

# Demonstration of the Oxford Corpus of Old Japanese

## オックスフォード上代日本語コーパスの デモンストレーション

Kerri L Russell  
[kerri.russell@orinst.ox.ac.uk](mailto:kerri.russell@orinst.ox.ac.uk)  
VSARPJ/NINJAL workshop

# Outline

- ▶ Online Oxford Corpus of Old Japanese (OCOJ)
- ▶ The Lexicon
- ▶ Search Engine

# OCOJ

# OCOJ

- ▶ <http://vsarpj.orinst.ox.ac.uk/corpus/index.html>

# OCOJ

- ▶ <http://vsarpj.orinst.ox.ac.uk/corpus/index.html>
  - Presents a description of our markup

# OCOJ

- ▶ <http://vsarpj.orinst.ox.ac.uk/corpus/index.html>
  - Presents a description of our markup
  - List of texts

# OCOJ

- ▶ <http://vsarpj.orinst.ox.ac.uk/corpus/index.html>
  - Presents a description of our markup
  - List of texts
  - Fully romanized version of ***all*** OJ poetic texts

## Texts

|                                                                                |                           |
|--------------------------------------------------------------------------------|---------------------------|
| <a href="#"><i>Kojiki kayō</i></a> (古事記歌; 712)                                 | (112 poems; 2527 words)   |
| <a href="#"><i>Nihon shoki kayō</i></a> (日本書紀歌謡; 720)                          | (133 poems; 2444 words)   |
| <a href="#"><i>Fudoki kayō</i></a> (風土記歌謡; 730s)                               | (20 poems; 271 words)     |
| <a href="#"><i>Bussokuseki-ka</i></a> (仏足石歌; after 753)                        | (21 poems; 337 words)     |
| <a href="#"><i>Man'yōshū</i></a> (万葉集; after 759)                              | (4685 poems; 83706 words) |
| <a href="#">Book 1</a>                                                         | (90 poems; 2150 words)    |
| <a href="#">Book 2</a>                                                         | (169 poems; 5179 words)   |
| <a href="#">Book 3</a>                                                         | (258 poems; 4958 words)   |
| <a href="#">Book 4</a>                                                         | (311 poems; 4891 words)   |
| <a href="#">Book 5</a>                                                         | (123 poems; 2811 words)   |
| <a href="#">Book 6</a>                                                         | (161 poems; 3420 words)   |
| <a href="#">Book 7</a>                                                         | (355 poems; 5095 words)   |
| <a href="#">Book 8</a>                                                         | (250 poems; 3885 words)   |
| <a href="#">Book 9</a>                                                         | (151 poems; 3360 words)   |
| <a href="#">Book 10</a>                                                        | (555 poems; 7809 words)   |
| <a href="#">Book 11</a>                                                        | (513 poems; 7278 words)   |
| <a href="#">Book 12</a>                                                        | (405 poems; 5681 words)   |
| <a href="#">Book 13</a>                                                        | (135 poems; 4845 words)   |
| <a href="#">Book 14</a>                                                        | (245 poems; 3517 words)   |
| <a href="#">Book 15</a>                                                        | (219 poems; 3399 words)   |
| <a href="#">Book 16</a>                                                        | (105 poems; 2145 words)   |
| <a href="#">Book 17</a>                                                        | (144 poems; 3460 words)   |
| <a href="#">Book 18</a>                                                        | (114 poems; 3460 words)   |
| <a href="#">Book 19</a>                                                        | (158 poems; 3368 words)   |
| <a href="#">Book 20</a>                                                        | (224 poems; 3720 words)   |
| <a href="#"><i>Shoku nihongi kayō</i></a> (続日本紀歌謡; 797)                        | (8 poems; 134 words)      |
| <a href="#"><i>Jōgū shōtoku hōō teisetsu</i></a> (上宮聖德法王帝説; 4 poems; 60 words) |                           |
| unknown (early Heian?)                                                         |                           |
| <a href="#"><i>Eastern Old Japanese</i></a>                                    | (240 poems; 3431 words)   |
| <a href="#"><i>Azuma-uta</i></a> (東歌)                                          | (174 poems; 2110 words)   |
| <a href="#"><i>Sakimori-uta</i></a> (防人歌)                                      | (66 poems; 1321 words)    |

---

Home

---

Online corpus

---

Pre-modern  
Japanese syntax  
research project

---

Research Centre  
for Japanese  
Language and  
Linguistics

---

Contact

---

## The Old Japanese Corpus

If you use this online version of the corpus for research, please cite it as described

Please report errors to: [vsarpj@orinst.ox.ac.uk](mailto:vsarpj@orinst.ox.ac.uk)

### *Kojiki kayō*

#### KK.1

夜久毛多都

*ya-kumwo tatu*

伊豆毛夜弊賀岐

*idumwo ya-pye-gaki*

都麻碁微爾

*tumagomwi ni*

夜弊賀岐都久流

*ya-pye-gaki tukuru*

曾能夜弊賀岐袁

*so no ya-pye-gaki wo*

#### KK.2

夜知富許能

*ya-tipoko no*

迦微能美許登波

*kamwi no mi-koto pa*

夜斯麻久爾

*ya-sima kuni*

都麻麻岐迦泥弓

*tuma maki-kanete*

夜久毛多都

*ya-kumwo tatu*

# The Lexicon

# The Lexicon

- ▶ A dictionary of OJ

# The Lexicon

- ▶ A dictionary of OJ
  - In its present state, contains all verbs and function words with basic glosses into English.

# The Lexicon

- ▶ A dictionary of OJ
  - In its present state, contains all verbs and function words with basic glosses into English.
  - It is also used as a research tool, which stores information about each lexical item.

```
<superEntry xml:id="L031386-main">
  <entry corresp="29226" xml:id="L031386a">
    <form type="stem">
      <orth stage="I">nom-</orth>
      <gramGrp>
        <pos>verb</pos>
        <iType type="QD"/>
      </gramGrp>
    </form>
    <sense n="1">
      <def>drink</def>
    </sense>
  </entry>
  <entry corresp="194192" xml:id="L031386b">
    <form type="noun">
      <orth stage="I">nomi</orth>
      <gramGrp>
        <pos>noun</pos>
```

# The Lexicon

**nom-**

L031386a

Part of speech: **verb**

Conjugation class: **quadrigrade**

1 Gloss: **drink**

[statistics](#) [attestations](#) [clause structure](#)

**nomi**

L031386b

Part of speech: **noun**

[statistics](#) [attestations](#)

# Search Engine

|                                                 |                                               |                                              |                                       |
|-------------------------------------------------|-----------------------------------------------|----------------------------------------------|---------------------------------------|
| lexical item                                    | <input type="text"/>                          | string search                                | <input type="text"/>                  |
| Orthography                                     | Source                                        | Constituent boundary                         |                                       |
| <input type="radio"/> phonographically attested | <input type="radio"/> cOJ                     | <input checked="" type="checkbox"/> sentence |                                       |
| <input type="radio"/> logographically attested  | <input type="radio"/> EOJ                     | <input checked="" type="checkbox"/> clause   |                                       |
| <input checked="" type="radio"/> all[default]   | <input checked="" type="radio"/> all[default] | <input checked="" type="checkbox"/> phrase   | <input type="button" value="search"/> |

# Search Engine

## Lexical Item Information

|                 |              |
|-----------------|--------------|
| Search lemma    | L031683a     |
| Lexical class   | verb         |
| Stem form       | <i>mise-</i> |
| Inflection type | LB           |
| Definition      | show         |

## Search Result Information

|                        |                              |
|------------------------|------------------------------|
| Source                 | all                          |
| Orthography            | all                          |
| Number of attestation  |                              |
| Constituent boundaries | sentence<br>clause<br>phrase |



|                                 |                 |                                                                                                   |
|---------------------------------|-----------------|---------------------------------------------------------------------------------------------------|
| omo-ye] <b>sasi</b> ]           | <b>misuru</b>   | } <i>madeni pa</i> ] { <b>tiri-kosu</b> <i>na</i> [ <i>yume</i> ] <i>to</i> } } <b>ipi tutu</b> ] |
| ▶                               |                 | < [ ]                                                                                             |
| akupa <i>no</i> yama            | <b>misureba</b> | } { { [ <i>wo-kamwi mo</i> ] { <b>yurusi-tamapi</b> } } { [ <i>me-kan</i>                         |
| ▶                               |                 | < [ ]                                                                                             |
| i-medo] [ { { { {               | <b>misu</b>     | <i>be-ki</i> } <i>kimi ga</i> ] <b>ari to</b> } <b>ipa-naku</b> } ] <i>ni</i> } »                 |
| mo] { <b>ara-mu</b> }           | <b>misetutu</b> | [ <i>motona</i> ] } »                                                                             |
| ▶                               |                 |                                                                                                   |
| o] { <b>ara-mu</b> } }          | <b>misetutu</b> | [ <i>motona</i> ] } »                                                                             |
| ▶                               |                 |                                                                                                   |
| ami <i>no</i> ] { <i>puru-i</i> | <b>misetutu</b> | } [ <i>motona</i> ] »                                                                             |
| ▶                               |                 |                                                                                                   |
| to] <b>amori-mas</b>            | <b>mise</b>     | - <b>koso</b> } { [ { [ <i>turugi-tati</i> ] <b>ipapi-matur-eru</b> ] <i>kamwi</i> ]              |
| ▶                               |                 |                                                                                                   |
| sira-nami] [ { <b>to</b>        | <b>mise</b>     | - <i>ma</i> <b>ku</b> } <i>posi-ki</i> ] [ <i>sira-nami</i> ] »                                   |
| ▶                               |                 |                                                                                                   |
| umi <i>no</i> <i>ariswo</i>     | <b>mise</b>     | - <i>masi mo nowo</i> } »                                                                         |
| ▶                               |                 |                                                                                                   |
| <b>iredo</b> } } [ <i>koko</i>  | <b>mise</b>     | - <i>masi mono wo</i> } } { { [ <i>utusemi no</i> { <b>kar-eru</b> } <i>mwi</i> ] <i>n</i>        |
| ▶                               |                 | < [ ]                                                                                             |
| araba] [ { [ <i>wa ga</i>       | <b>mise</b>     | - <i>masi mono wo</i> } »                                                                         |
| ▶                               |                 |                                                                                                   |
| ni] <i>yo-si</i> } <i>kunu</i>  | <b>mise</b>     | - <i>masi monowo</i> } »                                                                          |

◆ Search lemma

◆ After

|               |               |                                                                                                                                |
|---------------|---------------|--------------------------------------------------------------------------------------------------------------------------------|
|               | <b>misu</b>   | <i>be-ki</i> } <i>kimi ga</i> } <b>ari to</b> } <i>ipa-naku</i> } } <i>ni</i> } »                                              |
| <b>-mas</b>   | <b>mise</b>   | <b>-koso</b> } { [ { [ <i>turugi-tati</i> ] <b>ipapi-matur-eru</b> } <i>kamwi</i> ] <i>ni si maseba</i> } »                    |
| <b>asi</b> }  | <b>misuru</b> | } <i>madeni pa</i> ] { <b>tiri-kosu</b> <i>na</i> [ <i>yume</i> ] <i>to</i> } } <b>ipi tutu</b> } } } [ <i>kokodaku mo</i> ] { |
|               |               | ◀ [ III ] ▶                                                                                                                    |
| [ { <b>to</b> | <b>mise</b>   | <i>-maku</i> } <i>posi-ki</i> } [ <i>sira-nami</i> ] »                                                                         |
| <i>iSWO</i>   | <b>mise</b>   | <i>-masi mo nowo</i> } »                                                                                                       |

Questions and Comments Welcome

Kerri L Russell  
[kerri.russell@orinst.ox.ac.uk](mailto:kerri.russell@orinst.ox.ac.uk)