Comparing Classifier use in Chinese and Japanese

Yue Hui Ting and **Francis Bond**
Division of Linguistics and Multilingual Studies
Nanyang Technological University
{htyue1@e.ntu.edu.sg, bond@ieee.org}

Abstract

Numeral classifiers present a challenge to successful machine translation. We investigate two numeral classifier languages: Mandarin Chinese and Japanese. This paper presents a quantitative analysis of classifier translations between these two languages to better understand differences in classifier usage.

Keywords – numeral classifier, sortal, translation, Mandarin Chinese, Japanese, contrastive linguistics

1 Introduction

Mandarin Chinese (CMN) and Japanese (JPN) are numeral classifier languages. Numeral classifier languages express the quantity of referents by modifying a noun phrase (NP) with an obligatory numeral-classifier construction where the classifier denotes inherent referent attributes (Bond and Paik, 2000; Downing, 1996). Hence, for a numeral-classifier construction that is assigned to a noun, the numeral denotes the numerical quantity of the noun referent while the numeral classifier denotes the quality of the noun referent.

Bond and Paik (2000) identified five main types of classifiers. **Event** classifiers classify events (Japanese: -kai 回 'time'; Mandarin Chinese: -cì 次 'time'). **Mensural** classifiers are employed for the measurement of physical properties (Japanese: -sun 寸 'inches'; Mandarin Chinese: -cùn 寸 'inches'). **Group** classifiers classify groupings of referents (Japanese: -kumi 組 'pair, set'; Mandarin Chinese: -shuāng 双 'pair'). **Taxonomic** classifiers effect a generic interpretation of the noun phrase (Japanese: -shu 種 'kind, type'; Mandarin Chinese: -zhǒng 种 'kind, type'). Finally, when quantifying the noun, **Sortal** classifiers classify the type of referent that is being counted, as in (1) and (2)[*].

(1) JPN: *pen   2-hon*
            pen   2-CL (long, cylindrical)
            "2 pens"
(2) CMN: *6- zhāng   piào*
            6-CL(flat, broad) tickets
            "6 tickets"

The numeral classifier system is organized differently for different languages. Mok et al.'s (2012) parallel studies focusing on generating sortal classifiers found that there are differences in classifier usage for the same semantic hierarchy of noun classes, suggesting differing conceptual organization between Mandarin Chinese and Japanese.
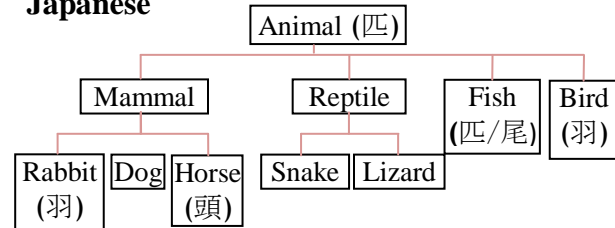
Chinese
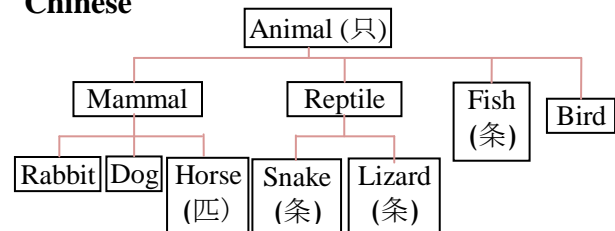
Japanese

Figure 1. Semantic Hierarchies in Chinese and Japanese

For example, the semantic hierarchies in Fig.1 show that there is not always one-to-one correspondence between the classifier characters of

---

[*] Abbr. used are CL for classifier, PTCL for particle, DET for determiner, num for numeral or numerative

these two languages. For example, although the same character 匹 exists in both classifier systems, it is used differently, as a general animal classifier in Japanese and as a specific classifier for horses in Mandarin Chinese. Japanese and Mandarin Chinese are an interesting pair of languages to compare for classifiers because they share to a limited extent the same Chinese character system and occasionally there are one-to-one correspondences for classifiers (e.g. 件 'case').

Because the classifier organization in both semantic hierarchies is different, the context in which a certain classifier is used may differ. Hence in Mandarin Chinese and Japanese, it is not the case that the same character for a classifier may be used as an equivalent translation in the same context. Classifiers have proven notoriously tricky to translate automatically with precision in various contexts.

Japanese and Mandarin Chinese classifiers also differ in terms of syntax (Ueda, 2009). Our approach to studying classifier use is to observe classifier phrases between hand-translated parallel sentences to search for predictable patterns in translation. There may also be lexical differences between the two numeral classifier systems. Understanding these differences offer us an insight into how the need for a classifier in certain semantic, grammatical, lexical or pragmatic contexts is negotiated within each language. Knowledge of classifier usage between Mandarin Chinese and Japanese will also be useful when considering crucial classifier features of each language to be addressed for classifier generation.

This paper will focus on sortal classifiers only and taps on the source data from Mok et al.'s (2012) parallel studies on Mandarin Chinese and Japanese classifiers.

The structure of the paper is as follows. In Section 2, we introduce the aims of our study. Section 3 presents a review of literature relevant to the grammar of both languages as well as their numeral classifier systems. In Section 4 we present our methodology and data for the pairwise comparison of sentences and our observations will be collated in Section 5 where we count and describe notable translations of categorized patterns. Section 6 discusses the implications of our findings and how they relate to existing literature. Finally we offer ideas for further research in our conclusion in Section 7.

2 Aims

This paper carries out pairwise comparison of parallel sentences to investigate the differences in sortal classifier usage between the two languages; Mandarin Chinese and Japanese. Based on our findings, we aim to come up with a better description of the use of classifiers in both Mandarin Chinese and Japanese.

3 Literature Review

In a numeral classifier phrase (consisting of the numeral, classifier, noun and the occasional particle), the numeral always occurs next to the classifier (Yamamoto, 2005, p. 5). The tighter constituent is hence composed of the classifier and numeral, as the noun constituent may occasionally occur distantly in cases of anaphora. Mok et al. (2012) listed classifier phrase combinations found from newspaper data. Combinations for Japanese include **num-CL-no-N** (where *no* 'of' is the adnomial particle), **N-PTCL-num-CL** (where PTCL can be case particles such as *ga, wo,* and *mo* which also appear in classifier phrases), and **N-num-CL**. For Mandarin Chinese, possible combinations are **DET-num-CL-N**, **DET-CL-N**, and **num-CL-N**.

There are several differences as to when classifiers can be omitted in Mandarin Chinese and Japanese.

One of these differences is the dropped or omitted numeral construction and non-numeral construction in Mandarin Chinese. The latter is an example of using a numeral classifier without a numeral in the classifier phrase. When a determiner precedes the classifier phrase, it gives rise to a *DET-CL* construction (Yamamoto, 2005, p. 6), (3).

(3)   CMN: *na   zhang zhi*
             that CL   paper
             "that piece of paper"

A dropped numeral construction occurs when the noun in question may be quantified as a single item, in which some cases the numeral *one* is dropped (Yamamoto, 2005, p. 23) from the usual indefinite use construction *1-CL-N*. This construction functions almost like an indefinite determiner when a verb precedes the numeral and classifier combination instead (4). It is not certain if the

dropping of the numeral *one* follows certain syntactic rules or if it simply serves as a shortening of the complete indefinite phrase.

(4)   CMN: *zhao zhang zhaopian*
            snap CL   photo
            "snap(or take) a photo"

Li and Thompson (1989) also describe numeral omission in Mandarin Chinese in determiner and numerative containing classifier constructions (Li and Thompson, 1989, p. 104).

Another difference between the numeral classifier systems of these two languages is the number of types of classifiers that exist in the system. To illustrate this, there is a phenomenon of "semantic split" (Hansen and Chen, 2001, p. 89) in classifier categories for Japanese where a group of nouns classified by a single classifier may be divided into smaller groups which are each classified by a different classifier in Mandarin Chinese, suggesting that nouns are classified in Japanese by a smaller number of classifiers.

Yin et al. (2006) came up with rules to translate classifiers from Mandarin Chinese to Japanese. These rules addressed the indefinite determiner and numerative classifier phrase in addition to the usual numeral-classifier phrase. However they did not seem to have addressed dropped numeral or demonstrative constructions.

4   Methodology

4.1   Data

The data for pairwise comparison were annotated sentences and classifiers done by Mok et al. (2012). These sentences were taken from the NICT Multilingual Corpus which is a Japanese-Chinese-English parallel corpus based on the Mainichi Newspaper (Zhang et al., 2005). 38,000 Japanese sentences from the Mainichi Shinbun (1984) have been translated into both Chinese and English by professional translators. Only 500 sentences were considered for analysis for this paper. The newspaper domain is a formal domain and the more formal the style of writing, the more variation and occurrence of classifiers the writing style exhibits, providing a rich pool of classifiers to work with (Craig, 1986, p. 8). Parallel sentences were compared with the help of equivalent English translations and the differences in classifier use in the sentences were analyzed.

A preliminary run-through of the data was done by hand on the first 100 parallel sentences to identify interesting and recurring observations and to classify them with a name (or tag). This would serve to make classification of observed patterns easier later. A program generated the sentence id and extracted parallel sentences, the English equivalent, as well as classifier information. For example, in a sentence without a classifier, *(N)* is generated to indicate that there was no classifier. Where there was a classifier, the character for the classifier was generated, such as *(回)*.

In the preliminary study, we noticed a few problems with the automatic tagging. Occasionally we had target NP mismatches where the classifier phrase in a sentence did not match any target NP phrase recognized by the program. Also, where one sentence had 2 classifier NPs and the other had only one, if the first classifier NP pair that was a correct match was not sortal, the next classifier was selected as a parallel match for the classifier in the sentence with only one classifier. This sometimes resulted in blatant errors. Additionally, we realized from our initial counts that the program did not consider the Japanese つ and Mandarin Chinese 人 classifiers in its tagging and hence missed out on those. These errors were later corrected.

Where both sentences have an equivalent classifier, they were considered aligned. In many cases, a classifier was present in only one language. We expected that the classifier would be more frequently absent from the Japanese sentence. The rationale for this expectation is Mandarin Chinese has more types of classifiers in its classifier system than Japanese (as addressed in Section 3). Also, Mandarin Chinese uses classifiers in one common construction that Japanese does not; the DET-CL-N combination. The preliminary observations identified a few categories (explained below); *non-classifier equivalents, omission of classifier, demonstrative, indefinite use,* and *aligned.*

We did not attempt to identify differences in classifier usage due to translator choice or judgment as the decisions of the translators are sometimes ambiguous and hence beyond the scope

of what we can hope to discuss extensively and satisfactorily.

4.2　Hand-annotation of sentence pairs

For the actual data analysis, a set of 243 sentence pairs was used. These were sentences in the original set of 500 that had a part-of-speech tagged numeral classifier in at least one sentence. (This means that there were no sentence pairs in which both sentences had no classifiers.) The parallel sentences were run through a program which generated the sentence id, the Japanese sentence, the parallel Chinese sentence, the English translation as well as additional information about the classifiers, (5).

(5)
* 95010108001 *
お正月 が 来る と 、 思い出す こと が あ
る 。
毎逢 新年 来临 , 我 就 会 想起 一 件_9
事 。
When a New Year comes, I remember one thing.
95010108001　N:-1　件:9　(N:sortal)

This program detects the presence of a classifier or classifiers and annotates to indicate the absence of a classifier or if otherwise, the classifier itself, as well as the word id which is the numerical position of the classifier in the sentence. Also annotated is the type of classifier in each sentence, whether it is *sortal*, *mensural,* or simply a non-classifier function; tagged as *not*. To compensate for any mistakes that might have been made in the automatic process as well as to enrich the information with the earlier identified tags, these 243 sentence pairs were hand-annotated to correct where needed, the automatically identified classifiers as well as the type of classifiers. In addition, the tags were added onto (5) to indicate if classifiers aligned or if it was a specific phenomenon if the classifier was found in one sentence only.

The tags used for the subsequent hand-annotation in the actual data analysis are as follows: *aligned, non-classifier equivalents* (jpn only)*, indefinite use* (cmn only)*, indefinite use no numeral* (cmn only)*, demonstrative* (cmn only)*,* and *omission* (jpn only)*.*

(a) Classifier present in one language only

Non-classifier equivalents:
e.g.　　JPN: ある　　(N)
　　　　　　a certain
　　　　CMN: 一位　　(位)
　　　　　　1-CL
Non-classifier equivalents in JPN do not employ the use of classifiers. In other words, these fixed, expressions convey roughly the same meaning without needing a classifier.

Omission of classifier:
e.g.　　JPN: 十五 の　　訓練所　　(N)
　　　　　　15　PTCL　training centre
　　　　CMN: 十五 个　训练所　　(个)
　　　　　　15　CL　training centre
Omission of classifier in JPN, with presence of の.
　　　　JPN: 二　　億　　缶　(N)
　　　　　　2 hundred million　can
　　　　CMN: 2　　亿多　　个　(个)
　　　　　　2 hundred million　CL
Omission of classifier in JPN due to a large, round number.

Demonstrative:
e.g.　　JPN: その 珊瑚　　(N)
　　　　　　that　coral
　　　　CMN: 那 串 珊瑚　　(串)
　　　　　　That　CL　coral
A demonstrative (this/that) alone suffices for reference in Japanese while in Mandarin Chinese a classifier is needed.

Indefinite use:
e.g.　　JPN:「X」 という　項目　　(N)
　　　　　　such a　question
　　　　CMN: "X" 一 项　提问　(项)
　　　　　　1　CL　question
(Where *X* represents a question.) The indefinite use of a classifier phrase includes the equivalent of the English 'a' used in Mandarin Chinese to introduce indefinite NPs.

Indefinite use no numeral:
e.g.　　JPN: 野蛮 人 に　　見えた (N)
　　　　　　wild person PTCL seen to be
　　　　CMN: 像 个 野 人　　　(个)
　　　　　　Like CL　wild-person

A variant of the above mentioned indefinite use where the numeral in the *1-CL* construction is dropped.

(b) Classifier present in both CMN and JPN

Aligned:
e.g.　JPN: 到着客　　約　　　百五十人　(人)
　　　　　　　passengers approx. 150　　CL
　　　CMN: 大约　一百五十　名　抵达旅客(名)
　　　　　　　a*pprox 150　　CL　passengers*
Equivalent classifiers exist in both languages.

　　　The above list of tags was refined in consideration of observations during the annotation process. The annotation was also revised where it was deemed needed due to revelations in the annotation process.

5　Results

Table 1. Automatic Classification (non-sortal included)

Scenario	No. of instances*
Classifier in JPN & CMN	101
Classifier in CMN only	177
Classifier in JPN only	29
Total	307

*Counts represent classifier comparisons, not sentences.

　　　Based on the counts in Table 1 above, we have found that numeral classifiers appear much more frequently in Mandarin Chinese only than in Japanese only. Looking at counts in Table 2 in the next column, most of these cases come from the use of demonstratives and indefinite use in Mandarin Chinese.

　　　The discrepancy between the 101 count for classifier in both languages in Table 1 and the 51 count for align in Table 2 is mostly due to alignment of non-sortal classifiers, most of which involve ordinal expressions (6) which formed an overwhelming proportion, and classifier characters not functioning as classifiers.

(6)　JPN: 第　　四百　回　定期 (回)
　　　　　ORD　400　CL　　　　season
　　　　　"the 400th season"
　　　CMN: 第　　四百　场 定期　(场)
　　　　　ORD　400　CL season
　　　　　"the 400th season"

Table 2. Manual Classification

Tag	No. of instances*	
Aligned	(CMN & JPN)	**51**
Non-classifier equivalents	(CMN & JPN)	**17**
Indefinite use	(CMN only)	**37**
Numeral present		30
Numeral absent		7
Demonstrative	(CMN only)	**17**
Numeral present		4
Numeral absent		13
Omission	(CMN only)	**22**
Other (*non-sortal* and *not*)	(CMN & JPN)	**156**
Total		**300**

*Counts represent classifier comparisons, not sentences.

5.1　Aligned (51)

　　　Cases of alignment were the most frequent. The bulk of the classifier alignment cases were for specific classifiers. Another sizeable portion were for sentences that involved the person classifier *hito* (人) in Japanese, which was translated to one of three person classifiers in Mandarin Chinese: *rén* (人)*, míng* (名) and *wèi* (位), which differ in terms of formality and pragmatic importance of the status of the people in question. The Mandarin Chinese general classifier *gè* (个) was used in translation for the Japanese classifier *tsu* (つ) (general inanimacy classifier), and some specific classifiers. In addition, *tsu* was on occasion translated to more specific classifiers in Mandarin Chinese.

5.2　Non-classifier equivalents (17)

　　　In most cases of non-classifier equivalents, the Japanese sentence employed an expression that did not contain a classifier but whose translated equivalent required a classifier. Consistent observations were in the counting of months and countries where the Japanese expressions following a number are *ka-getsu (ヶ月)* and *ka-kkoku (各国)* respectively and these may be known as fused classifier nouns. Hence, these count nouns are directly modified by the numeral. More interesting expressions were *ikutsuka no* 'a few of', where the classifier *tsu* is included in the lexical item *ikutsu,* 'a few' and *to iu*, which is an expression concluding a description that corresponds to an indefinite determiner classifier phrase when translated to

Mandarin Chinese, as well as noun and verb non-classifier equivalents.

e.g.　**Noun non-classifier expression**
　　　JPN: 片手　　　　　(N)
　　　　katate
　　　　"single-hand"
　　　CMN: 一 只 手　　　(只)
　　　　1 CL shǒu
　　　　"one hand"
　　Verb non-classifier expression
　　　JPN: ボーッとして　(N)
　　　　boottoshite
　　　　"to be in a daze"
　　　CMN: 一 片 空白　　(片)
　　　　1 CL kòngbái
　　　　"a sheet of blankness"

5.3　Indefinite Use (37)

Indefinite use of classifier phrases in Mandarin Chinese was common; *1-CL-N*, where no determiners precede the *1-CL-N* construction and where no expression that renders definiteness on the noun precedes the construction as well (e.g. *zuì hòu* (最后) 'final/last'). The equivalent Japanese sentences did not employ the use of classifiers or numerals. The English translations involved indefinite expressions, such as involving "a" or "an". The preceding environment of such Mandarin Chinese phrases were mostly verbs (with *shì* (是) 'is/be' coming up repeatedly), and some few cases were the spatial preposition *nèi* (内) 'within'. Dropped numerals were observed in this category under *indefinite use no numeral* where the preceding environment is a verb but the construction is simply *V-CL* where there is no numeral. *CL-N* classifier phrases with no preceding determiner were always judged to have a singular interpretation; that the numeral is *one* and can be omitted.

There was one exception where the sortal *1-CL-N* classifier construction as defined in this sub-section was translated in English to a definite expression involving the determiner "one".

e.g.　JPN: 思い出すこと　が　　ある (N)
　　　　Recall　　matter PTCL　exist
　　　CMN: 想起 一件 事　　　　　(件)
　　　　Recall 1 CL matter
　　　ENG: I remember one thing

This was the only relevant example that involved the sortal use of classifiers and where the English translation was faithful to the CMN expression. An example of a 'non-faithful' translation was where the CMN expression was *V-1-CL-N* (*there was-1-CL-television*) but was translated as *possessive-N* (*their television*).

5.4　Demonstrative (17)

Not all demonstrative classifier constructions omit numerals. The construction *DET-num-CL-N* was present for both the numeral *one* (*DET-1-CL-N*) and *two* (*DET-2-CL-N*) and was unlikely to be limited to just those numbers. The majority of the demonstrative classifier constructions (13 out of 17) omitted the numeral and the nouns in these expressions were interpreted as singular. The use of determiners *zhè* (这) 'this', *nà* (那) 'that', *cǐ* (此) 'this', and *gāi* (该) 'this, that' before the classifiers, as well as the lack of numerals seem to point to an interpretation of singularity.

5.5　Omission (22)

Straightforward cases of classifier omission occurred in Japanese where it seemed possible for a classifier to be present but it was not. The Mandarin Chinese translation however still required a classifier. In this case the numeral directly modifies a count noun. Some of these cases occurred when the Japanese numeral was a large, round number such as *800* or *50*. However in most cases the numeral was under ten.

e.g.　JPN: 四 都市　　　　　(N)
　　　　4 toshi
　　　　"four cities"
　　　CMN: 四 个 城市　　(个)
　　　　4 CL chéngshì
　　　　"four cities"

5.6　Other (156)

These are made up of non-sortal classifiers such as *event, mensural, group,* and ordinal expression classifiers (7) with or without aligned classifiers, as well as classifier characters appearing in non-classifier uses (7) and hence not analyzed.

(7)   JPN: 七 番 勝負   （番）
            7 CL match
            "seven-game match"
      CMN: 七 盘 比赛   （盘）
            7 CL match
            "seven-game match"

6 Discussion and Future Work

In cases of classifier alignment, the earlier mentioned phenomenon of "semantic split" (Greenberg, 1990, p. 89) observed in primary research with speakers is observed here. This is manifested when a classifier character that appears twice in the same Japanese sentence is translated to different classifier characters in Mandarin Chinese (8), suggesting the existence of more specific classifier categories in Mandarin Chinese.

(8)

JPN: はがき 約   二千 通...   （通）
        Postcards approx. 2000 CL
        ...郵便物 の   約   千 通 （通）
        Mail   PTCL approx 1000 CL

CMN: 两千 枚 贺年片...   （枚）
        2000 CL   new year postcards
        ...一千 封 普通 邮件   （封）
        1000 CL mail

Also, it seems that there are plenty of Japanese non-classifier noun and verb equivalents corresponding to classifier-including expressions in Mandarin Chinese, doing away with the need for a classifier phrase, further reducing the frequency of classifiers appearing in Japanese.

With regards to omission, the newspaper is a formal and impersonal domain and the omission of classifiers in Japanese seems to reflect this as it seems characteristic to drop classifiers in impersonal presentations of quantity, resulting in the construction *num-N*. This however does not occur in our Mandarin Chinese data. Also, if the characteristic of the hand-translation process is that translators tend to translate into a less rigid form of language, it might explain why there are many cases of *num-N* in Japanese being translated to a longer and more natural expression in Mandarin Chinese. If however, both the newspaper domain

and translator behaviour are not the reasons for such an observation, it is possible that Japanese is moving towards allowing counting with no classifiers (compare Align 51 and Omission 22) and is getting a small class of fully countable nouns such as *shou* 勝 'victory' and *hai* 敗 'loss' which can be counted simply by having a numeral precede it, 三 勝 and 三敗 (*san* 三 '3').

For demonstrative classifier constructions in Mandarin Chinese, if the numeral is a number other than *one*, it logically cannot be omitted. It is also possible that wherever a noun is referred to, its classifier must come up as well though not performing a numeral classifier function but simply a noun classifier function instead.

For indefinite use in Mandarin Chinese however, it is unclear from our findings if there are rules governing the dropping of the numeral *one*. In most cases it is not dropped. Pragmatic choices or phonological reduction may solely be at play here (Chen, 2003, p. 1171).

Based on our findings for demonstratives and indefinite use, where the numeral is omitted, it seems that Mandarin Chinese uses classifiers in phrases that appear to function like determiners, basically showing information structure by indicating whether a piece of information is old (by using a demonstrative) or whether it is new (by indefinite use with a classifier). Chen (2003) offers an interesting discussion on the indefinite use of classifiers in Mandarin Chinese and mentions a "presentative use" of the indefinite article in the *yi 'one'-CL-N* construction that may be used for new and stressed information (Chen, 2003, p. 1171) but also talks about a tendency towards non-referential use when the numeral *yi* is omitted. It is also possible to divide the indefinite use respectively into (i) numeral use and (ii) the English equivalent 'a' according to whether *yi* is stressed or unstressed (Rullmann and You, 2006), giving rise to implications for presenting new and old information. Two further questions we would like to answer by comparing the Japanese and Chinese to English are: (i) Are the indefinite uses always translated with an indefinite article? And (ii) Are the demonstratives always translated as demonstratives or also with the definite article?

Greenberg (1990, p. 253) proposes that the demonstrative is the most common starting point of the development of a definite article (known as

Stage I). Further development then sees it offering both definite and indefinite uses (Stage II). Our findings on how demonstratives and indefinite use in classifier phrases act as determiners seem to suggest that Mandarin Chinese is in the process of evolving articles.

Finally, for future research, translation comparison for less impersonal domains (e.g. editorials) might shed light on whether certain classifier usage differences may be due to pragmatic factors. With regards to cross-linguistic interests, the NTU multilingual Corpus (Tan and Bond, 2011) contains more corpora linked to other classifier languages such as Thai, Vietnamese, Indonesian and Korean. These resources may be exploited in future studies to observe classifier usage patterns and a comparison may be done later between the studied languages to determine if similar (or dissimilar) phenomena and patterns exist.

7   Conclusion

In this paper, we identified categories of classifier translations from Japanese to Mandarin Chinese and looked at notable translations that have implications for understanding lexical, syntactic and pragmatic differences. The analysis of classifier translations reveals that it will be tricky to translate non-classifier expressions from Japanese to classifier-including expressions in Mandarin Chinese, posing this as a noteworthy problem to overcome.

References

Bond, F., & Paik, K. (2000). Re-using an ontology to generate numeral classifiers. *Proceeding s of the 19th International Conference on Computational Linguistics: COLING-2000:* Saarbrucken.

Chen, P. (2003). Indefinite determiner introducing definite referent: a special use of '*yi* 'one' + classifier' in Chinese. *Lingua, 113,* 1169-1184.

Craig, C.G. (1986). *Noun Classes and Categorization.* Philadelphia: John Benjamins.

Downing, P. (1996). *Numeral classifier systems: The case of Japanese.* Philadelphia: John Benjamins.

Greenberg, H. (1990). How Does a Language Acquire Gender Markers? In K. Denning and S. Kemmer (Eds.) *On Language: Selected Writings of Joseph H. Greenberg*, pp. 241-270. Stanford: Stanford University Press

Hansen, L., & Chen, Y.L. (2001). What counts in the acquisition and attrition of numeral classifiers? *JALT Journal, 23*(1), 83-100.

Li, C.N., & Thompson, S.A. (1989). *Mandarin Chinese: A Functional References Grammar.* London, England: University of California Press.

Mok, H.S.W., Gao, H.N.E., & Bond, F. (2012) Using WordNet to predict numeral classifiers in Chinese and Japanese. In *Proceedings of the 6th Global WordNet Conference (GWC 2012)*:Matsue.

Rullmann, H., & You, A.L. (2006). General Number and the Semantics and Pragmatics of Indefinite Bare Nouns in Mandarin Chinese. In K. von Heusinger and K. P. Turner (Eds.) *Where Semantics Meets Pragmatics*, pp. 175-196. Amsterdam: Elsevier.

Tan, L., & Bond, F. (2011). Building and annotating the linguistically diverse NTU-MC (NTU-multilingual corpus). In *Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation (PACLIC 25),* pp. 367–376. Singapore.

Ueda, Y. (2009). Number in Japanese and Chinese. *Nanzan Linguistics, 5,* 105-130.

Yamamoto, K. (2005). The Acquisition of Numeral Classifiers: The Case of Japanese Children. In P. Jordens (Ed.), *Studies on Language Acquisition* (Vol. 27). Berlin, Germany: Walter de Gruyter.

Yin, D.P., Shao, M., Jiang, P.L., Ren, F., & Kuroiwa, S. (2006). Proceedings from ICCOMP-2006: *Rule-based translation of quantifiers for Chinese-Japanese machine translation*, pp. 558-563. Athens, Greece.

Zhang, Y., Uchimoto, K., Ma, Q., & Isahara, H. (2005). Proceedings from MTS-2010: *Building an Annotated Japanese-Chinese Parallel Corpus – A Part of NICT Multilingual Corpora*, pp. 71-78. Phuket.