Numeral Classifiers and Labeling: A Preliminary Study *

Masao Ochi

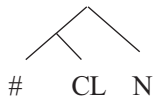
1. Introduction

This short paper reports some preliminary results of an on-going investigation of the syntactic status of numeral classifiers (NCs) in Japanese. Adopting the hypothesis put forth by Huang and Ochi (henceforth, H&O) (2014), I will explore the idea that two classifier systems co-exist in Japanese (see also Oho (2021)). It will be argued that they differ with respect to the strong vs. weak status under the recent labeling theory of Chomsky (2013, 2015): the pre-nominal NC arises when the classifier head is weak whereas the post-nominal pattern is obtained when the classifier head is strong.

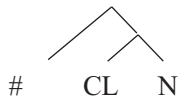
2. Two Classifier Systems

Numeral classifiers have been approached from two contrasting perspectives. One line, due to Krifka (1995), holds that a classifier is for numerals. The idea is that numerals in classifier languages, unlike those in non-classifier languages, lack an inherent measure function, and thus require a classifier which supplies a measure function, $\mu_{\#}$. This line of hypothesis is best illustrated by a structure like (1a), in which a numeral and a classifier form a constituent and serve as a modifier of the head noun. Another prominent line of approach, due to Chierchia (1998), holds that a classifier is needed for nouns. Based on the idea that nouns in classifier languages are mass-like, Chierchia proposes that a classifier plays a role of atomizing such nouns so that they can be the target of numeral measurement. The line of analysis can be represented by (1b), in which a classifier and a noun form a constituent, to which a numeral is added.

(1) a.



b.



Now, H&O propose that the pre-nominal NC (see (2a)) and the post-nominal NC (see (2b)) in Japanese have distinct underlying structures, roughly corresponding to those in (1). For H&O, the pre-nominal NC has (1a) as its underlying structure. As for the post-nominal NC, H&O postulate something like (1b) as the underlying structure, following Watanabe (2006): a classifier and a noun form a unit.

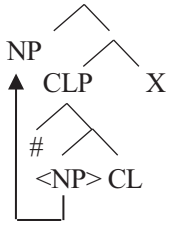
*The research presented here is financially supported by the Grants-in-Aid for Scientific Research (C) (No. 20K00679), the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

(2) a. Boku-wa [san-satsu-no hon]-o yonda. (pre-nominal NC)
I-TOP three-CL-LK book-ACC read
'I read three books.' (Note: *-no* is assumed to be a linker, inserted at PF)

b. Boku-wa [hon san-satsu]-o yonda. (post-nominal NC)
I-TOP book three-CL-ACC read
'I read three books.'

Further, H&O propose, again following Watanabe (2006), that the post-nominal NC construction involves movement of the NP complement of CL to the edge of the nominal domain, thus giving rise to the word order in which the head noun precedes the numeral classifier sequence (see (3)).

(3)



As for pre-nominal NCs (see (2a)), H&O follow Saito et al. (2008) and take them to be adnominal modifiers. This is supported by the well-known observation that pre-nominal NCs can be ordered freely among the modifiers in the pre-nominal field (see (4)), which is expected as modifiers do not have a fixed syntactic slot to occupy.

(4) a. Taro-no san-mai-no syasin
Taro-GEN three-CL-GEN photo
'Taro's three photos'

b. San-mai-no Taro-no syasin
three-CL-GEN Taro-GEN photo
'(lit.) three Taro's photos'

And yet, there is something peculiar about pre-nominal quantifiers that sets them apart from other adnominal modifiers of this language. Consider (5), which contain a universal quantifier, *subete* 'all.' As observed by Watanabe (2017) (see also Ochi (2019)), (5b), which has a post-nominal universal quantifier, is ambiguous with respect to the number of books that the speaker read. On the other hand, (5a), in which *subete* appears pre-nominally, is unambiguous. In this case, no singular reading of the head noun, *hon* 'book,' is available.

(5) a. Boku-wa subete-no hon-o yonda.
I-TOP all-GEN book-ACC read
'*I read all of the book. / 'I read all of the books.'
b. Boku-wa hon subete-o yonda.
I- TOP book all-ACC read
'I read all of the book. / 'I read all of the books.'

Ochi (2019) took this observation as an indication that a pre-nominal quantifier enters into number (i.e., plural) agreement with the head noun. Furthermore, pre-nominal quantifiers, unlike post-nominal quantifiers, cannot be stacked, which is uncharacteristic of Japanese.[1]

(6) a. *subete-no hyaku-satsu-no hon
all-GEN 100-CL-GEN book
'all 100 books'
b. *hyaku-satsu-no subete-no hon
100-CL-GEN all-GEN book
'all 100 books'
c. hon hyaku-satsu subete
book 100-CL all
'all 100 books'

The ungrammaticality of (6a) and (6b) can be explained if (i) a pre-nominal quantifier, be it a universal quantifier or a numeral quantifier, necessarily establishes an agreement relation with the head noun in Japanese, (ii) such an agreement relation needs to be uniquely established within a single nominal domain, and (iii) no such requirement holds of post-nominal quantifiers (see Ochi (2019)).

These points clearly show that the pre-nominal classifier and the post-nominal classifier should be given a separate treatment (the reader is referred to H&O (2014) for more discussion). In the light of these considerations, I would like to explore the idea that two classifier systems co-exist in a single language by sketching a possible way in which this idea can be syntactically implemented.

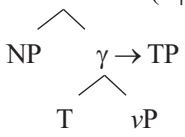
3. Classifiers as Weak/Strong Heads

The discussion in this section is based on Saito (2016, 2018), who presents an analysis of how labeling works in a language without φ-feature agreement. Imagine a point in the derivation at which T and *v*P merge, thus creating $\gamma = \{T, vP\}$, as shown in (7a). One of the members, T, is a head, so it should provide

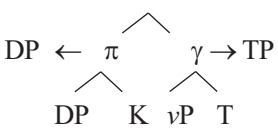
[1] (6a) is grammatical with the reading "all sets of 100 books." This reading is set aside (see Ochi (2019)).

a label for γ. However, T in English is assumed to be too weak to provide a label (see Chomsky 2013, 2015). In this situation, internally or externally merging a syntactic object (NP in (7a)) with γ would help avoid the problem of labeling failure. Search into NP and γ identifies two heads, N and T, respectively. These two heads share ϕ-features, which enables the mother node λ to obtain <ϕ, ϕ> as its label. Saito (2016, 2018) addresses the question of how this type of syntactic configuration can be labeled in a language like Japanese that lacks agreement (see (7b)). Saito (2016) proposes that case particles (as well as other grammatical devices such as adnominal and adverbial inflections) serves as anti-labeling devices, which Saito (2018) recasts in terms of 'weak' heads, as summarized in (8).

(7) a. $\lambda \rightarrow$ TP (<ϕ, ϕ>)



b. $\lambda \rightarrow$?



(8) Search {α, β} for a label. If α is a weak head or search into α yields a weak head, then search on the α side is suspended and it continues only on the β side.

Consider (7b), which is the TP structure for Japanese. T in Japanese is assumed to be strong, so T provides a label for γ. Now consider the DP side. When DP and K (case particle) are merged, search on the K side is suspended. Consequently, π's label is DP. When π (DP) and γ (TP) are merged, we have an {XP, YP} configuration, and unlike in (7a), there is no ϕ-agreement in Japanese. Saito argues that (8) allows λ to be labeled. Search on the side of π is suspended once K, a weak head, is found. Thus T provides a label for λ.

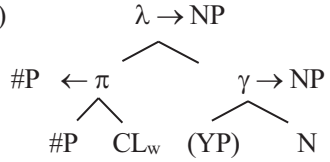
I will adopt this line of analysis and propose (9) for classifiers.

(9) In Japanese, a classifier head may be weak or strong.

Recall that there are two combinatorial possibilities for a classifier. I will argue below that the combinatory possibility in (1a) arises when CL is a weak head.[2] And (1b) is possible only when the CL head is strong.

Let us start with (1a). Consider (10).

(10)



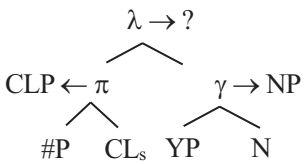
(Note: YP is an optional adnominal modifier)

[2] This configuration is in fact possible when CL is strong as well, but that option will be shown to be empirically inadequate for Japanese pre-nominal NC configurations. See the discussion below.

Note that I will assume without any discussion that numerals are phrasal (see Tatsumi (2021)). Now imagine a point in the derivation at which #P and CL are merged, where CL is a weak head. The label search on the side of CL is suspended. As a result, labeling Algorithm (LA) identifies #P as the label of π. When π = {#P, CL} and γ (= NP) are merged, the label of λ becomes NP because, once again, CL is a weak head: search on the side of π is suspended and continues only on the side of γ (= NP). Notice now that π (whose label is #P) and γ (= NP) are in a local relation in (10). This, I suggest, is the source of the number agreement (i.e., plural agreement) seen in (5a) (and, more importantly, in (6a)).

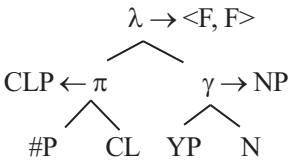
Now suppose that CL is strong in the same configuration. We then obtain the following.

(11)



When π (= CLP) and γ (=NP) are merged, the resulting configuration λ gives rise to a potential labeling problem. This is because we have an {XP, YP} configuration, assuming that γ = {YP, NP} (if there is no YP, no labeling problem would arise). Consequently, the label of λ cannot be determined. However, Chomsky suggests that this type of structure can be labeled if, for instance, XP and YP share a prominent feature, such as ϕ-features. Although Japanese is standardly assumed to lack ϕ-agreement, it is quite conceivable that CL and N agree: after all, the choice of a classifier is determined by the nature of a noun. We could thus say that γ is successfully labeled as a result of this agreement relation.

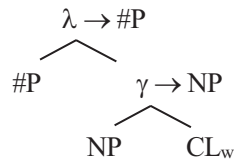
(12)



Nevertheless, (12) does not seem to be appropriate for Japanese pre-nominal NC constructions. Recall that the pre-nominal numeral and the head noun agree in number (see the discussion around (5a) and (6a)). But this relation cannot be captured by (12) because #P and NP are not in a local relation here. Thus, although (12) may be a possible manifestation of (1a), it is not a suitable structure for Japanese pre-nominal NCs. In sum, as far as Japanese is concerned, the pre-nominal NC structure can be constructed with a weak CL head, but not with a strong CL head. Why that is so remains unclear. One possibility is that 'weak' is the default value for CL (and K(ase)). For Saito, Kase is weak in Japanese, and it is unclear whether there is a language that possesses strong Kase. For this reason, I will tentatively adopt the hypothesis that the weak head option is the norm for items like case particles and classifiers: the 'strong' option for such heads arises as a last resort (see below).

I will now turn to the other combinatory possibility in (1b), in which a classifier is merged with a noun.

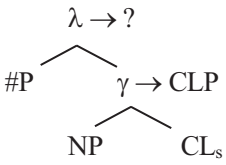
(13)



Let us suppose that CL is weak. When NP and CL are merged, we obtain γ, whose label is NP, given that CL is a weak head. Next, #P and γ (= {NP, CL}) are merged. Search on the side of γ comes to a halt as soon as CL, a weak head, is located. As a result, search continues only on the side of #P, which provides the label for λ. But this would lead to a problem with respect to selection. We would like λ to be NP, not #P, since λ should serve as an argument of a predicate that selects a certain type of nominals.

Let us therefore consider the possibility that CL in the same configuration is strong.

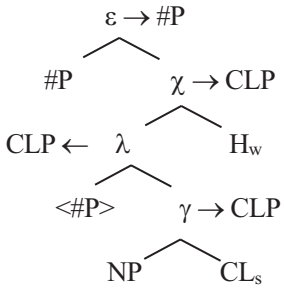
(14)



This time, the strong CL provides a label of γ because this is an {X, YP} configuration. When γ is merged with #P, we obtain an {XP, YP} configuration (= {#P, CLP}). Chomsky suggests that this type of configuration would need to be modified in order to avoid labeling failure: either (i) XP and YP share a prominent feature, or (ii) one of the members of the set moves out. If, for instance, XP of α = {XP, YP} moves out, YP provides the label of α, since the trace of XP is invisible for LA.

The first option is immediately discarded: there is no obvious sense in which # and CL share a prominent feature. Let us therefore consider the possibility that something moves out of λ. Let us assume that there is a weak head, H_w, that merges with λ.[3]

(15)

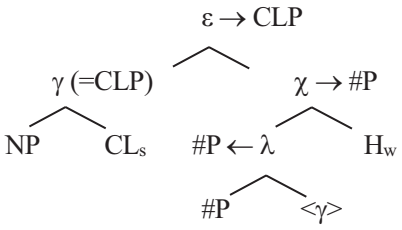


[3] Due to space limitation, I cannot discuss derivations in which H is strong, but they fail to yield a desired result with respect to the label of the top mode ε.

Suppose #P moves out of λ as shown in (15). The label of λ is CLP as a result of this movement. Due to the weak status of H (see (8)), the label of χ becomes CLP, and the top node, ε, has #P as its label. Again, we have a problem regarding selection.

Suppose instead that γ moves out of λ. This time, CL provides a label for ε, again an undesired result.

(16)

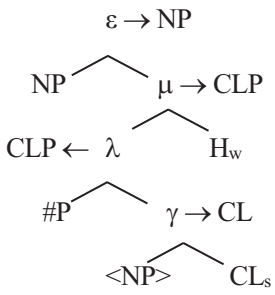


Here I suggest that there is yet another possibility to modify (14) to avoid labeling failure: NP moves out of λ. As schematically shown in (17), moving QP out of λ effectively turns an {XP, YP} configuration into an {X, YP} configuration.

(17) QP .... {λ {ZP, Y}, {<QP>, X}}

Consider (18). When NP and CL are merged, CL provides a label for γ, as before. But, crucially, γ in this case dominates only CL after NP moves. This effectively turns λ into an {X, YP} configuration.

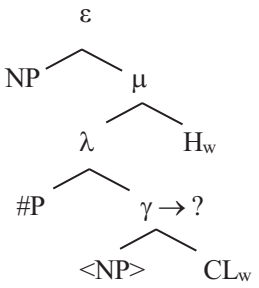
(18)



As a result, CL provides a label for λ. Then H_w, a weak head, is merged with λ (whose label is CLP) and accommodates the displaced NP. The label of the top node ε becomes NP, due to the weak nature of H (see (8)). We thus obtain a desired result this time. Now recall that in (10), which is the proposed structure for the pre-nominal NC, π (= #P) and γ (= NP) are sisters. That is not the case in (18). Although NP and #P occur within the same extended nominal projection, they are not sisters. This would explain why the post-nominal CL configuration does not trigger agreement between a numeral and a noun (see (5b) and (6b)).

Importantly, the type of derivation sketched above is not possible when CL is weak. This is because γ fails to be labeled once NP moves out of it, as shown in (19). The only visible element in γ is CL, which cannot provide a label because it is weak.

(19)



We thus conclude that the structure (1b), which underlies the post-nominal NC in Japanese, is not possible when CL is weak. It is possible only when CL is strong, which reinforces the remark made earlier that the strong head option may be a last resort option for CL.

4. Final Remarks

This paper has argued that Japanese utilizes two classifier systems that differ with respect to the strong/weak status that is relevant for labeling. But many questions arise. First, if there are indeed two ways in which a classifier in Japanese is combined with another syntactic object, how does the interpretation work? Second, what does it mean to say that a single grammar accommodates two distinct measurement systems? For the first question, I refer the reader to Oho (2021), who argues that Japanese has two classifier systems and, correspondingly, two numeral systems. One type of classifiers introduces a measure function $\mu_{\#}$ and combines with a numeral. The other type is an atomizer and combines with a noun. On the assumption that nouns in Japanese are number-neutral, Oho argues that the two combinatory possibilities end up with the same denotation for the entire nominal expression.

As for the second question, we should recognize that the issue is much broader. Take English, which, at least on the surface (see Borer (2005)), does not use a classifier and allows a numeral to directly combine with a noun (e.g., *three books*). And yet, some English measurement constructions also employ terms such as *grams* and *items*.

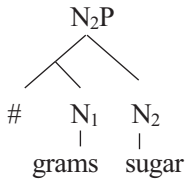
(20) a. I put 500 *grams* of sugar in the pot.
b. I put four *items* of furniture in the hallway. (Bale et al. 2019)

Importantly, as Bale et al. (2019) note, such examples of measurement do not form a homogeneous group. For instance, the two cases in (20) show distinct agreement patterns. (21a), with *grams*, favors the singular form of the predicate whereas (21b), which uses *items*, sounds more natural with plural agreement.

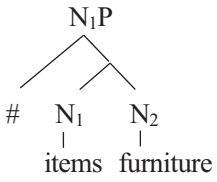
(21) a. The 500 grams of sugar we added to the sauce {gives / ?give} it a sweet aftertaste.
b. The 500 items of furniture that we bought at the flea-market {were / # was} in good condition.

Let us follow Bale et al. and assume that the head of the subject noun phrase is *sugar* (and not *grams*) in (20a), and *items* (and not *furniture*) in (20b). The combination of a numeral and *grams* serves us a measurement modifier of the head noun in (20a). In (20b), *items* combines with a mass noun, *sugar*, and creates partitions over the denotation of that mass noun. The trees in (22a) and (22b) illustrate these two patterns. I will follow Bale et al. (2019) and call them *measuring* structure and *partitioning* structure, respectively.

(22) a.



b.



Importantly, the structures in (1) and those in (22) are nearly identical, except for the category label of the element that merges with # or N (i.e., CL in (1) and N_1 in (22)). It is therefore conceivable that languages make use of two combinatorial strategies for measurement regardless of whether or not classifiers are employed.

But if there are two distinct measurement structures shown in (1) (and (22)), do they give rise to distinct interpretations? As Bale et al. (2019) note, many measurement terms give rise to potential ambiguity. For example, *Mary put four cups of water on the table* can literally mean that there were four different cups filled with water that Mary placed on the table ('partitioned' reading), or it can mean that Mary filled a cup with water and put that water into a bowl etc. on the table, and she repeated this action four times ('measured' reading). One could say that this ambiguity is structural in nature in that the partitioned reading comes from a structure like (22b) and the measured reading comes from (22a).

The issue is far from settled, however. Bale et al. show that Ch'ol, a Mayan language, uses the 'measuring' structure in (22a) for both measured and partitioned readings, and conclude that the measured reading and the partitioned reading should not be tied to (22a) and (22b), respectively. Also, in Japanese, *guramu* 'gram', which according to the discussion above typically yields the measured reading, can be used pre-nominally or post-nominally, which again indicates that there is no one-to-one correspondence between the measured/partitioned readings and the two structures in (1).

(23) a. 50 guramu-no satou-o nabe-ni ireta.
50 gram-GEN sugar-ACC pan-DAT put
'I put 50 grams of sugar into a pan.'

b. satou 50 guramu-o nabe-ni ireta.
sugar 50 gram-ACC pan-DAT put
'I put 50 grams of sugar into a pan.'

Finally, we have seen that (1a) manifests itself in Japanese when the classifier head is weak. In principle, (1a) could be built with a strong classifier head, but the the presence of the number agreement (see (5a) and (6a)) would not be captured. Thus, Japanese always uses a weak classifier head when constructing the pre-nominal NC structure. But there may a classifier language that employs a strong classifier head upon building (1a). An investigation of this issue needs to be kept for another occasion. We have also seen that (1b) can be realized in Japanese only when CL is strong (see (18)), since a weak CL head would not allow its NP complement to move to the nominal edge, and, as a result, the entire nominal 'projection' would not be properly labeled (see (19)). But what about (22b), which has an almost identical structure but, unlike (18), does not employ any nominal-internal movement? In this case, *items* is sufficiently nominal and shows number agreement with a numeral. Accordingly, it can provide a label (= NP) for the top node in (22b).

References

Bale, Alan, Jessica Coon and Nicolás Arcos López (2019) Classifiers, partitions, and measurements: Exploring the Syntax and semantics of sortal classifiers. *Glossa* 4(1): 1–30.

Borer, Hagit (2005) *Structuring sense*. Oxford: Oxford University Press.

Chierchia, Gennaro (1998) Reference to kinds across languages," *Natural Language Semantics* 6: 339–405.

Chomsky, Noam (2013) Problems of projection. *Lingua* 130: 33–49.

Chomsky, Noam (2015) Problems of projection: Extensions. In: Elisa Di Domenico, Cornelia Hamann, and Simona Matteini (eds.) *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 3–16. Amsterdam: John Benjamins.

Huang, C.-T. James and Masao Ochi (2014) Remarks on classifiers and nominal structure in East Asian. In: C.-T. James Huang and Feng-hsi Liu (eds.) *Peaches and plums*, 53–74. Taipei: Language and Linguistics Monograph Series, Academia Sinica.

Krifka, Manfred (1995) Common nouns: A contrastive analysis of English and Chinese. In: Gregory N. Carlson and Francis Jeffry Pelletier (eds.) *The generic book*, 398–411. University of Chicago Press.

Ochi, Masao (2019) Labeling Algorithm, agreement, and pro-form *no* in Japanese. *Proceedings of WECOL 2018*, 151–161. California State University, Fresno,

Oho, Atsushi (2021) Two kinds of numeral classifiers in Japanese. Paper presented at the 163rd Meeting of the Linguistic Society of Japan. 20 November, 2021 (via Zoom).

Saito, Mamoru (2016) (A) case for labeling: labeling in languages without phi-feature agreement. *The Linguistic Review* 33: 129–175.

Saito, Mamoru (2018) Kase as a weak head. *McGill Working Papers in Linguistics* 25.1 *(Special Issue in Honour of Lisa Travis)*, 382–391.

Saito, Mamoru, T.-H. Jonah Lin and Keiko Murasugi (2008) N'-ellipsis and the structure of noun phrases

in Chinese and Japanese. *Journal of East Asian Linguistics* 17: 247–271.

Tatsumi, Yuta (2021) *Linguistic realization of measuring and counting in the nominal domain: A cross-linguistic study of syntactic and semantic variations*. Doctoral dissertation, University of Connecticut.

Watanabe, Akira (2006) Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language & Linguistic Theory* 24: 241–306.

Watanabe, Akira (2017) The mass/count distinction in Japanese from the perspective of partitivity. *Glossa* 2(1): 1–26.