

Supplementing field work with corpora:  
An investigation of ergativity and anti-passives in Inuktitut  
Ling 488 Independent Study 1

Louisa Bielig

## Contents

<b>1</b>	<b>Project Abstract</b>	<b>1</b>
<b>A</b>	<b>Corpus Creation</b>	<b>1</b>

## 1 Project Abstract

## A Corpus Creation

Sample Output:

-----

The end result is an html file which automatically makes an alligned corpus and ap

\* Raw Aligned Text

```


```

1co:9:6 āŘĚāŤīāŤZāSŸ āŘyāŤŘāŠGāŘZāSřāSŮ āŘĚāŤZāSŇāŠČ āŘČāSĚāSęāŤlāsŮāSŇāSČāŘLāŬĚāS

1co:9:6 Â£Wa ca tucultique'ex chÃ'len teen yÃ'letel BernabÃ' unaj c meyaj yÃ'letel Ãa

1co:9:6 Eller hafver jag och Barnabas allena icke magt sammaledes gÃra?

1co:9:6 Or is it only Barnabas and I who have to work to support ourselves?

1co:9:7 áŞĠáŞŤáŘŁáŚę áŘĚáŞĠáŮĚáŚŘáŚČáŞťáŘĚáŞŤáŞŮáŞĆ áŘŁáŚŘáŤŘáŮŤáŞĆáŘŁáŮĚáŞŤ

1co:9:7 Â£MÃax cu beetic u soldadoil yÃ'etel cu tojoltic ti' xan ba'ax cu xupic?

1co:9:7 Ho tjenar till krig pÃ€ sin egen sold nÃ€gon tid? Ho planterar en vingÃ€r

1co:9:7 What soldier has to pay his own expenses? What farmer plants a vineyard and

```


```

\* XML

```
'''xml
```

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

<xml>

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

&lt;chapters&gt;

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

<verses>

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

<inuktitut>áŘĚáŦĭáŦŽáŠŸ áŘŷáŦŘáŞĜáŘžáŞráŞŮ áŘĚáŦŽáŠŇáŠĈ áŘĈáŞĚáŠ

<yucatec>ÂfWa ca tucultique'ex chÃ'en teen yÃ'etel BernabÃ' una.

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

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

</verse6>

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

<inuktitut>áſĜáſłáſŢáŔĹáſę áŔĚáſĜáſŢáŨĚáſŔáſĈáſŢáŔĚáſłáſŦáſĈ áŔĹ

<yucatec>Â£MÃqax cu beetic u soldadoil yÃl'etel cu tojoltic ti' x

<swedish>Ho tjenar till krig pÃ€ sin egen sold nÃ€gon tid? Ho pl

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

</verse7>

&lt;/verses&gt;

&lt;/chapter9&gt;

&lt;/chapters&gt;

&lt;/book&gt;

&lt;/xml&gt;

\* JSON

```
““json
{
  "book":{
    "_book": "1co",
    "chapters":{
      "chapter9":{
        "_chapterNumber": "9",
        "verses":{
          "verse6":{
            "_verseNumber": "6",
            "inuktitut": "áŘĚáŦiáŦŽáŞŸ áŘŷáŦŘáŞĞáŘžáŞráŞŮ áŘĚáŦŽáŞŦáŞČ áŘČáŞĚ
            "yucatec": "ÂſWa ca tucultique'ex chÃlen teen yÃletel BernabÃl 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 áŘĚáŞĞáŞŦáŮĚáŞŘáŞČáŞŦáŘĚáŞſáŞŮáŞĆ á
            "yucatec": "ÂſMÃqax cu beetic u soldadoil yÃletel cu tojoltic ti
            "swedish": "Ho tjenar till krig pÃě sin egen sold nÃěgon 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!
```

6

```
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.
```