Compilation Techniques for Pedagogically Effective Bilingual Learners' Dictionaries

presented by
Jack Halpern, CEO
The CJK Dictionary Institute
日中韓辭典研究所

ASIALEX 2016 PHILIPPINES
The 10th International Conference of the Asian Association of Lexicography
June 1-3, 2016

Summary

- Background and aims
- Presentation of senses and logico-semantic ordering
- Semantic interrelatedness and core meanings
- Semantic transparency and morphological productivity
- Lexical categories in Chinese
- Arabic headword selection criteria
- Multiword expressions in Arabic
- Pedagogical selection criteria
- Interactive Parallel Text – a new paradigm
- Future Work

2

Shortcomings of existing dictionaries

- Archaic headwords and senses
- Inappropriate examples
- Overly prescriptive approach
- Learner-unfriendly sense ordering
- Omission of important multiword expressions

3

Pedagogically effective dictionaries by CJKI

- The CJKI Chinese Learner's Dictionary

- The Kodansha Kanji Learner's Dictionary

- The CJK Arabic-English Learner's Dictionary

4

Ordering of senses

- Chronological
- Frequency-based
- Logico-semantic

Chronological ordering

A Learner's Chinese Dictionary   *The Oxford Chinese Dictionary*

Frequency ordering

Entry for 本 *běn* in *A Chinese-English Dictionary: a Basic Dictionary for Chinese Language Learning* (Foreign Languages Press, 2007)

本 *běn*

1. <名 *n.*> booklet
2. (～儿) <名*n.*> capital; principal
3. <名*n.*> root or stem of a plant; generally referring to the origin of the matter
4. <名*n.*> version; edition and book binding
5. <量 *meas.*> used to indicate the copy of books
6. <代 *pron.*> one's own; native

Logico-semantic ordering

▲ bound morpheme ⬞ bound or free morpheme △ free morpheme

Interrelatedness of senses

List A

- 1 a KEEP in place, ask to stay
-   b KEEP for future use, reserve, leave behind
-   c KEEP in custody, detain
-   d KEEP one's mind on, concentrate on

List B

- 1 keep in place, ask to stay
- 2 reserve, leave behind
- 3 detain
- 4 concentrate on

9

Semantic transparency in Japanese

English	Etymology	Kanji	Roman	Meaning
hydrophobia	Greek	恐水病	*Kyōsuibyō*	fear+water+illness
aquarium	Latin	水族館	*Suizokukan*	water+family+building
waterwheel	Anglo-Saxon	水車	*Mizuguruma*	water+wheel

10

Transparency and productivity

11

POS in Chinese Dictionaries

- Lack of POS codes

- Limited POS codes

- Incomplete POS codes

12

POS codes for 留 in CCLD

❸ ⓐ (remain in a given condition) **STAY, remain** *VB*▲
　　ⓑ **STAY abroad to study**▲

a 大战后，她一直留在中国。 *Dàzhànhòu tā yìzhí liú zài Zhōngguó EH* She has remained in China since the war.
留级 *liújí VO* fail to advance in school grade
留学 *liúxué VO* study abroad
停留 *tíngliú VB* stay temporarily, stop over

b 留美 *liúMěi VO* study in the U.S.
留日 *liúRì VO* study in Japan

VO resultative verb　*VB* verb transitive & intransitive　*EH* phrase
▲ bound morpheme　▵ bound or free morpheme　△ free morpheme

13

Character Meanings

- The Equivalent

- Core Meaning

- Interrelatedness of Senses

14

CCLD Entry for 留

留
1646
■2-5-5

▶ KEEP ▶ STAY
liú
① 留　① 留

田	HSK-4	S10-5-5	GB3384
102	B0741	Ⓝ2580	U7559

❶ ⓐ [original meaning] **KEEP** in place, ask (someone) to **STAY** *VB*▲
　　ⓑ **KEEP for future use, reserve, leave behind**▲
　　ⓒ **KEEP in custody, detain**▲
　　ⓓ **KEEP one's mind on, concentrate on**▲

a 我们留她吃晚饭。 *Wǒmen liú tā chī wǎnfàn EH* We asked her to stay for dinner.
留步 *liúbù EA* don't bother to see me out (said by departing guests to host)
挽留 *wǎnliú VB* urge [persuade] someone to stay

b 留言 *liúyán VO* leave a message [comments]; *NC* short message
留念 *liúniàn VB* keep[accept] as a souvenir; *NC* souvenir
保留 *bǎoliú VB* reserve, retain; *NC* reservation

15

KKLD entry for 留

■5-5
2

玄
立
卯
宍
世
�

留
1646
■2-5-5

▶ KEEP
▶ STAY
リュウ ル と（める）-と（める）-ど（め）
と（まる）と（まる）▲ ⑧とめ

田	Jōyō-5	S10-5-5	K4617
102	B0741	Ⓝ2580	U7559

COMPOUNDS

❶ cause to remain in a given place or condition:
　ⓐ [original meaning] **KEEP in place, keep from moving, keep in position**
　ⓑ **KEEP in custody, detain**
　ⓒ **KEEP for future use, leave behind**
a 係留する けいりゅうする moor, anchor
慰留する いりゅうする dissuade from resigning

b 留置 りゅうち detention, custody, retention
抑留する よくりゅうする detain, intern, seize, arrest
c 留保する りゅうほする reserve, withhold, keep back
保留 ほりゅう reservation
遺留品 いりゅうひん article left behind, lost property
❷ ⓐ (remain in a given condition) **STAY, remain, continue**
　ⓑ (remain in a given place) **STAY, sojourn, reside**
a 留任する りゅうにんする remain [stay] in office
留年する りゅうねんする stay more than two years in the same class
残留する ざんりゅうする stay behind
停留所 ていりゅうじょ (bus) stop, station
留学生 りゅうがくせい student studying abroad
留守 るす absence (from home); caretaking; defending when the lord is absent
駐留する ちゅうりゅうする be stationed at,

17

Traditional criteria for selecting Arabic headwords

- Quran and classical literature
- Copying from past dictionaries
- Analyzing authentic texts found in corpora
- Subjective judgment of lexicographer
- 'Importance' based on frequency and lexicographer's discretion

18

Multiword expressions in Arabic

- Omission of multiword expressions
  (اَلشَّرْقُ ٱلْأَوْسَطُ 'assharqu l'awSaTu 'the Middle East')
- Omission of derived words
  (بِسُرْعَةٍ bisur'atin 'quickly')
- Inclusion of infrequent words
  (حَمْد Hamd 'praise')
- Exclusion of common words
  (مُسَامَحَة musaamaHa 'forgiveness')

19

Pedagogical selection criteria

- Potential headwords drawn from corpora, textbooks, learners' dictionaries, websites, and other sources
- Special efforts made to include modern words such as *blog* and *ATM*
- Rare/archaic words omitted by relying on modern corpora and sources
- Dialectal words omitted regardless of their corpus frequency
- MWEs systematically included as main entries or subentries rather than as examples
- Important derived words, especially adverbs, systematically included
- Headword/senses culled by lexicographers to maximize relevance

20

Future Work

- Lexical categories
- Headword selection
- Multiword expressions
- Pedegogically effective tools

21

Slide 22

Learner's Interactive Bitext
Electronic Reading Application

Libera

Slide 23

Interactive Parallel Text (IPT)

- Allows learners to effortlessly read vast amounts of L2 texts
- New medium for accessing comprehensible L2 texts
- Libera a cutting edge platform for reading enjoyably
- Texts precisely linked on segment level
- Interactive links highlight L2 segment and activate glosses, grammar notes, audio, and transcriptions

---

Slide 24

Four-Panel Sentence Mode

- Texts precisely linked at segment (word or phrase) level
- Tapping a segment in any one text simultaneously highlights linked segment in all other texts

Chapter **1** Sentence **3**▼   DEMO VERSION   ▷ + ᵃ↕

Japanese
カシムは金持ちの商人でし なぜなら彼は金持ちの商
人の娘と結婚し、彼と商売を
していたからです。

Idiomatic English
Qasim was a rich merchant because he had married the daughter of a rich merchant and he worked with him in trade.

Customized Dictionary
商売をする　しょうばい をする
phrase
do business, engage in business, deal trade

W Word info

Romanized Japanese
Kashimu wa kanemochi no shōnin deshita. Nazenara kare wa kanemochi no shōnin no musume to kekkon shi, kare to shōbai o shite ita kara desu.

◁　商売をしていた　　しょうばい を して いた　　he worked...in trade　▷

Segment Bar

Slide 25

Thank You
Salamat po

The CJK Dictionary Institute
Niiza-shi, Saitama, Japan

www.cjk.org