不足は無い。本発表では、UD による依存文法記述の助けを借りて、『狄島夜話記』における言語的「構造」を探る。

2 Universal Dependencies の概要

UD は、書写言語における品詞・形態素属性・依存構造 (係り受け関係) を、言語に関わらず記述する手法である。句構造を考慮せずに係り受け関係を記述することで、言語横断

[3]Marie-Catherine de Marneffe, Christopher D. Manning, Joakim Nivre, Daniel Zeman: Universal Dependencies, Computational Linguistics, Vol.47, No.2 (June 2021), pp.255-308.

表 1: CoNLL-U の各フィールド

1. ID: 単語ごとに付与されたインデックスで、文ごとに 1 から始まる整数。縮約語に対しては、単語の範囲を示すのも可。
2. FORM: 語、または、句読記号。
3. LEMMA: 基底形、語幹。
4. UPOS: UD で規定された言語普遍な品詞タグ (表 2)。
5. XPOS: 言語固有の品詞タグ。
6. FEATS: UD で規定された言語普遍な形態素属性のリスト。言語固有の拡張も可。
7. HEAD: 当該の単語の係り受け元 ID。係り受け元が無い場合は 0 とする。
8. DEPREL: UD で規定された言語普遍な係り受けタグ (表 3)。HEAD が 0 の場合は root とする。言語固有の拡張も可。
9. DEPS: 複数の係り受け元を持つ場合、全ての HEAD:DEPREL ペア。
10. MISC: その他のアノテーション。

表 2: UD 品詞タグ (UPOS)

Open class words	**Closed class words**	**Other**
ADJ 形容詞	ADP 側置詞	PUNCT 句読点
ADV 副詞	AUX 助動詞	SYM 記号
INTJ 感嘆詞	CCONJ 並列接続詞	X その他
NOUN 名詞	DET 限定詞	
PROPN 固有名詞	NUM 数詞	
VERB 動詞	PART 接辞	
	PRON 代名詞	
	SCONJ 従属接続詞	

表 3: UD 係り受けタグ (DEPREL)

	Nominals	**Clauses**	**Modifier Words**	**Function Words**
Core arguments	nsubj 主語 obj 目的語 iobj 間接目的語	csubj 節主語 ccomp 節目的語 xcomp 節補語		
Non-core dependents	obl 斜格補語 vocative 呼称語 expl 形式語 dislocated 外置語	advcl 連用修飾節	advmod 連用修飾語 discourse 談話要素	aux 動詞補助成分 cop 繋辞 mark 標識
Nominal dependents	nmod 体言による連体修飾語 appos 同格 nummod 数量による修飾語	acl 連体修飾節	amod 用言による連体修飾語	det 決定語 clf 類別語 case 格表示
Coordination	**MWE**	**Loose**	**Special**	**Other**
conj 接続 cc 接続語	fixed 固着 flat 並列 compound 複合	list 細目 parataxis 隣接表現	orphan 親なし goeswith 泣き別れ reparandum 言い損じ	punct 句読点 root 親 dep 未定義

性を高めており、全ての文法構造を単語間のリンクで記述するのが特徴である。

依存構造解析それ自体は、Tesnière の構造的統語論[4]に源を発し、Мельчук の有向グラフ記述[5]によって、一応の完成を見た手法である。その最大の特長は、いわゆる動詞中心主義によって言語横断的な記述が可能だという点にあり、Мельчук 依存文法をコンピュータ向けに洗練した UD においても、言語に関わらない記述、という特長が前面に押し出されている。UD における文法構造記述は、句構造を考慮せず、全てを単語間のリンクとして表現する。これにより、言語横断的な文法構造記述を可能としている。

UD 係り受けコーパスの交換用フォーマットとして、CoNLL-U と呼ばれるタブ区切りテキスト (文字コードは UTF-8) が規定されている。CoNLL-U の各行は各単語に対応しており、表 1 に示す 10 個のタブ区切りフィールドで構成される。ID・FORM・LEMMA は、単語そのものに関するフィールドである。UPOS・XPOS・FEATS は、単語の品詞と形態素属性に関するフィールドである。HEAD・DEPREL・DEPS は、単語の係り受けに関するフィールドである。

UD における係り受け関係は、単語間の有向グラフを HEAD と DEPREL で記述する。HEAD は、その単語に入る有向枝のリンク元 ID を示しており、DEPREL は、その有向枝における係り受けタグである。ただし、HEAD が 0 の場合、その枝に入るリンク元は存在しない。リンクの本数は単語の個数に等しく、各リンクのリンク先は、全て互いに異なっている。すなわち、各単語から出るリンクは複数の可能性があるが、各単語に入るリンクは 1 つだけである。なお、リンクはループしない。

UD の係り受けリンクは、Мельчук 依存文法の後裔にあたり、いわゆる動詞中心主義である。動詞をリンク元として、主語や目的語へとリンクする。修飾関係においては、被修飾語から修飾語へとリンクする。ただし、側置詞 (前置詞や後置詞) を体言の修飾語だとみなす点[6]が、Мельчук とは異なっている。また、コピュラ文においては動詞中心主義を採らず、補語をリンク元として、主語や繋辞へとリンクする。

なお、UD は単語長を規定しておらず、各言語ごとに、自由に単語長を決めることができる。古典中国語 UD[7]では、基本的に漢字 1 文字を 1 単語とみなしており、固有名詞等に限って複数の漢字からなる単語を用いている。アイヌ語 UD[8]では『アイヌ語沙流方言辞典』[9]を、作業上の単語認定に用いている。

---

[4]Lucien Tesnière: Éléments de Syntaxe Structurale, Paris: C. Klincksieck (1959).

[5]Igor A. Mel'čuk: Dependency Syntax: Theory and Practice, State University of New York Press (1988).

[6]Joakim Nivre: Towards a Universal Grammar for Natural Language Processing, CICLing 2015: 16th International Conference on Intelligent Text Processing and Computational Linguistics (April 2015), pp.3-16.

[7]安岡孝一, ウィッテルン クリスティアン, 守岡知彦, 池田巧, 山崎直樹, 二階堂善弘, 鈴木慎吾, 師茂樹, 藤田一乗: 古典中国語 (漢文) Universal Dependencies とその応用, 情報処理学会論文誌, Vol.63, No.2 (2022 年 2 月), pp.355-363.

[8]安岡孝一: Universal Dependencies によるアイヌ語テキストコーパス, 情報処理学会研究報告, Vol.2021-CH-127『人文科学とコンピュータ』, No.5 (2021 年 8 月 28 日), pp.1-8.

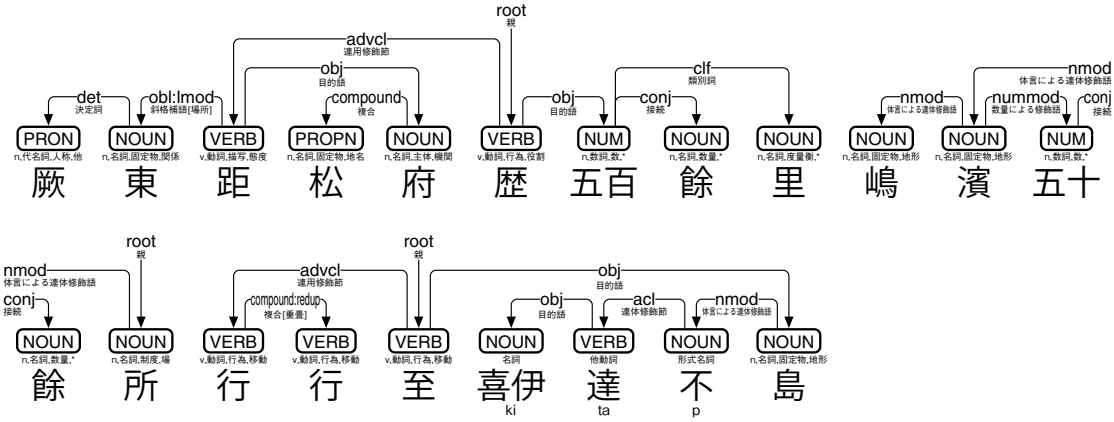
[9]田村すず子: アイヌ語沙流方言辞典, 東京: 草風館 (1996 年 9 月).

3 『狄島夜話記』とアイヌ語

『狄島夜話記』は、大和當麻寺の巡国僧「梅氏」が、松島瑞巖寺から松前府 (現在の松前町松城) 経由で有珠善光寺に行き、また松島瑞巖寺に戻ってきた際に語った内容を、釋天嶺が記録させた、という形式で書かれている。「梅」は「楳」(木へんに某) の異体字であり、巡国僧の名はあえて秘されている。「梅氏」の用務に関しては書かれておらず、松前府や有珠善光寺で見聞したというアイヌの特産物や風俗が、源義経北行伝説とともに書かれている。

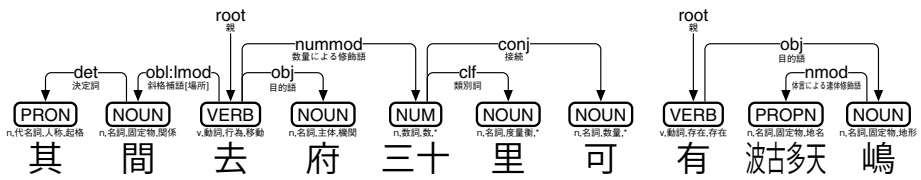
『狄島夜話記』に出てくるアイヌ語は、佐藤知己[2]によれば、志也茂、加茂伊、部素利、阿以農、免農古之、志也末奴、阿末茂、部幾利、勢礼保の 9 つである。ただし地名を含めると、もう少し増えるようだ。UD で解析しつつ、順に見ていくことにしよう。作業手順としては、最初に『狄島夜話記』の全文テキストを準備し、次にこのテキストを SuPar-Kanbun[7] で UD へ変換し、古典中国語 UD エディターで係り受けを訂正しつつ、漢文としてうまく読めないところを中心に、アイヌ語 UD として解釈可能かどうか検討した。

3.1 喜伊達不



松前の東 500 里あまり、最果ての地が喜伊達不である。現在の浜中町霧多布だとすると、せいぜい 700km しか離れていないので 500 里は言い過ぎだが、「梅氏」にとっては伝聞なので仕方ないだろう。上記の UD では「ki ta p」(茅を刈るところ)[10]に比定した。

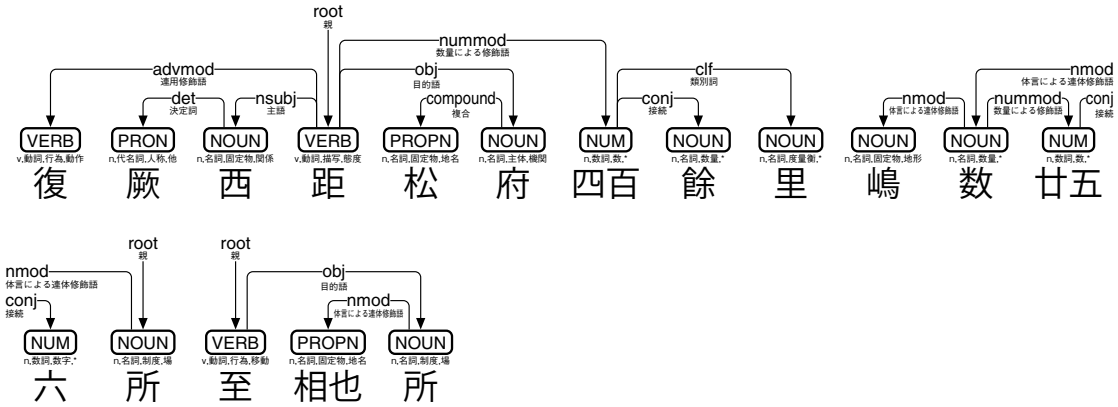
3.2 波古多天



[10]吉田東伍: 大日本地名辭書, 續篇, 東京: 冨山房 (1909 年 12 月).

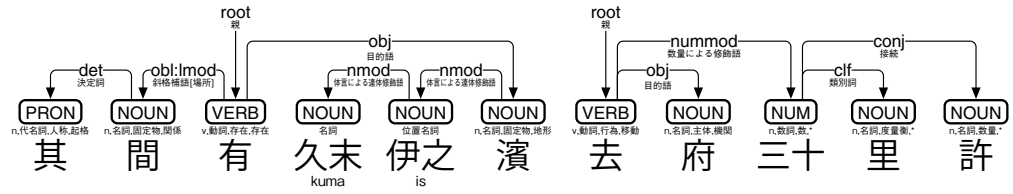
松前から喜伊達不の方向に 30 里ほど、松前藩による商業地の一つが波古多天である。現在の函館市弁天町にあった高龍寺[11]に「梅氏」が立ち寄ったなら、この地が「ハコタテ」と呼ばれているのを耳にしただろう。ただ、筆者としては「ハコタテ」がアイヌ語とは考えにくく、上記の UD ではアイヌ語に比定しなかった。

3.3 相也



松前の西 400 里あまり、最果ての地が相也である。現在の稚内市宗谷村宗谷だとすると、北に 700km 程度なので、西 400 里は方角からしてズレている。筆者としては「ソウヤ」がアイヌ語として理解しにくく、上記の UD ではアイヌ語に比定しなかった。

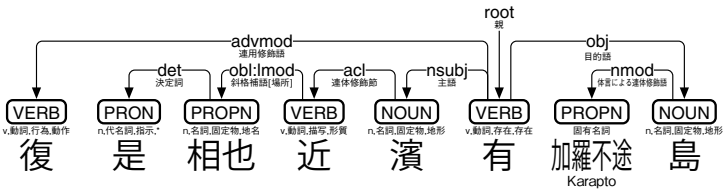
3.4 久末伊之



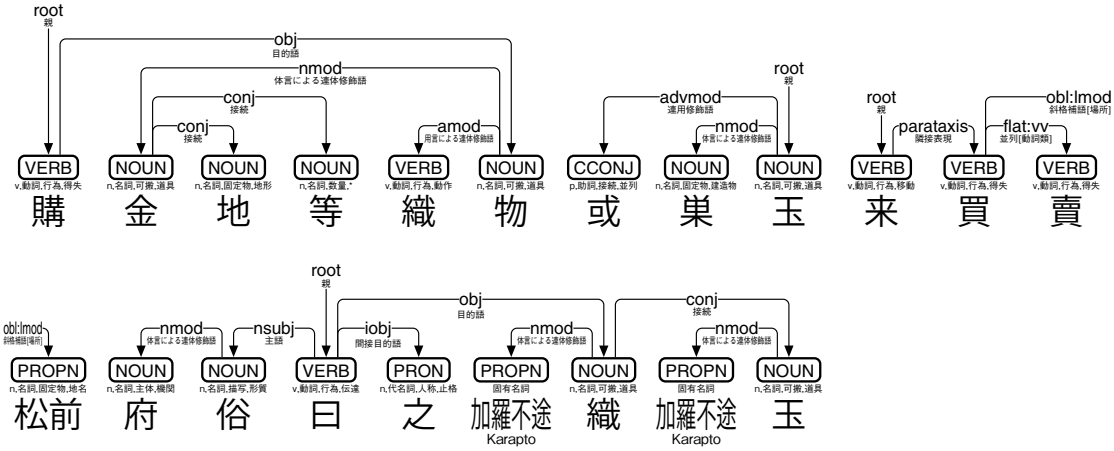
松前から相也の方向に 30 里ほど、松前藩による商業地の一つが久末伊之である。現在の八雲町相沼海岸[12]だとして、「kuma is」の「kuma」(干し竿) は理解できるが、位置名詞「is」が何を指しているのかわからない。「pis」(浜) が「is」に訛った可能性も考えられるが、とりあえず「is」のままにしておく。

[11]高龍寺は現在、函館市船見町に移っている。
[12]2005 年 9 月まで熊石町相沼海岸だった。

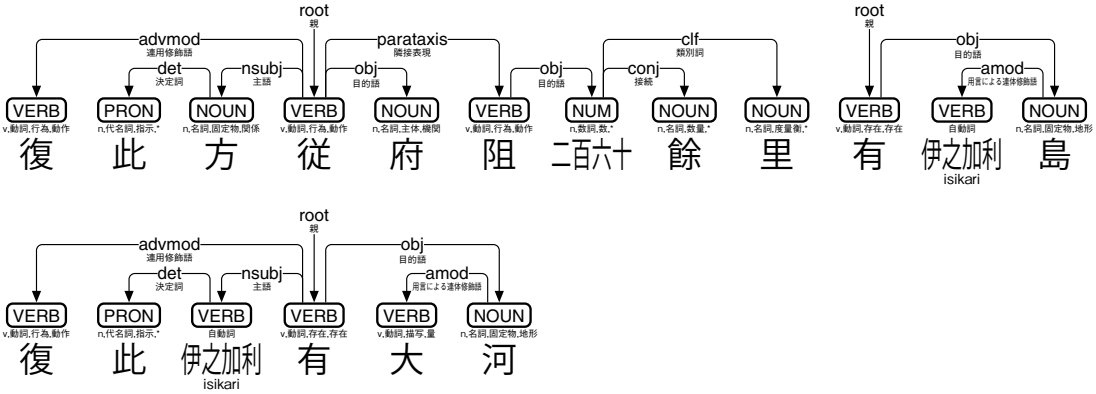
3.5 加羅不途



相也の対岸が、加羅不途である。いわゆる山丹交易[13]によって、加羅不途から相也へ織物や玉が入ってきており、これらの輸入品に対し、松前では加羅不途の名を付けて売買していたようである。なお、これらの UD では「Karapto」を 1 語[14]とした。



3.6 伊之加利



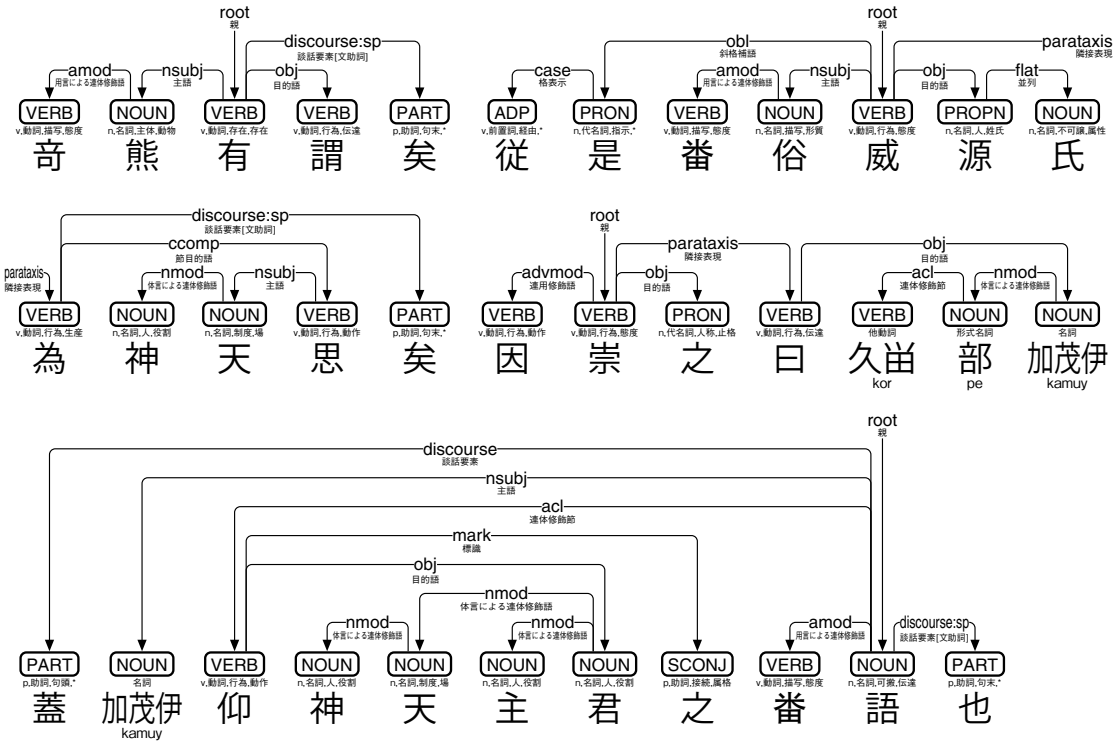
松前から相也の方向へ 260 里あまり、大きな川の河口に伊之加利がある。石狩川の河口だと考えられるが、石狩川は流域が絶えず変化しており、現在の河口とは異なる場所だと思われる。上記の UD では「isikari」(ぐずぐずする)[15]に比定した。アイヌ語の動詞が，固有名詞として使われていることになるため，UD の依存構造が多少，気持ち悪い．

[13]肥塚貴正: 蝦夷風俗彙纂, 東京: 稲田佐兵衛, 東京: 吉川半七, 安達郡二本松: 太田勘助 (1882 年 2 月), 後編巻 3.

[14]筆者としては「kar apto」(吹きつける雨) に比定するのも魅力的だったが、採用しなかった。

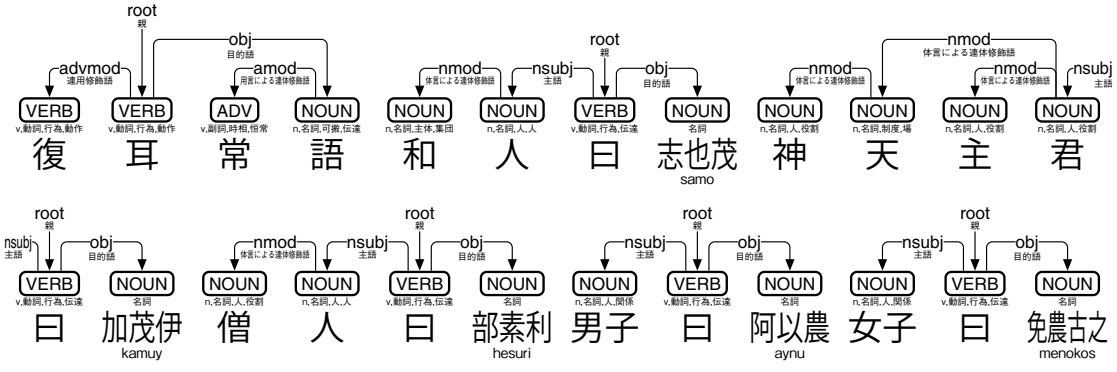
[15]М. М. Добротворский: Аинско-русский словарь, Казань: Университецкая типография (1875).

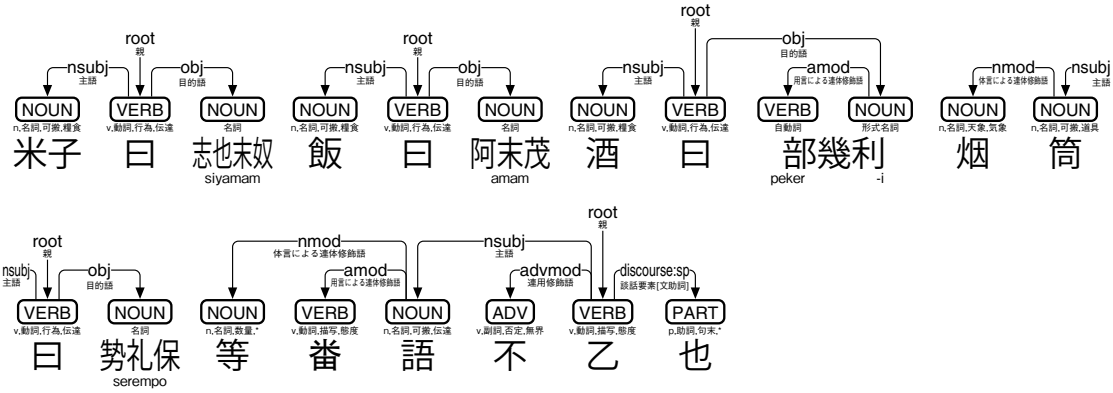
3.7 久甾部加茂伊



伊之加利の周辺では、熊を信仰の対象としており、久甾部加茂伊と呼ばれている。この信仰は源義経北行伝説に関係する、という不思議な説が開陳される。加茂伊については、神天の主君を仰ぐという説明があり「kamuy」に比定できるが、久甾部については説明がない。音の響きから「kor pe kamuy」としてみたが、筆者としては正直あまり自信がない。

3.8 志也茂、部素利、阿以農、免農古之、志也末奴、阿末茂、部幾利、勢礼保





「X 曰 Y」の形式で、漢語 X とアイヌ語 Y を対比している。加茂伊については、久甾部加茂伊での説明 (前節参照) の再掲だが、他の 8 つが現れるのはここだけなので、これらの対比をそのまま採用する[2]しかない。ただし、部幾利については「peker-i」(清らかなもの) と読めることから、清酒を指していると思われる。

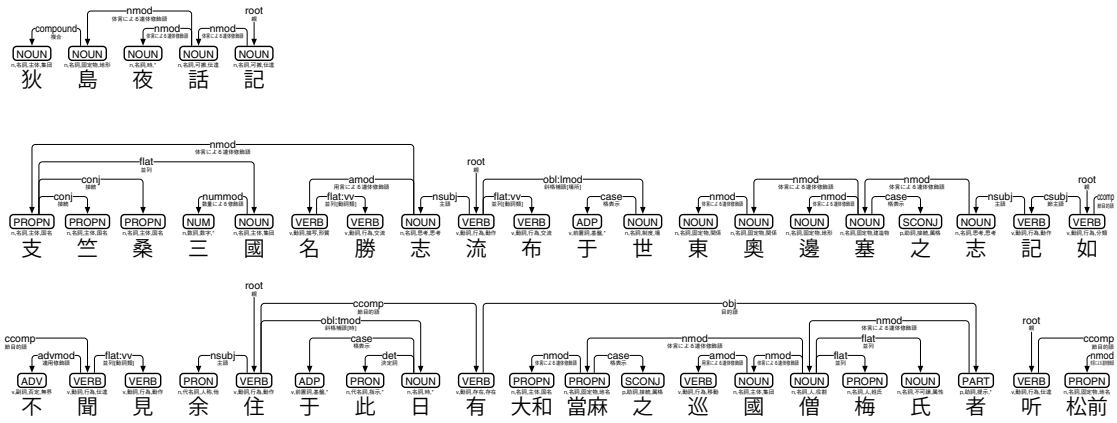
4 おわりに

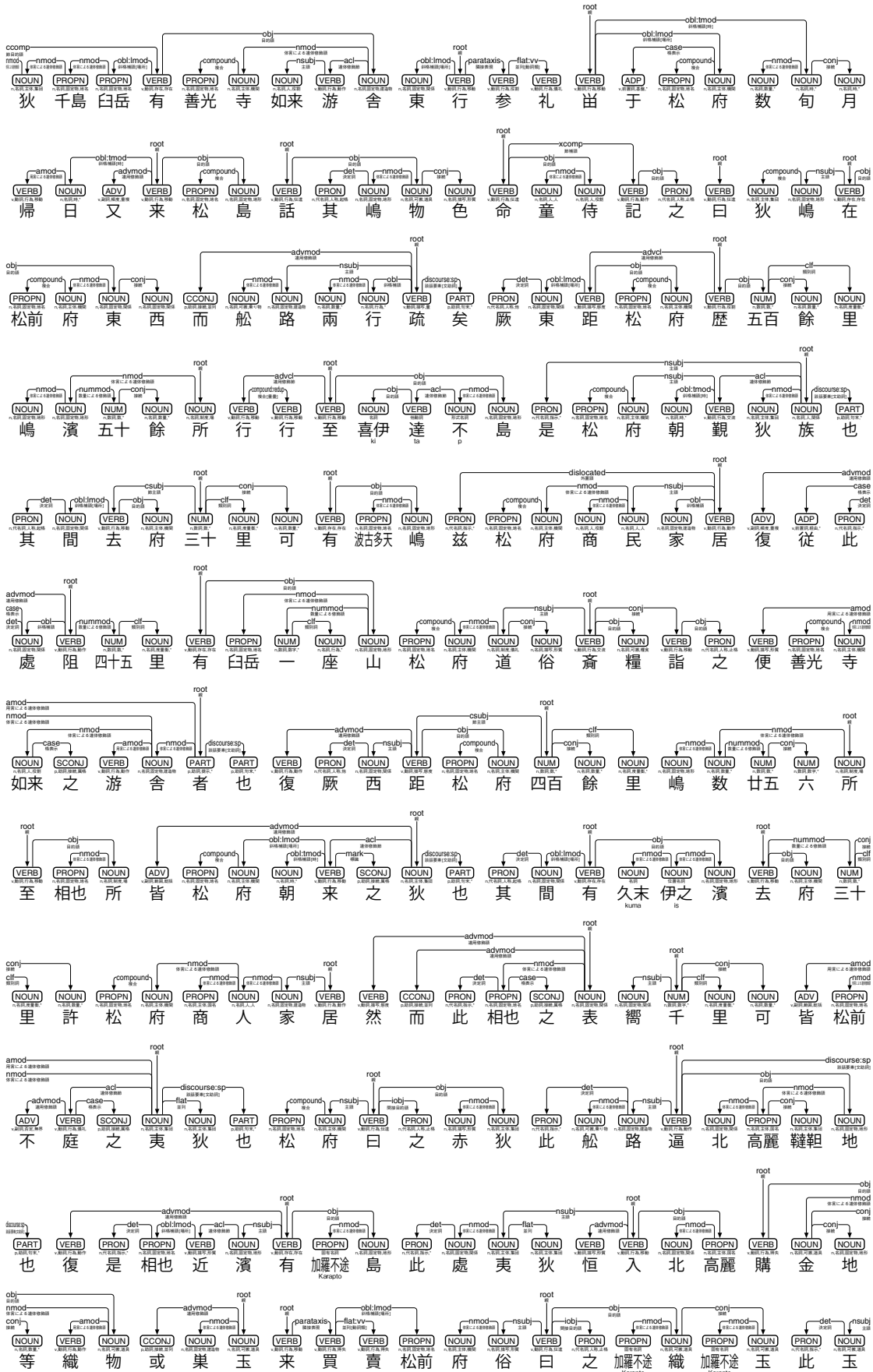
『狄島夜話記』に現れるアイヌ語に対し、UD の助けを借りつつ、依存構造解析をおこなった。喜伊達不、伊之加利、久甾部加茂伊、部幾利については、アイヌ語の動詞を語中に含んでいることから、UD として面白いものとなった。ただし、これらの語の自動解析は、かなりの困難が予想される。なお、本稿で作成した『狄島夜話記』全文 UD を、付録に示す。古典中国語 UD エディターとともに

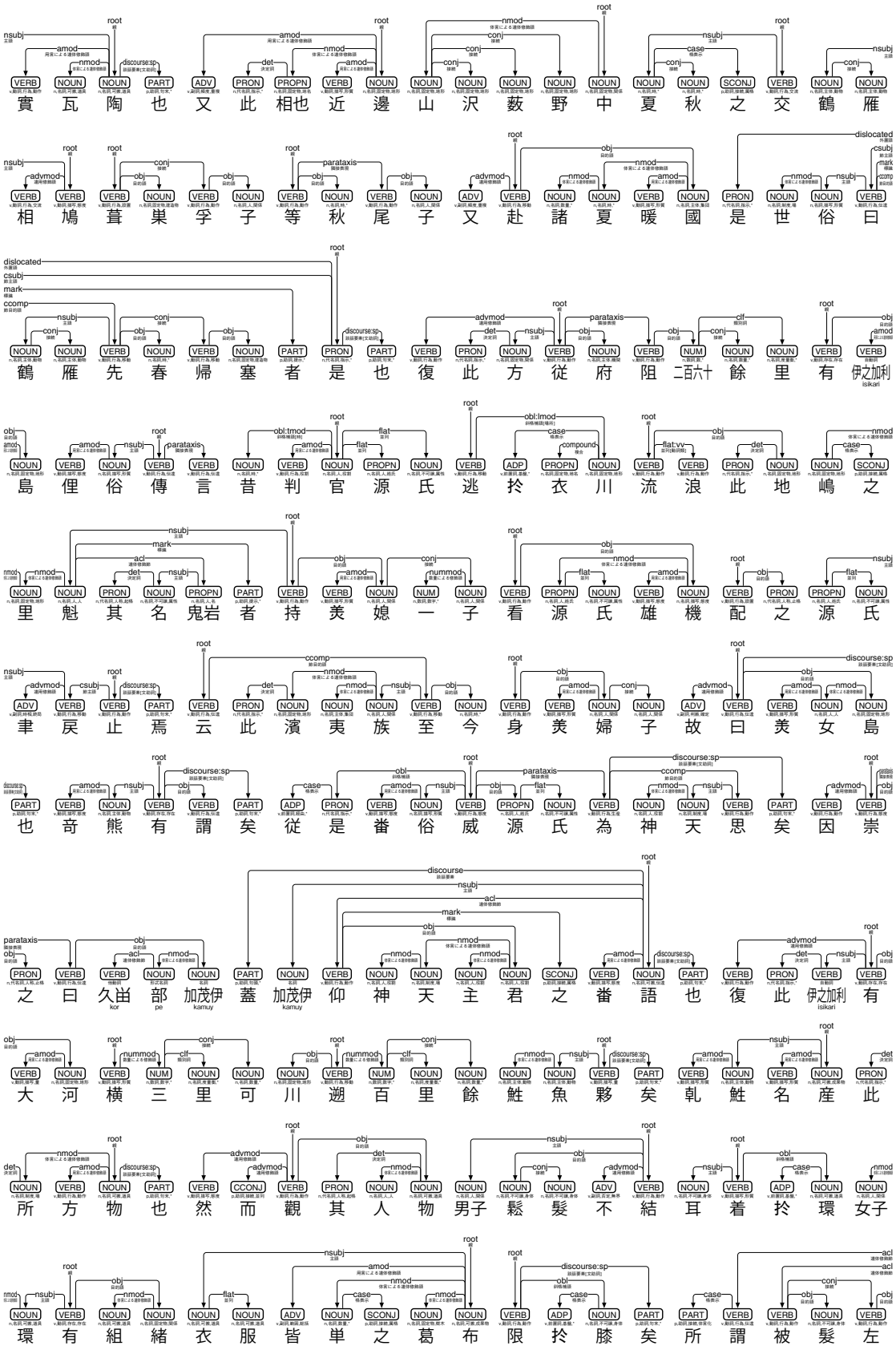
https://corpus.kanji.zinbun.kyoto-u.ac.jp/db-machine/~yasuoka/conllusvg/Ezogashima.html

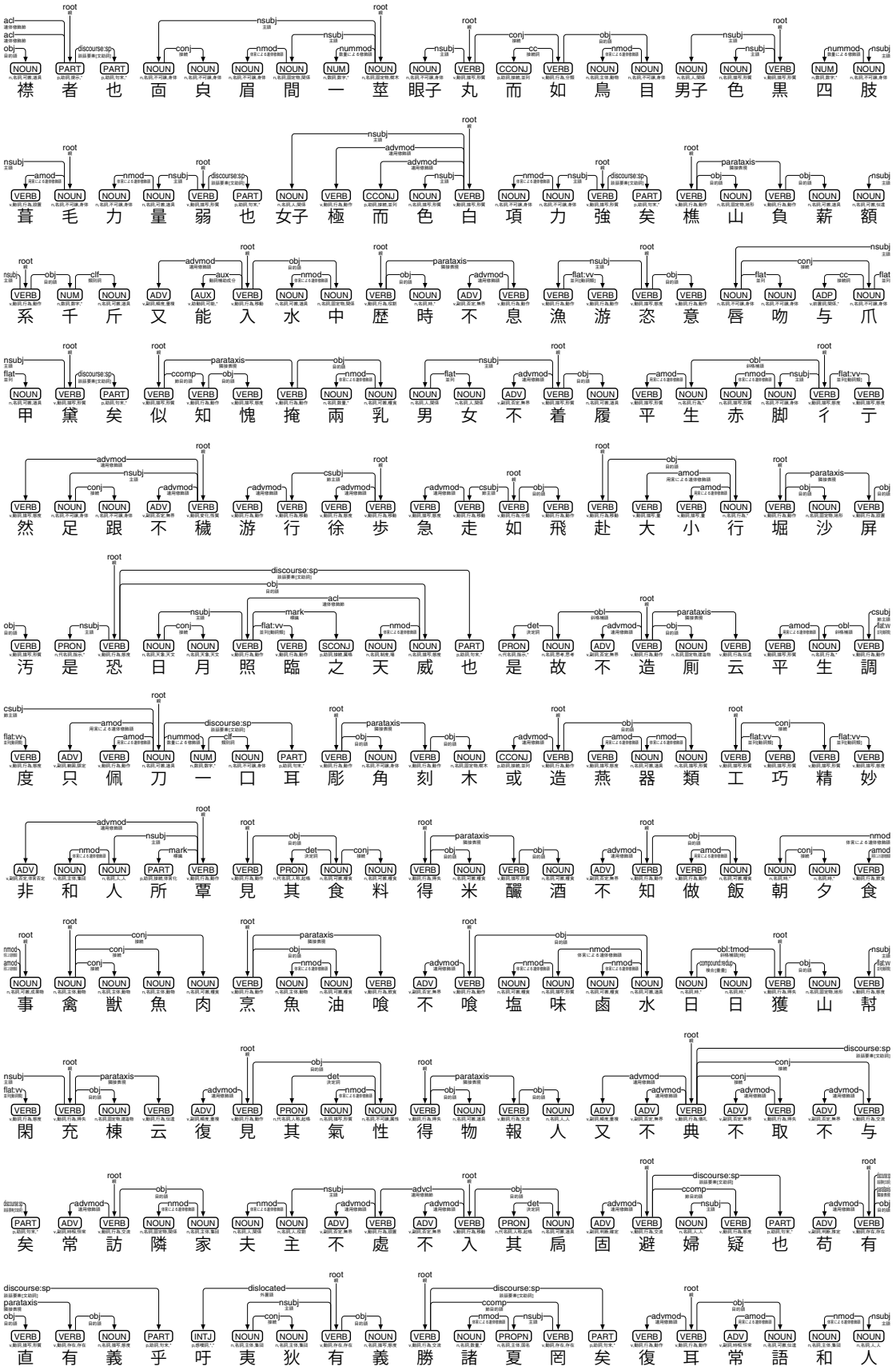
で公開しているので、ぜひアクセスしてみてほしい。

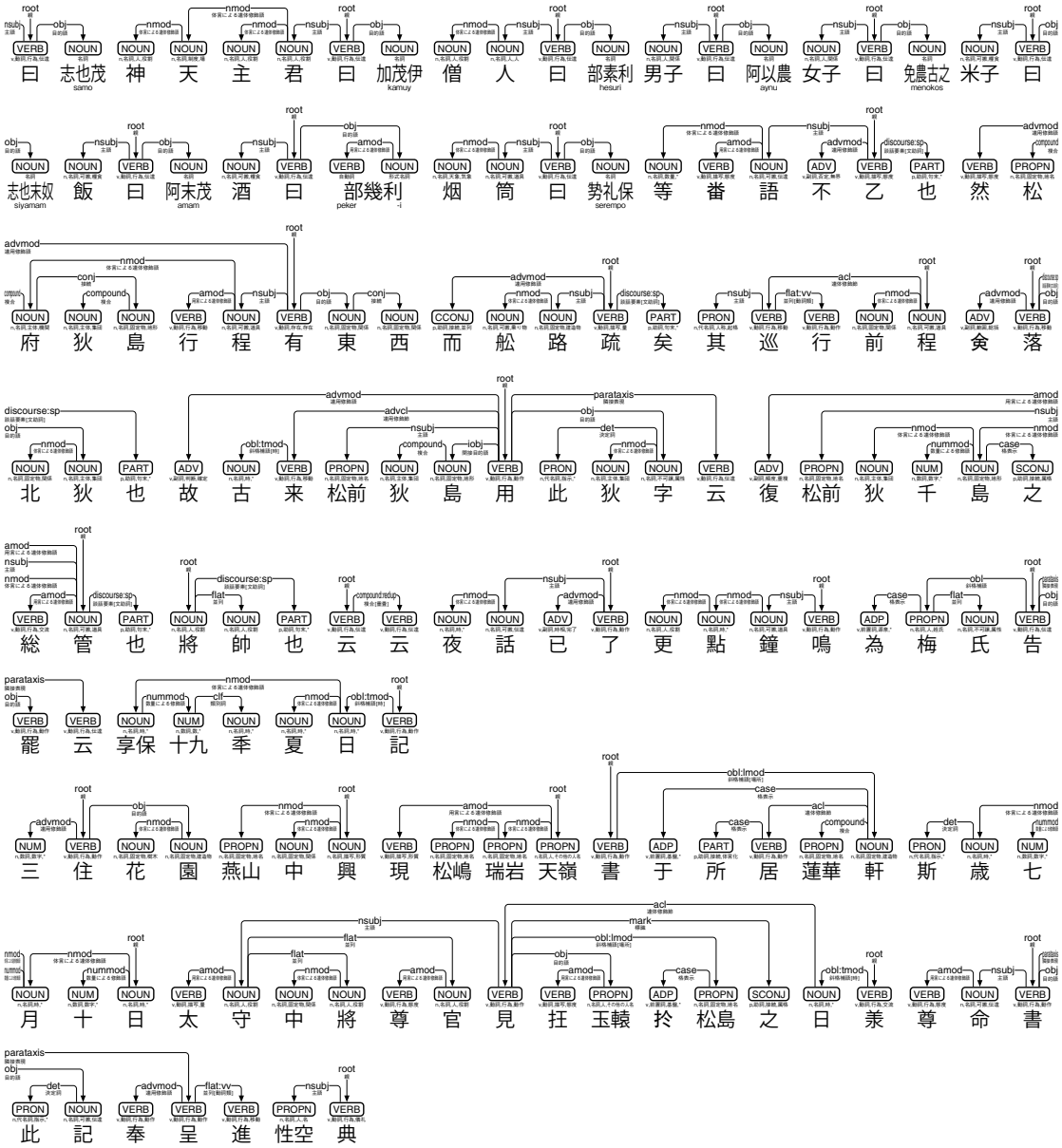
付録 『狄島夜話記』全文 Universal Dependencies




狄 千島 臼岳 有 善光 寺 如来 游 舎 東 行 参 礼 出 于 松 府 数 旬 月
帰 日 又 来 松 島 話 其 嶋 物 色 命 童 侍 記 之 曰 狄 嶋 在
松前 府 東 西 而 舩 路 兩 行 疏 矣 厥 東 距 松 府 歴 五百 餘 里
嶋 濱 五十 餘 所 行 行 至 喜伊 達 不 島 是 松 府 朝 覲 狄 族 也
其 間 去 府 三十 里 可 有 波古多天 嶋 茲 松 府 商 民 家 居 復 從 此
處 阻 四十五 里 有 臼岳 一 座 山 松 府 道 俗 斎 糧 詣 之 便 善光 寺
如来 之 游 舎 者 也 復 厥 西 距 松 府 四百 餘 里 嶋 数 廿五 六 所
至 相也 所 皆 松 府 朝 来 之 狄 也 其 間 有 久末 伊之 濱 去 府 三十
里 許 松 府 商 人 家 居 然 而 此 相也 之 表 嚮 千 里 可 皆 松前
不 庭 之 夷 狄 也 松 府 曰 之 赤 狄 此 舩 路 逼 北 高麗 韃靼 地
也 復 是 相也 近 濱 有 加羅不途 島 此 處 夷 狄 恒 入 北 高麗 購 金 地
等 織 物 或 巣 玉 来 買 賣 松前 府 俗 曰 之 加羅不途 織 加羅不途 玉 此 玉






曰 志也茂 神 天 主 君 曰 加茂伊 僧 人 曰 部素利 男子 曰 阿以農 女子 曰 鬼ヶ乃 米子 曰
志也末奴 飯 曰 阿末茂 酒 曰 部幾利 烟 筒 曰 勢礼保 等 雀 語 不 乙 也 然 松
府 狭 島 行 程 有 東 西 而 舩 路 疏 矣 其 巡 行 前 程 奄 落
北 狭 也 故 古 来 松前 狭 島 用 此 狭 宇 云 復 松前 狭 千 島 之
総 管 也 將 帥 也 云 云 夜 話 已 了 更 點 鐘 鳴 為 梅 氏 告
罷 云 享保 十九 季 夏 日 記
三 住 花 園 燕山 中 興 現 松嶋 瑞岩 天嶺 書 于 所 居 蓮華 軒 斯 歳 七
月 十 日 太 守 中 將 尊 官 見 扞 玉轅 於 松島 之 日 兼 尊 命 書
此 記 奉 呈 進 性空 典