The corpus of English loanwords in Hong Kong Cantonese

Cathy S.P. Wong
Department of English
The Hong Kong Polytechnic University
INTRODUCTION

- English loanwords have had a long history in Cantonese.
- The earliest documentation dates back to Morrison (1828) in which words such as
  - 「波球」bo1 kau4 for “ball”
  - 「啤酒」be1 zau2 for “beer”
- Loanwords that have fully integrated into HK Cantonese include:
  - 「的士」dik1 si2 for “taxi”
  - 「奄列」aml1 lit6 for “omlette”
A lot of interesting issues are found:

A. Some loanwords have been nativized
   - e.g. baa1 si2 巴士 for “bus” ➔ no native term for it

B. Some loanwords go through “death”
   - e.g. daai6 got1 lau4 大嘗樓 “high court”

C. Some English words are “borrowed” twice
   - e.g. “cream” was first borrowed as gei4 lim1 忌廉 in food items;
   - e.g. “cream” is recently borrowed as kwim1 but only referring to “cream” in hand or face cream
The Etymological Dictionary Project (2)

- This is what prompted the etymological project.
- The purpose is to charter the “life cycle” of the majority of English loanwords found in Cantonese.
- The aim of the etymological dictionary project is to assemble an exhaustive list of English loanwords found in Cantonese with full annotation in terms of their pronunciation and origin, etc. and to compile a searchable database.
The Etymological Dictionary Project (3)

- The initial product of this project is a small corpus: a collection of 694 English loanwords (*Bauer and Wong 2008*) documented in Cantonese over a time span of 180 years, from 1828 to 2008.

- Sources:

<table>
<thead>
<tr>
<th>Academic</th>
<th>Non-academic</th>
<th>Internet</th>
</tr>
</thead>
<tbody>
<tr>
<td>dictionaries</td>
<td>local newspapers</td>
<td>forums</td>
</tr>
<tr>
<td>monographs</td>
<td>tabloid magazines</td>
<td>discussion groups</td>
</tr>
<tr>
<td>journal articles</td>
<td>youth magazines</td>
<td></td>
</tr>
<tr>
<td>student theses, dissertations</td>
<td>comic books</td>
<td></td>
</tr>
</tbody>
</table>
Of these 693 items

- 85%: only phonetically-transliterated syllables
  - e.g. baa1 si2 巴士 (“bus”)
  - transliterated from the English /bʌs/.
- 15%: at least one phonetically-transliterated syllable
  - e.g. be1 zau2 啤酒 (“beer”)
  - the first syllable is transliterated from “beer” and the second syllable is a Cantonese word meaning “wine”.

The Etymological Dictionary Project (4)

2014/3/11 loanword corpus - cathy wong
In terms of word categories, the majority of these loanwords are nouns (80%). Only 12% of them are verbs.

The semantic categories are quite diversified.

<table>
<thead>
<tr>
<th>food and drinks</th>
<th>recreation</th>
</tr>
</thead>
<tbody>
<tr>
<td>academic terminology</td>
<td>fashion</td>
</tr>
<tr>
<td>technology</td>
<td>occupations</td>
</tr>
<tr>
<td>police jargon</td>
<td>music</td>
</tr>
</tbody>
</table>
The Corpus (1)

- In order to facilitate the eventual publication of this corpus as a dictionary, the corpus has the following features:
  - a text file in a particular format
  - each lexical entry consists of “bands” of information
  - the next slide illustrates what the entry belzau2 啤酒 (“beer”) looks like in the original corpus
  - the slide after that shows what the entry will look like in the printed dictionary.
The Corpus (2)

- .hw bel zau2
- char 啤酒
- ps N.
- en fm Eng. beer
- seealso 啤be1
- et first recorded as 卑酒 pay tsaw in RM1828:14 and CHA1907:21
- stch 啤酒 pi2 jiu3
- df beer
- exchar 佢一次過飲咗五罐啤酒
- exrom keoi5 jat1 ci3 gwo3 jam2 zo2 ng5 gun3 bel zau2
- exeng He has taken five cans of beer at a time
- ser 1000090179
<table>
<thead>
<tr>
<th>baai³ naa² fan¹ laap⁵</th>
<th>巴拿芬蠟 n.  《巴拿芬》 baai³ naa² fan¹ fm. Eng. paraffin》石蠟油 shi² laa² you² paraffin wax; often used in cosmetics</th>
<th>Paraffin 有 多 種 用 途 baai³ naa² fan¹ laap⁵ jau² do¹ zung² jung² tou⁴ Paraffin can be used in various ways [WTH2004APP2:8]</th>
</tr>
</thead>
<tbody>
<tr>
<td>baai³ neo³²2  吧女 n. 《baai³ fm Eng. bar》 酒吧女郎 jiu³ baai³ no³ lang² bar girl  吧女有年際化趨勢 baai³ neo³²2 jau² nin⁵ hing¹ faa³ ge³ ceoi² sei² There is a trend for younger girls to become bar girls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>baai³ jik¹  班戟 n. (M. 個 go³; 件 gin³) 薄皮艇 baai³ jik¹ baa³ jik¹ Mango pancake [CHN1972:218, ZN1999:6]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>baai³ si¹ dak⁵ siu¹ duk⁶ 巴斯消毒 《baai³ si¹ dak⁵ siu¹ duk⁶ fm. Eng. pasteurization》 巴斯消毒 baai³ si¹ de² xiao³ duk⁶ n. pasteurization  巴斯消毒能致死奶製品中微生物 baai³ si¹ dak⁵ siu¹ duk⁶ nan⁵ saai³ sei³ baai³ zung¹ go³ moe¹ sao⁵ mat⁶ Pasteurization can kill all microorganisms in dairy products ① to pasteurize baai³</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Currently, the content of the etymological dictionary is still being carefully revised and meticulously edited.

The corpus has now been put on the web.

First, the original WORD document was transferred to an EXCEL file.

A website was constructed using the EXCEL file.

“English Loanwords in Hong Kong Cantonese” is available at

http://funstuff.engl.polyu.edu.hk/loanwords/
The Online Version (2)

- The website is easy to use.
- A simple search can be done by typing in a word in English, Chinese, or romanized Cantonese under the “Search” button on the front page.
- The functions under the “Advanced Search” tap include three choices: search by
  - (1) Chinese characters/words,
  - (2) English words, and
  - (3) Jyutping (i.e. Cantonese romanization),
Your Input is Most Welcome!

- Currently, the team is still working on the editing functions of the website.
- The idea is to update the corpus by adding newly found loanword items.
- Please try out the online version of our loanword corpus and give us comments and feedback.
- For enquiries and comments, please email me:
  - cathy.wong.engl@polyu.edu.hk