K-Romanizer
– Windows Application for ALA-LC Korean Romanization –

Presented at Annual CKM Session
(Toronto, March 16, 2017)

Hyoungbae Lee
Korean Studies Librarian
Princeton University
Programming

• Based on AutoIt, a freeware automation language for Microsoft Windows

  o Easy to learn BASIC-like syntax
  o Simulate keystrokes and mouse movements
  o Manipulate windows and processes
  o Interact with all standard windows controls
  o Scripts can be compiled into standalone executables
Working Group Timeline

- Oct. 2015 - Mar. 2016: Developing and testing (demo version)
- Mar. 2016 (CEAL Meeting at Seattle): Demonstration (demo version)
- Mar. 2017 (CEAL Meeting at Toronto): Demonstration (release version)
Approach and Result

• Demo version tested with examples in ALA-LC Korean Romanization Table (December, 2015)
  o The ALA-LC Table is basically rules and exceptions
  o Rules ⇒ Write automation scripts
  o Exceptions ⇒ Build dictionaries

• Release version tested with random book titles (April-December, 2016)
  o 1,600 titles (4,000 words)
  o Error rate: less than 5%
  o Error rate excluding capitalization: less than 3%
Working Group Members

- Hyoungbae Lee (Princeton, Chair)
- Hyunjoo Eom (UBC)
- Eunseung Oh (Stanford)
- Jee-Young Park (Chicago)

Proudly Present ...
K-Romanizer ver. 4.2
Main Key Combo 1

- Ctrl + Shift + K romanizes Personal names

이 세상 (李世祥, 1671-1719) ⇒ Yi Se-sang
박 마리아 ⇒ Pak Maria
김 산각 ⇒ Kim Satkat
선우 휘 ⇒ Sŏnu Hwi
김 조 광수 ⇒ Kim Cho Kwang-su

신 립, 류 성롱, 이 하늬, 김 초롱¹, 심 마르시아스 ⇒
Sin Rip, Yu Sŏng-nyong, Yi Hanŭi, Kim Ch'o-rong, Sim Marūsiasū

¹ K-Romanizer cannot tell whether “초롱” is of Han'gŭl origin or a native Korean name, thus inserting a hyphen by default. On the other hand, “하늬” in “하늬” cannot be of Han'gŭl origin, and K-Romanizer regards it as a non-Han'gŭl name by removing a hyphen.
Main Key Combo 2

- **Ctrl + Alt + K** romanizes general words and phrases

이 세상 ("this world") ➞ I sesang
오늘의 사건 사고 ➞ Onũl ŭi sakũn sago
한미 경제 협력 ➞ Han-Mi kyōngje hyōmnyŏk
토론토 CEAL 연례 회의 ➞ T'oront'o CEAL yŏllye hoeũi
2017 년 3 월 16 일 ➞ 2017-yŏn 3-wŏl 16-il

†a 한국 신문 방송 매크 글 변천사 : †b 독립 신문^2 100 돌 기념 ➞
†a Han'guk sinmun pangsong malgŭl pyŏnch'ŏnsa : †b tongnip sinmun 100-tol kinyŏm

^2 "독립 신문" has to be manually capitalized as “Tongnip sinmun.”
Special Features 1

• Recognizes geographical names (國, 道, 市, 區, 郡)

서울 특별시 ⇔ Sŏul T'ŭkpyŏlsi
서울시 종로구 ⇔ Sŏul-si Chongno-gu
전라북도 고창군 ⇔ Chŏlla-bukto Koch'ang-gun
대한 민국, 중국, 일본 ⇔ Taehan Min'guk, Chungguk, Ilbon

ǂa 실크 로드 문명 기행 : ǂb 일본 나라³ 에서 이탈리아 로마 까지 ⇔
ǂa Silk’ŭ rodŭ munmyŏng kihaeng : ǂb Ilbon nara esŏ It’allia Roma kkaji

---
³ "실크 로드" and "나라" has to be manually capitalized as "Silk’ŭ Rodŭ" and "Nara" respectively.
Special Features 2

• Recognizes names of Korean kings (宗, 祖, 王, 君)

선조 와 광해군 ⇒ Sŏnjo wa Kwanghae Kun
고구려 광개토 대왕 ⇒ Koguryŏ Kwanggaet'o Taewang
의 친왕 과 영 친왕 ⇒ Úi Ch‘inwang kwa Yŏng Ch‘inwang

†a 숙종 대왕 민 중전⁴ 덕행록 : †b 숙종 대왕 어제 ⇒
†a Sukchong Taewang min chungjŏn tŏkhaengnok : †b Sukchong Taewang ōje

⁴ "민 중전," the queen of King Sukchong, has to be manually capitalized as “Min Chungjŏn.”
• Romanizes Hancha or Hancha-Hangul mixture

李世祥 ⇔ Yi Se-sang
高句麗 廣開土大王 ⇔ Koguryŏ Kwanggaet'o Taewang
全羅北道 高敞郡 ⇔ Chŏlla-bukto Koch'ang-gun
합천: 陜川 or 陜川? ⇔ 陜川 = Sŏmch'ŏn / 陜川 = Hapch'ŏn

ǂa 高麗時代의研究: ǂb 特히圖讖思想의發展을中心으로

ǂa Koryŏ sidae ŭi yŏn'gu: ǂb t'ŭkhi toch'am sasang ŭi palchŏn ŭl chungsim ŭro
Other Useful Features 1

- Ctrl + Shift + Alt + K will not capitalize the first letter

- Shift + Alt + K capitalizes the first letters of all words

A warning message will pop up for these words due to ambiguity: 조선=Chosŏn, 리더=lidŏ, Chejudo, 리더(Leader)=lidŏ, 리더(reader)=ridŏ, 濟州島=Chejudo, 濟州道=Cheju-do. Users have to modify them as necessary.
Other Useful Features 2

- Entering special characters for cataloging

\[\text{Ctrl} + \text{Alt} + \text{O}\] enters a lower-case \(\ddot{o}\)

\[\text{Shift} + \text{Alt} + \text{O}\] enters an upper-case \(\ddot{O}\)

\[\text{Ctrl} + \text{Alt} + \text{U}\] enters a lower-case \(\ddot{u}\)

\[\text{Shift} + \text{Alt} + \text{U}\] enters an upper-case \(\ddot{U}\)

\[\text{Ctrl} + \text{Alt} + \text{\textbackslash}\] enters a double dagger \(\ddagger\)
Other Useful Features 3

- Ctrl + Shift + Alt + C adds Hangul reading to Hancha texts.

春雨細不滴, 夜中微有聲, 雪盡南溪漲, 草芽多少生 / 鄭夢周

⇒

春雨細不滴, 夜中微有聲, 雪盡南溪漲, 草芽多少生 / 鄭夢周 (춘우세부적, 야중미유성, 설진남계창, 초야다소생 / 정몽주)

不恨吾生晩 只恨吾有眼 無復見斯人 危途悌空濁 / 李安訥

⇒

不恨吾生晩 只恨吾有眼 無復見斯人 危途悌空濁 / 李安訥 (불한오생만 지한오유안 무복견사인 위도제공산 / 이안눌)

---

6 A warning message will pop up for 復 and 見, as these characters have more than one reading. Users have to modify them as necessary: 復 should be manually changed into 부 in this poem. On the other hand, 不 has two readings as 부 or 불 depending on the following sound, and the sound change will be automatically applied by K-Romanizer.
Download

- **One Big CAVEAT!**
  Since word division has to be first done manually, it is strongly advised that users get familiarized with the ALA-LC Korean Romanization rules before using K-Romanizer.

- **Download Link available at:**
  [http://www.eastasianlib.org/ckm/subcommittees.html](http://www.eastasianlib.org/ckm/subcommittees.html)

- Download a new version of K-Romanizer at least **every 3 months** for updated dictionaries. Each version will expire after 90 days from the last update.