

『蝦夷島奇観』における漢訳アイヌ語について

安岡孝一(やすおかこういち)*・安岡素子(やすおかもとこ)**

1 はじめに

秦檜磨(村上島之允)『蝦夷島奇観』(東京国立博物館 QB-10015、寛政12年序、文化4年成立)には、18世紀末のアイヌ語の文章がカタカナで示されている。これらのアイヌ語の文章には、漢文が並行して示されており、いわば漢訳アイヌ語となっている。ただし、これらの文章は、アイヌ語としても、漢文(古典中国語)としても、語順が多少変であり、互いに影響を受けているように見える。『蝦夷島奇観』の祖型^[1]とされる『蝦夷見聞記』(北海道大学附属図書館 旧記061、寛政10年序、昭和13年敷写)では、漢文ではなく和文が並行して示されており、あるいはアイヌ語から和文へ翻訳されたのちに、漢文へと重訳された可能性も考えられる。

本稿では、Universal Dependencies (UD) の助けを借りて、『蝦夷島奇観』の漢訳アイヌ語の背後にある言語的「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^[2]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^[3]に源を発し、Мельчук の有向グラフ記述^[4]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

*京都大学人文科学研究所附属人文情報学創新センター

**京都外国語大学外国語学部日本語学科

^[1]佐々木利和:「蝦夷島奇観」について, 蝦夷島奇観, 東京: 雄峰社 (1982年4月), pp.227-244.

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

^[3]Lucien Tesnière: Éléments de Syntaxe Structurale, Paris: C. Klincksieck (1959).

^[4]Igor A. Mel'čuk: Dependency Syntax: Theory and Practice, New York: State University of New York Press (1988).

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的な品詞タグ (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的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的な係り受けタグ (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 繫辞 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

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^[5]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、主語や繫辞へとリンクする。

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。本発表のアイヌ語UD^[6]では『アイヌ語沙流方言辞典』^[7]を、作業上の単語認定に用いた。古典中国語UD^[8]では、漢字1文字を1単語とみなしつつ、固有名詞等に限って複数の漢字からなる単語を用いた。近代日本語UD^[9]では、国語研短単位^[10]を単語長として用いた。

3 『蝦夷島奇観』の漢訳アイヌ語

秦億磨(村上島之允)『蝦夷島奇観』(東京国立博物館QB-10015)は重要文化財指定を受けており、カラー画像131枚がWWW公開^[11]されている。これらの画像のうち、アイヌ語の文章とその漢訳を含む4枚(画像番号C0007129^[12]・画像番号C0012761・画像番号

^[5]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.

^[6]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.

^[7]田村すず子: アイヌ語沙流方言辞典, 東京: 草風館 (1996年9月).

^[8]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.

^[9]安岡孝一: 形態素解析部の付け替えによる近代日本語(旧字旧仮名)の係り受け解析, 情報処理学会研究報告, Vol.2020-CH-124 『人文科学とコンピュータ』, No.3 (2020年9月5日), pp.1-8.

^[10]近藤明日子: 近代文語 UniDic 短単位規程集, Ver.1.1, 立川: 国立国語研究所コーパス開発センター (2016年3月).

^[11]<https://webarchives.tnm.jp/imgsearch/search?invnum=QB-10015>

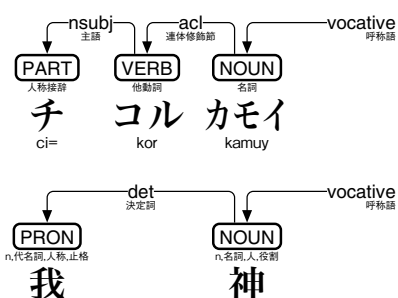
^[12]画像番号C0007129と画像番号C0012808は、ほぼ同一部分を撮影している。

C0012791・画像番号 C0012821) に対し、アイヌ語・古典中国語・近代日本語 UD コーパスを作成した。アイヌ語については、係り受け解析エンジン *esupar*^[13] のアイヌ語 DeBERTa モードで仮コーパスを作成し、その結果をアイヌ語 UD エディターで編集する、という手順^[14]で、アイヌ語 UD コーパスを作成した。古典中国語(漢文)については、係り受け解析エンジン *SuPar-Kanbun*^[15] で仮コーパスを作成し、その結果を古典中国語 UD エディターで編集する、という手順^[18]で、古典中国語 UD コーパスを作成した。近代日本語については、係り受け解析エンジン *SuPar-UniDic*^[16] の近代文語モードで仮コーパスを作成し、その結果を近代日本語 UD エディターで編集する、という手順^[17]で、近代日本語 UD コーパスを作成した。作成結果^[18]を付録 A に示す。

さらに、アイヌ語の各文とその漢訳について、UD の助けを借りて検討を加えた。検討過程と結果を、以下に述べる。

3.1 我神至今為神今日送兄

画像番号 C0007129 中の「チコルカモイ タ子バツクノ カモイニマヌワ タンドアナキ子ヲ マンデエアンルイタバン」(ci=kor kamuy tane pakno kamuy ne manu wa tan to anakne omante e=an ruwe tapan) に対する漢訳「我神至今為神今日送兄」を、順に見てみよう。



アイヌ語の「ci=kor kamuy」と漢文の「我神」は、いずれも呼びかけとして妥当だと考えられる。ただ、一人称(除外的複数)主格の「ci=」と一人称(単複同形)止格の「我」を、同一視してよいかどうかは、議論のあるところだろう。

^[13]<https://github.com/KoichiYasuoka/esupar>

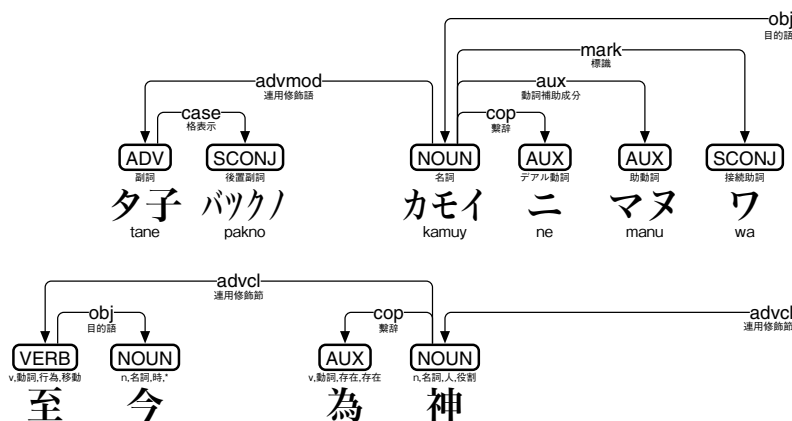
^[14]安岡孝一・安岡素子: ローマ字・カタカナ・キリル文字によるアイヌ語 Universal Dependencies の可能性, Evidence-based Linguistics Workshop 2023 発表論文集 (2023 年 9 月), pp.47-60.

^[15]<https://github.com/KoichiYasuoka/SuPar-Kanbun>

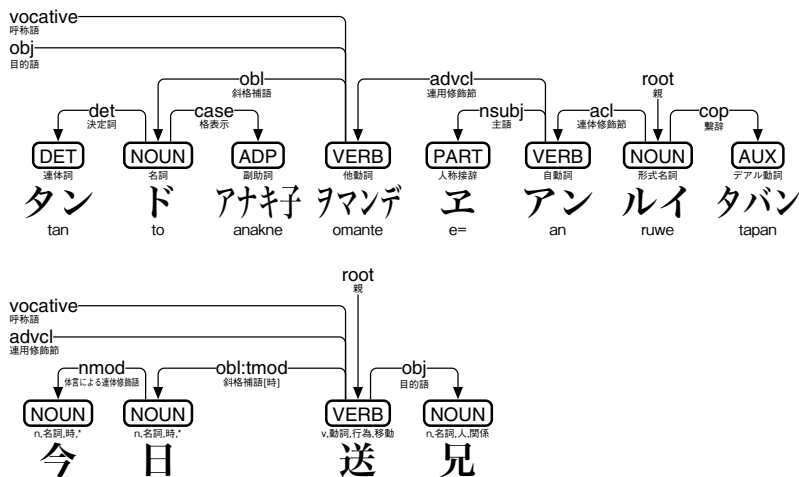
^[16]<https://github.com/KoichiYasuoka/SuPar-UniDic>

^[17]安岡孝一, 安岡素子: アイヌ語訳『五倫名義解』 Universal Dependencies 並行コーパスへの挑戦, 東洋学へのコンピュータ利用, 第 36 回研究セミナー (2023 年 7 月 28 日), pp.3-37.

^[18]UD コーパス作成に用いた係り受け解析エンジンは、学際大規模情報基盤共同利用・共同研究拠点公募型共同研究『単語間に区切りのない書写言語における係り受け解析エンジンの開発』の成果である。また、各言語 UD エディターと並行コーパス管理システムの開発、およびそれらを用いたコーパス作成作業は、文部科学省『AI 等の活用を推進する研究データエコシステム構築事業』の支援を受けている。



「tane pakno」を「至今」と訳すのは、語順の入替という点で妥当だと考えられる。「kamuy ne manu wa」を「為神」と訳すのは、「為」をコピュラ動詞とみなす場合において、語順の入替を含め妥当だと考えられる。

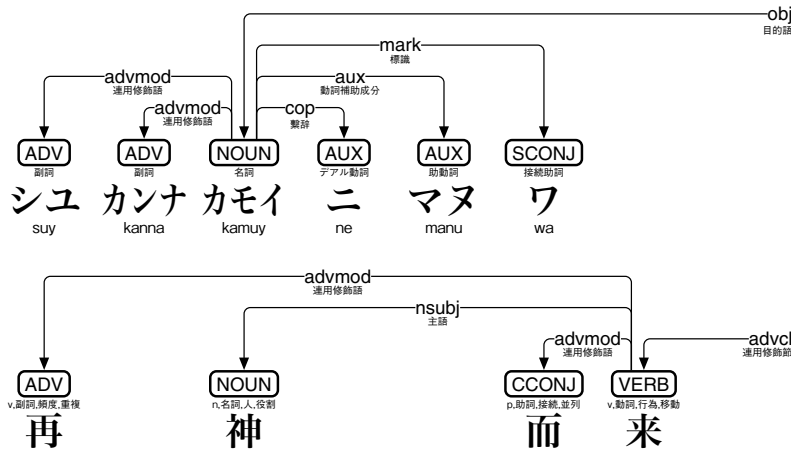


「tan to anakne」を「今日」と訳すのは、まあ妥当だと考えられる。一方、「omante e=an」はアイヌ語としては語順がおかしく、「e=omante an」の方が適切^[19]である。漢文の「送兄」に、アイヌ語の語順が引きずられた可能性が高い。なお、文末の「ruwe tapan」は、漢訳されていない。

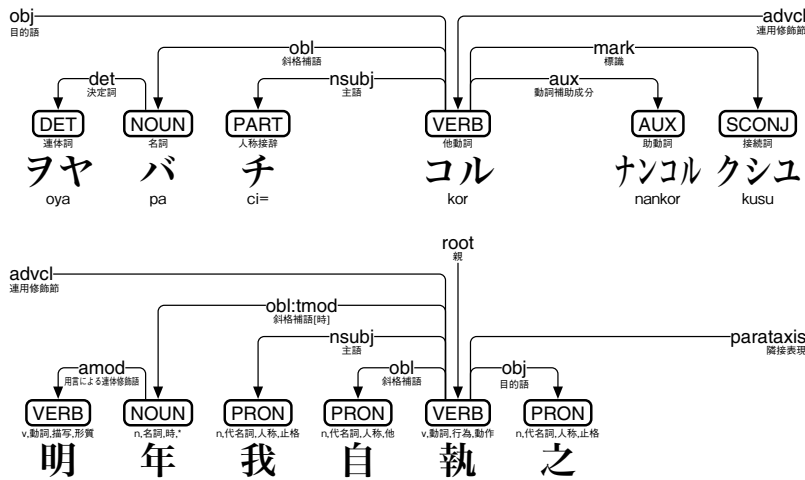
3.2 再神而来明年我自執之今兄敢辞

画像番号 C0007129 中の「シユカンナカモイニマヌワ ヲヤバチコルナンコルクシユ タ子 アナキ子 エオロタ サランバクキナンコンナ」(suy kannā kamuy ne manu wa oya pa ci=kor nankor kusu tane anakne e=or ta saranpa ku=ki nankor na) に対する漢訳「再神而来明年我自執之今兄敢辞」を、順に見てみよう。

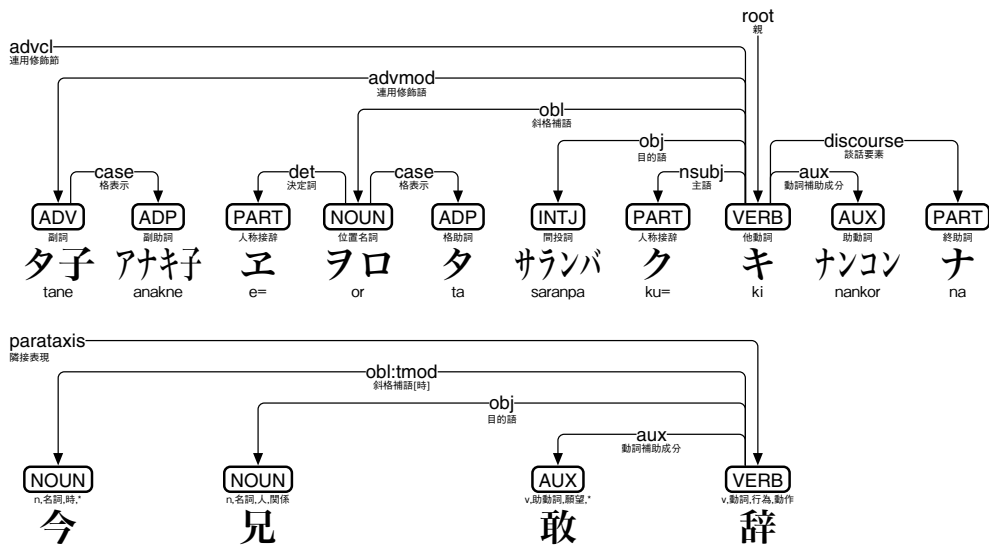
[19] 『蝦夷見聞記』十六丁ウでは、正しく「エヲマンデアン」となっている(付録 B 参照)。



アイヌ語の「suy kanna kamuy ne manu wa」も、漢文の「再神而来」も、いずれも語順は正しい。しかし、意味が対応していない。端的には、アイヌ語のコピュラ動詞「ne」を、漢文では誤って「来」と訳してしまっている。ここは3.1節の漢訳にしたがって、せめて「再為神也」と訳さなければ、「suy kanna kamuy ne manu wa」の意味を反映しきれないと思われる。



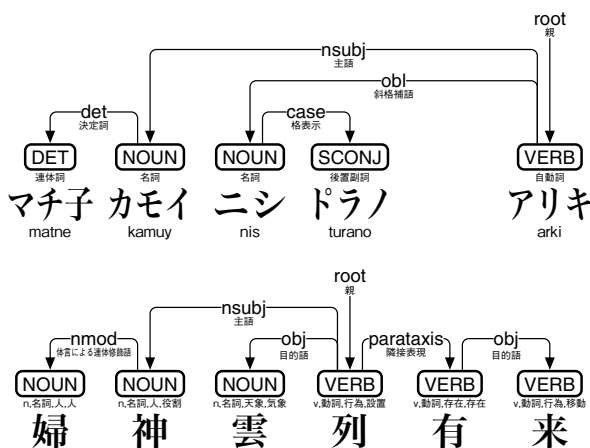
「oya pa」を「明年」と訳すのは、妥当だと考えられる。しかし「ci=kor nankor kusu」と「我自執之」は、意味が全く対応していない。「kor」は漢文の「有」にあたる動詞なので、目的語を補って「明年我等有兄矣」とでも訳さなければ、「oya pa ci=kor nankor kusu」の意味を反映しきれないと思われる。



「tane anakne」を「今」と訳すのは、まあ妥当だと考えられる。「saranpa ku=ki nankor na」を「敢辞」と訳すのは、「辞」を「さよならする」とみなす場合において妥当だが、一人称(単数)主格の「ku=」が反映できているかどうかは、議論のあるところだろう。また、「e=or ta」を「兄」とだけ訳すのは、「or ta」を訳し忘れている。ここは「於兄」として、せめて「今敢辞於兄」と訳さなければ、「tane anakne e=or ta saranpa ku=ki nankor na」の意味を反映しきれないと思われる。

3.3 婦神雲列有来

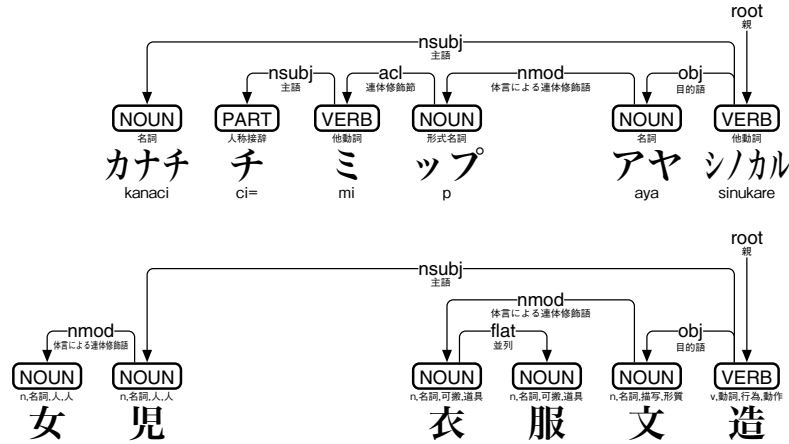
画像番号 C0012761 中の「マチ子カモイ ニシドラノアリキ」(matne kamuy nis turano arki) に対する漢訳「婦神雲列有来」を見てみよう。



「nis turano」と「雲列」が同一の語順となっている。これはアイヌ語としては正しい語順だが、漢文としては語順がおかしく、「列雲」の方が適切である。アイヌ語の「nis turano」もしくは日本語の「雲を列ねて」に、漢文の語順が引きずられたと考えられる。

3.4 女兒衣服文造

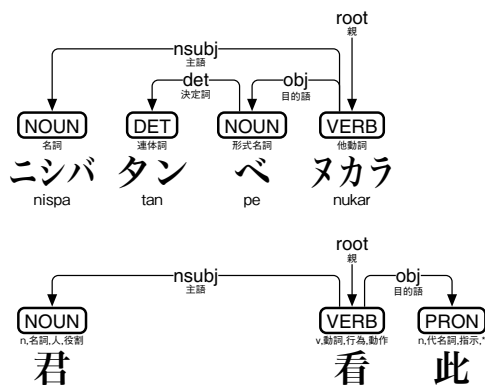
画像番号 C0012791 の標題部分「カナチ チミップ アヤシノカル」(kanaci ci=mi p aya sinukare) に対する漢訳「女兒衣服文造」を見てみよう。



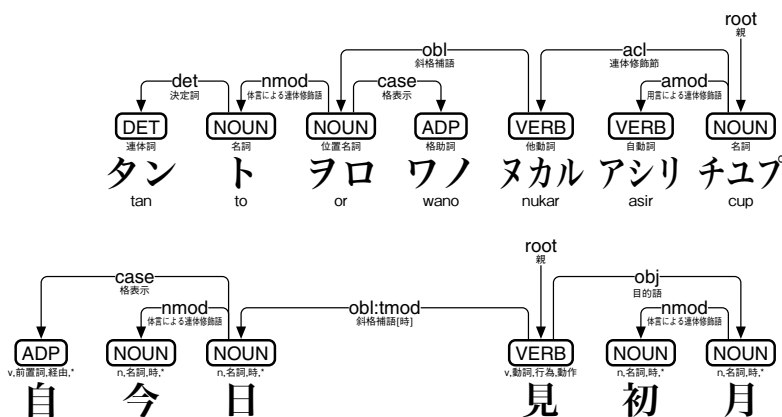
「aya sinukare」と「文造」が同一の語順となっている。これはアイヌ語としては正しい語順だが、漢文としては語順がおかしい。しかも「sinukare」(見てもらう)と「kar」(造る)を取り違えている。これらを考えあわせると、せめて「女兒請看衣服之文」と訳さなければ、「kanaci ci=mi p aya sinukare」の意味を反映しきれないと思われる。

3.5 君看此自今日見初月

画像番号 C0012821 中の「ニシバ タンベヌカラ タントヲロワノ ヌカル アシリ チュプ」(nispa tan pe nukar tan to or wano nukar asir cup) に対する漢訳「君看此自今日見初月」を、順に見てみよう。



アイヌ語の「nispa tan pe nukar」は、目的語「tan pe」が動詞「nukar」の直前にある。漢文の「君看此」は、目的語「此」が動詞「看」の直後にある。アイヌ語と漢文で語順が入れ替わっており、いずれも妥当だと考えられる。



「tan to or wano」を「自今日」と訳すのは、語順の入替という点で妥当だと考えられる。一方、「nukar asir cup」と「見初月」が同一の語順となっている。これは漢文としては正しい語順だが、アイヌ語としては語順がおかしく、「asir cup a=nukar」の方が適切^[20]である。漢文の「見初月」に、アイヌ語の語順が引きずられたと考えられる。

4 まとめ

ここまでの議論を踏まえ、『蝦夷島奇観』における漢訳アイヌ語について、まとめよう。

- 「ci=kor kamuy tane pakno kamuy ne manu wa tan to anakne omante e=an ruwe tapan」と「我神至今為神今日送兄」では、アイヌ語「omante e=an」の語順が漢文に引きずられており、「e=omante an」の方が適切である。
- 「suy kanna kamuy ne manu wa oya pa ci=kor nankor kusu tane anakne e=or ta saranpa ku=ki nankor na」と「再神而来明年我自執之今兄敢辞」では、漢訳が不十分でアイヌ語に対応しきれていない。
- 「matne kamuy nis turano arki」と「婦神雲列有来」では、漢文「雲列」の語順がアイヌ語(もしくは日本語)に引きずられており、「列雲」の方が適切である。
- 「kanaci ci=mi p aya sinukare」と「女兒衣服文造」では、漢訳が不十分でアイヌ語に対応しきれていない。
- 「nispā tan pe nukar tan to or wano nukar asir cup」と「君看此自今日見初月」では、アイヌ語「nukar asir cup」の語順が漢文に引きずられており、「asir cup a=nukar」の方が適切である。

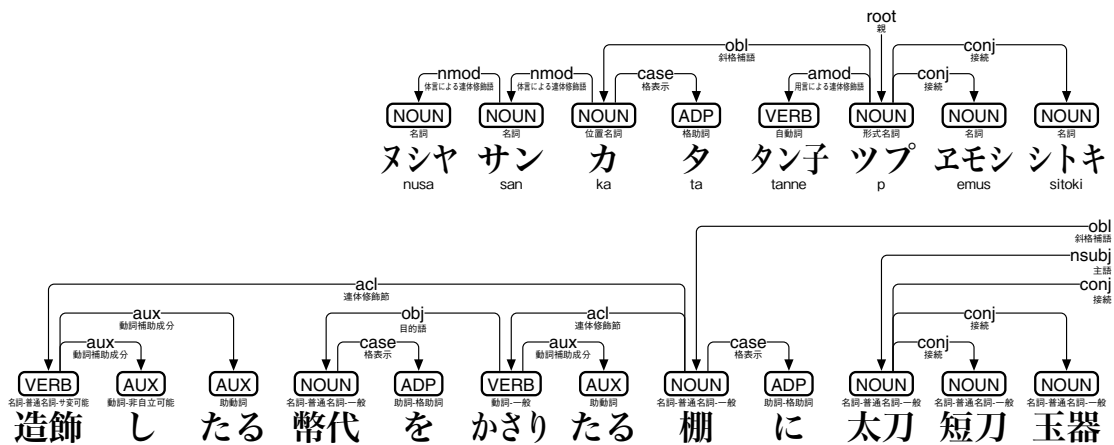
『蝦夷島奇観』は、冒頭の蝦夷諸島圖(画像番号 C0012760)で『日本書記』を漢文のまま引用しており、その点で、アイヌ語の漢訳に挑戦する理由も意味もあつたに違いない。しか

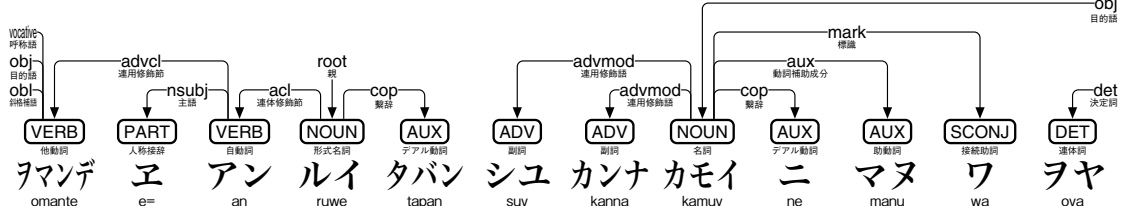
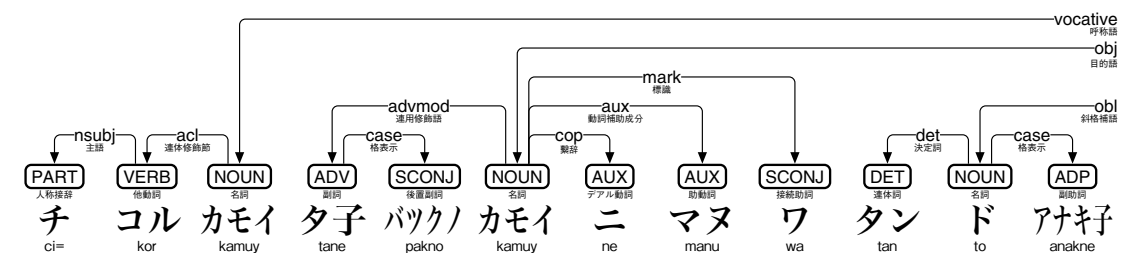
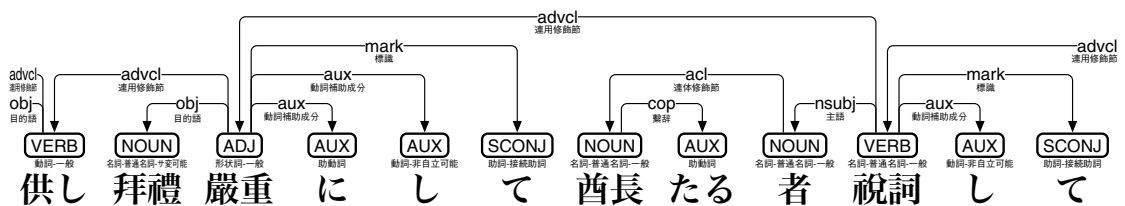
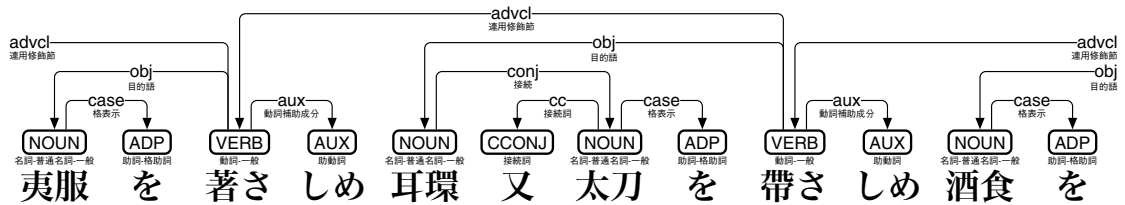
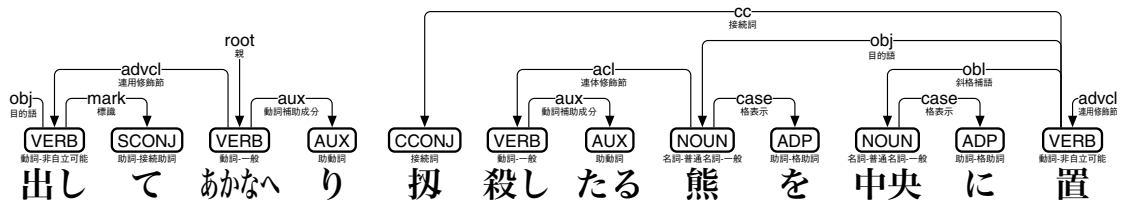
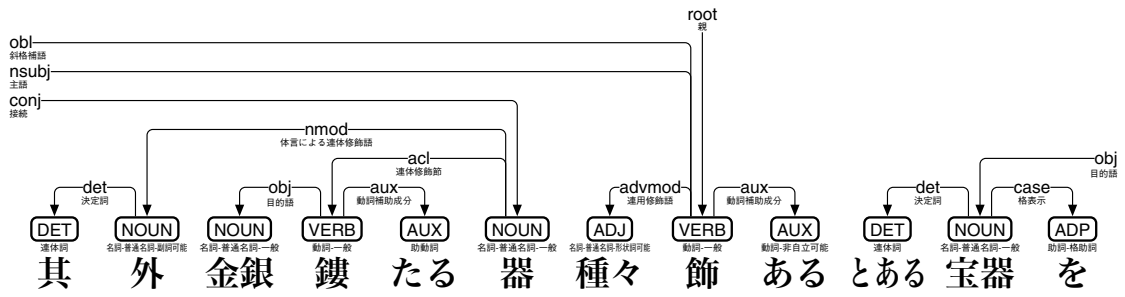
^[20] 『蝦夷見聞記』三丁ウでは「アシリチユフナル」となっている(付録 B 参照)。

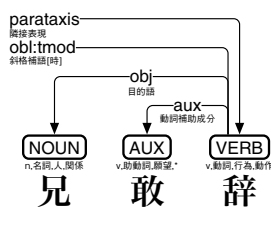
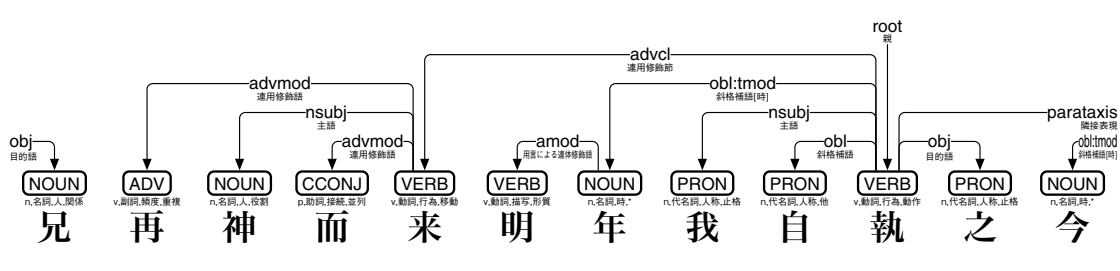
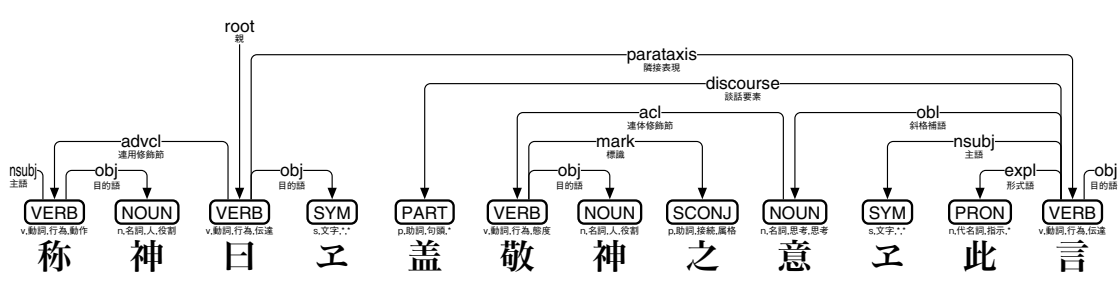
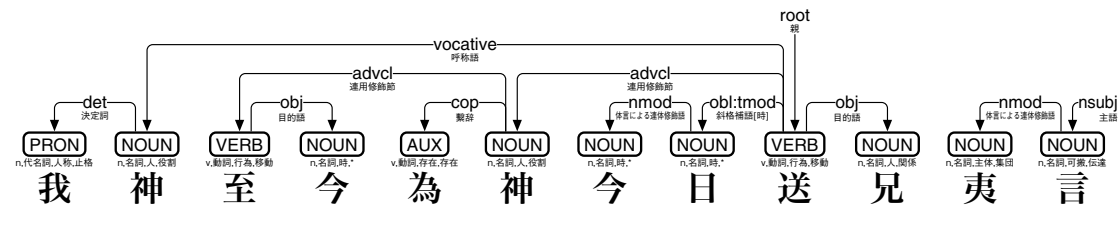
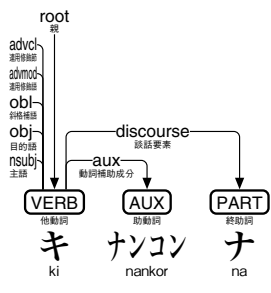
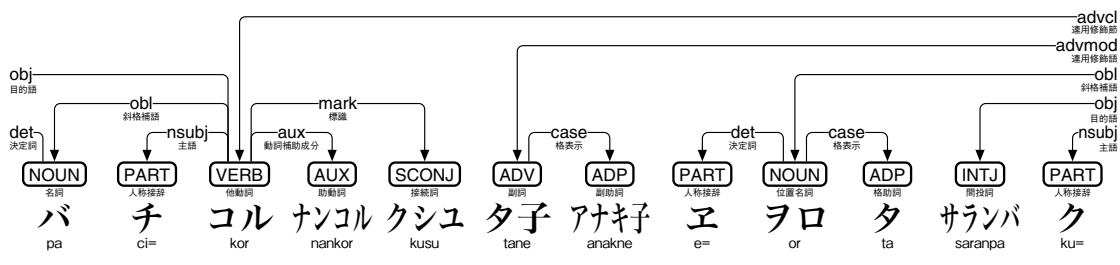
しながら、アイヌ語の語順(つまり文法規則)が解明されていない時代において、漢訳に挑戦するのは、あまりに無謀な試みだった。結果として、アイヌ語の方の語順を入れ替えてしまう、という致命的ミスに至っているのは、非常に残念な結果だと言わざるを得ない。無理な漢訳アイヌ語ではなく、漢字による逐語訳程度にとどめて、せめて語順は入れ替えないでほしかった、と思うのは、筆者たちだけだろうか。

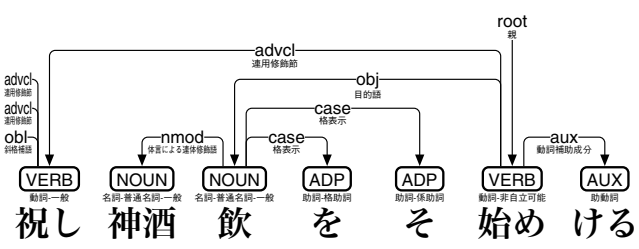
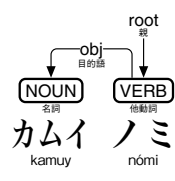
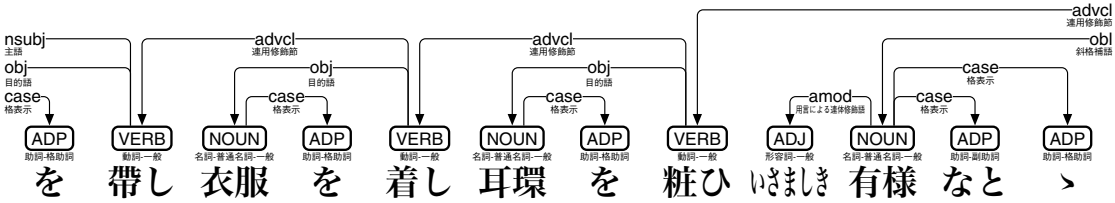
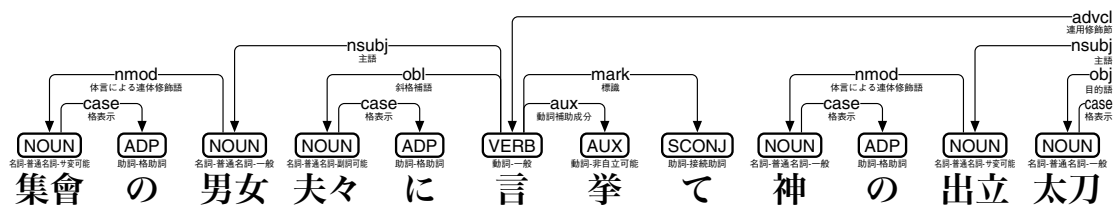
付録A 『蝦夷島奇観』 Universal Dependencies

画像番号 C0007129

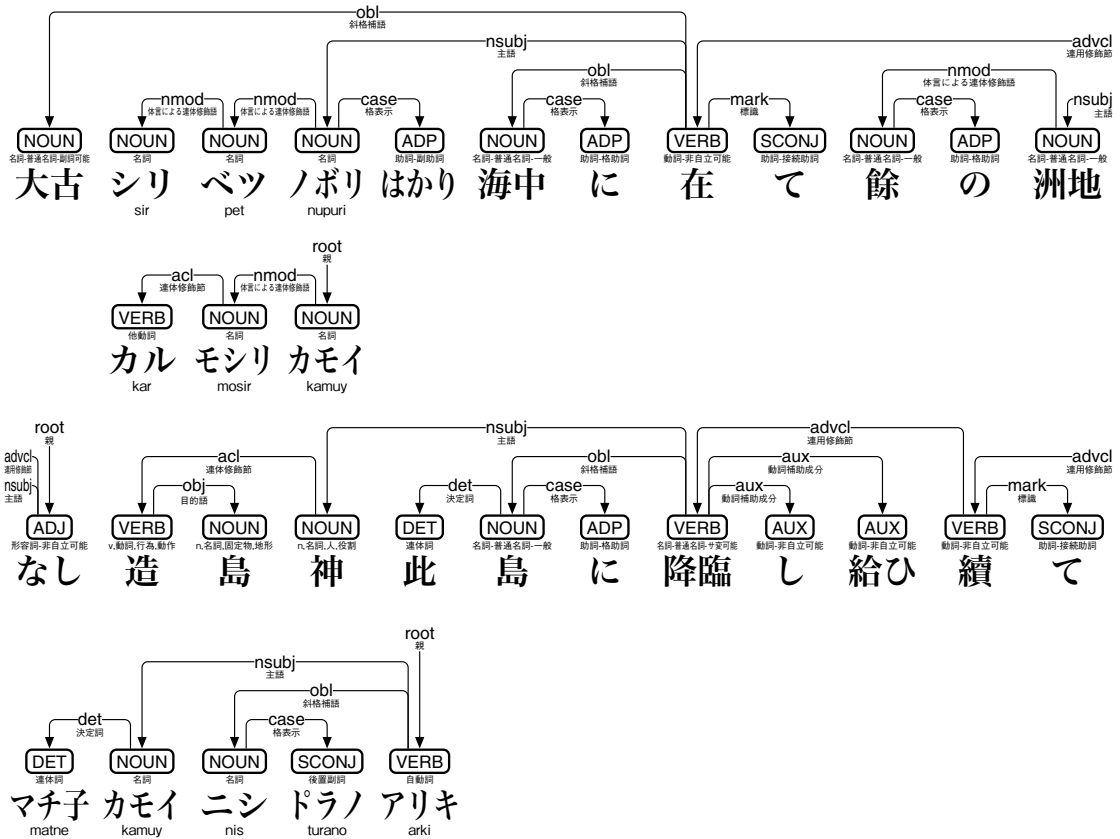


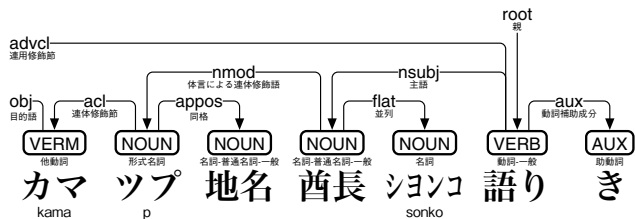
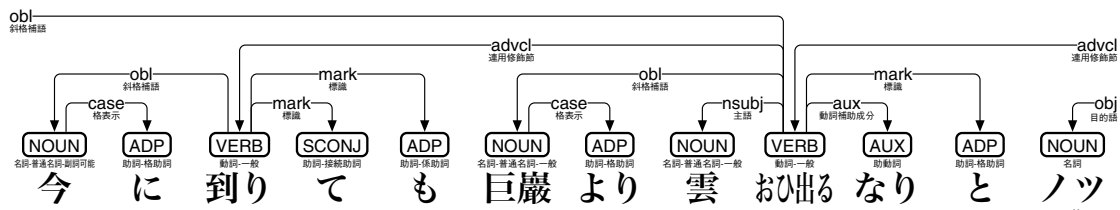
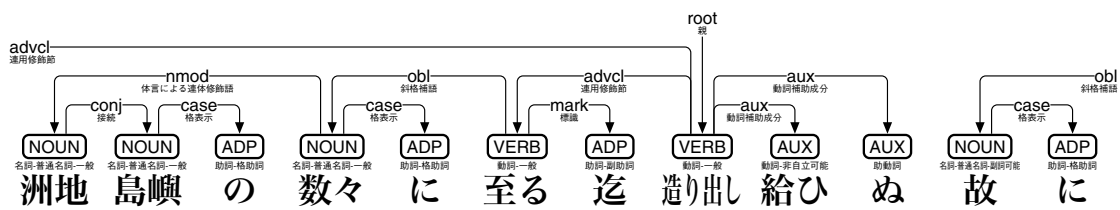
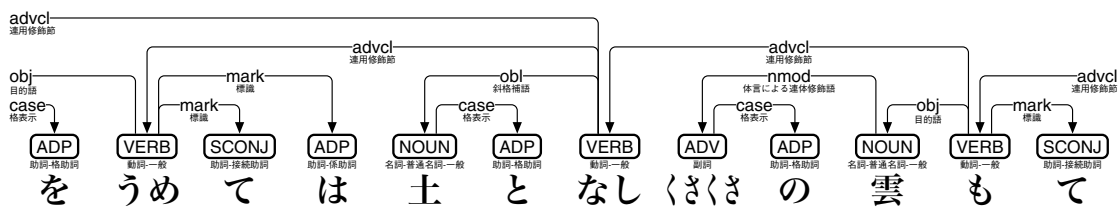
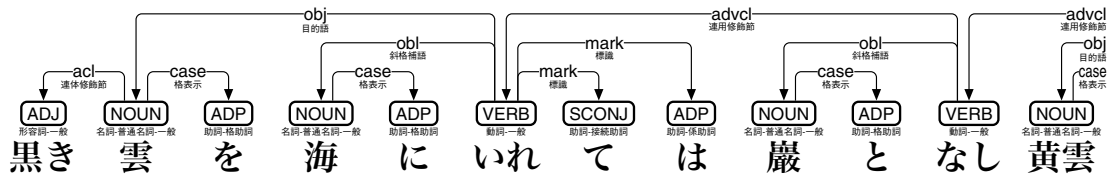
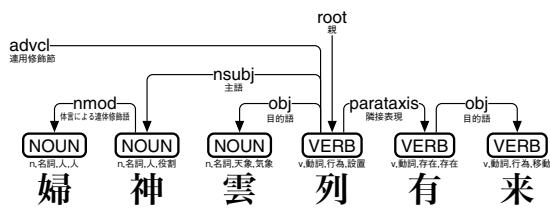






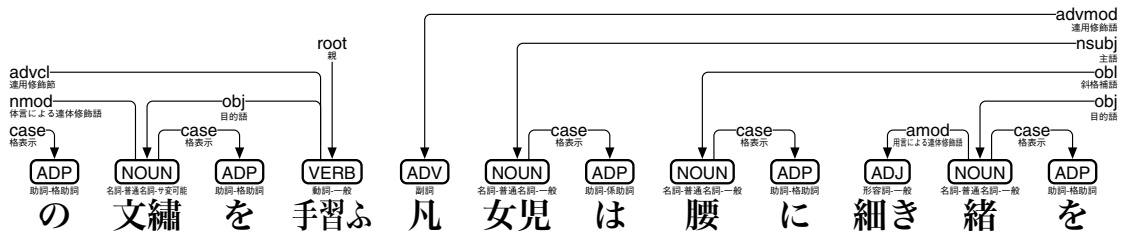
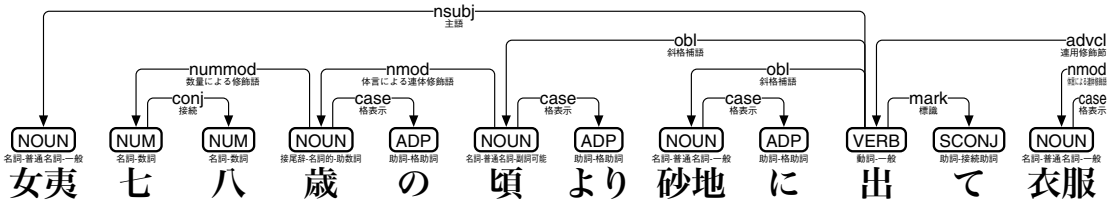
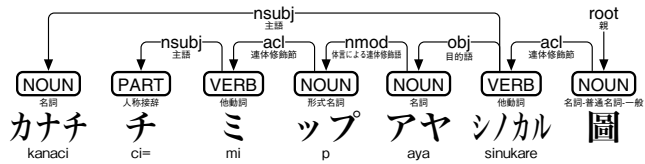
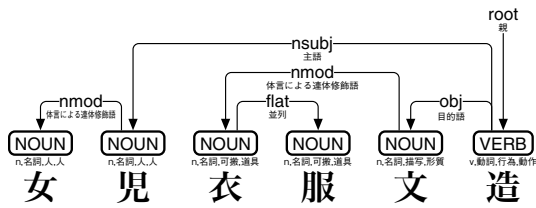
※ 「nusa san ka ta」「tanne p」「emusi」「sitoki」は、それぞれ「幣代をかきりたる棚に」「太刀」「短刀」「玉器」に和訳されている。「ci=kor kamuy tane pakno kamuy ne manu wa tan to anakne omante e=an ruwe tapan」は 3.1 節を、「suy kanna kamuy ne manu wa oya pa ci=kor nankor kusu tane anakne e=or ta saranpa ku=ki nankor na」は 3.2 節を参照。「kamuy nómi」は「神酒飲」に和訳されているが、「nómi」(祈る)を「酒飲」では誤訳である。

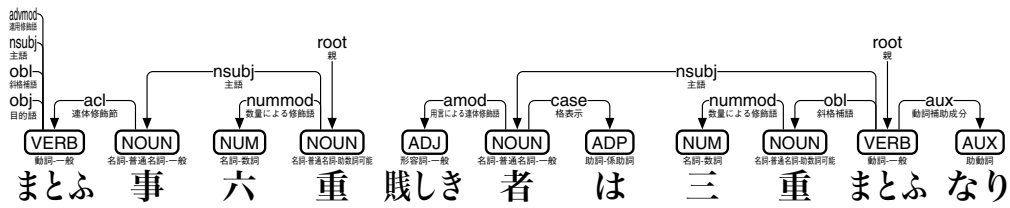




※「kar mosir kamuy」は「造島神」に和訳^[21]されているが、アイヌ語の語順は「mosir kar kamuy」の方が適切^[22]だと考えられる。「matne kamuy nis turano arki」は3.3節を参照。「sir pet nupuri」「nottu kama p」「sonko」は、固有名詞とみなされているせいか、訳が付与されていない。

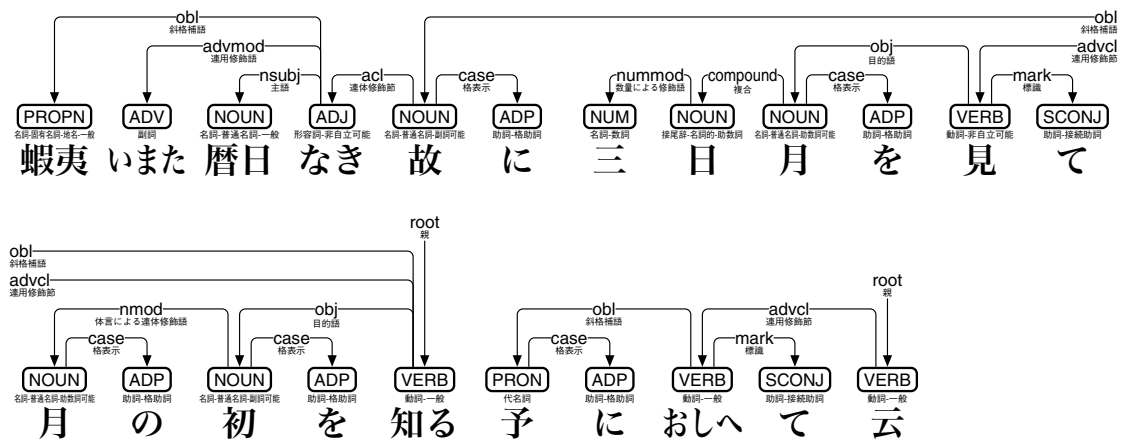
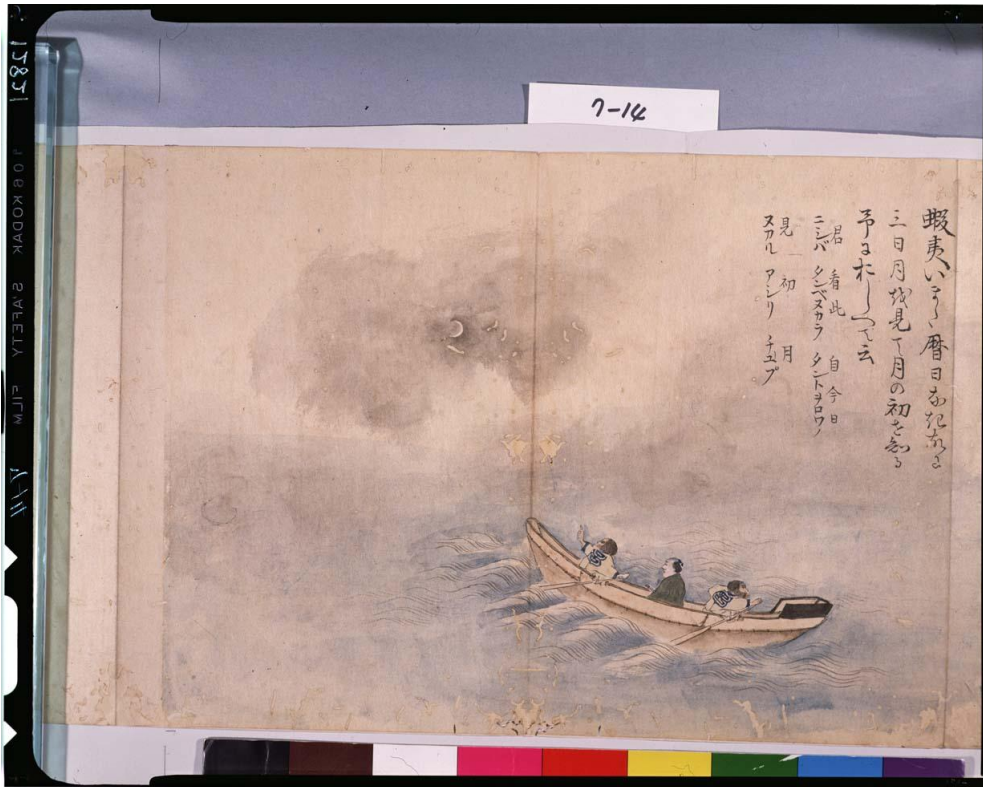
[21] 漢訳ならば「造島神」ではなく「造島之神」となる。
 [22] 吉田巖: アイヌの英雄神話, 人類学雑誌, 第29巻, 第3号 (1914年3月), pp.95-103.

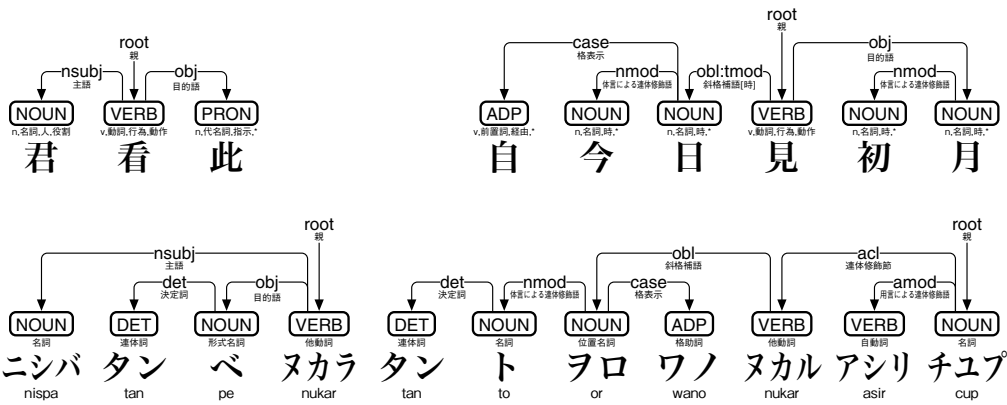




※ 「kanaci ci=mi p aya sinukare」 は 3.4 節を参照。

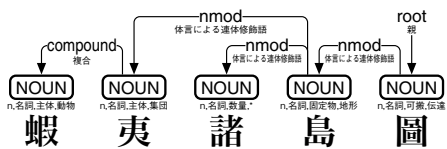
画像番号 C0012821

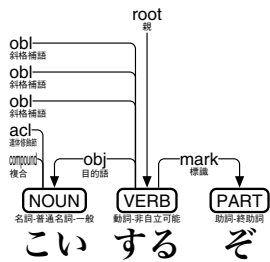
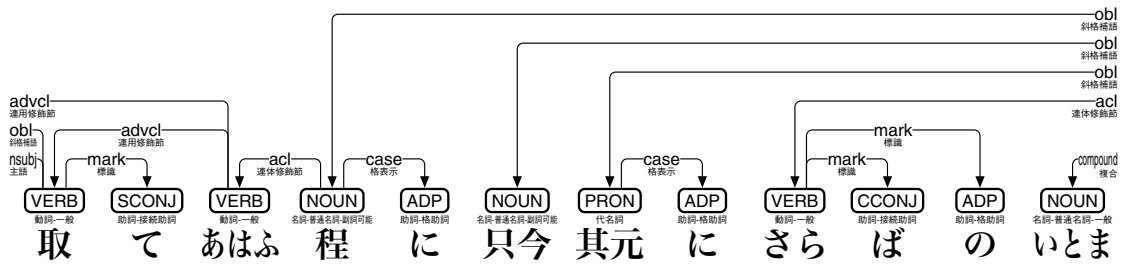
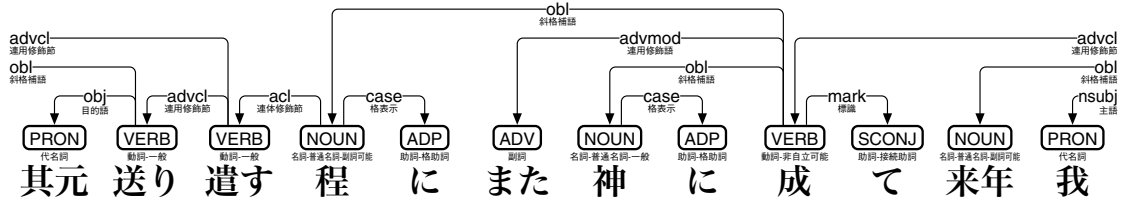
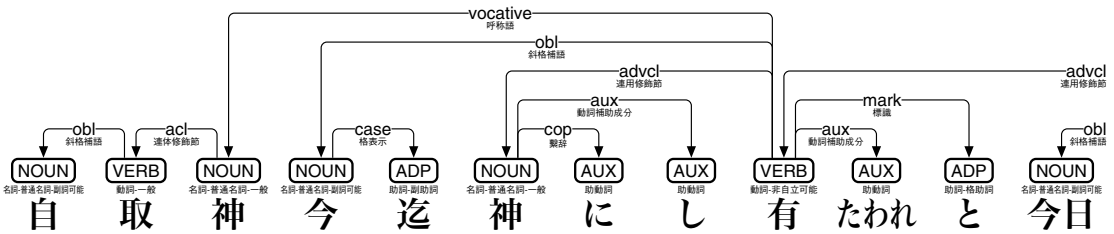
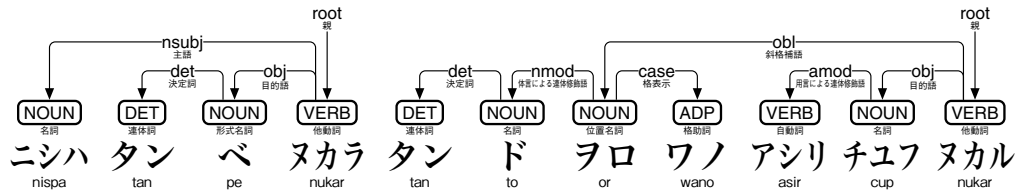
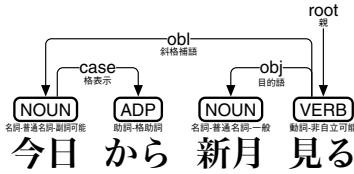
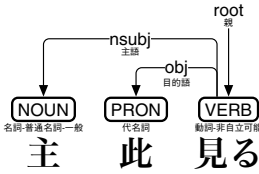
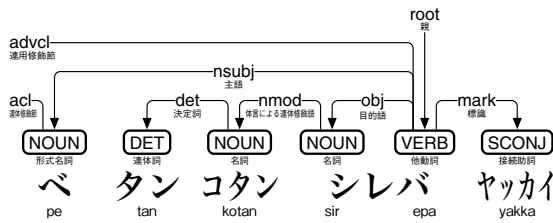


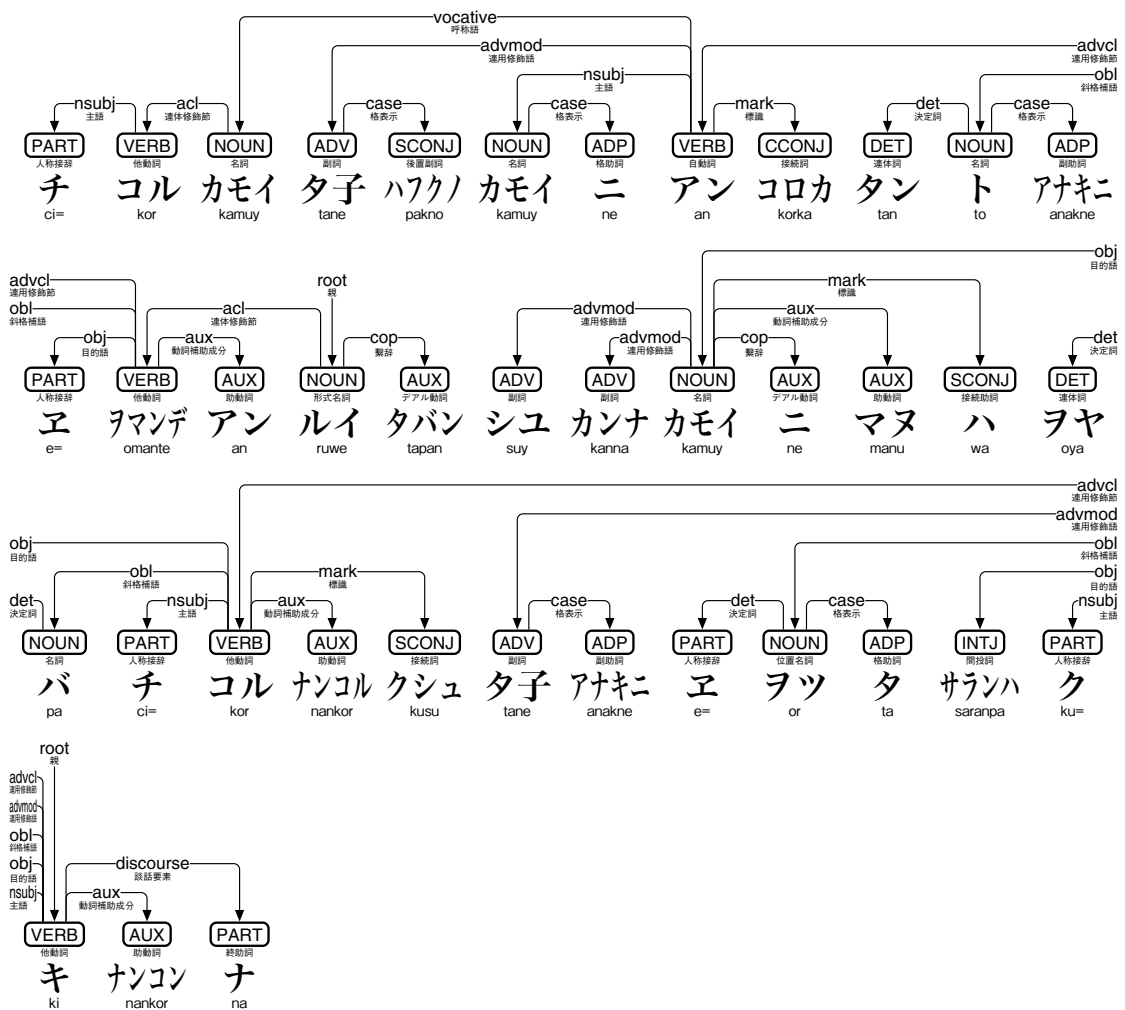


※ 「nispa tan pe nukar tan to or wano nukar asir cup」は3.5節を参照。

画像番号 C0012760







※ 「nispata tan pe nukar tan to or wano asir cup nukar」は、3.5 節と語順が異なっている。「ci=kor kamuy tane pakno kamuy ne an korka tan to anakne e=omante an ruwe tapan」は、文の一部が 3.1 節と異なっている。「suy kanna kamuy ne manu wa oya pa ci=kor nankor kusu tane anakne e=or ta saranpa ku=ki nankor na」は、3.2 節と同一である。