The Structure of Chinese Loanwords

Feng Zhiwei
KORTERM KAIST (Korea Advanced Institute of Science and Technology)
IAL MOE (Institute of Applied Linguistics, Ministry of Education, China)
zwfeng@world.kaist.ac.kr

The loanwords are very popular in every language. There are six different loanwords (bold letters) in following short English text:

"Jane saw a **baby squirrel** eating **ketchup** left out after yesterday **barbeque**. Although she was still wearing her **cotton pajamas**, she hurried outside to chase the creature away."

Here, "baby" is borrowed from Dutch, "squirrel" from French, "ketchup" from Malay, "barbeque" from Indian, "cotton" from Arabic, "pajamas" from Urdu.

These 6 loanwords are phonemic loanwords, but there a lot of semantic loanwords. So the loanwords can be divided into two types: phonemic loanwords and semantic loanwords.

The phonemic loanwords and semantic loanwords have their respectively advantages and disadvantages:

The advantages of phonemic loanwords:

-- The pronunciation of phonemic loanword is similar with the pronunciation of word in donor language, so it is convenient for international exchange. For example, the pronunciation of Chinese phonemic loanwords "喇嘛/lama/(Lama)" "柠檬/ningmeng/(lemon)" "咖啡/kafei/(coffee)" are very similar with the pronunciation in donor languages. It is convenient for international exchange.

-- The phonemic loanword is easy accepted by the persons who know donor languages or who were well educated.

The disadvantages of phonemic loanwords:

--The pronunciation of phonemic loanword is not so easy to learn and memorize by common people who didn't know the donor languages.

-- The transcription of phonemic loanword is not so strictly, there will be a lot of variation of phonemic loanwords for the same word in donor language. E.g. SARS in English can be

transcribed in different forms in Chinese language: 萨斯/sasi/ or 沙斯/shasi/.

The advantages of semantic loanwords:

　　-- The semantic loanword is translated from donor language with the original words or morphemes of recipient language, so it is easy to learn and to memorize for common people of recipient language, it is also easy to popularize in the society of recipient language.

　　-- If we use Chinese character to record the semantic loanwords, because Chinese character is ideogram, it can express semantic sense directly, so the semantic word written in Chinese characters is easy be accepted by common people who can recognize Chinese characters.

Disadvantages of semantic loanwords:

　　-- It is not convenient for international exchange.

　　Therefore we have to harmonize the relation between the phonemic loanwords and the semantic loanwords. The analysis of the structure of loanwords will be helpful for the harmonization.

　　In this paper, we shall discuss the structure of phonemic loanwords and semantic loanwords of Chinese language.

　　The most of phonemic loanwords in Chinese is monosyllabic. The monosyllabic phonemic loanwords include following types:

　　① Loanwords expressed with newly created Chinese characters: To create new Chinese character to translate the words in donor language.

　　　- The names of 109 chemical elements are all monosyllabic Chinese characters, the most of chemical elements have to be expressed with newly created Chinese characters. E.g. metal element: 锝/de/(technetium),铬/ge/(chrome, chromium),锇/e/(osmium); non-metal element: 碲/di/(tellurium),砷/shen/(arsenic),硒/xi/(selenium); gas element: 氟/fu/(fluorin),氪/ke/(krypton),氖/nai/(neon). They are the phonemic loanwords.

　　　- The name of chemical compound:　The newly created Chinese character "蒽/ēn/" expressed the chemical compound "anthracene".

　　② Loanwords expressed with original Chinese characters: Because the original Chinese characters have already definite meaning, so this type of loanwords became the words with multiple meanings (polysemy).

　　　E.g. The loanwords "埃、钫、阀、开、米" are all polysemy words.

Chinese character	Original meaning	new meaning of loanword
埃/ai/	dust	angstrom (scientific unit for length)
钫/fang/	ancientry vessel	francium (chemical element name)
阀/fa/	warlord	valve
开/kai/	open	karat
米/mi/	rice	meter

　　In order to differentiate the different meanings for same Chinese character, some times the pronunciation of the same Chinese character needs to be changed:

　　For example,

　　禅－original meaning is "demise"（禅让）, and it is pronounced as /shàn/.The loanword is pronounced as /chán/, its meaning is Buddhist [from "dhyāna" of Sanskrit]

　　钯－Original meaning is "prong, rake"(a variant form of "耙"), it is pronounced as/pá/.The loanword is pronounced as / bǎ /, its meaning is palladium, the name of metal element.

刹－original meaning is "brake", and it is pronounced as /shā /. The loanword is pronounced as /chà/, its meaning is "Buddhist temple" (from "ksetra" of Sanskrit).

打－original meaning is "beat", and it pronounced as /dǎ /(third tone). The semantic loanword is pronounced as /dá/ (second tone), its meaning is "dozen".

③ Abbreviation of phonetic loanwords: For example,

伏/fu/:from 伏特(volt)

赫/he/: from 赫兹 (hertz)

卡/ka/: from 卡路里 (calorie)

瓦/wa/: from 瓦特 (watt)

欧/ou/: from 欧姆 (ohm)

安/an/: from 安培 (ampere)

帕/pa/: from 帕斯卡 (Pascal)

Polysyllabic phonemic loanwords: E.g. 盘尼西林 /pannixilin/（penicillin），密陀僧 /mitouseng/(mudarasingu，from Sanskrit)，模特儿/mote'er/(model)，尼古丁/niguding/（nicotin），比特/bite/（bit），山农/shnanong/（Shannon），哈特莱/hatelai/（Hartley），摩托/motuo/（motor），锆/gao/（zirconium[English] or Zirkon[German]），马达/mada/（motor），雷达/leida/（radar），拓扑/tuopu/（topology）.

In the process of translation of phonemic loanwords, the pronunciation of word in donor language may be changed.

There is a special type of phonemic loanword in Chinese language -- 字母词/zimuci/ (alphabet word). The alphabet word kept the original form in donor language, but it already becomes a part of Chinese vocabulary and was included in <现代汉语词典> （Modern Chinese Dictionary）. For example,

ASCII: American Standard Code for information Interchange

KTV: Kara OK and television

DVD: Digital Video Disc

MTV: Music Television

OA: Office Automation

IT: Information Technique

UFO: Unidentified Flying Object

The alphabet word can be regarded as a special type of phonemic loanwords.

However, the most of loanwords in Chinese are semantic loanwords.

The structure of semantic loanwords is as follows:

-- Simple loanwords:

Some names of chemical elements are translated with the original Chinese characters. E.g. "金/jin/(aurum, gold),银/yin/(argentums, silver),铜/tong/(copper),铁/tie/(iron),铝/lǚ/(aluminum),铅/qian/(plumbum, lead),锡/xi/(stannum, tin),汞/gong/(hydrargrum, mercury),磷/lin/(phosphorus),硫/liu/(sulphur)". They come from Chinese original name of materials, and the sense is similar with the correspondent words in donor language, they are semantic loanwords;

A few of gas elements expressed by newly created Chinese characters are semantic loanwords. E.g. 氧/yang/(oxygen)、氢/qing/(hydrogen)、氮/dan/(nitrogen).

-- affixation:

The affix in Chinese semantic loanwords:

Prefix:

    1. 反/fan/: 反作用 /fan-zuoyong/(reaction)，反时针/fan-shizhen/(in reverse)，反弹道导弹 /fan-dandaodaodan/ (rebound).

    2. 超/chao/: 超音波/chao-yinbo/ (supersound)，超低温/chao-diwen/ (ultralow temperature)，超导体/chao-daoti/ (superconductor).

    3. 非/fei/: 非金属/fei-jinshu/ (nonmetal)，非偏振光/ei-pianzhenguang/ (non-polarization ray)，非电解质/fei-dianjiezhi/ (non electrolyte)

    4. 子/zi/: 子程序/zi-chengxu/ (sub-program)，子系统/zi-xitong/ (sub-system)，子公司 /zi-gongsi/ (sub-company).

    5. 相/xiang/: 相差/xiang-cha/ (discrepancy)，相等/xiang-deng/ (equality),相似/xiang-si/ (resemble).

Suffix:

    1. 性/xing/: 惯性/guan-xing/(inertia)，弹性/tan-xing/ (elasticity),酸性/suan-xing/(acidity)，碱性/jian-xing/(alkalescence), 腐蚀性/fushixing/(causticity)，挥发性/huifaxing/ (volatilization)，模糊性/mohuxing/(faintness)，概率性/gailúxing/(probability)，离散性/lisanxing/(disperse)，任意性/renyixing/ (randomicity).

    2. 度/du/: 速度/su-du/ (velocity)，，角度/jiaodu/ (point of view)，弧度/hudu/(radian)，梯度 /tidu/(grads)，密度/midu/(dencity)，溶解度/rongjiedu/(solubility)，密度/mi-du/(desity)，灵敏度 /lingmin-du/ (sensibility).

    3. 率/lǜ/: 几率/ji-lǜ/ (probability)，频率/pin-lǜ/(frequency)，功率/gong-lǜ/ (power), 折射率 /zheshelǜ/(refractive index).

    4. 化/hua/: 液化/ye-hua/(fluidify)，汽化/qi-hua/(boiloff)，氧化/yang-hua/ (oxidation)，风化 /feng-hua/ (airslake)，熔化/ronng-hua/(melt)，硬化/ying-hua/ (rigidification)，极化/ji-hua/ (polarization).

    5. 体/ti/: 气体/qi-ti/(gas)，液体/ye-ti/ (liquid)，固体/gu-ti/ (solid)，刚体/gang-ti/ (rigid body)，导体/dao-ti/ (conductor)，晶体/jin-ti/ (crystal)，天体/tian-ti/ (celestial body), 绝缘体 /jueyuanti/(insulator).

  Compounding formation: There is a correspondent relationship between the word formation in semantic loanword and the syntactic structure in sentence or phrase, so we can use the syntactic rules to translate the word in donor language to semantic loanword. The different compounding formations of semantic loanwords are as following:

-- coordinating structure: 信息/xinxi/ (information)，误差/wucha/ (error)，宇宙/yuzhou/ (cosmos).

-- endocentric structure: 硬件/yingjian/ (hardware)，软件/ruanjian/ (software), 算法/suanfa/ (algorithm)，参数/canshu// (parameter)，向量/xiangliang/ (vector).

-- verb-object structure: 作业/zuoye/ (job)，比重/bizhong/（specific gravity），积分/jifen/ (integral).

-- subject-predicate structure: 电流(electric current)，脉搏(pulse)，血沉(erythrocyle sedimentation rate, ERS).

-- verb-complement structure: 扩散(diffuse)，吸附(absorption)，稀释(dilution).

In addition to phonemic loanwords and semantic loanwords, there are semi-semantic & semi-phonemic loanwords. For example, 蒙特卡罗法/mengtekaluo fa/（Monte Carlo method）, first part "蒙特卡罗"/mengtekaluo/ is phonemic loanword, second part "法"/fa/ is semantic loanword, they combines to "蒙特卡罗法" /mengtekaluo fa/.

Other examples pf phonemic & semantic loanwords: 斐波那契搜索/feibonaqi sousuo/（Fibonacci search）, 布尔运算/bu'er yunsuan/（Boolean operation）, 伦琴射线/lunqin shexian/（Rontgen ray）, 本生灯/bensheng deng/（Bunsen burner）, 居里点//juli dian（Curie point）, 夫琅和费谱线/fulanhefei puxian/（Fraunhofer lines）, 伏打效应/fuda xiaoying/（Volta effect）, 伏打电池/fuda dianchi/（Voltaic cell）, 安培计/anpei ji/（Anpere meter）, 莱顿瓶/laidun ping/（Leyden jar）, 马赫数/mahe shu/（Mach number）, 巴黎绿/bali lù/（Paris green）, 普鲁士蓝/pulushi lan/（Prussian blue）, etc.

The majority of loanwords in Chinese are scientific and technical terms. How to extract the terms in the large-scale and authentic corpus? It is an important aspect for the automatic knowledge mining and the automatic knowledge extract. The structure of loanwords can help us to recognize the features of the terms and to harmonize the relation between the phonemic loanwords and semantic loanwords. It will be very valuable for the automatic knowledge mining and the automatic knowledge extract.

BIBLIOGRAPHY

1. 冯志伟，自然语言的计算机处理，1996 年。上海，上海外语教育出版社。
2. 冯志伟，现代术语学引论，1997 年，北京，语文出版社。
3. 冯志伟，应用语言学综论，1999 年，广州，广东教育出版社。
4. 冯志伟，术语浅说，2000 年，北京，语文出版社。