

Supplementing field work with corpora:

An investigation of ergativity and anti-passives in Inuktitut

Ling 488 Independent Study 1

Louisa Bielig

Contents

1 Project Abstract	1
2 Introduction	2
3 Methodology and Issues	2
3.1 Corpus Creation	3
4 Discussion	3
A Corpus Creation Scripts	4

1 Project Abstract

blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla
 blablabla blablabla blablabla blablabla blablabla blablabla blablabla blablabla


```
</pre>
```

```
* XML
```

```
‘ ‘ ‘xml
```

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<xml>
```

```
  <book book="1co">
```

```
    <chapters>
```

```
      <chapter9 chapterNumber="9">
```

```
        <verses>
```

```
          <verse6 verseNumber="6">
```

```
            <inuktitut>áŘĚáŦiáŦžáŞŸ áŘŷáŦŘáŞĞáŘžáŞráŞŦ áŘĚáŦžáŞŦáŞČ áŘČáŞĚáŞ
```

```
            <yucatec>ÂŦWa ca tucultique'ex chĀĹen teen yĀĹetel BernabĀĹ unaŷ
```

```
            <swedish>Eller hafver jag och Barnabas allena icke magt sammaled
```

```
            <english>Or is it only Barnabas and I who have to work to suppor
```

```
          </verse6>
```

```
          <verse7 verseNumber="7">
```

```
            <inuktitut>áŞĞáŞĹáŞŦáŘĹáŞĕ áŘĚáŞĞáŞŦáŦĚáŞŘáŞČáŞŦáŘĚáŞĹáŞŦáŞČ áŘĹ
```

```
            <yucatec>ÂŦMĀax cu beetic u soldadoil yĀĹetel cu tojoltic ti' x
```

```
            <swedish>Ho tjenar till krig pĀĚ sin egen sold nĀĚgon tid? Ho pl
```

```
            <english>What soldier has to pay his own expenses? What farmer p
```

```
          </verse7>
```

```
        </verses>
```

```
      </chapter9>
```

```
    </chapters>
```

```
  </book>
```

```
</xml>
```

* JSON

```
““json
{
  "book":{
    "_book":"1co",
    "chapters":{
      "chapter9":{
        "_chapterNumber":"9",
        "verses":{
          "verse6":{
            "_verseNumber":"6",
            "inuktitut":"áŘĚáŦiáŦžáŞŸ áŘŷáŦŦáŞĞáŘžáŞŕáŞŰ áŘĚáŦžáŞŦáŞČ áŘČáŞĤ
            "yucatec":"ÁſWa ca tucultique'ex chÁſlen teen yÁſletel BernabÁſl un
            "swedish":"Eller hafver jag och Barnabas allena icke magt sammal
            "english":"Or is it only Barnabas and I who have to work to supp
          },
          "verse7":{
            "_verseNumber":"7",
            "inuktitut":"áŞĞáŞſáŞŦáŘŁáŞę áŘĚáŞĞáŞŦáŰĚáŞŦáŞČáŞŦáŘĚáŞſáŞŰáŞĆ á
            "yucatec":"ÁſMÁqax cu beetic u soldadoil yÁſletel cu tojoltic ti
            "swedish":"Ho tjenar till krig pÁĚ sin egen sold nÁĚgon tid? Ho
            "english":"What soldier has to pay his own expenses? What farmer
          }
        }
      }
    }
  }
}
```

Install:

1. Download [Node.js](<http://nodejs.org/>) if you don't already have it
2. Then download this project

```
““bash
$ wget https://github.com/louisa-bielig/MultilingualCorporaExtractor/archive/master.zip
$ unzip master.zip
$ cd MultilingualCorporaExtractor
$ npm install
““
```

Usage:

Here is a sample use of the interactive script:

<pre>

```
$ ./createdata.sh
```

```
Enter the three character code for the book you want to use for your corpus  
e.g. gen for Genesis and press [ENTER]: lco
```

```
Enter the starting chapter number for lco and press [ENTER]: 9
```

```
Enter the ending chapter number for lco and press [ENTER]: 9
```

```
Enter the language number code and press [ENTER]: 455
```

```
Enter the language text code and press [ENTER]: inuktitut
```

```
Working...
```

```
9 9 lco 455 inuktitut lco-9-9-1370897837.html
```

```
Chapter 9 downloaded.
```

```
Finished!
```

```
Enter the language number code and press [ENTER]: 455
```

```
Enter the language text code and press [ENTER]: inuktitut
```

```
Working...
```

```
9 9 lco 455 inuktitut lco-9-9-1370897837.html
```

```
Chapter 9 downloaded.
```

```
Finished!
```

```
Enter the language number code and press [ENTER]: 324
```

```
Enter the language text code and press [ENTER]: yucatec
```

```
Working...
```

```
9 9 lco 324 yucatec lco-9-9-1370897837.html
```

```
Chapter 9 downloaded.
```

```
Finished!
```

```
Enter the language number code and press [ENTER]: 161
```

```
Enter the language text code and press [ENTER]: swedish
```

```
Working...
```

```
9 9 lco 161 swedish lco-9-9-1370897938.html
```

```
Chapter 9 downloaded.
```

```
Finished!
```

8

```
Enter the language number code and press [ENTER]: 116
```

```
Enter the language text code and press [ENTER]: english
```

```
Working...
```

```
9 9 lco 116 english lco-9-9-1370897938.html
```

```
Chapter 9 downloaded.
```