

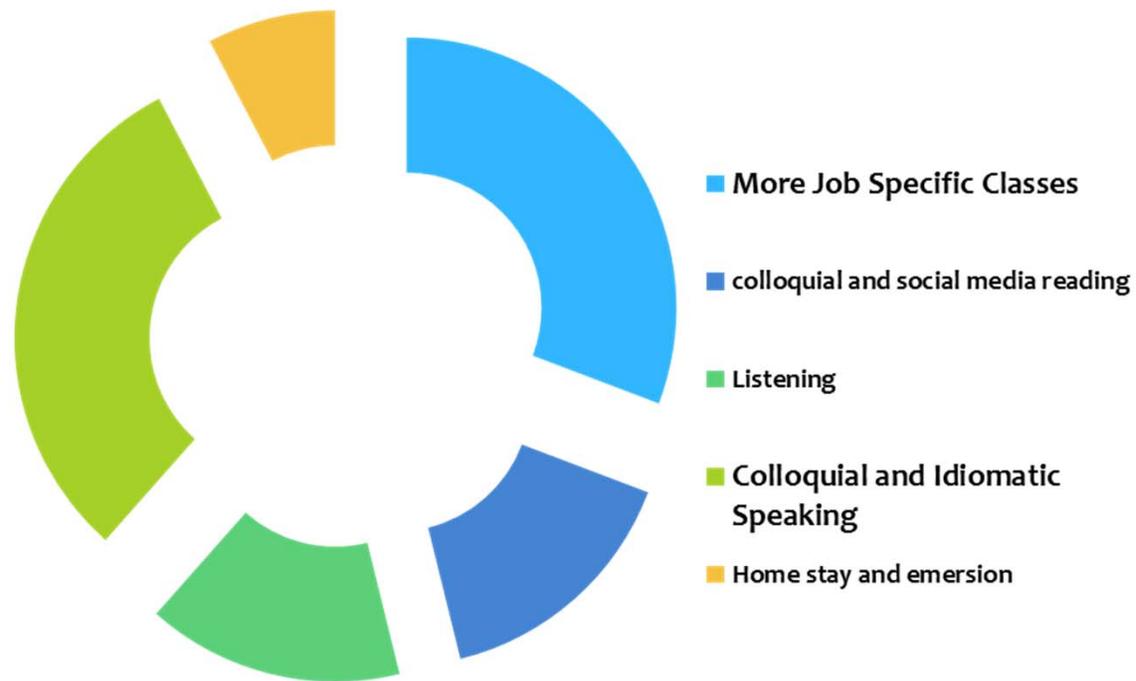
Burmese Project

LEARN
April 23, 2018

Ye Min Tun
Burmese, LCI

Memo 1

* Post-Visit Report:



Design Thinking

12-Week Workshop

- Challenges
- Prototyping
- Solutions
- Memo-2: A Project Proposal

The Analysis and Design Process

Stage I: Technical Challenges for Burmese

Burmese Unicode Fonts and Lack of Available Keyboard

- * Is there any feasible Burmese font that we could use with any language tools available on the internet?
- * Do we have a Burmese keyboard for Unicode font?

End of the Stage II of Design Thinking

Major Challenges Identified and Solved

We were able to:

- (1) Identify a Burmese Unicode font that can be used for vocabulary development with tools and technology available on the internet**
- (2) Design and create new Burmese Unicode keyboard software at FSI**
- (3) Find an appropriate tool for Burmese word segmentation**
- (4) Find appropriate tools to create a word frequency list**
- (5) Prototype a Burmese word frequency list**

Making Word Lists

“Word lists lie at the heart of good vocabulary course design, the development of graded materials for extensive listening and extensive reading, research on vocabulary load, and vocabulary test development.”

Paul Nation

*Making and Using Word Lists for Language Learning and Testing
(2016)*

Scope and Sequence of the Burmese Curriculum Project

Phase I: Vocabulary frequency list (7 - months)

- (i) 1000 to 3000 high-frequency words

Phase II: Texts and Prototype-lessons (5 - months)

- (i) Reading texts
- (ii) Lesson prototypes
- (iii) New ways of teaching the Burmese script and sound system
- (iv) Prototype lessons to be tested in class

Phase III: Curriculum (9 - months)

- (i) A vocab-frequency list based curriculum
- (ii) Job related materials
- (iii) An extensive reading program
- (iv) Grammar instruction and materials

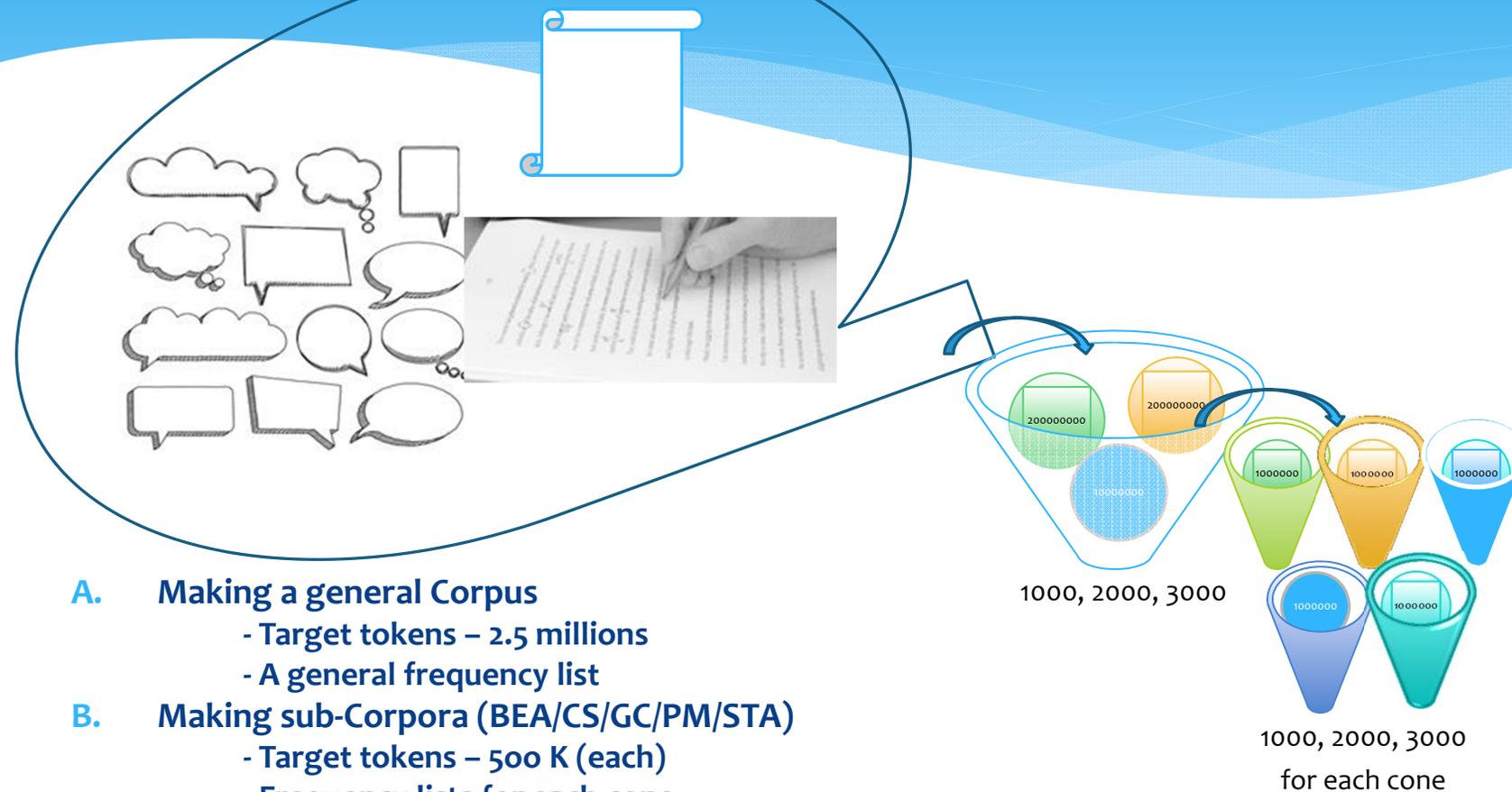
Phase I

Purpose for making a particular list has a strong effect on the decisions and procedures that need to be followed.

Paul Nation

Making and Using Word Lists for Language Learning and Testing (2016)

Phase I: Creation of Vocabulary Frequency List for the Burmese Project



Phase I: How and What

Step 2. Convert text into Unicode if needed

<http://burglish.my-mm.org/latest/trunk/web/fontconv.htm>

The screenshot displays the 'Burmese Font Converter' web application. The browser address bar shows the URL `/latest/trunk/web/fontconv.htm`. The page title is `::Burmese Font Converter::`. The interface features two main text input areas:

- Left Input (Zawgyi-One):** Contains Burmese text in Zawgyi font, including phrases like 'အလုပ်အကိုင်အခွင့်အလမ်း ဖန်တီးနိုင်ရေးအတွက် နားလည်မှုစာချွန်လွှာ တစ်ရပ်ကို နိုင်ငံဘာ ၂၂ ရက်တွင် လက်မှတ်ရေးထိုးခဲ့သည်။' and 'နားလည်မှုစာချွန်လွှာအရ ကန်တော်ကြီးဟိုတယ်နှင့် ခရီးသွားလုပ်ငန်းသင်တန်းကျောင်းသည် မြန်မာနိုင်ငံတွင်ဖြန့်ချိနေသော Unilever ၏ကုန်ပစ္စည်း ၁၅ မျိုးကိုသင်ထောက်ကူ ပစ္စည်းအဖြစ် အသုံးပြုနိုင်မှာဖြစ်ပြီး Unilever မှစားဖိုမှူးကြီးများကိုယ်တိုင် စာတွေ့လက်တွေ့ သင်ကြား ပြသပေးမည်ဟု Unilever ၏ ကြော်ငြာမန်နေဂျာ ဒေါ်ဆုလင်းကိုက ပြောသည်။'
- Right Input (Myanmar3):** Contains the converted Myanmar3 text, such as 'လွှာတစ်ရပ်ကို နိုင်ငံဘာ ၂၂ ရက်တွင် လက်မှတ်ရေးထိုးခဲ့သည်။' and 'နားလည်မှုစာချွန်လွှာအရ ကန်တော်ကြီးဟိုတယ်နှင့် ခရီးသွားလုပ်ငန်းသင်တန်းကျောင်းသည် မြန်မာနိုင်ငံတွင်ဖြန့်ချိနေသော Unilever ၏ကုန်ပစ္စည်း ၁၅ မျိုးကို သင်ထောက်ကူ ပစ္စည်းအဖြစ် အသုံးပြုနိုင်မှာဖြစ်ပြီး Unilever မှစားဖိုမှူးကြီးများကိုယ်တိုင် စာတွေ့လက်တွေ့ သင်ကြား ပြသပေးမည်ဟု Unilever ၏ ကြော်ငြာမန်နေဂျာ ဒေါ်ဆုလင်းကိုက ပြောသည်။'

At the bottom of the interface, there are four buttons: 'Convert ->', 'Char Code->', 'Escape Unicode->', and 'Convert Beta'.

Phase I: How and What

Step 3. Segmentation

<http://www.nlpresearch.com>

The screenshot shows a web application interface for Myanmar Word Segmentation. At the top, there is a navigation bar with links for 'Home', 'About Us', and 'Contact Us', and a search box. A sidebar on the left contains a menu with options: 'Optical Character Recognition', 'Machine Translation', 'English to Myanmar', 'Syntactic Transfer (Reordering)', 'Transliteration', 'Publications', 'Text-to-Speech Processing', 'Dictionary', 'Myanmar Text Processing', 'English Text Processing', and 'Contact Details'. The main content area is titled 'Myanmar Word Segmentation Version 1.0' and features a section 'Enter Myanmar Sentence to Segment'. Below this, there is a 'Choose font' dropdown menu set to 'Myanmar3'. A text input field contains a Burmese sentence: 'တစ်ရပ်ကို နိုင်ဘာ ၂၂ ရက်တွင် လက်မှတ်ရေးထိုးခဲ့သည်။နားလည်မှုစာချွန်လွှာအရ ကန်တော်ကြီးဟိုတယ်နှင့် ခရီးသွားလုပ်ငန်းသင်တန်းကျောင်းသည် မြန်မာနိုင်ငံတွင်ဖြန့်ချိနေသော Unilever ၏ ကုန်ပစ္စည်း ၁၅ မျိုးကိုသင် ထောက်ကူ ပစ္စည်းအဖြစ် အသုံးပြုနိုင်မှာဖြစ်ပြီး Unilever မှစားဖိုမှူးကြီးများကိုယ်တိုင် စာတွေ့လက်တွေ့ သင်ကြား ပြသပေးမည်ဟု Unilever ၏ ကြော်ငြာမန်နေဂျာ ဒေါ်ဆုလင်းကိုက ပြောသည်။' Below the input field is a 'Segment' button. The output section, titled 'Segmented Text', displays the same sentence with spaces inserted between words: 'ကန်တော်ကြီး_ ဟိုတယ်_ သင်တန်းကျောင်း_ နှင့်_ Unilever Food Solutions_ ပူးပေါင်း_ ဆောင်ရွက်ရေး_ နားလည်မှု_ စာ_ ချွန်_ လွှာ_ လက်မှတ်_ ရေးထိုး_ Print_ အိ_ အိ_ သူ_ | Wednesday_ ။ 23 November 2016Share_ ကန်တော်ကြီး_ ဟိုတယ်_ နှင့်_ ခရီးသွားလုပ်ငန်း_ သင်တန်း_ ကျောင်း_ နှင့်_ Unilever Food Solutions_ ပူးပေါင်း_ ပြီး_ ကမ္ဘာ့_ အဆင့်မီ_ စားဖိုမှူး_ များ_ ပေါ်ပေါက်_ လာ_ ရေး_ နှင့်_ သင်တန်း_ ကျောင်းသားများ_ အတွက်_ အလုပ်အကိုင်_ အခွင့်အလမ်း_ ဖန်တီး_ နိုင်_ ရေး_ အတွက်_ နားလည်မှု_ စာ_ ချွန်_ လွှာ_ တစ်ရပ်_ ကို_ နိုင်ဘာ_ ၂၂_ ရက်_ တွင်_ လက်မှတ်ရေးထိုးခဲ့သည်_ ။ နားလည်မှု_ စာ_ ချွန်_ လွှာ_ အရ_ ကန်တော်ကြီး_ ဟိုတယ်_ နှင့်_ ခရီးသွားလုပ်ငန်း_ သင်တန်း_ ကျောင်း_ သည်_ မြန်မာနိုင်ငံ_ တွင်_ ဖြန့်ချိ_ နေ_ သော_ Unilever_ ၏_ ကုန်ပစ္စည်း_ ၁၅_ မျိုး_ ကို_ သင်_ ထောက်_ ကူ_ ပစ္စည်း_ အဖြစ်_ အသုံးပြု_ နိုင်_ မှာ_ ဖြစ်ပြီး_ Unilever_ မှ_ စားဖိုမှူး_ ကြီး_ များ_ ကိုယ်တိုင်_ စာတွေ့_ လက်တွေ့_ သင်ကြား_ ပြသ_ ပေး_ မည်_ ဟု_ Unilever_ ၏_ ကြော်ငြာ_ မန်နေဂျာ_ ဒေါ်_ ဆု_ လင်း_ ကို_ က_ ပြောသည်_ ။' At the bottom left of the sidebar, the phone number '062667' is displayed.

Phase I: How and What

Step 4A. Cleaning the noises

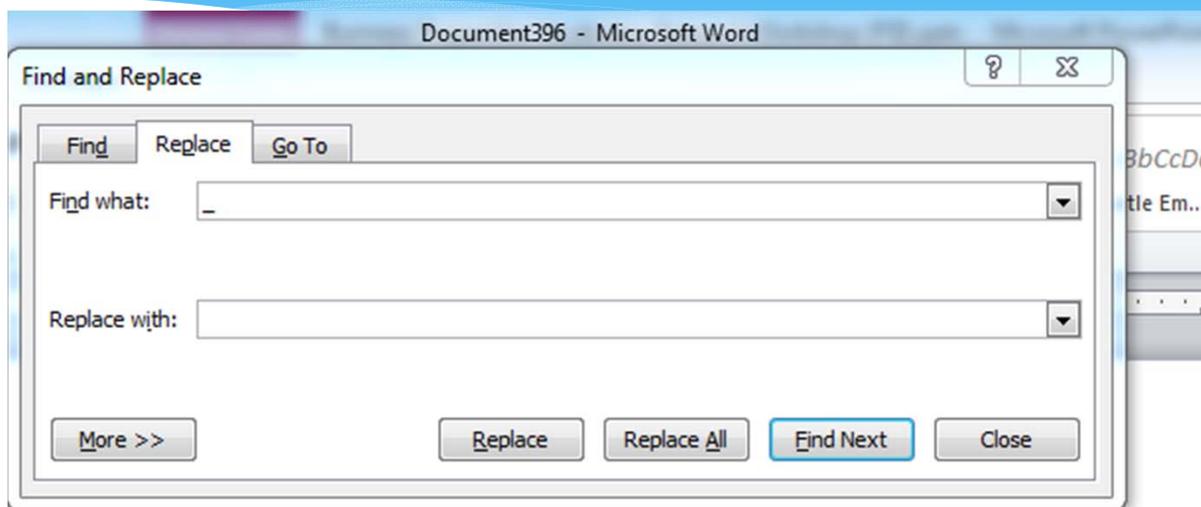
ကန်တော်ကြီး_ ဟိုတယ်_ သင်တန်းကျောင်း_ နှင့်_ Unilever Food Solutions_ ပူးပေါင်းဆောင်ရွက်ရေး_ နားလည်မှု_ စာ_ ချွန်_ လွှာ_ လက်မှတ်_ ရေးထိုး_ Print_ အိ_ အိ_ သူ_ | Wednesday_ ။

23 November 2016Share_ ကန်တော်ကြီး_ ဟိုတယ်_ နှင့်_ ခရီးသွားလုပ်ငန်း_ သင်တန်းကျောင်း_ နှင့်_ Unilever Food Solutions_ ပူးပေါင်း_ ပြီး_ ကမ္ဘာ့_ အဆင့်မီ_ စားဖိုမှူး_ များ_ ပေါ်ပေါက်_ လာ_ ရေး_ နှင့်_ သင်တန်း_ ကျောင်းသားများ_ အတွက်_ အလုပ်အကိုင်_ အခွင့်အလမ်း_ ဖန်တီး_ နိုင်_ ရေး_ အတွက်_ နားလည်မှု_ စာ_ ချွန်_ လွှာ_ တစ်ရပ်_ ကို_ နိုဝင်ဘာ_ ၂၂_ ရက်_ တွင်_ လက်မှတ်ရေးထိုးခဲ့သည်_ ။

နားလည်မှု_ စာ_ ချွန်_ လွှာ_ အရ_ ကန်တော်ကြီး_ ဟိုတယ်_ နှင့်_ ခရီးသွားလုပ်ငန်း_ သင်တန်းကျောင်း_ သည်_ မြန်မာနိုင်ငံ_ တွင်_ ဖြန့်ချိ_ နေ_ သော_ Unilever_ ၏_ ကုန်ပစ္စည်း_ ၁၅_ မျိုး_ ကို_ သင်_ ထောက်_ ကူ_ ပစ္စည်း_ အဖြစ်_ အသုံးပြု_ နိုင်_ မှာ_ ဖြစ်ပြီး_ Unilever_ မှ_ စားဖိုမှူး_ ကြီး_ များ_ ကိုယ်တိုင်_ စာတွေ့_ လက်တွေ့_ သင်ကြား_ ပြသ_ ပေး_ မည်_ ဟု_ Unilever_ ၏_ ကြော်ငြာ_ မန်နေဂျာ_ ဒေါ်_ ဆု_ လင်း_ ကို_ က_ ပြောသည်_ ။

Phase I: How and What

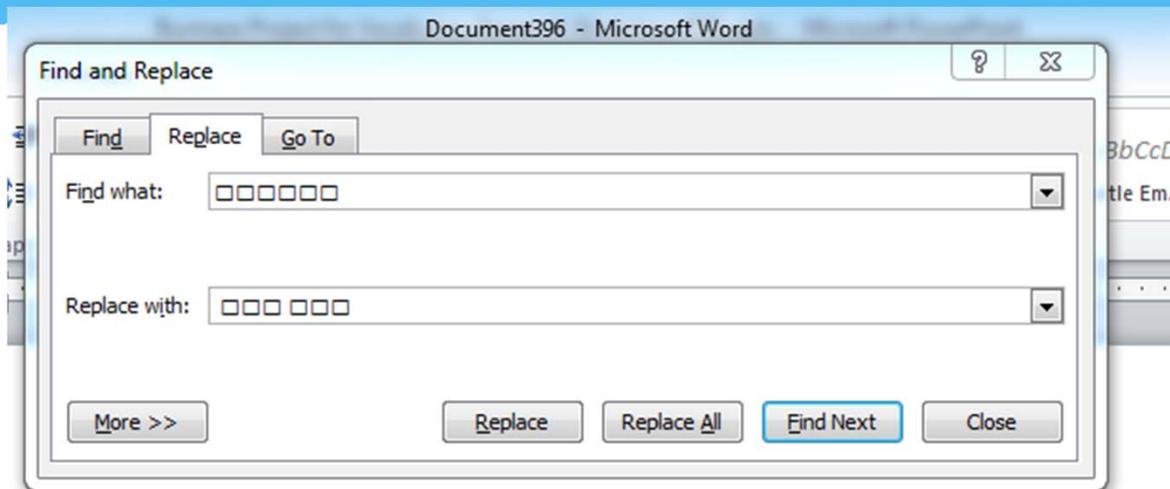
Step 4B. Cleaning the noises



ကန်တော်ကြီး ဟိုတယ် သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်းဆောင်ရွက်ရေး နားလည်မှု စာ ချွန် လွှာ လက်မှတ် ရေးထိုး Print အိအိသူ | Wednesday 23 November 2016Share ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်း ပြီး ကမ္ဘာ့ အဆင့်မီ စားဖိုမှူး များ ပေါ်ပေါက် လာ ရေး နှင့် သင်တန်း ကျောင်းသားများ အတွက် အလုပ်အကိုင် အခွင့်အလမ်း ဖန်တီး နိုင် ရေး အတွက် နားလည်မှု စာ ချွန် လွှာ တစ်ရပ် ကို နိုဝင်ဘာ ၂၂ ရက် တွင် လက်မှတ်ရေးထိုးခဲ့သည် ။ နားလည်မှု စာ ချွန် လွှာ အရ ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း သည် မြန်မာနိုင်ငံ တွင် ဖြန့်ချိ နေ သော Unilever ၏ ကုန်ပစ္စည်း ၁၅ မျိုး ကို သင် ထောက် ကူ ပစ္စည်း အဖြစ် အသုံးပြု နိုင် မှာ ဖြစ်ပြီး Unilever မှ စားဖိုမှူး ကြီး များ ကိုယ်တိုင် စာတွေ့ လက်တွေ့ သင်ကြား ပြသ ပေး မည် ဟု Unilever ၏ ကြော်ငြာ မန်နေဂျာ ဒေါ် ဆု လင်း ကို က ပြောသည် ။

Phase I: How and What

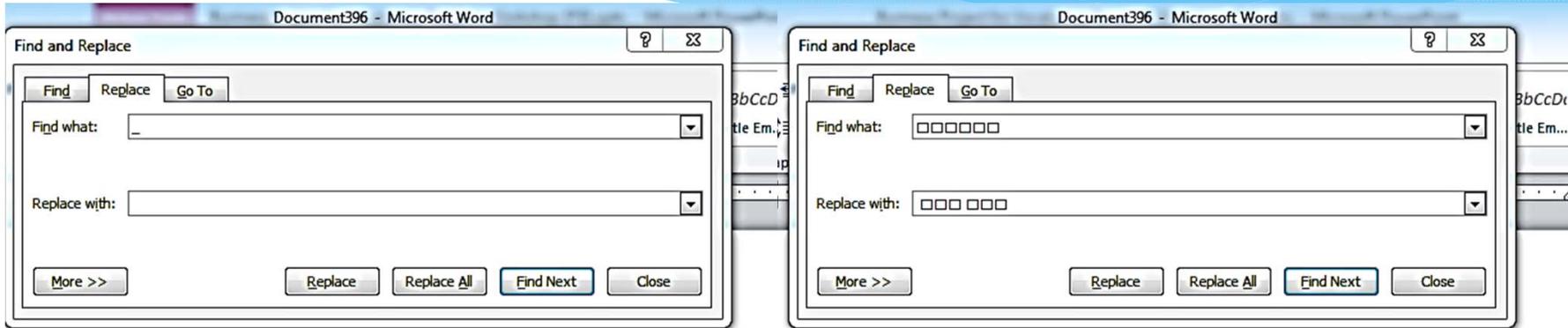
Step 4C. Improving accuracy



ကန်တော်ကြီး ဟိုတယ် သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်းဆောင်ရွက်ရေး နားလည်မှု စာ ချွန် လွှာ လက်မှတ် ရေးထိုး Print အိအိသူ | Wednesday ။
23 November 2016Share ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်း ပြီး ကမ္ဘာ့ အဆင့်မီ စားဖိုမှူး များ ပေါ်ပေါက် လာ ရေး နှင့် သင်တန်း ကျောင်းသားများ အတွက် အလုပ်အကိုင် အခွင့်အလမ်း ဖန်တီး နိုင် ရေး အတွက် နားလည်မှု စာ ချွန် လွှာ တစ်ရပ် ကို နိုဝင်ဘာ ၂၂ ရက် တွင် လက်မှတ်ရေးထိုး ခဲ့ သည် ။
နားလည်မှု စာ ချွန် လွှာ အရ ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း သည် မြန်မာနိုင်ငံ တွင် ဖြန့်ချိ နေ သော Unilever ၏ ကုန်ပစ္စည်း ၁၅ မျိုး ကို သင် ထောက် ကူ ပစ္စည်း အဖြစ် အသုံးပြု နိုင် မှာ ဖြစ်ပြီး Unilever မှ စားဖိုမှူး ကြီး များ ကိုယ်တိုင် စာတွေ့ လက်တွေ့ သင်ကြား ပြသ ပေး မည် ဟု Unilever ၏ ကြော်ငြာ မန်နေဂျာ ဒေါ် ဆု လင်း ကို က ပြော သည် ။

Phase I: How and What

Step 4C. Improving accuracy



ကန်တော်ကြီး ဟိုတယ် သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်းဆောင်ရွက်ရေး နားလည်မှု စာ ချွန် လွှာ လက်မှတ် ရေးထိုး Print အိအိသူ | Wednesday ။
 23 November 2016Share ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်း ပြီး ကမ္ဘာ့ အဆင့်မီ စားဖိုမှူး များ ပေါ်ပေါက် လာ ရေး နှင့် သင်တန်း ကျောင်းသားများ အတွက် အလုပ်အကိုင် အခွင့်အလမ်း ဖန်တီး နိုင် ရေး အတွက် နားလည်မှု စာ ချွန် လွှာ တစ်ရပ် ကို နိုဝင်ဘာ ၂၂ ရက် တွင် လက်မှတ်ရေးထိုးခဲ့သည် ။
 နားလည်မှု စာ ချွန် လွှာ အရ ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း သည် မြန်မာနိုင်ငံ တွင် ဖြန့်ချိ နေ သော Unilever ၏ ကုန်ပစ္စည်း ၁၅ မျိုး ကို သင် ထောက် ကူ ပစ္စည်း အဖြစ် အသုံးပြု နိုင် မှာ ဖြစ်ပြီး Unilever မှ စားဖိုမှူး ကြီး များ ကိုယ်တိုင် စာတွေ့ လက်တွေ့ သင်ကြား ပြသ ပေး မည် ဟု Unilever ၏ ကြော်ငြာ မန်နေဂျာ ဒေါ် ဆု လင်း ကို က ပြောသည် ။

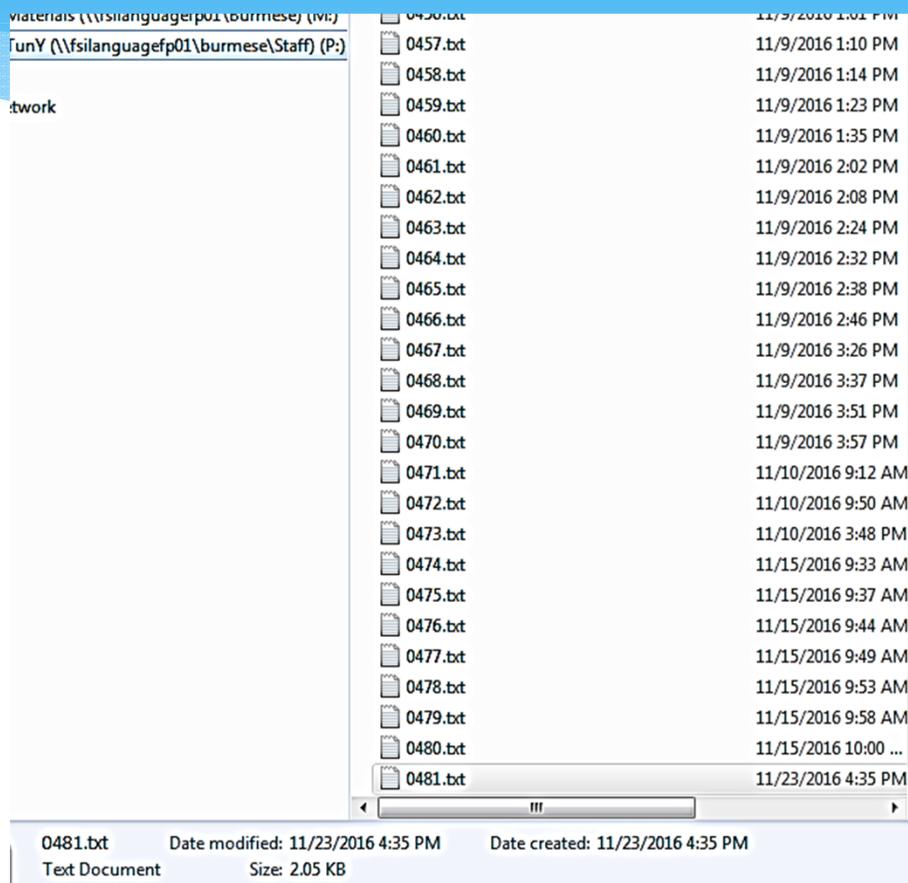
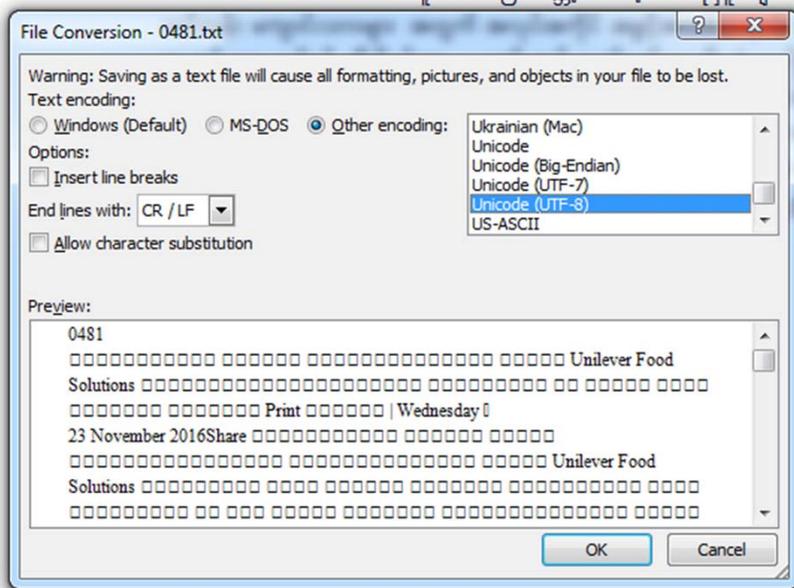
ကန်တော်ကြီး ဟိုတယ် သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်းဆောင်ရွက်ရေး နားလည်မှု စာ ချွန် လွှာ လက်မှတ် ရေးထိုး Print အိအိသူ | Wednesday ။
 23 November 2016Share ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူးပေါင်း ပြီး ကမ္ဘာ့ အဆင့်မီ စားဖိုမှူး များ ပေါ်ပေါက် လာ ရေး နှင့် သင်တန်း ကျောင်းသားများ အတွက် အလုပ်အကိုင် အခွင့်အလမ်း ဖန်တီး နိုင် ရေး အတွက် နားလည်မှု စာ ချွန် လွှာ တစ်ရပ် ကို နိုဝင်ဘာ ၂၂ ရက် တွင် လက်မှတ်ရေးထိုး ခဲ့ သည် ။
 နားလည်မှု စာ ချွန် လွှာ အရ ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်တန်းကျောင်း သည် မြန်မာနိုင်ငံ တွင် ဖြန့်ချိ နေ သော Unilever ၏ ကုန်ပစ္စည်း ၁၅ မျိုး ကို သင် ထောက် ကူ ပစ္စည်း အဖြစ် အသုံးပြု နိုင် မှာ ဖြစ်ပြီး Unilever မှ စားဖိုမှူး ကြီး များ ကိုယ်တိုင် စာတွေ့ လက်တွေ့ သင်ကြား ပြသ ပေး မည် ဟု Unilever ၏ ကြော်ငြာ မန်နေဂျာ ဒေါ် ဆု လင်း ကို က ပြောသည် ။

Phase I: How and What

Step 5. Save a text file

0481

ကန်တော်ကြီး ဟိုတယ် သင်တန်းကျောင်း နှင့် Unilever Food Solutions ပူး
နားလည်မှု စာ ချွန် လွှာ လက်မှတ် ရေးထိုး Print အိအိသူ | Wednesday ||
23 November 2016Share ကန်တော်ကြီး ဟိုတယ် နှင့် ခရီးသွားလုပ်ငန်း သင်
Unilever Food Solutions ပူးပေါင်း ပြီး ကမ္ဘာ့ အဆင့်မီ စားဖိုမှူး များ ပေါ်ပေါ်



Phase I: How and What

Step 6A. Spreadsheet Record for General Corpus

B482 MOU signed between a KDG Hotel Training School & Unilever Food Solution													
In	Title	CPR	F/I	Category	Theme	Sub-then	Genre	Posted Date	Capture Date	Source	Link	Words	Grand
445	444 Children Poems	N	I	CS		poem		2012	11/8/2016	Ministry of Information	https://docs.google.co	3731	
446	445 Pounding the Drums by Zin Min (Tha Main Htaw)	N	I	CS		Play		1/27/2016	11/8/2016	Ministry of Information	http://www.moi.gov.n	53271	
447	446 Offsprings and Traditional Festivals	N	I	CS				6/30/2015	11/8/2016	Ministry of Information	http://www.moi.gov.n	2770	
448	447 Natural Disaster Management by Lin Ko Ko	N	F	STA				2/21/2014	11/8/2016	Ministry of Information	http://www.moi.gov.n	26732	
449	448 Lily under the Storm	N	F/I	CS		Story		3/14/2014	11/8/2016	Ministry of Information	http://www.moi.gov.n	50537	
450	449 Water for Living Creatures	N	F	STA				7/30/2015	11/8/2016	Ministry of Information	http://www.moi.gov.n	3177	
451	450 3 Kinds of Fruits Everybody should eat	N	F	STA	Agri/Livestock/Health			7/30/2015	11/8/2016	Ministry of Information	http://www.moi.gov.n	44106	
452	451 Golden Short Stories	N	F	CS	Culture/Religion	Story		8/7/2015	11/8/2016	Ministry of Information	http://www.moi.gov.n	67420	
453	452 Nuclear Energy, Environment and Threat of Radiation	N	F	STA	Nuclear			10/25/2016	11/8/2016	Ministry of Information	http://www.moi.gov.n	25337	
454	453 Greenland under the Sky	N	F/I	STA	Agriculture	Story		1/27/2016	11/8/2016	Ministry of Information	http://www.moi.gov.n	41618	
455	454 Water Pot Shelf Journal (Science and Technology)	N	F/I	STA	Science and Technology			10/13/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	11353	
456	455 Water Pot Shelf Journal (Science and Technology)	N	F/I	STA	Science and Technology			9/9/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	10693	
457	456 Water Pot Shelf Journal (Agri and Livestock)	N	F/I	STA	Agri/livestock			10/13/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	9117	
458	457 Water Pot Shelf Journal (Agri and Livestock)	N	F/I	STA	Agri/livestock			9/9/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	7762	
459	458 Water Pot Shelf Journal (Culture)	N	F/I	CS	Culture			9/9/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	13656	
460	459 Water Pot Shelf Journal (Social)	N	F/I	CS	Social			9/9/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	12778	
461	460 Water Pot Shelf Journal (Decorum)	N	F/I	CS				9/9/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	13000	
462	461 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			10/13/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	10481	
463	462 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			9/9/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	9730	
464	463 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			8/12/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	10904	
465	464 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			7/15/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	11798	
466	465 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			6/17/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	11603	
467	466 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			5/4/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	11459	
468	467 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			4/5/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	10854	
469	468 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			3/8/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	9924	
470	469 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			2/5/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	10361	
471	470 Water Pot Shelf Journal (Economy)	N	F/I	BEA	Business & Economics			1/6/2016	11/9/2016	Ministry of Information	http://www.moi.gov.n	10425	
472	471 Wikipedia (Myanmar)	N	F	BEA	Business & Economics			11/10/2016	11/10/2016	Myanmar Wikipedia	https://my.wikipedia.c	1514	
473	472 Myanmar Small Business Development	N	I	BEA	Business & Economics			9/20/2015	11/10/2016	Myanmar Small Business D	http://www.myanmar-	11572	
474	473 View on Burma's Economic Development and Its Strategy	N	F	BEA	Business & Economics			10/31/2009	11/10/2016	scribd	https://www.scribd.co	165446	
475	474 Land price and electricity problems must be solved to become an i	Y	F	BEA	Business & Economics			11/15/2016	11/15/2016	Zaygywet	http://www.zaygywet.c	1343	
476	475 Not to make mistakes when you buy houses	Y	F	BEA	Business & Economics			11/15/2016	11/15/2016	Zaygywet	http://www.zaygywet.c	1512	
477	476 Business and Capital Investment	N	I	BEA	Business & Economics			11/15/2016	11/15/2016	Ministry of Information	http://www.moi.gov.n	1441	
478	477 Multi Level Marketing	N	F	BEA	Business & Economics			10/18/2016	11/15/2016	Ministry of Information	http://www.moi.gov.n	2166	
479	478 Unemployment Issue in Burma	N	F	BEA	Business & Economics			11/15/2016	11/15/2016	Ministry of Information	http://www.moi.gov.n	1420	
480	479 Reforestation	N	F	STA	Environment			8/13/2016	11/15/2016	Ministry of Information	http://www.moi.gov.n	1404	
481	480 Tax	N	F	BEA	Business & Economics			7/27/2016	11/15/2016	Ministry of Information	http://www.moi.gov.n	1616	
482	481 MOU signed between a KDG Hotel Training School & Unilever Food	Y	F	BEA	Business & Economics			11/23/2016	11/23/2016	Myanmar Times	http://myanmar.mmti	121	
483	482												

Why do we need a record?

1. Future use of materials
2. Copy rights issue
3. Sources of materials
4. Spoken Vs. Written
5. Range of the Corpora (How many tokens acquired for each category)

Phase I: How and What

Step 6A. Spreadsheet Record for General Corpus

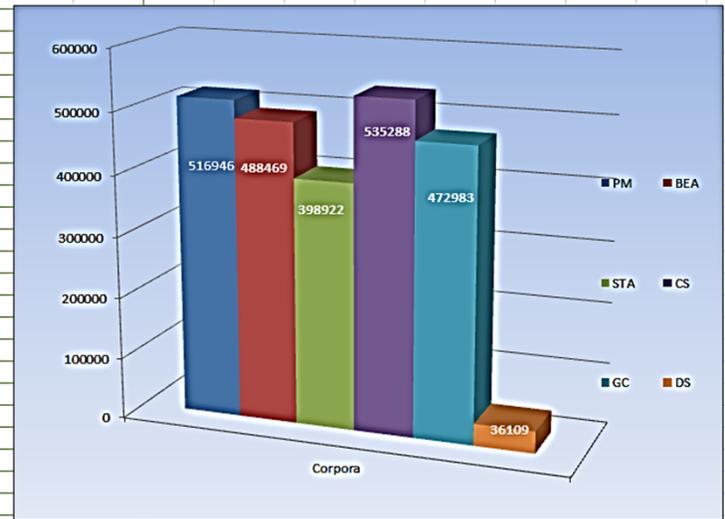
B482		fx MOU signed between a KDG Hotel Training School & Unilever Food Solution				
In	Title	CPR	F/I	Category	Theme	
445	444 Children Poems	N	I	CS		
446	445 Pounding the Drums by Zin Min (Tha Main Htaw)	N	I	CS		
447	446 Offsprings and Traditional Festivals	N	I	CS		

Sub-them	Genre	Posted Da	Captured	Source	Link	Words	Grand
poem		2012	11/8/2016	Ministry of Information	https://docs.google.co	3731	
Play		1/27/2016	11/8/2016	Ministry of Information	http://www.moi.gov.n	53271	
		6/30/2015	11/8/2016	Ministry of Information	http://www.moi.gov.n	2770	

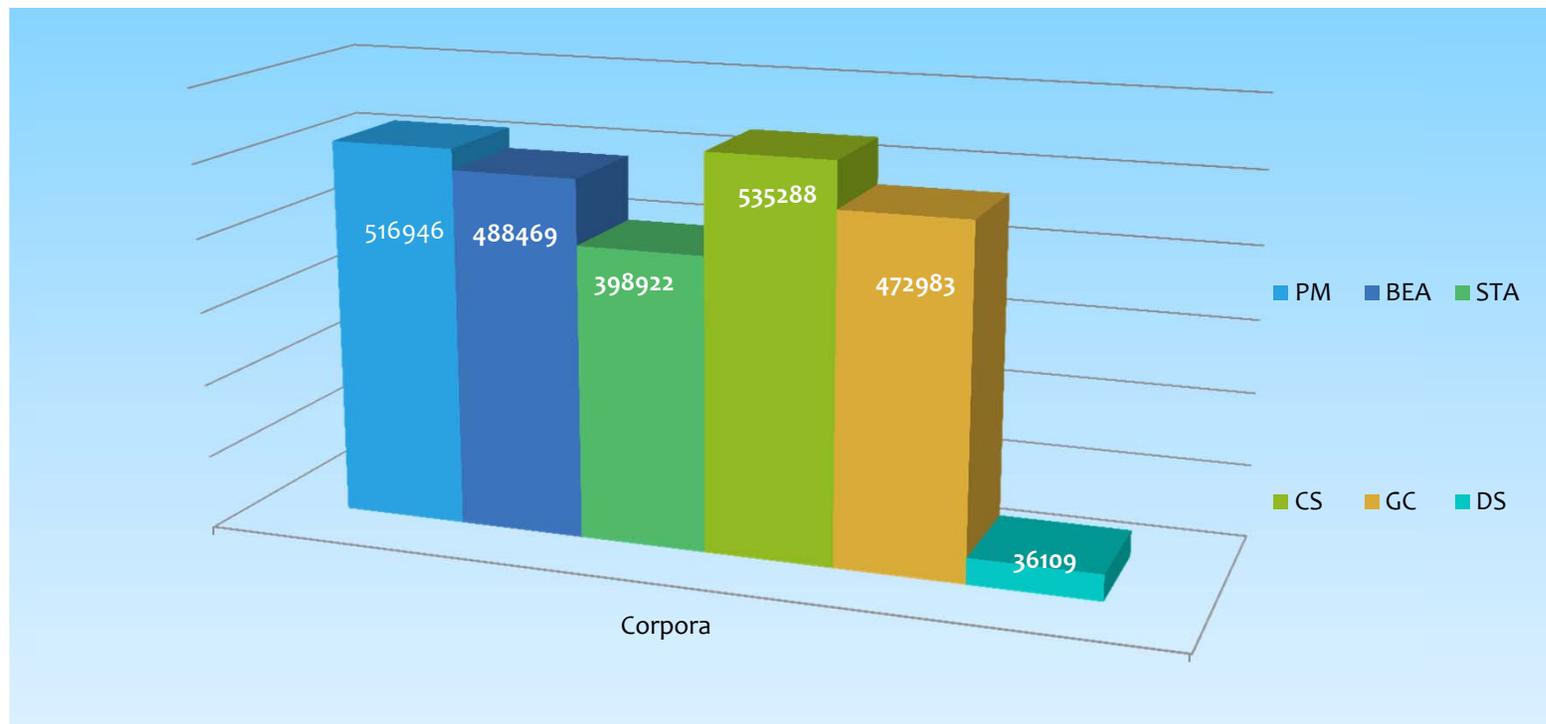
Phase I: How and What

Step 6A. Spreadsheet Record for Sub-Corpora

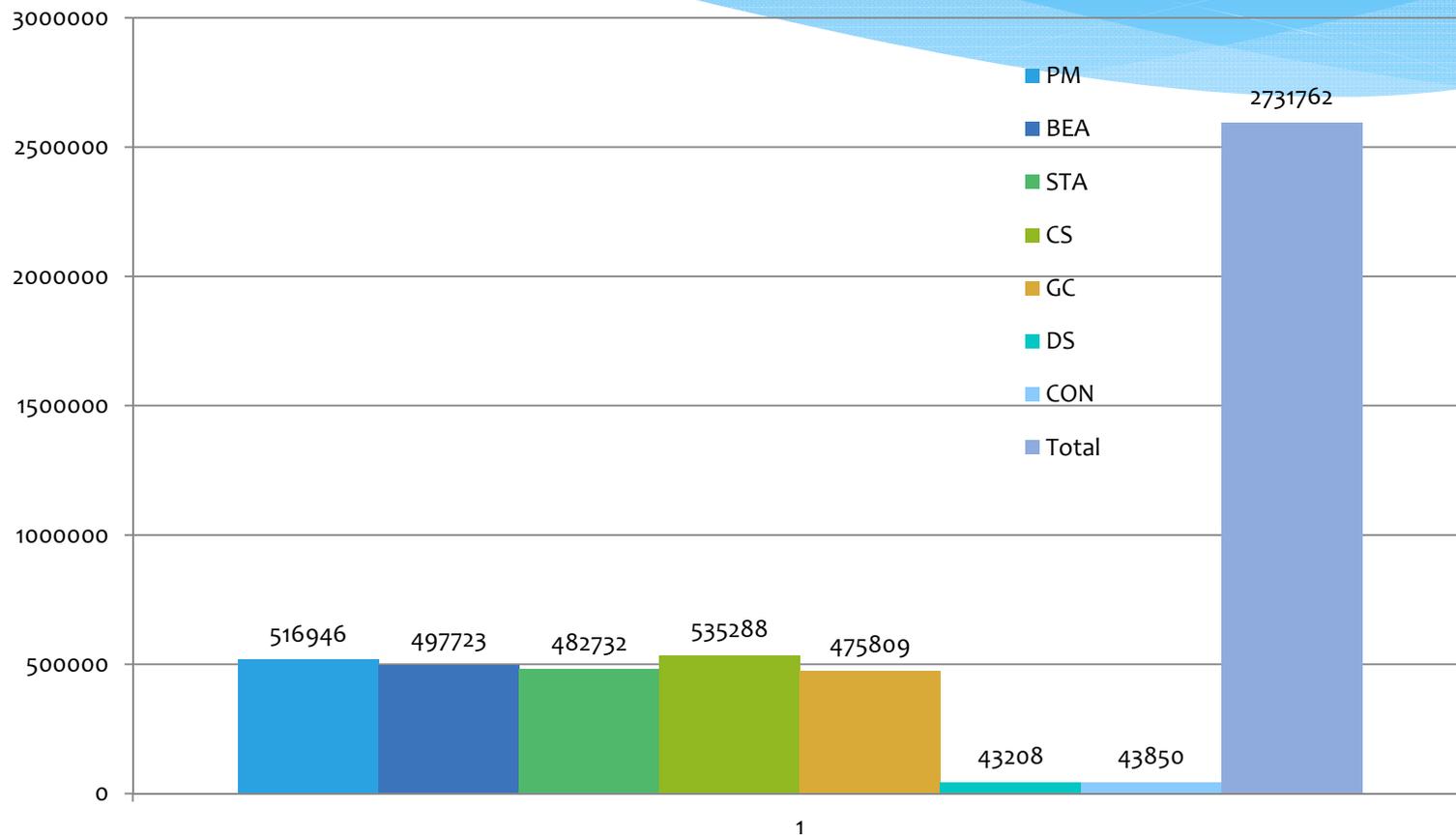
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
	PM	Total PM	BEA	Total BEA	STA	Total STA	CS	Total CS	GC	Total GC	DS	Total DS	Grand								
1	527	516946	1328	488469	1686	400326	1550	535288	262	472983	570	36109	2450121								
2	926		915		1866		573		1808		2586										
3	1365		563		1559		1626		1632		494										
4	244		1301		1713		1254		730		460										
5	398		1158		1615		1046		655		1398										
6	3176		2234		985		2125		145		5957										
7	952		4845		13413		1051		171		3803										
8	249		1992		1116		1812		114		1491										
9	5093		5093		1468		1028		282		1439										
10	533		918		1188		1849		1571		1251										
11	533		1333		1582		1243		1074		760										
12	472		503		732		1334		1179		374										
13	646		624		2361		1566		1728		15526										
14	618		538		520		1278		785		36109										
15	488		1070		578		1009		970												
16	840		880		994		857		1019												
17	2538		1397		1389		671		1544												
18	840		284		834		2659		2211												
19	338		1310		624		1026		905												
20	2298		1666		2025		896		1016												
21	1988		1635		749		1446		2163												
22	1159		1672		783		1158		3167												
23	1012		1330		1237		830		1080												
24	1015		877		3326		992		76098												
25	1667		993		2009		1708		5873												
26	1677		503		1405		20809		5123												
27	1376		1624		883		3603		3520												
28	2059		1258		1582		1787		4098												
29	449		1079		1967		2976		3451												
30	575		995		2008		2458		338												
31	520		684		1751		1343		1063												
32	443		1661		829		1497		932												
33	609		1304		1769		2254		1215												
34	363		1266		2934		1612		362												
35	557		824		653		3099		634												
36	552		1183		447		1537		618												
37	632		830		1151		1246		588												
38	1880		980		1129		2624		560												



of acquired tokens as of December 2016 =
2,451,217
(Sub-Corpora)

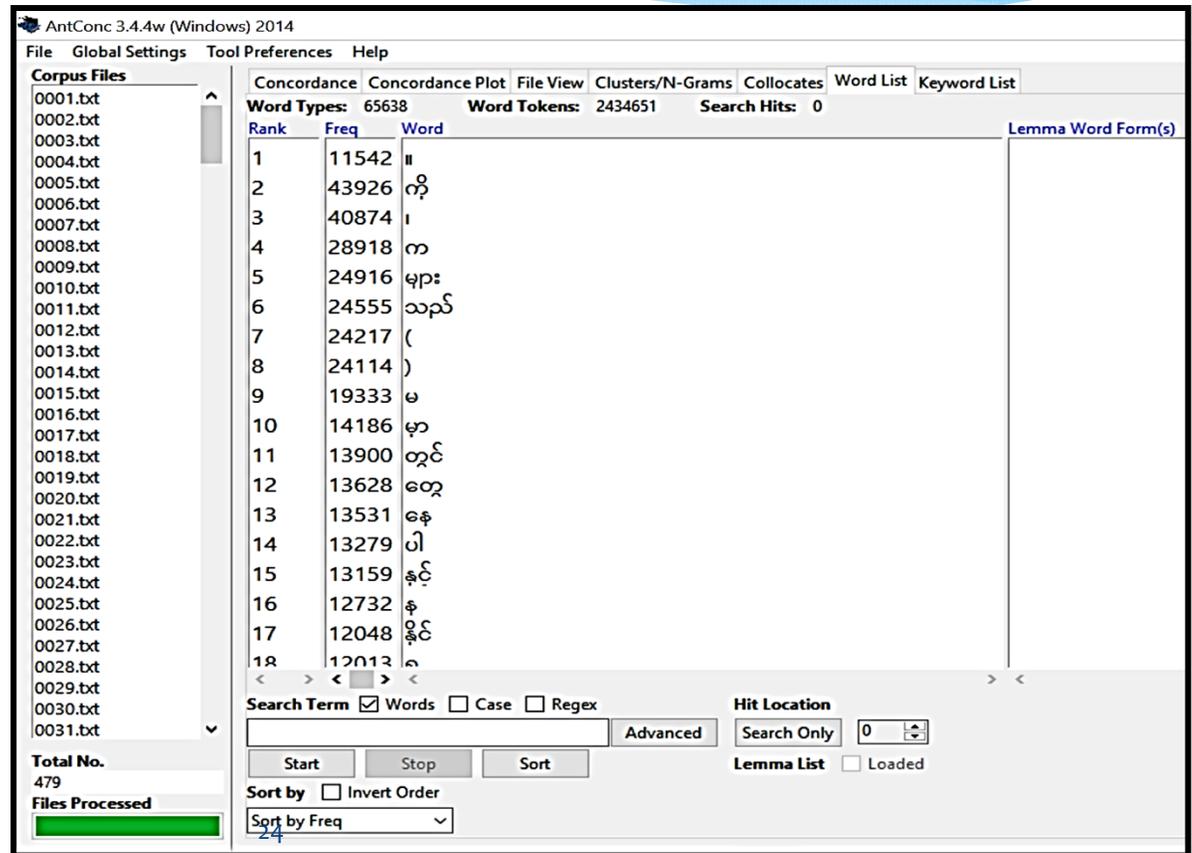
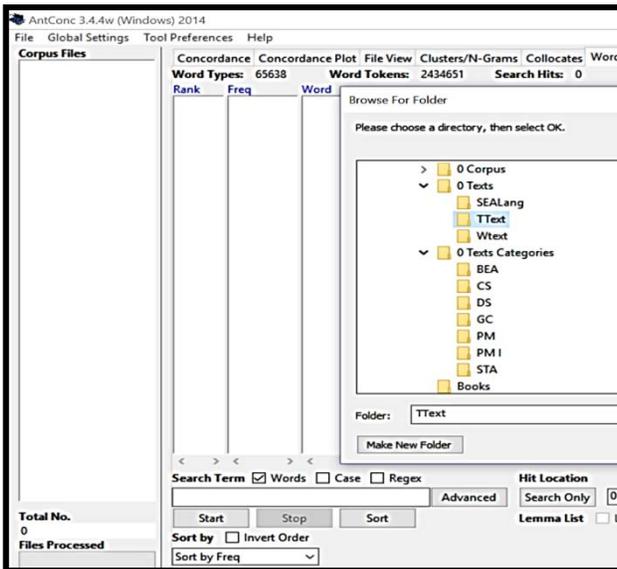
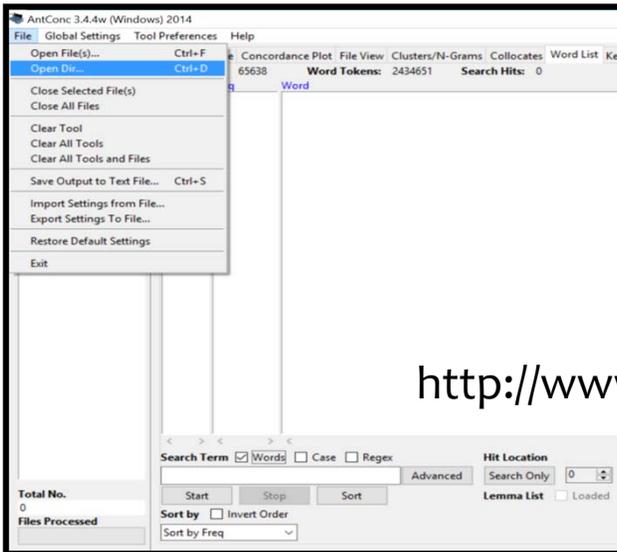


Total # of acquired tokens = 2,731,762 (Sub-Corpora)



Creating High Frequency Lists Testing with Tools (AntWordProfiler, AntConc, TextFixer) <http://www.laurenceanthony.net/>

<http://www.laurenceanthony.net/>



AntConc: A tool to do many things

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Corpus Files

- 0001.txt
- 0002.txt
- 0003.txt
- 0004.txt
- 0005.txt
- 0006.txt
- 0007.txt
- 0008.txt
- 0009.txt

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Word Types: 65638 Word Tokens: 2434651 Search Hits: 0

Rank	Freq	Word	Lemma Word Form(s)
1	11542		
2	43926	0	
3	40874		
4	28918	0	

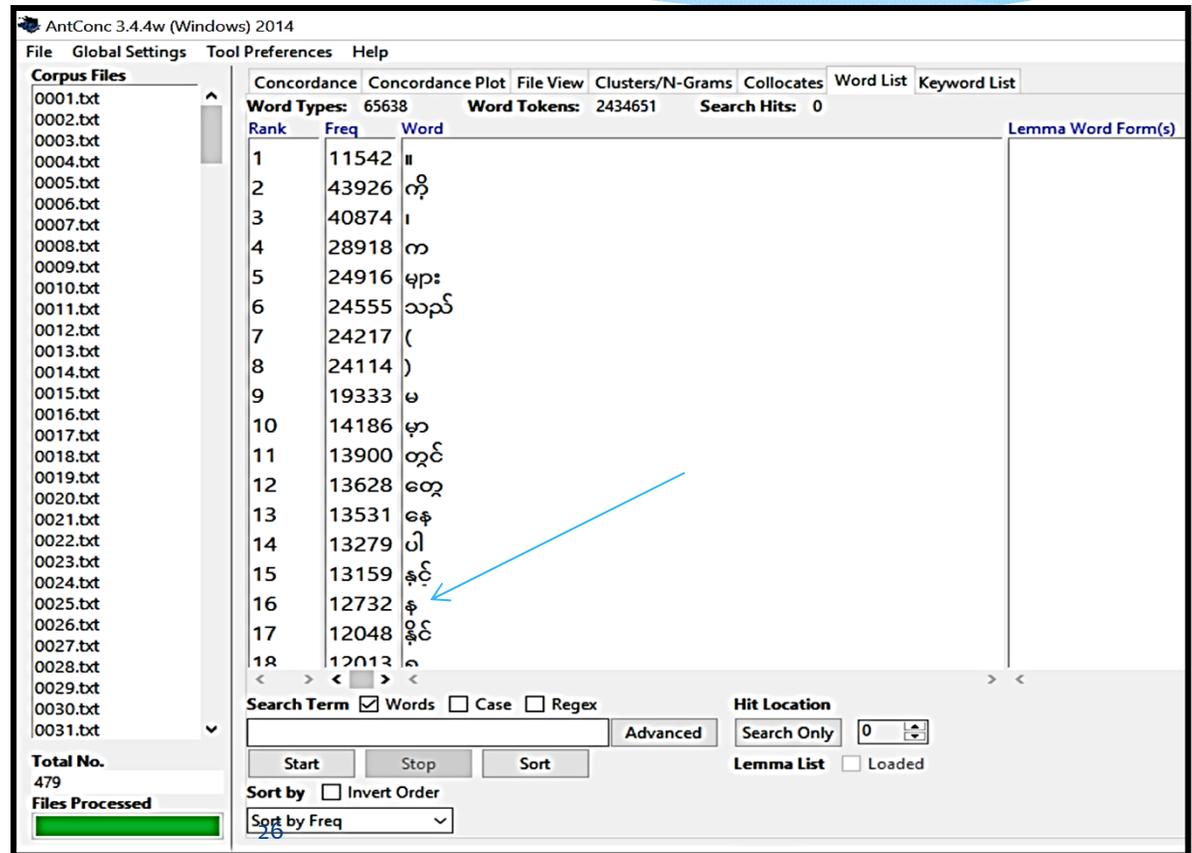
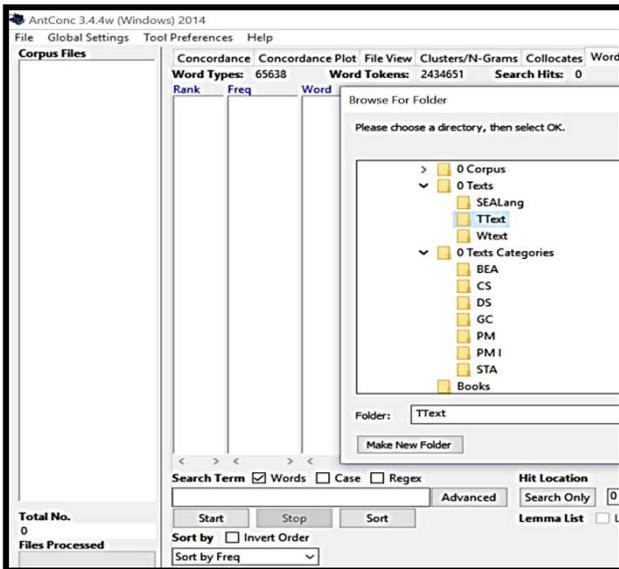
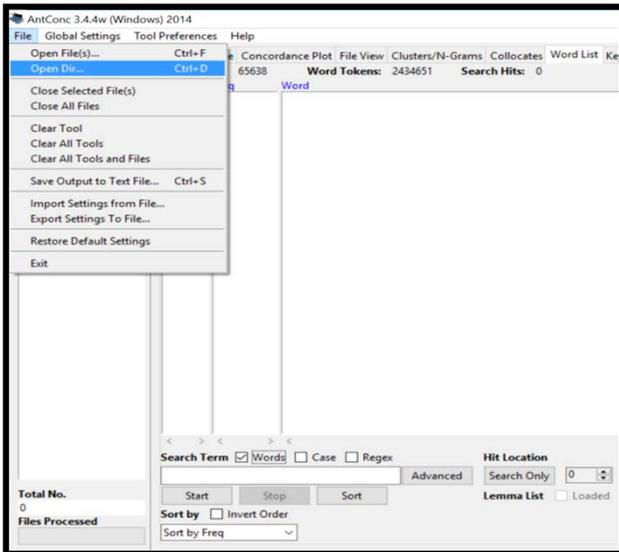
Please choose a directory, then select OK.

- > 0 Corpus
- ✓ 0 Texts
 - SEALang
 - TText
 - Wtext
- ✓ 0 Texts Categories
 - BEA
 - CS
 - DS
 - GC
 - PM
 - PM I
 - STA
 - Books

Folder: TText

Make New Folder

Creating High Frequency Lists Testing with Tools (AntWordProfiler, AntConc, TextFixer)



Creating High Frequency Lists Testing with Tools (AntWordProfiler, AntConc, TextFixer)

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Concordance Hits 13670

Hit	KWIC	File
24	ပြောကြား ခဲ့ပါသည်။ ယင်းသို့ ဖြန့် ဝှံ့ဖြူး မှု သည် ပမာဏ အနည်းငယ် သာ	0013.txt
25	ကာလ တွင် လျှပ်စစ် စွမ်းအား တိုးမြှင့် ဖြန့် ဝှံ့ဖြူး နိုင် ရေး အတွက် ဖြန့်	0013.txt
26	န ဝှံ့ဖြူး နိုင် ရေး အတွက် ဖြန့် ဝှံ့ဖြူး မှု ကုန်ယက် အခြေခံ အဆောက်အအုံ အဆင်	0013.txt
27	ပြည်ဆိုင်ရာ နေရောင်ခြည် စွမ်းအင် မိတ်ဖက် စု ဝှံ့ဖြူး ဦး လုပ်ဆောင် မှု (International Solar	0013.txt
28	။ ၂၆။ အားလုံး ပါဝင် သည့် ကျယ် ပြန့် ဝှံ့ဖြူး သော အမျိုးသား ပြန်လည် စည်းလုံး ရေး လုပ်	0013.txt
29	သဘောတူ ပီ ခဲ့ ကြ ပါသည်။ န ယူဒေး လီ မြို့ ။ ရက်စွဲ ၊	0013.txt
30	ထွန်း ဦး ကို တိုး မြှင့် ခ ဝှံ့ဖြူး အပ် လိုက် ပြီး အခြား ဒု ဗိုလ်ချုပ်ကြီး	0014.txt
31	တဲ့ အဆင့် မြင့် ရာထူး တွေ ကို ခ ဝှံ့ဖြူး အပ် ရာမှာ နိုင်ငံတကာ စံချိန် နဲ့ မ	0014.txt
32	ပီ တဲ့ အပြင် ဘယ် သူတွေ ကရွေးချယ် ခ ဝှံ့ဖြူး အပ် နေ သလဲ ဆိုတာ မေးခွန်း ထုတ်	0014.txt
33	ရာထူး နေရာတွေ ကို ဘယ် သူတွေ ကရွေးချယ် ခ ဝှံ့ဖြူး အပ် နေ သလဲ ဆိုတာကိုတော့ မေးခွန်း ထုတ်	0014.txt
34	မင်း အောင် လိုင် အပါအဝင် ကို ဘယ်သူ ခ ဝှံ့ဖြူး လဲ ဆိုတဲ့ ပစ္စာ က မေး	0014.txt
35	တဲ့ လူကြီးတွေ နဲ့ တိုင်ပင် ပြီးတော့ မှ ခ ဝှံ့ဖြူး ထား တာ ရှိပါတယ် ။ "။	0014.txt
36	တွင်း ရာထူး အသစ်တွေ တိုး မြှင့် ခ ဝှံ့ဖြူး အပ် ခံ ရ သူတွေ ကို ကြ	0014.txt
37	မြင်တယ် ။ တပ်မတော် အနေနဲ့ ဒီ အမိ ဝှံ့ဖြူး နာခံ တယ် ဆိုတဲ့ နေရာ မှာလည်း	0014.txt
38	မှာလည်း ပဲ တရားပဒေ နဲ့ ပီညွတ် တဲ့ အမိ ဝှံ့ဖြူး နာခံ မှု မျိုး ဖြစ် ပို့အတွက် ကို	0014.txt
39	ဦး နေရာမှာ လူ သစ် တွေ ကို ခ ဝှံ့ဖြူး အပ်ကြောင်း ဖော် ပြ ထား ပါတယ်	0014.txt
40	0031 န ရှိ မင်း ။ ပဲ့ခင်း ဆုတောင်း ။ (0031.txt
41	။ ဗိုက် ကလည်း ဆာ လှ ပြီ ။ န ရှိ မင်း ။ ၀-၉-၂၀၀၆။ (114)	0031.txt

Total No. 581
Files Processed

Search Term Words Case Regex Search Window Size 50

Start Stop Sort

Kwic Sort Level 1 | 1R Level 2 | 2R Level 3 | 3R

Creating High Frequency Lists Testing with Tools (AntWordProfiler, AntConc, TextFixer)

AntConc 3.4.4w (Windows) 2014

EN English (United States)

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Corpus Files
Combined ALL.txt

Word Types: 27917 Word Tokens: 2945097 Search Hits: 0

Rank	Freq	Word	Lemma Word Form(s)
14	21900	ဝေသာ	
15	21560	ဖြစ်	
16	21388	ဲ့	
17	19713	နဲ	
18	18326	နင်	
19	17660	တင်	
20	17260	မှာ	
21	16008	ရီ	
22	15982	-	
23	15441	ရင်း	
24	15430	ပိုး	
25	15273	နင်	
26	15257	၏	
27	14396	က	
28	13940	မှ	
29	13928	တဲ့	
30	13851	နဲ	
31	13410	မှ	

Search Term Words Case Regex Hit Location
ဝေသာ Advanced Search Only 0
Start Stop Sort Lemma List Loaded
Sort by Invert Order
Sort by Freq Clone Results

Total No.
1
Files Processed

Creating High Frequency Lists Testing with Tools (AntWordProfiler, AntConc, TextFixer)

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Corpus Files

0001.txt
0002.txt
0003.txt
0004.txt
0005.txt
0006.txt
0007.txt
0008.txt
0009.txt
0010.txt
0011.txt
0012.txt
0013.txt
0014.txt
0015.txt
0016.txt
0017.txt
0018.txt
0019.txt
0020.txt
0021.txt
0022.txt
0023.txt
0024.txt
0025.txt
0026.txt
0027.txt
0028.txt
0029.txt
0030.txt
0031.txt

Word Types: 38118 Word Tokens: 2935779 Search Hits: 0

Rank	Freq	Word
156	2305	အထိ
157	2299	ရေးသား
158	2295	ပါက
159	2292	စနစ်
160	2280	ချက်
161	2280	ယ
162	2258	ခွင့်
163	2245	ျါ
164	2227	ဖော်ပြ
165	2227	မြို့
166	2226	တင်
167	2221	စာ
168	2211	သင့်
169	2208	သို့မဟုတ်
170	2207	ကောင်း
171	2180	ပါဝင်
172	2178	ဗ
173	2176	ေက

Search Term Words Case Regex Hit Location Search Only 0

Start Stop Sort

Sort by Invert Order Sort by Freq

Total No. 581 Files Processed

AntConc 3.4.4w (Windows) 2014

File Global Settings Tool Preferences Help

Concordance Concordance Plot File View Clusters/N-Grams Collocates Word List Keyword List

Corpus Files

Combined ALL.txt

Word Types: 27917 Word Tokens: 2945097 Search Hits: 0

Rank	Freq	Word
193	2034	စိုက်ပျိုး
194	2032	ပြည်ထောင်စု
195	1999	တစ်ဦး
196	1982	အရေးအခန်း
197	1982	ါ
198	1981	ကလေး
199	1974	ထိန်းသိမ်း
200	1969	အချိန်
201	1965	မရှိ
202	1950	မှတ်
203	1948	တင်
204	1946	တောင်
205	1944	သတို့
206	1930	/
207	1929	ရွာ
208	1922	ပစ္စည်း
209	1921	အခြား

Search Term Words Case Regex Hit Location Search Only 0

Start Stop Sort

Sort by Invert Order Sort by Freq

Total No. 1 Files Processed

What I am currently doing ...

Using AntConc to make a stable list of high frequency words

Next ...

Using AntWordProfiler to profile the texts

What does this project mean for similar languages?



Burmese Project



Discussion:



One-Page Description

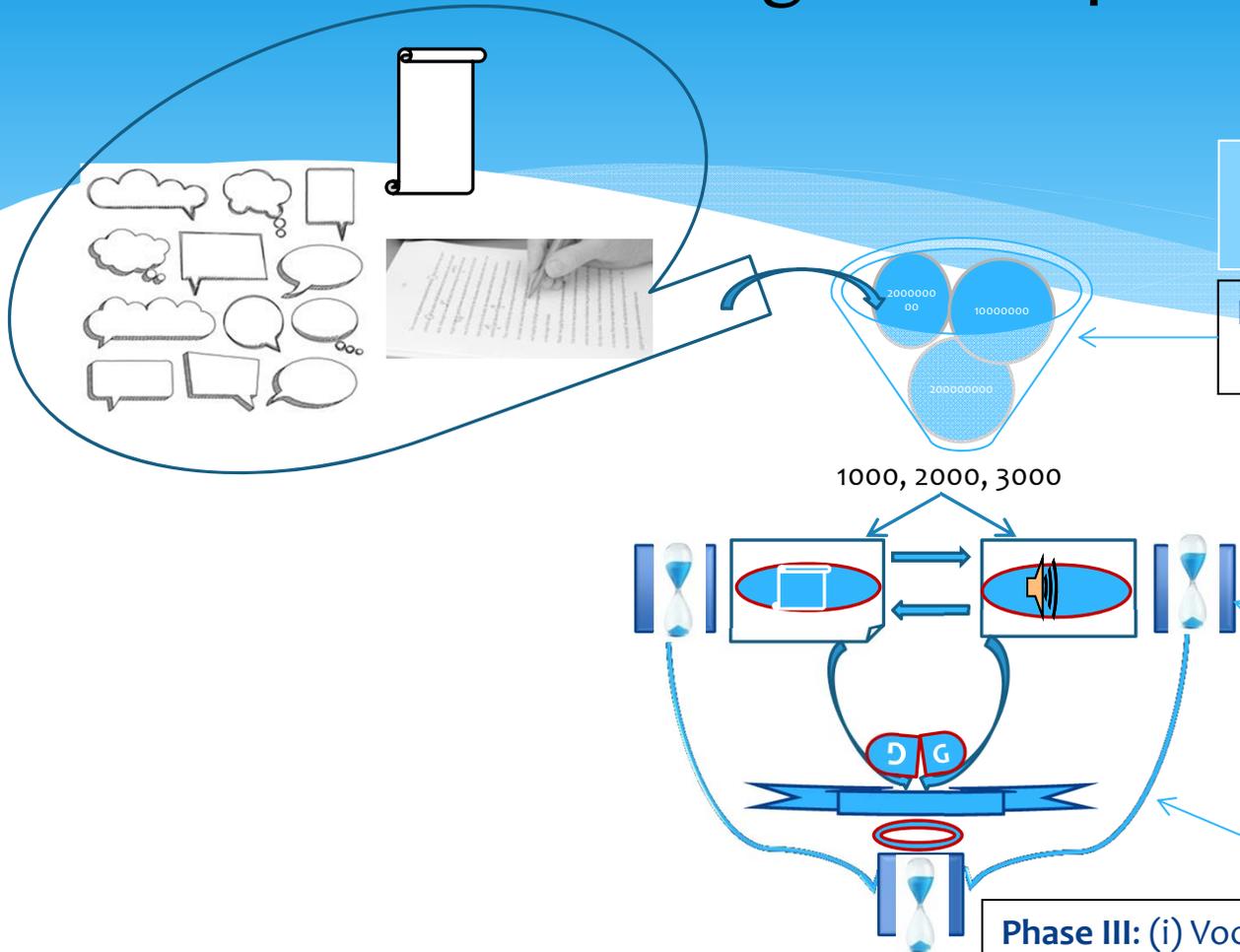
**Road Map: Three Phases of
Burmese Project**

**Phase I: Creation of
Vocabulary List**

**Phase II: Texts and Prototype-
lessons**

- (i) Reading texts
- (ii) Lesson prototypes
- (iii) New ways of
teaching the
Burmese script and
sound system
- (iv) Prototype lessons
to be tested in class

**Phase III: (i) Vocabulary based Curriculum with
“Four Strands” principle
(ii) Development of Grammar
Instructions and Integrate in
Curriculum**



Reference