Compositional Mechanisms of Japanese Numeral Classifiers

Miho Mano
Naruto University of Education
748 Nakashima, Takashima, Naruto-cho
Naruto-shi, Tokushima, JAPAN
`mmano@naruto-u.ac.jp`

Abstract

This paper suggests that Generative Lexicon Theory (Pustejovsky, 1995, 2006, 2011) offers a new analysis of numeral classifiers, focusing on Japanese having various kinds of classifiers. It is often said that classifiers agree with quantified nouns, that is, the nouns have to match the semantic requirements of the classifiers. This paper examines their lexical structures and compositional mechanisms. Though Huang and Ahrens (2003) explain the compositional mechanisms between the classifiers and the quantified nouns using "coercion" instead of the agreement, this paper indicates that other mechanisms including Type Matching (Pustejovsky, 2011) also occur in Japanese depending on the type required by the classifier and the source type of the quantified noun, following Mano and Yonezawa's (to appear) suggestion.

1 Introduction

Japanese has various counters, called *josuushi* in Japanese, including so-called "numeral classifiers (Aikhenvald, 2000)," as well as other East Asian languages such as Chinese, Indonesian, Korean, and Thai. In Japanese, nouns cannot be directly modified by numerals but must be quantified by counters, as shown in (1).

(1) a. *ni-{inu/kuruma}   (Japanese)
    2-dog/car
    'two dogs/cars'

b. ni-hiki-no inu/ ni-dai-no kuruma
   2-CL-GEN dog/ 2-CL-GEN car[1]
   'two dogs/two cars'

The counters are morphemes used together with numerals, and each of them has semantic restrictions on its objects. For example, a classifier -匹 *hiki* in (1b) requires its objects to be nonhuman animals, and -台 *dai* mainly selects for machines[2].

Most of the previous studies on classifiers assume that the modified nouns agree with the classifiers, because the classifiers can only count nouns which have particular meanings. Many studies have been done on the semantic restrictions of numeral classifiers (concerning Japanese, see Matsumoto (1991, 1993); Downing (1996), Iida (1999), and Nishimitsu and Mizuguchi (2004).).

Huang and Ahrens (2003), however, suggest that the classifiers do not simply agree with the quantified nouns but coerce particular meanings to them, focusing on the numeral classifiers of Mandarin Chinese. This paper shares the view that classifiers can coerce the nouns to refer to particular types, but as Mano and Yonezawa (to appear) point out, it seems that they may agree with the nouns without changing their source types.

There are some contradictory examples in Japanese, however. Taking *chuusha* 'injection,' for example, which is a polysemous noun that means a

---

[1] The abbreviations used in this paper are as follows: ACC= accusative case, CL=classifier, GEN=genitive case, NOM= nominative case, phys=physical object, PRES=Present tense, PROG=progressive, PST=past tense, TOP=topic marker
[2] See Matsumoto (1991, 1993) and Iida (1999) for more detailed restrictions of *-dai*.

physical object 'syringe' and also an event 'injection.' It can be counted by the classifiers, -本 *hon* and -回 *kai*. *-Hon* requires one dimensional (i.e. long and thin) physical objects (*phys*), and *-kai* is a classifier for events. Given that classifiers coerce the quantified nouns to be required types, *ni-hon-no chuusha* (2-CL-GEN injection) should be of type *phys* meaning 'two syringes,' while *ni-kai-no chuusha* should mean an event 'having injection(s) two times.' But that is not the case. The verb, *owaru* 'end,' is the predicate that selects for *event* as its complement, so it is predicted that only *ni-kai-no chuusha* is allowed. In fact, *-hon* is also allowed as in (2), however, contrary to the coercion analysis.

(2) ni-{hon/kai}-no chuusha-ga owat-ta.
    2-CL/CL-GEN injection-NOM end-PST
    '(I) had {two injections/injection(s) two times}.'

With regard to this issue, we examine the lexical structures of Japanese numeral classifiers and the compositional mechanisms more closely. The lexical structures of classifiers are examined in Section 2, and the compositional mechanisms are demonstrated in Section 3. Section 4 shows a conclusion and further issues.

2 Classifiers and their Lexical Structures

According to Iida (1999), there are about 360 counters in Japanese, and Kageyama et al. (2011) divide them into "numeral classifiers" and "measure specifiers" depending on their functions, which will be shown in 2.1. This paper focuses on only numeral classifiers. Their lexical structures will be examined in 2.2 and 2.3.

2.1 Classifiers and Measure Specifiers

Some categorizations of Japanese counters have been proposed (cf. Matsumoto, 1991, 1993; Downing, 1996; Iida, 1999; Nishimitsu and Mizuguchi, 2004). This paper adopts Kageyama et al.'s (2011) categorization, which divides them into "numeral classifiers" and "measure specifiers" according to their functions[3]. Numeral classifiers (classifiers, henceforth) classify and count limited and specific groups of nouns, which means their

function is "categorization (cf. Bisang, 1993)" of objects. On the other hand, measure specifiers can be used as measures for a wide variety of nouns as in (4b), and their function is considered to be "individuation (cf. Bisang, 1993)" of objects. Some examples are shown in (3-4) (the simplified semantic restrictions of each classifier are in round brackets[4].).

(3) a. classifiers: -回 *kai* (events),-人 *nin* (human), -匹 *hiki* (animals), -個 *ko* (3D phys), -枚 *mai* (2D phys), -串 *kushi* (skewered foods), -台 *dai* (machines), -機 *ki* (planes), -基 *ki* (placed artifacts)

b. measure specifiers: -束 *taba* 'bundle,' -杯 *hai* 'cup,' -箱 *hako* 'box,' -切れ *kire* 'slice,' -キロ *kiro* 'kilogram'

(4) a. ni-hiki-no {ikita okiami/*himono/*su}
    2-CL-GEN living.krill/dried.fish/vinegar
    'two {living krills/dried fish/water}'

b. ni-{kiro/hai}-no {ikita okiami/himono/su}
    2-kilogram/cup-GEN
    'two kilograms/cups of {living krill/dried fish/vinegar}'

This paper focuses on classifiers because they have more semantic restrictions on the quantified nouns than measure specifiers, which enables us to examine their compositional mechanisms more clearly.

2.2 Previous studies on the Lexical Structures of Classifiers

Only a few Generative Lexicon approaches have so far been attempted on classifiers (cf. Bond and Paik, 1997; Huang and Ahrens, 2003; Kageyama et al., 2011; Mano and Yonezawa, to appear), and there seems to be still room for argument.

Bond and Paik (1997) propose a basic lexical structure for Japanese sortal classifiers, assuming that the Formal qualia are allowed to take at least two values: a sortal typing of the argument and a feature of dimensionality. (5) is a lexical structure for -個 *ko* (3D phys). There are two variables in the argument structure: one is a numeral+ (which includes numerals, quantifiers, and interrogatives), and another is a quantified noun. The latter is a

---

[3] Bisang (1993) suggests four functions of classifiers: individuation, categorization, referentializaion, and relationalization.

[4] See Matsumoto (1991, 1993), Downing (1996), and Iida (1999), for more information on the semantic restrictions.

default argument, because it is not necessarily expressed overtly in Japanese, as in (6).

(5) -*ko* "3D" $\begin{bmatrix} \text{ARGSTR} \begin{bmatrix} \text{ARG1} & \text{x: numeral+} \\ \text{D-ARG1} & \text{y: inanimate} \\ & \text{DIMEN 3D} \end{bmatrix} \\ \text{QUALIA [FORMAL quantifies (x, y)]} \end{bmatrix}$
(Bond and Paik, 1997)

(6) san-ko(-no   hako-ga)   aru.
    3-CL(-GEN   box-NOM)   be.PRES
    'There are three (boxes).'

Huang and Ahrens (2003) and Kageyama et al. (2011) develop the idea further and show that classifiers may have some requirements also in the Constitutive, Telic, and Agentive roles in addition to the Formal role. (7) shows some examples of classifiers that have semantic requirements in the qualia structures of their objects pointed out by Kageyama et al. (2011).

(7) a.   Formal role: -人 *nin* (humans)*, -*匹 *hiki* (animals), -本 *hon* (1D phys), -枚 *mai* (2D phys), -個 *ko* (3D phys)

  b.   Constitutive role: -戸 *ko* (residences)*, -*串 *kushi* (skewered foods)*, -*体 *tai* (bodies)

  c.   Telic role: -機 *ki* (planes to fly)*, -*着 *chaku* (clothing to wear), -軒 *ken* (buildings to live in)

  d.   Agentive role: -揃え *soroe* (coordinated ones), -基 *ki* (placed large artifacts)

2.3   Semantic Restrictions by Classifiers

Following these studies, Mano and Yonezawa (to appear) propose more detailed lexical structures for Japanese classifiers. The basic lexical structure of classifiers suggested by them is given in (8). Each classifier is considered to have specifications on the quantified noun, i.e. D-ARG1.

(8) $\begin{bmatrix} \text{CL -}\alpha \\ \text{ARGSTR} = \begin{bmatrix} \text{ARG1 = x: quantity}^5 \\ \text{D-ARG1 = y: entity} \end{bmatrix} \\ \text{QUALIA} = \text{[FORMAL = quantify (x, y)]} \end{bmatrix}$
(Mano and Yonezawa, to appear)

They assume that the most basic type of classifiers is the one that specifies the Formal role of their objects, i.e. (7a). For example, (9) shows the lexical structures of the classifiers for animates, and -*kai*, a classifier to count events, is also considered to have a similar structure, as in (10) (It should be noted that only the relevant parts of lexical structures are shown in this paper.). We agree with them with regard to this type.

(9) a. -人 *nin*   D-ARG1=y: human   (ibid.)
                  [FORMAL=human (y)]

  b. -匹 *hiki*   D-ARG1=y: animal
                  [FORMAL=animal (y)]

  c. -羽 *wa*   D-ARG1=y: bird
                  [FORMAL=bird (y)]

(10) -回 *kai* D-ARG1=y: event   (ibid.)
              [FORMAL = event]

Next, we will review other types, (7b-d), which have semantic requirements on the roles other than the Formal role. The lexical structure for -串 *kushi*, a classifier to count skewered foods, is shown in (11). It requires skewers to be included in the Constitutive role of the objects. They also point out that the Formal, Agentive, and Telic roles are also specified because -*kushi* counts foods only as shown in (12).

(11) -串 *kushi* ARGSTR=D-ARG1=y: food
$\begin{bmatrix} \text{QL= FORMAL=food (y)} \\ \text{CONST=consist\_of (y, \{\underline{skewer}\ldots\})} \\ \text{TELIC=eat (e, z, y)} \\ \text{AGENT=skewer (e, w, y)} \end{bmatrix}$
(ibid.)

(12) san-kushi-no   {sate/*kanzashi/*nendo}
    3-CL-GEN   satay/hair.stick/clay
    'three sticks of satay/hair stick/clay'

For -機 *ki,* a classifier to count planes focusing on the Telic role, they also assume multiple specifications as in (13).

(13) -機 *ki*   ARGSTR=D-ARG1=y: machine
$\begin{bmatrix} \text{QL= FORMAL=machine (y)} \\ \text{TELIC=fly (e, y)} \\ \text{AGENT=make (e, z, y)} \end{bmatrix}$
(ibid.)

(14) san-ki-no   {hikooki/*kami-hikooki/*tori}
    3-CL-GEN   plane/paper-plane/bird

---

[5] This is considered to be equivalent to "numeral+" in Bond and Paik (1997).

'three planes/paper planes/birds'

There are two problems with their analysis, however. One is that the specifications are sometimes redundant (e.g. *artifacts* must have particular Agentive and/or Telic roles), though it might be true that they have multiple specifications. Another is that they ignore the importance of the Formal role. It should be noted that all Japanese classifiers have particular type requirements on the Formal role of their objects (the animacy is strictly restricted in Japanese classifiers[6]). In addition to it, note that the Formal role is considered a head type, and the additional qualia values can be seen as structural complementation to it in Pustejovsky (2011:1409).

Following Pustejovsky (2011), we suggest the following simplified representations in (15-16) for the lexical requirements of classifiers in order to solve the problems above. These representations are consistent with the characteristics of Japanese classifiers: though some classifiers have multiple requirements, they usually focus on "one" role in addition to the Formal role (cf. (7)).

(15) a. -人 *nin* is of type *human*→t   cf.(9)
　　b. -匹 *hiki* is of type *animal*→t
　　c. -羽 *wa* is of type *bird*→t
　　d. -回 *kai* is of type *event*→t   (10)
(16) a. -串 *kushi* is of type *food*$\otimes_C$ *skewer*→t (11)
　　b. -機 *ki* is of type *machine*$\otimes_T$ *fly*→t   (13)
　　c. -基 *ki* is of type *artifact*$\otimes_A$ *place*→t

3　Compositional Mechanisms

3.1　Problems of the Previous Studies

Huang and Ahrens (2003) analyze the Mandarin classifier system and suggest that classifiers coerce nominal semantic types:

　"…classifiers can coerce nouns to have a particular individual reading depending on the information entailed in the classifier itself. The classifier can vary in the Constitutive, Formal, Telic or Agentive roles that it carries (p.361)."

　The situation in Japanese, however, is considered to be more complex because there are some

cases where type coercion does not seem to take place, as pointed out in Section 1. Therefore, we adapt the four mechanisms suggested by Pustejovsky (2011) and show that all of the mechanisms occur when classifiers count nouns in Japanese. Mano and Yonezawa (to appear) also take the same view, but their discussion is limited.

　Pustejovsky (2011:1411) suggests the following mechanisms in (17)[7] for the selection of an argument, which allow for modulation of types during semantic composition.

(17) a. SELECTION (Type Matching): The target type for a predicate, F, is directly satisfied by the source type of its argument, A: $F(A_\alpha)_\alpha$

　　b. ACCOMMODATION SUBTYPING: The target type a function requires is inherited through the type of argument, A: $F(A_\beta)_\alpha, \beta \sqsubseteq \alpha$

　　c. COERCION BY INTRODUCTION: the type a function requires is imposed on the argument type. This is accomplished by wrapping the argument with the type required by the function:
　　　$F(A_\alpha)_{\odot\sigma}, \alpha \sqsubseteq \beta$ (domain-preserving)
　　　$F(A_\alpha)_\beta, \alpha{\to}\beta$ (domain-shifting)

　　d. COERCION BY EXPLOITATION: the type a function requires is imposed on the argument type. This is accomplished by taking a part of the argument's type to satisfy the function: $F(A_{\alpha\odot\tau})_\beta, \tau \sqsubseteq \beta$

3.2　Classifiers and the Compositional Mechanisms

Here we will show that all four mechanisms in (17) are observed when classifiers modify and count nouns in Japanese. What is the most crucial is whether the head type (i.e. the Formal role) of the quantified noun is changed or not.

　First, Selection (Type Matching: TM) will take place when the type required by a classifier is directly satisfied by the quantified noun. It is predicted that the quantified source noun stays the same type in this case because the operation does not change the type. In (18a), the noun, *hito* 'human being,' satisfies the type required by the classifier, *-nin*, because both the source and target

---

[6] See Matsumoto (1991, 1993), Downing (1996), Iida (1999, 2004), and Nishimitsu and Mizuguchi (2004), for example.

[7] $\odot$ represents the disjunction of the two type constructors, $\otimes$ and ·.

types are *human* (cf. (15a)) [8]. The same is true for (18b), in which both the source and target types are *event* (cf. (15d)).

(18) a. san-nin-no hito-ga iru.
    3-CL-GEN human-NOM be.PRES
    'There are three men.'
  b. san-kai-no ensoo-ga owat-ta.
    3-CL-GEN performance-NOM end-PST
    '(lit.) Three performances were over.'

Second, we will show an example of Accommodation Subtyping (AS). As shown in (19), *tsubame* 'swallow' is counted by both -*wa* and -*hiki*.

(19) ni-{wa/hiki}-no tsubame-ga tondeiru.
    2-CL/CL-GEN swallow-NOM fly.PROG
    'Two swallows are flying.'

It is assumed that the TM applies when it is quantified by -*wa* (15c), because the Formal role of *tsubame* is typed as *bird*. The AS takes place, however, when it is counted by -*hiki* (15b), because the type *bird* is a subtype of the type *animal* (*bird* ⊆ *animal*). Actually, all the nouns counted by -*wa* can also be counted by -*hiki*, but not vice versa.

type requirement of the classifier	mechanism
-*wa*: *bird*→t	TM
-*hiki*: *animal*→t	AS

Table 1. Compositional mechanisms of *tsubame*

Third, we will see cases where coercion takes place. *Dango* 'rice dumpling,' for example, is an artificial type, and its type structure is considered to be "food⊗$_T$ eat." It is not necessarily skewered, but it is interpreted to be skewered when counted by -*kushi* as in (20). As -*kushi* is of type "*food*⊗$_C$ *skewer*→t" (16a), Coercion by Qualia Introduction (CI-Q) applies to *dango*, adding the Constitutive value *skewer*[9].

(20) san-kushi-no dango
    three-CL-GEN rice.dumpling
    'three sticks of rice dumpling'

Dotted objects can be good illustrations of occurrence of more than one generative mechanism. For example, *supiichi* 'speech', whose lexical structure is shown in (21), seems to be a dotted type (event·information), even though it is quantified by the classifiers, -*kai* and -*hon*. This is because its eventive meaning can be modified by *nagai* 'long,' and its content meaning (info) can be modified by *omoshiroi* 'interesting' regardless of the existence of the classifiers, as shown in (22).

(21) *supiichi* 'speech'

event·information_lcp
QL= FORMAL=information (x)
    AGENT=speak (e, z, x)
    TELIC=communicate_to (e, z, x, w)
        (Mano and Yonezawa, to appear)

(22) a. nagakute omoshiroi supiichi
    long interesting speech
    'long and interesting speech'
  b. kare-no (ni-{kai/hon}-no) supiichi-wa
    3SG-GEN 2-{CL/CL}-GEN speech-TOP
    dochira-mo nagakat-ta-ga omoshirokat-ta.
    both-also long-PST-but interesting-PST
    'Both of his speeches were long but interesting.'

Now we examine the compositional mechanisms occurring in (22b). As shown above, *supiichi* is a complex type (e·i), and the classifier, -*kai*, is a classifier for *event* as in (15d), so Coercion by Dot Exploitation (CE-·) applies. The classifier, -*hon*[10], in (22b) seems to focus on the informational aspect of *supiichi* as a way of communication. This kind of -*hon* is considered to be of type "*information*⊗$_T$ *communicate*→t," which means that after the CE-·, CE-Q occurs in *ni-hon-no supiichi*, exploiting the Telic value *communicate*.

---

[8] *Kashu* 'singer' (*human*⊗$_T$ *sing*) is also counted by -*nin*. We assume that CE occurs, making *kashu* be of type *human*. As pointed out by one reviewer, however, we should compare this with the case of *hito* and make clear whether there is any (syntactic and semantic) difference between them or not.
[9] We assume that TM applies when -*kushi* counts *sate* 'satay: grilled meat stick' as in (12), because the Constitutive value *skewer* is included in the lexical structure of *sate* as in (i).

(i) *sate*   QL=FORMAL=food (y)
        CONST=consist_of (y, {skewer, meat…})
        TELIC=eat (e, x, y)
[10] -*Hon* is a shape classifier for inanimate one-dimensional physical objects, but it is well known to have several extended usages, counting other than physical objects: for example, hits (baseball), movies, letters, phone calls, etc. See Lakoff (1987), Matsumoto (1993), and Iida (1999).

Lastly, we will show a case in which Coercion by Introduction (CI) takes place. As shown in (23a), *supiichi* can also be counted by *-mai* which selects for a type *phys (2D)*.

(23) a. ni-mai-no supiichi-wo yabut-ta.
　　　　2-CL-GEN speech-ACC tear-PST
　　　　'(I) tore two sheets of speech.'
　　b. ni-{kai/hon}-*(bun)-no supiichi-wo yabutta.
　　　　2-CL-quantity-GEN
　　　　'(I) tore two speeches.'

The CI is considered to apply to *supiichi* (e·i) here, resulting in the noun being a type *phys*. This is confirmed in (23a), because it can be the argument of the predicate, *yaburu* 'tear,' which selects for *phys*. This is impossible when it is counted by *-kai* and *-hon*, as shown in (23b).

Table 2 summarizes the compositional mechanisms observed with regard to *supiichi*.

type requirement of the classifier	mechanism
-kai: *event*→t	CE-·
-hon: *information*$\otimes_T$*communicate*→t	CE-Q
-mai: *phys(2D)*→t	CI

Table 2. Compositional mechanisms of *supiichi*

It follows from what has been shown in this section that all four mechanisms are observed when classifiers select their arguments.

4　Conclusion

By using the Generative Lexicon Theory, this paper suggests the formalization of semantic requirements of classifiers in Japanese. It is also shown that all four compositional mechanisms in (17) are observed between classifiers and the quantified nouns. It is reasonable to say that the Generative Lexicon approach can propose a new analysis of classifiers.

It needs further investigation, however. As space is limited, we have concentrated on limited classifiers, but the lexical structures of other classifiers should be examined. In addition to this, I have not addressed the issues of the quantifier floating. Their syntactic structures should be more carefully examined.

Acknowledgments

References

Alexandra Y Aikhenvald. 2000. *Classifiers: A Typology of Noun Categorization Devices.* Oxford University Press, Oxford.

Walter Bisang. 1993. Classifiers, Quantifiers and Class Nouns in Hmong. *Studies in Language*, 17:1-51.

Francis Bond and Kyonghee Paik. 1997. Classifying Correspondence in Japanese and Korean. *PACLING-97*:58-67.

Pamela A Downing. 1996. *Numeral Classifier Systems: The Case of Japanese.* John Benjamins, Amsterdam.

Chu-Ren Huang and Kathleen Ahrens. 2003. Individuals, Kinds and Events: Classifier Coercion of Nouns. *Language Sciences*, 25:353-373.

Asako Iida. 1999. *Nihongo Shuyoo-josuushi-no Imi-to Yoohoo*. Ph.D. Dissertation, University of Tokyo.

Asako Iida. 2004. *Kazoekata-no Jiten*. Shoogakkan, Tokyo.

Taro Kageyama, Miho Mano, Yu Yonezawa, and Takayuki Tohno. 2011. Meishi-no Seishitsu-to Kazu-no Kazoekata. Taro Kageyama ed. *Nichi-ei-taishoo Meishi-no Imi-to Koobun*, 10-35, Taishukan, Tokyo.

George Lakoff. 1987. *Woman, Fire, and Dangerous Things.* The University of Chicago Press, Chicago.

Miho Mano and Yu Yonezawa. to appear. Lexical Semantics of Japanese Counters in the Generative Lexicon Theory. *Lexicon Forum*, 6. (In Japanese)

Yo Matsumoto. 1991. The Semantic Structures and System of Japanese Classifiers. *Gengo Kenkyu*, 99:82-106. (In Japanese)

Yo Matsumoto. 1993. Japanese Numeral Classifiers: A Study of Semantic Categories and Lexical Organization. *Linguistics*, 31:667-713.

Yoshihiro Nishimitsu and Shinobu Mizuguchi eds. 2004. *Ruibetsushi-no Taishoo*. Kuroshio, Tokyo.

James Pustejovsky. 1995. *The Generative Lexicon.* MIT Press, Cambridge, MA.

James Pustejovsky. 2006. Type Theory and Lexical Decomposition. *Journal of Cognitive Science*, 6:39-76.

James Pustejovsky. 2011. Coercion in a General Theory of Argument Selection. *Linguistics*, 49:1401-1431.