

The 3rd International Workshop on Advanced Learning Sciences

01/08/2015

Comparative Analysis of English, Japanese and Chinese Based on Cross-Linguistic Learner Corpora

Keiko Mochizuki, Ya-Ming Shen, Weilun Yu

Chinese data sources :

Tokyo University of Foreign Studies (TUFS)

National Taiwan Normal University (NTNU)

	TUFS	NTNU
Number of learners	165	Over 600
Mother tongue	Japanese	English
Learner data	College students in their 2nd, 3rd or 4th year	Students taking TOCFL*
Learner level (CEFR** equivalent)	A2-B2	A2-B2

* TOCFL: the Test Of Chinese as Foreign Language

** CEFR: The Common European Framework of Reference for Languages

Chinese data sources :

	TUFS	NTNU
Type of data	Composition as take-home task	Composition as test
Size of data	216,727 characters	237,475 characters
Number of compositions	471	691

Data processing:

1. Correct compositions
2. Add tags to instances of misuse
3. Convert files into XML format
4. Upload to online research platform



Case study:

Yi (一) + classifier usage in Chinese

Japanese Native Speakers: underuse of *yi+CL*

English Native Speakers: overuse of *yi+CL*

For example: *yi ge* (一個)

B1 and B2 level	number of Chinese characters	occurrence of <i>yi ge</i> (characters)
TUFS (JNS)	134,094	1,490
NTNU (ENS)	151,823	385

$X^2(1) = 150.03; p < .001$

Case study:

Yi (一) + classifier usage in Chinese

Japanese Native Speakers: underuse of yi+CL

- (1) a. 生孩子 [一件] 是不簡單的事。
- b. 京都是 [一個] 很美麗的地方。
- c. 東大和有 [一個] 很大的公園，附近也有 [一條] 小河。
- d. 最近他在車站附近開了 [一家] 中餐館。

Case study:

Yi (一) + classifier usage in Chinese

English Native Speakers: overuse of yi+CL

- (2) a. 有什麼問題就跟我打 **一通** 電話吧。
- b. 我們游完泳我計劃我們去電影院看 **一部** 電影。
- c. 你開 **一個** 慶祝會的時候，我不能參加.....
- d. 我在台北沒有發生 **一個** 大問題。

Case study:

Yi (一) + classifier usage in Chinese

bare nouns like cotton, 玻璃 → mass noun

Classifier functions as a(n)

individualizer (Okouchi 1985)

unitizer (Lucy 1992)

(3) a. *two cottons

b. two *balls* of cotton

(4) a. ??打破玻璃

b. 打破兩塊玻璃

Case study:

Yi (一) + *classifier* usage in Chinese

CL + noun is also called

counted noun (Langacker 1990)

bounded (noun) (Shen 1995)

atomic (Wilhem 2008)

Case study:

Yi (—) + classifier usage in Chinese

	Number Singular/ Plural	Classifier system	Degree of individualization
E	obligatory	few	High
C	not obligatory	rich	Moderate
J	not obligatory	not as rich as Chinese	Low

Mochizuki et. al (2015)

References

Langacker, R. W. 1990. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin/New York: Mouton de Gruyter.

Lucy, J. 1992. *Grammatical categories and cognition: A case study of linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

大河内康憲(1985)「量詞の個体化機能」『中国語学』232

沈家煊. (1995)「有界与無界」『中国語文』5

Talmy, L. 2000. *Toward a cognitive semantics. Vol. 1: Concept structuring systems*. Cambridge, Mass.: MIT Press.

Wilhelm A. 2008. Bare nouns and number in Dene Suline. *Natural language semantics* 16: 39-68.