1. Introduction

Those numeral classifier languages like Korean, Japanese, and Chinese mobilize classifiers together with numerals in naming and counting objects in the world. Classifiers serve to classify objects according to their shape, function and other semantic features often culturally defined. Discrete (or fuzzy) numerals are used to count the number of objects.\footnote{This paper reflects some ideas I presented at the 1991 LSA Institute (UC Santa Cruz) Korean Workshop and at the 1990 ICKL(Osaka) as well as recent investigations on quantized Themes and aspect, based on the initial ideas developed in Lee (1989), which was presented at the 1989 LSA Institute Workshop on Cross-linguistic Quantification sponsored by Barbara Partee and others. In the current version, I elaborated further on quantization, mass/count, and aspect. I would like to express my gratitude to those who asked questions and made comments at the meetings and particularly to an anonymous reviewer for ITALLC98 for his/her helpful comments.} In Korean and Japanese, plural marking also occurs but it is not syntactically obligatory and is often semantically/pragmatically oriented. In Chinese, a plural NP is definite. Plural marking in classifier languages is different from that in Indo-European.
languages. It is not a simple matter of distinguishing between singular and plural. The numeral classifier construction is analogous to the numeral measure unit construction in Indo-European languages.

This paper is concerned with characterizing the numeral classifier construction in Korean (in Japanese, and in Chinese, where relevant) in connection with Incremental Theme effects particularly on object arguments, as well as specificity/scope, count/mass, bare common nouns and implicature effects. Particularly, we are interested in showing how Incremental (Dowty 1991) or quantized (Krifka 1987, 1992, 1997) Theme is achieved by means of numeral classifiers, which can be syntactically indefinite, and how it is correlated with telicity in aspect, a salient feature of event structure.

Children acquire the construction a few months after age two, even though they utter some general classifier (kae ‘small object’ in Korean and analogous classifiers in other languages) more often than other specific classifiers before they become accustomed and sensitive to different kinds of classifiers. This early occurrence indicates the importance and salience of the construction in cognitive and linguistic development. Even babies are understood to be born with a number line in the brain that allows them to grasp small numbers, up to three or four.

2. Indefiniteness/Definiteness and Incremental Theme

2.1 Indefinite Article, Mass/Count, Incremental Theme and Aspect

In English, indefiniteness is represented typically by bare plurals/mass nouns. The indefinite article, on the other hand, has two faces. First, it is used to denote quantization or Incremental Theme event with change of state (creation/consumption) verbs, as follows:

(1) Mary ate apples for thirty minutes. (‘durative’)
(2) Bill ate an apple in five minutes. (‘perfective’)
(3) Sue drank a glass of wine in five minutes. (‘perfective’)

In (1), the object is indefinite and non-quantized as a bare plural or mass, thus constituting an activity, combined with the verb. Therefore, the event description is modified by durative time adverbial expressions such as ‘for thirty minutes.’ On the other hand, the object ‘an apple’ in (2) is syntactically indefinite but, together with a verb of state change such as eat and build, semantically constitutes a quantized or incremental theme and makes a big difference in aspect. It denotes one whole apple, i.e., number plus atomic individual. It must be noted that the indefinite article a here has the reading of number ‘one.’ It is an Incremental Theme (Dowty 1991), a quantized Theme (Krifka 1987, 1989, 1997), or a ‘measured out’ Theme (Tenny 1994). The
activity of ‘eating apples’ may also be claimed to have been ‘measured’ in a sense even in (1) by means of ‘for thirty minutes’; the event has the beginning and can have the end (or beginnings and ends if we view the situation as ‘eating apple after apple’ in subevents), but the terminal point is not duly or event-internally represented, i.e. the event is protracted and non-telic. The ‘measurement’ of duration involved here is simply external. And the event is not guaranteed to have come to an end in (1) and even when it came to an end, it is not clearly and naturally bounded in correlation with the consumption of individual apples. The last apple could be just a slice in (1). The duration expression ‘for ---’ (or tongan in Korean), compatible with activities, has to do with the notion of ‘continue,’’ which is in turn claimed to have an affinity with ‘mass’ (Hirtle 1982). When the predicate is telic in its semantic nature, as in (2) through (3), it involves a homomorphism from its Theme argument denotations into a domain of events. The changing state of parts of the apple and their part-whole relationships is mapped in the parts of the event of eating it and its part-whole relationships (Dowty 1991). If it is likewise quantized, with numeral expressions or definite descriptions, with full individuations, then no proper subevents (of eating subparts of an apple) of an event of eating an apple can be described by the same sentence that includes ‘an apple’ or other quantized Theme expressions like ‘three apples’ and ‘the apple.’ Observe an instance of definite description.

(4) Mary ate the apple/the apples/the two apples in three minutes.

The definite article ‘the’ has the ‘unique being’ presupposition and is already quantized as ‘only one.’ In the case of definite plural, the same mass algebra as done by means of a maximality operator \( \iota \) applied to any common noun (or predicate) to denote the greatest aggregate of the common noun, whether plural or mass, can be employed. So, the maximal set in the given context as a whole is also quantized and the number or individualization of its members does not particularly matter even in the case of count nouns. Similarly, ‘a glass of wine’ is quantized by a measure function applied to mass. In contrast, ‘apples’ and ‘wine’ are not quantized and are both mass. The bare nominal ‘apples’ is also treated as mass (cumulative)(or kind) in contexts like ‘three pounds of APPLES,’ even though quantization has been achieved by a quantization function applied to the mass (weight) of apples to yield ‘three pounds.’ Although I am basically in line with Link (1983) and Chierchia (1998), I adopt neither the mass/count dichotomy by the former nor the count-driven approach by the latter. My approach is mass/kind-driven, even though I take important counting and quantization in particular.

On the other hand, the indefinite article cannot be used for quantization in a situation where an indefinite NP plus a light verb constitutes an activity. Consider the following:
(5) I took a walk for thirty minutes.

In (5), ‘take a walk’ as a whole is an activity (cumulative) and the indefinite article ‘a’ is not used for quantization in this context, and, therefore, can be modified by a durative time adverbial (‘for thirty minutes’).

Still another important kind of predicates that do not take a quantized object argument, though with an indefinite article, numerals, or definite article, includes verbs such as ‘push’ and its Korean equivalent mil-, as shown in the following examples:

(6) a. Mary pushed a/two cart(s) for/*in ten minutes.
    b. Mary-nun kaht han/two tae -rul sip pun tongan/*man-e mil-oss-ta
    Top cart one/two Cl Acc 10 mnts for in pushed
    ‘Mary pushed a/two cart(s) for/*in ten minutes.’

In the act of ‘Pushing,’ the agent’s initial exertion of force counts but the object’s change (of state or of location to a terminal point) in its event structure is not designed. Therefore, location change to a terminal point must be expressed by means of ‘(pushed a cart) to the car,’ to be associated with ‘in ten minutes.’

The range of verbs involving alternations between Locative and ‘Derived’ Theme such as load and paint show a clear contrast in quantization. Observe:

(7) *Mary loaded hay onto two trucks in an hour.
(8) Mary loaded two trucks with hay in/for an hour.
(9) Mary loaded two tons of hay onto the truck in/???for an hour.

As shown in (7), the numeral in the PP does not contribute to any quantization in the event involved; the object figure is mass and the Locative ground does not form a telic event with that object. The same numeral NP in the object position in (8), however, can constitute a quantized Theme (in the implicated sense of the trucks becoming ‘full’) and take the ‘in’ time-span PP. Devoid of this ‘fullness’ or global affectedness sense, the event remains as an activity and can be modified by the ‘for’ duration PP. Unlike (7), (9) has a numeral measure NP object figure and, together with the Goal location, it contributes to quantization. In Korean, on the other hand, the verb corresponding to ‘load,’ sit-, cannot have a derived Theme argument, thus only forming the frame of [object figure] + [Locative PP ground], constituting an activity, or alternatively, a quantized figure plus ground as an accomplishment just like (9).

The Korean verb chaewu- ‘fill’ constitutes an accomplishment, with its quantized ground (underlyingly Goal-Locative) object and its Instrumental PP (‘with water’), as in English. But unlike in English, it can constitute an activity,
with its figure object (‘water’) plus Goal-Locative PP. In child English, the latter frame is reported to occur. English is sensitive to global affectedness (being full) and Korean is sensitive both to global affectedness and to locomotive activity (as in ‘pour water into the basket’).

We can notice that there are two sorts of quantization: a strong sort of quantization involving change of state or affectedness (as with consumption verbs like ‘eat,’ ‘destroy,’ ‘drink’ and creation verbs like ‘write,’ ‘build,’ etc.) and a weaker sort of quantization involving locomotion (as with ‘run a mile,’ ‘go to,’ ‘push-(¬mass)-to,’ ‘load-(¬mass)-onto,’ etc.). A derived Theme from Locative seems underspecific in terms of quantization in the event involved, as witnessed in (8). Therefore, Krifka’s definition of quantization as prohibition of a proper part of an individuated part is not enough to characterize the strong type, as in (10), and some notion of incremental Theme must be incorporated, as done separately in (11) for ‘drink two cups of tea’ (Krifka 1997, cf. Verkuyl 1993):

\begin{align}
A \text{ Predicate } P \text{ is quantized iff } \forall x, y [P(x) \land P(y)] \rightarrow (y < x) \nonumber \\
\{ e | \exists x [DRINK(e) \land TWO-CUPS-OF-TEA(x) \land TH(e, x)] \} \nonumber
\end{align}

Then, how to treat the weaker sort of quantization involving Goal, etc. can be problematic, but the Goal NP may be viewed as Theme and the verb + preposition (‘go to,’ or -e ka- ‘go to’) as a transitive V-Particle construction. However, this cannot be passivized (though the NP can be relativized in Korean), whereas somewhat affected Locative Theme construction (such as the object of ‘sleep in, march on,’ such as in (18), a bare noun occurs in English (e.g., This bed was slept in by Jackson).

Likewise, quantized NPs such as definite NPs, numeral NPs and cardinality ‘a’ NPs constitute accomplishments (or achievement) in aspect with their associated telic verbs and thus occur with the time-span adverbial ‘in ten minutes.’

2.2. In Numeral Classifier Languages

In no-article numeral classifier languages like Korean, a bare common noun has many faces: an indefinite mass noun face, a definite NP face, a kind name face, etc. In a cumulative situation like ‘at apples,’ as in (18), a bare noun occurs in Korean. In a quantized context like ‘eat an apple,’ a numeral classifier construction occurs. Observe:

\begin{align}
\text{(12) Mary -nun (samsip-pun tongan) sakwa -rul mok -oss -ta} 
\end{align}
Top 30 minute for apple Acc eat Past Dec
‘Mary ate apples (for thirty minutes).’
(13) Mary -nun (sam-pun man-e) sakwa han kae -rul mok -oss -ta
Top 3 minute in apple one Cl Acc eat Past Dec
‘Mary ate an apple (in three minutes).’

Sentence (12) denotes an activity and cannot be modified by the ‘in’ time adverbial, whereas (13) denotes an accomplishment and cannot be modified by the ‘for’ time adverbial. On the other hand, if some completive (perfective) auxiliary or adverb accompanies the VP, a bare nominal object in it gets a definite NP interpretation and therefore is quantized. Observe:

(14) Mary -nun sakwa -rul (ta) mek-e peri/chiu -ess -ta
-Top apple -Acc exhaustively eat up -Past -Dec
‘Mary ate up the apple(s).’ (cf. *Mary ate up apples/an apple.)
(15) Mary -nun sakwa se kae -rul (ta) mek-e peri/chiu -ess -ta
-Top apple 3 Cl -Acc exhaustively ate up
‘Mary ate up the three apples.’

If (14) is followed by a denial of definiteness, such as ‘but I don’t know which apple(s),’ it turns out to be infelicitous. Even without such perfective elements, a bare nominal can be definite depending on the context. With no articles, definiteness is often contextually determined rather than grammatically marked in Korean. Even if the common noun is followed by the numeral classifier construction, like ‘sakwa[apple] se[three] kae[Cl]-rul[Acc],’ as in (15), the whole NP is still definite (meaning he three apples’). For (14) and (15), kouy ‘almost,’ the set-proportional adverbial, can be put right after the accusative marked object and show the incremental nature of the quantized Theme. The adverbial can be applied to the total set (i.e., more than two and half) or distributively to the members of the quantized set (most part of each apple). It is underspecified. The same must hold in Japanese and Chinese, although possible syntactic combinations are somewhat different from each other. In English, definiteness must be marked by the definite article ‘the’ and the numeral noun construction without it like ‘three apples’ is grammatically indefinite, even though it is quantized. Sentences (14) and (15) can be modified by the time adverbial TIME UNIT man-e ‘in TIME UNIT,’ not the duration adverbial TIME UNIT tongan ‘for TIME UNIT.’

Chinese exhibits a parallel contrast in quantization: a non-specific object NP of a state-change verb cannot occur with the ‘in’ time adverbial, whereas a numeral classifier, some NP can occur with it plus the perfective marker, as follows (Taiwanese Mandarin, consulted by W-T D. Tsai):

(16) Akiu (*zai shi-fenzhong-nei) you chi dongxi
at 10 minute in have eat thing
'Akiu ate things (*in ten minutes).'

(17) Akiu zai shi-fenzhong-nei chi-le mianbao/san ge mianbao
at 10 minute in eat Perf bread 3 Cl bread
'Akiu ate (up) the bread/(the) three loaves of bread in ten minutes.'

In (16), the bare common noun and the numeral classifier NP must be interpreted as definite and the latter may be specific.

If not aspectually telic or contextually definite, the numeral classifier construction is typically indefinite even though it is counted in these languages as well. Numerical quantification is aspectually significant and crucially related to the event structures of linguistic expressions involving it. The indefiniteness of the construction is well testified in existential sentences like the following:

(18) yennal-e onu maul-e nongbu han saram -i sal-ko iss-oss-ta
in old days a village in farmer one Cl Nom living be
'In old days, in a village, there lived a farmer.'

3. Numeral Classifier Construction and Specificity

3.1 Syntactic and Selectional Constraints

In Korean, the combination of Numeral + Classifier can either follow or precede its associated common noun, as follows:

(19) sakwa (???-tul) se kae (-ka/-rul)
apple Pl three Cl Nom/Acc
(20) se kae -uy sakwa (?*tul) (-ka/-rul)
three Cl Gen apple Pl Nom/Acc
(19-20) ‘three apples’

The pre-N Num-Cl construction (19) is predominant, i.e. more than 60%, and seems to be increasing in speech, and the post-N construction (20), mediated by the possessive marker -uy, is less than 40% in speech (Park and Kim 1996). The distribution is half and half in novels and the ratio of the pre-N pattern is a little higher in non-novels and was far more higher (almost 80%) in classical writings. Previously (Lee 1989) I proposed that the construction be derived from a modifying clause of [Subject N + Predicative Numeral Cl] preceding the head N. The subject N (a null category here) and the head N are co-referential. Then the post-N construction can be obtained by switching the order of the clausal construction to the right of the head N (as a case of an underlying pre-nominal clausal modification, but if we can treat and justify the post-N pattern as basic,
it would be better). This predication treatment can be justified by the following facts: First, Korean has the same kind of clause type (21), second, there is (honorification) agreement between the head N and the classifier, third, the classifier selects a possible head N and the violation of selection restriction creates graded sortal incorrectness. Consider the following examples in turn:

(21) yoca -ka se myong -i -ta [Subj-Predicate Relation]
    woman Nom three Cl[Hum] be Dec
    ‘The number of women is three.’ (Lit. Women are three.)
(22) sondaegnim tu pun / ?*myong [Honorification Agreement]
    teacher two Cl[Hon] Cl[non-Hon]
    ‘two teachers’
(23) a. ai se myong / *tae --- [Selection Restriction]
    child three Cl[Hum] Cl[machine]
    ‘three children’
    b. mal tu mari / ?*kae (small object class)
    horse two Cl[Anim] Cl
    ‘two horses’
    c. kkangphae se myong/ ??mari (animal class)
    gangster three Cl
    ‘three gangsters’

In (22), if the speaker chooses an honorable person head N, then an honorable person classifier ‘pun’ must be chosen at the same time. If this syntactic honorification agreement is violated by choosing a non-honorable person classifier, say, ‘myong,’ then the utterance becomes not only syntactically but also pragmatically/socially unacceptable at the same time.

There is a degree of acceptability with mismatches between classifiers and head Ns; people applied to a machine classifier as in (23a) is a serious violation, an animal to an object classifier less serious (23b) and bad people applied to a non-human animal classifier still less serious (23c). Young children of age two tend to use a ‘general’ classifier, ‘kae’[small sized objects] in Korean, to refer to all kinds of things before they become familiar with all different classifiers. Even adult salespersons tend to use the same general classifier, particularly among themselves. The last example (23c) can be heard occasionally and we feel some metaphorical sensation from it. The metaphorical implication involved is: ‘gangsters are animals.’ The predication treatment of the Num Cl construction can handle such discrepancies between Cl and head N as violation of selection restriction, as shown in (23) above. To show the close semantic agreement relation between Cl and head N, we may want to place Cl in Spec of NP, while placing Num(erat) in D of DP, leaving Spec of DP for a choice-function variable or ‘specific’ Num, as in Reinhart (1997). Otherwise, we can treat Num as a modifier (predicate) of an N, as already discussed, solving
semantic well-formedness by type shifting.

As for the plural marker, it is more natural with animate beings (e.g., ai-tul ‘child-Pl’ *CHILD) basically in the interpretation of plural sum (Link 1983, Landman 1996). However, singular nouns can denote plural individuals, often inviting a kind/group (or collective). For some nouns, sub-sort reading is found in inanimate and animate beings (e.g., kaku-tul ‘different sorts of furniture,’ kacok-tul ‘family members’). However, the plural marker -tul is not or rarely attached to the head noun in the post-nominal Num + Cl construction, as shown in (19 and 20), but it is slightly better when attached to a human noun in the prenominal Num + Cl construction like yel myong-ay sonyo-tul ‘ten Cl-Poss girl-Pl.’ A distributive reading is always possible with the post-/pre-N Num + Cl construction lacking the plural marker. To ascertain distributivity, however, a distributive marker like -ssik (distributive marker) or kakkak ‘each’ is attached, as follows:

(24) sonyen tu myong-i norae-rul se kok -ssik pul-ess-ta
    boy  2 Cl  Nom  song Acc 3 Cl  distr  sang
    ‘Two boys each sang three songs’/’Two boys iterated singing three songs at one time.’

It must be noted that the distributive marker -ssik is only attached to Num Cl or Num. The plural marker associated with the human subject can be copied to any constituent of a sentence, as in the following:

(25) nohi -tul oso -tul tul-o-tul w-a sul han can-tul ha-ko-tul ka
    you -Pl quickly-Pl enter-Pl and wine 1 glass-Pl drink-and-Pl go
    ‘You come in quickly and drink a glass of wine and go.’

In consequence, Pl(ural) is attached to the singular expression of ‘one glass’ like [sul ‘wine’ han ‘one’ can ‘glass’ -tul] and likewise to a mass noun like ‘wine’ and ‘water,’ e.g., sul-tul ‘wine-Pl,’ mul-tul ‘water-Pl’ in the surface, contrary to Kang (1994). Intuitively, the plural marker comes from the plurality of the subject argument but it might be rather based on the plurality of function or event, as hinted by Keenan and Faltz (1985).

Another post-N construction is the so-called ‘floated’ numeral construction, i.e. N + Case Marker + Num + Cl. Floating from within the structural case markers but not locative/dative markers is free. Dative is a little better than locative. Observe:

(26) sakwa -ka/-rul /*-e se kae.
    apple Nom/Acc Loc/Dat three Cl

Still another possibility is the construction of [Num + N], which is not
productive and the kinds of N that can take this form are rather limited to things that are very familiar to us in every day life. This type of construction is similar to the one in English, though without plural agreement. Observe:

(27) a. han saram ‘one person’
   b. tu cip ‘two houses’ or rather ‘two households’
   c. *se pyol ‘three stars’ (cf. pyol se kae ‘star three Cl’; pyol set ‘star three’ = three stars)

In (27), the function of the classifier and that of the head N can be said to have been merged. Chinese happens to show the pre-N [Num + Cl + N] construction, and Japanese has all the three varieties corresponding to the Korean counterparts (19-20, 26). Japanese, however, cannot add a double case marker (the same Nom/Acc case marker as the head N) to the right of (26). In (33b), if ‘house’ but not ‘household’ is meant, the normal [N + Num + Cl] construction, cip tu chae ‘house two Cl,’ must be used. A few collective nouns such as nara ‘country’ and tongne ‘village’ can take the [Num + Cl] construction, and some container names such as glass, bottle, bowl, spoon, sack, and cup are employed as measure unit classifiers. Then, a question arises as to whether the construction [Num + N] can be distinguished from [Num + Cl]. The former is limited to certain nouns, and measure unit classifiers coming from container names are exclusively used as classifiers, not as container names, when occurring with numerals as in han can ‘one glass’ (sometimes a glass of wine or an event of drinking of wine metonymically) and tu pyong ‘two bottles (of something).’ In (27c), pyol set ‘star three = three stars,’ without the classifier kae following, is all right, though not so often used. Here, I would say that Cl is elliptical in limited contexts, rather than optional, viewing the fact that se kae-tu pyol ‘three Cl-Poss(of) star’ is felicitous, whereas ses-tu pyol ‘three-Poss star’ is impossible along with (27c) itself (hana-tu pyol ‘one Cl-Poss(of) star’ is exceptionally or idiomatically OK in the appositive sense, applying only to hana ‘one’ but not higher numerals). Furthermore, its ellipsis is possible with numerals for low numbers such as ‘one,’ ‘two,’ ‘three’ and maybe up to ‘four,’ i.e., hana, tul, set, net, but not easy with numerals for higher numbers. So, pyol paek ‘one hundred stars,’ and *pyol yong ‘zero star.’ The classifier kae must follow them such as pyol sumu-kae ‘star 20-Cl.’

In Korean, scrambling a [numeral + classifier] from the direct or indirect object position to the front is forbidden in general, but in Japanese, scrambling one from the object position is allowed, whereas scrambling one from the indirect position is not (‘floating’ from there is neither allowed). Consider:

(28) *se kwon haksan -i chaek -ul sa-ass -ta
      3 volume student  Nom book  Acc bought
      ‘A student bought three books.’
Some unaccusativity effect of numeral classifier floating appears in Korean, as follows (Lee 1989):

(31) yoca se myong -i/ yoca -ka se myong pawi -rul mil -oss -ta
    woman 3 person Nom woman Nom 3 person rock Acc pushed
    ‘Three women pushed a rock.’

(32) yoca se myong -i/ yoca -ka se myong tochak -ha -yoss -ta
    woman 3 person Nom woman Nom 3 person arrived
    ‘Three women arrived.’

The numeral classifier ‘se myong’ cannot be floated from the subject NP of an agentive verb, either transitive or intransitive, as in (31), but it can from the subject of an unaccusative verb, as in (32). The plural marker -tul, forbidden before floating, can occasionally follow the common noun before the nominative marker after floating to give the sum individuals effect. I claim that floating is a sort of adverbialization, getting closer to the verbal part and the operation should be easier with unaccusative verbs, which are more closely related to the surface subject or underlying object. Floating from the object position has no such effect. A similar effect was observed in Japanese (Miyagawa 1989).

3.2 Specificity

Then, what would be the consequences of positional variations with respect to information status in (19), (20), (26) and (27)? First of all, the Numeral+Cl ‘floated’ construction (32), i.e., N+Nom/Acc Num+Cl, is nonspecific in information status. If the speaker (but not the addressee) has some individual(s) represented by an expression in mind, the expression is specific but indefinite. This notion is either semantically or pragmatically treated but has some syntactic relevance. Consider:

(33) a. Joe-nun Sue -eke chaek -ul tu kwon ponae -oss -ta
    Top to book Acc two Cl send Past Dec
    ‘Joe sent Sue two books.’

   b. ‘Joe-nun chaek -ul tu kwon Sue -eke ponae -oss -ta
Top book Acc two Cl to send Past Dec

‘Joe sent two books to Sue.’

c. Joe-nun chaek tu kwon-ul Sue-eke ponae -oss -ta

Top book two Cl Acc to send Past Dec

‘Joe sent two books to Sue.’

The expression which has the ‘floated’ numeral classifier in (33a) is non-specific and, therefore, its scrambling toward the sentence initial position as in (33b) is not optimal. Scrambling it to the front of the sentence is worse. As already mentioned, the floated numeral classifier is closer to the verb, the final element of the sentence, and is focal in the pre-verbal default focus position.

Secondly, the NP containing its internal post-N numeral classifier as in (20) is rather specific in information status. Therefore, it can be scrambled to the left from the preverbal position. Scrambling it to the initial position of the sentence is also quite all right. If the head noun is definite explicitly or contextually and the numeral classifier part is specific by being appositive to the head noun, the Num Cl cannot float out of the whole NP.

(34) ku chaek tu kwon (-i/ -ul)
    the book two Cl Nom Acc
‘the/those two books’

(35) ku chaek -ul tu kwon sa-ass -ta\(^2\)
    the book two Cl bought
‘(I) bought two copies of the book (same type/subspecies)’

(36) sul tu can -ul/ sul -ul tu can mashi -oss -ta
    wine two glass Acc wine Acc two glass drank
‘I drank two glasses of wine.’

(37) ku sul -ul tu can / ku sul tu can -ul mashi -oss -ta
    the wine Acc two glass the wine two glass Acc drank
‘I drank two glasses of the wine.’

If we examine the above examples carefully, we can notice that the head noun

\(^2\) There is a similar construction which behaves a little differently and reflects part-whole relations, as follows:

Mary-nun sommul-lo pat-un chaek-ul kyou tu page ilk-oss-ta

Top gift as receive book Acc only two read

‘Mary read only two pages of the book she received as a gift.

‘The head noun here must be definite or specific to form a part-whole inalienable relation between it and its part. It is different from ‘kul[writing] myo[a few]
cul[line]’=a few lines of writing. The head N ‘kul’ is regarded as mass.
can be definitized (and therefore quantized) and its part can be quantized by numeral classifier expressions. The relation can be either appositional (equal) or part-whole. In other words, the part can be the whole or a proper subpart of the whole. If the head noun is definite (or specific) likewise in any language, numeral quantization is subpart quantization.

If we pay more attention to indefinite head noun expressions, we can notice that the head noun is plural-marked in non-classifier plural-marking languages like English Atomic individuals are paid more attention (as in ‘three apples’). However, in classifier languages, the head noun is bare and singular-marked, being regarded perhaps more as mass than as atomic individuals or their groups. We can have not only ‘sakwa se kae’[apple+3+Cl] but also ‘sakwa se sangca’[apple+3+box] and ‘sakwa se kwan’[apple+3+Measure]. The latter two have the original reading of mass because of the measure phrases. But they don’t necessarily denote pure apple meat contained. The numeral classifier expression and the numeral measure unit expression have no difference in structure and they may be analogously conceptualized. The only difference is that countables have their own inherent individuating bounds. We can exploit this feature by asking ‘How many apples are in the three boxes?’ in Korean, English and other languages. ‘Rice’ is mass if its grains on some occasion are not counted or measured by weight or bulk units. As long as they are not quantized, they behave as mass and form atelic or nontelic activity or process when associated with verbs in aspect and event structure, as we have seen previously. Therefore, Chierchia’s (1997) suggestion that mass nouns come out of the lexicon already pluralized’ (‘furniture,’ ‘water’) might be Indo-European-based. If not quantized, things are mass, even if they are mereologically structured or structurable. For more than one kind, rather than pieces of the same kind, of ‘furniture,’ the plural marker can be attached optionally in Korean: *kaku-tul* (‘furniture-Pl’), but for ‘water,’ only a floated plural marker can be attached, if it were not different kinds of water. ‘Furniture’ is a super-level category kind encompassing lower basic-level category kinds like ‘chair,’ ‘table,’ and ‘lamp,’ and is a little different from ‘substance’ mass. It is also a little different from mass that has naturally and potentially individuated parts like hair and rice.

Things in the world must be masses initially and, on the basis of measurability and individuation/countability, measure or cardinality function applies and numeral.

4. More on Semantic Constraints

Some classifiers denote object individuals but some other classifiers can denote kinds and still other classifiers denote events. We already saw some different kinds of classifier examples. Let us take some additional examples as well:
(38) a. Individuals: kae [small-sized objects], mari [animal], myong [person], thol [grain]
   b. Measure: li [distance], kwon [weight], sanga [box], pyong [bottle], can [glass], caru [sack]

(39) Kinds: kaci [kind], congryu, cong [kind]

(40) Events: pon [iteration/frequency], thonghwa, thong [phone call], hoe [iteration], kkon
     [achievement, love affair], thang [bad deed as a goal], thok [treat], charae [turn]

Some similar taxonomy was also observed in a single common noun in Chinese
that can be associated with different classifiers by Ahrens and Huang (1996).
Their example was ‘telephone’ and similar semantic types can be found in
Korean, as illustrated below:

(41) conhwa tu tae wo telephones’ (machinery)
    telephone two Cl[tool, machinery, artifact]

(42) conhwa tu kaci wo kinds of telephones’
    telephone two Cl[kind]

(43) conhwa tu thonghwa wo phone calls’
    telephone two Cl[event]

Event types can be predicted by the telic (purpose) feature of the original
artifact ‘telephone’ (Pustejovsky 1995). The numeral classifier in (43) simply
reflects the frequency of the arbitrarily bounded events. So, we can further
quantize the spatio-temporal restrictions by means of time and distance units,
e.g. (conhwa) samsip-pun thonghwa ‘thirty-minute call’ or cang-kori conhwa
‘long distance call,’ which constitute sub-kinds of the event classifier thonghwa.
The expression han kkon olli-ta ‘succeed in a pick-up/round-up’ is often used to
denote one successful pick-up or round-up.

The kind-individual relation is not easy to capture and numeral classifier
quantization is normally based upon common-sensically or folk taxonomically
homogenous intra-kind/species, not upon heterogenous inter-kind/species
membership. For instance, even if the classifier mari applies to all the kinds of
animals ranging from a bacteria to a horse, we do not or rarely apply the
numeral classifier [Numeral mari] to a mixture of cows and cats and bacteria.
We apply it to cows, cats, or bacteria, separately in Korean and in any other
conceivable languages. So, we can say that paekma yol mari ‘10 white horses’
and mal yol mari ‘10 horses’ are quite different. The latter can include horses
belonging to different subspecies of the same intra-horse species, but not the
former. This way, Kung-sun Lung’s paradoxical pai ma fei ma ‘white horses
are not horses’ may be justified in one reading in which ‘white horse’ is a
subkind of a horse. The classifier may be represented, in a way analogous to
how a measure expression was represented (Krifka 1997), but with the head
predicate [subkind animal] particularly noted, as:

\[(44) \, mari = \lambda n \, \{x \mid \text{Card}(x) = n \wedge \text{animal}_{\text{subkind}}(x)\}\]

The part ‘\text{Card}(x)’ here refers to the number of atomic individuals belonging to
some animal subkind, say, ‘horse.’ A classifier alone is totally ungrammatical
(45a), and it must be combined with a numeral. Still, a Num-Cl expression
alone is incomplete or anomalous in meaning if a matching common noun is
missing and it must be combined with a common noun compatible with the
classifier from the context for an interpretation. Observe:

\[(45)\]  
\[a. \quad \text{yoki mari -ka} \quad \text{iss} \quad \text{-ta} \]  
\[\text{here Cl Nom exist Dec} \]  
\[\text{‘Here is mari.’} \]  
\[b. \quad \text{yoki (mal) se mari -ka iss} \quad \text{-ta} \]  
\[\text{here horse three Cl Nom exist Dec} \]  
\[\text{‘Here are three mari (horses).’} \]  
\[c. \quad \text{yoki ses-i} \quad \text{iss} \quad \text{-ta} \]  
\[\text{here 3 Nom exist Dec} \]  

A sentence with the expression se mari, Num + Cl, without an associated
common noun such as mal ‘horse’ is simply elliptical and semantically and
syntactically all right in an appropriate discourse context. Such a Num + Cl
expression can function as an indefinite pronoun, e.g., mal-i manh-ass-nun-te,
han mari-ka opso-ci-oss-ta ‘There were many horses, but one disappeared.’ But
unlike in English (this one), demonstrative + Cl is impossible, e.g., *i mari ‘this
Cl.’ On the other hand, a bare numeral with neither a classifier nor a common
noun (just like three lifted fingers) may be understood in an appropriate
discourse context but is even syntactically incomplete in Korean. It sounds like
baby syntax. Even in English, a numeral alone for a classifier-requiring noun is
syntactically incomplete; if someone says, ‘how me some specimens of rice,’”
the response “Here are three” is inadequate (Sean Fulop, p.c.). The wh-word
corresponding to how many is a single word myot in Korean and this word takes
exactly the in-situ position of the numeral in the construction whether it is
floated or not. Observe:

\[(46)\]  
\[\text{no sakwa -rul myot kae (-rul) mok -oss} \quad \text{-ni?} \]  
\[\text{you apple Acc how many Cl Acc eat Past Q} \]  
\[\text{‘How many apples did you eat?’} \]
All the natural animal sub-kinds such as horses and bacteria and their sub-species except human beings take the classifier mari. Thus, for all x in domain, if x is a horse, then x belongs to a non-human animal sub-kind. The classifier kuru, on the other hand, applies only to trees. So trees and kuru are co-extensive. Therefore, it appears common noun denotations are subsets of classifier denotations. But common nouns denoting human beings take not a unique classifier but different classifiers (myong ‘non-honorific’ vs. pun ‘honorific’) depending on the social honorific relations between the speaker and the referent of the common noun employed (22). The semantics of natural kinds can be relevant here to a certain extent, but more importantly folk-taxonomic criteria are enforced. This is where language and thought meet. How people view and classify things of the world is reflected in their language in a cognitively interesting way (e.g., one and the same classifier covers women, fire, and dangerous things in an aboriginal Australian language and the same classifier for paper is applied to bricks such as pyoktol han cang ‘one sheet of brick’ in Korean [(mechanically) flat, rectangular, fairly thin], the same classifier tiao applied to ‘rope,’ ‘snake,’ ‘pants,’ and ‘dish’ in Chinese [long, thin, cylindrical, and flexible]). I would claim that it is easier and simpler to perceive all the natural and artifact kinds as mass initially, and to differentiate between masses and identify members of one mass as belonging to the same mass, the notion of category/kind should evolve. The mass ‘hair’ should be easier to perceive than the count ‘hairs.’ That may be why children acquire the form of mass, i.e., bare common noun with no plurality marking first cross-linguistically. In Korean, the originally quantity-describing adjective manh-ta ‘much’ is applied to both mass quantity and count. Perceptually, humans are reported to be perceived as cylinders initially, and then gross parts like limbs, and then their minor parts like hands and then fingers are gradually perceived.

Mass is analogous to activities as count is to telic events (achievement or accomplishment). Mass as opposed to count is continue and its ingredients are not bounded and not individuated. Similarly activities as opposed to telic events are basically continue and not bounded. In children acquisition of tense, the present tense is associated with activities/states and the past tense that tends to be perfective is pre-dominantly associated with telic events cross-linguistically. Children acquire bare nouns first in any language, and acquire rather late determiner and plural in Indo-European and numeral classifier in Korean and other classifier languages. Children begin to be able to use the general classifier kae ‘small object’ at about two in Korean. Similar general classifiers are also first acquired in other classifier languages. Children make errors of applying classifiers confusedly to mass and count, and make other selection errors before they become seven (K. Lee 1997). Those who acquire Korean in a non-Korean speaking society such as the U.S. typically make an error of applying the general classifier to all possible things and being not sensitive to different classifiers even in adulthood. It is not easy to internalize the Num-Cl
construction in one's grammar, even though the numeral part (and the general classifier) is acquired. A child can be said to acquire mass/kind common nouns and gradually develop ways of applying a measure/cardinality function in relation to the thematic roles of different verbs and tense that require such.

The following examples of kind-individual interaction well justify the need of dual quantification for the same NP. Consider:

(47) khokkiri -ka (ta) kurim -ul kuri -ne
    elephant Nom even drawing Acc draw Surprise
    ‘Even an elephant is drawing a picture, I am surprised.’

(48) *khokkiri han mari -ka (ta) kurim -ul kuri -ne
    elephant Nom even drawing Acc draw Surprise
    ‘Even an elephant is drawing a picture, I am surprised.’

The surprise marker can be applied to the perception of the individual instance of an elephant drawing a picture. And the exhaustion (completive) marker can be applied to the elephant kind as a whole, say, in contrast with mankind, as in (47), but not to any numeral classifier construction as in (48). The elephant-kind is least likely to draw a picture but an elephant is in fact drawing a picture to my surprise. We may be able to apply some type shifting to handle the relations among e(individual extension), GQ<<e,t>,t>, and <e,t>(predicative), as suggested by Partee et al (1987) and Chierchia (1997), but without introduction of Topic-Comment information structure and presupposition, it cannot be as fruitful and explanatory as expected. (47) has a Topic sentence presupposition such that 'An elephant does not draw a picture,' of which the predicate is an individual-level one and the subject, then, necessarily becomes a generic Topic. Because of the spontaneous surprise marker -ne, (48) as a whole is a focus unit with the stage-level predicate on the surface. But the bare common noun subject functioning as an instance of an elephant at the perceptual level is interwoven with the generic (kind) Topic denotation in the presuppositional structure. This is not simply a matter of contrastive focus. (48) cannot be generically associated. In (48), if ta 'even' is placed after kurim-ul 'picture-Acc,' the sentence sounds better in the interpretation of 'one elephant is even drawing a picture,' talking about a particular elephant doing picture-drawing among other things.

In the subject position, the numeral classifier construction, as in (48), contributes to 'counting' the subject referent, but it does not constitute a 'quantized' or 'incremental' argument with activity-accomplishment underspecification verbs. However, if the object gets a Num-Cl expression in (48), like kurim han phok-ul 'one piece of picture,' the event readily becomes telic and the object is limited to the use of 'quantized' argument. A kind-individual duality can also occur in the object position with some bare singular common noun, as can be observed in a sentence similar to 'A child riding even on an elephant, surprisingly.'
However, Num-Cl subjects of achievement (and passive accomplishment) verbs constitute quantized arguments, as shown below:

(49) thokki tu mari-ka se sikan man-e/tongan-e cuk-oss-ta  
    rabbit two Cl Nom three hour in in died  
    'Two rabbits died in three hours.'

(50) thokki tu mari-ka se si-e cuk-oss-ta  
    rabbit two Cl Nom three o'clock at died  
    'Two rabbits died at three o'clock.'

Sentence (49) implies the time span from the onset of causation (or waiting or watching) to the terminal point of death and sentence (50) has the time point expression, which, I believe, is an instance of the time-span 'in' expression. Therefore, the time point expression is more compatible with perfective aspect/quantized event than with activity/process ('He was dying at three o'clock'). Unlike sentences with a quantized object argument, typically representing an accomplishment, consisting of a causation process and a resulting terminal point, sentences with a quantized subject argument can have a starting point, if specified, different from that of a causing process in the accomplishment.

Let us consider the referential use and the conditional use of the following example sentences of Topic (Lee 1996):

(51) tu namnyo -nun tinggul-ko iss-ta  
    two man-woman Top rolling around  
    'The two, man and woman, are rolling around.'

(52) tu namnyo -nun nul haengbokha-ta  
    two man-woman Top always blessed  
    'The two, man and woman, are always blessed.'

The utterance of (51) is about a (visible) couple who are rolling around and it can be preceded by the demonstrative co ‘that.’ The Topic here is referentially used. The predicate is a stage-level one and, therefore, the Topic tends to be contrastive. The utterance of (52), on the other hand, can be about a particular couple referentially or about any couple consisting of man and woman. In the latter conditional sense, if the condition of having a couple of man and wife is met generically, the couple is always blessed. The relation between ‘two’ and ‘man-woman’ is appositional here. There is an arbitrary coupling operation involved for this Topic sense. Because of the variable binding, conditional sense, some quantificational or modal expression usually accompany the utterance. For instance, (52) can have haengbokha-ke marion-i-ta ‘is destined to be blessed’ as its predicate instead. Li’s (1996) examples in Chinese that require modals seem to be all this type of Topic constructions. Otherwise, if a numeral
classifier expression gets the topic marker -nun in Korean, whether sentence initially or mid-sententially, the expression necessarily becomes a Contrastive Topic because of the focal force on number expressions.

If focus markers such as -to ‘even’ and -man ‘only’ interact with number expressions, interesting implicatures arise. One of the most interesting phenomena is the occurrence of negative polarity, based on quantificational scale, as illustrated below:

(53) a. kaemi -ka han mari -to op-ta
   ant Nom one Cl even not exist
   ‘There is not even a single ant.’ ‘There are no ants.’

b. kaemi han mari -to op-ta
   ant one Cl even not exist
   ‘There is not even an ant,’ i.e. ‘There is no trace of humans.’

The scope of the focus inducer of concession -to ‘even’ in (53a) is the numeral classifier and particularly the ‘floated’ numeral is extraordinarily stressed. The nominative marker is attached to the common noun and the utterance is concerned with the common noun denotation ‘ant.’ The quantificational scale is a number scale on which 1 is the lowest. This indefinite Num-Cl + to functions as an NPI, requiring a negative licensor. However, the scope of the same focus inducer in (53b) is the whole NP and the common noun head is stressed. Then, the utterance triggers a quantificational scale on which the lowest (most easily expected) is the existence of an ant. If even such a small and trivial animal cannot be found, then a large and important creature like a human cannot be found either. If -to ‘even’ is attached to a common noun without Num-Cl directly it cannot constitute an NPI. If -to ‘even’ is replaced by the Nom marker -ka in (53a), its meaning becomes ‘ant is missing.’ If a dative common noun gets a focus inducer -man ‘only,’ -to ‘also,’ or -nun ‘as for’ (contrastive) a Num-Cl can follow it as a result of ‘floating,’ even though ‘floating’ is not possible with dative otherwise, as yoca[woman]-eke[Dat]-man[only] se[3] -myong[Cl]-eke[Dat] ‘only to three women.’

5. Quantized/non-quantized Theme with Aspecreal Verb.

When an aspectual verb like ‘begin’ takes an NP directly as its complement, the NP must be definite interestingly and otherwise it is claimed to be ungrammatical, as cited in the following (Pustejovsky and Bouillon 1994):

(54) a. John began the cheese/*cheese. John began the book/*books

b. Jean a commence’ le fromage/*du fromage. Jean a commence’ le libre/*des libles.

(53) a. kaemi -ka han mari -to op-ta
   ant  Nom one Cl even not exist
   ‘There is not even a single ant.’ ‘There are no ants.’

b. kaemi han mari -to op-ta
   ant one Cl even not exist
   ‘There is not even an ant,’ i.e. ‘There is no trace of humans.’

The scope of the focus inducer of concession -to ‘even’ in (53a) is the numeral classifier and particularly the ‘floated’ numeral is extraordinarily stressed. The nominative marker is attached to the common noun and the utterance is concerned with the common noun denotation ‘ant.’ The quantificational scale is a number scale on which 1 is the lowest. This indefinite Num-Cl + to functions as an NPI, requiring a negative licensor. However, the scope of the same focus inducer in (53b) is the whole NP and the common noun head is stressed. Then, the utterance triggers a quantificational scale on which the lowest (most easily expected) is the existence of an ant. If even such a small and trivial animal cannot be found, then a large and important creature like a human cannot be found either. If -to ‘even’ is attached to a common noun without Num-Cl directly it cannot constitute an NPI. If -to ‘even’ is replaced by the Nom marker -ka in (53a), its meaning becomes ‘ant is missing.’ If a dative common noun gets a focus inducer -man ‘only,’ -to ‘also,’ or -nun ‘as for’ (contrastive) a Num-Cl can follow it as a result of ‘floating,’ even though ‘floating’ is not possible with dative otherwise, as yoca[woman]-eke[Dat]-man[only] se[3] -myong[Cl]-eke[Dat] ‘only to three women.’

5. Quantized/non-quantized Theme with Aspecreal Verb.

When an aspectual verb like ‘begin’ takes an NP directly as its complement, the NP must be definite interestingly and otherwise it is claimed to be ungrammatical, as cited in the following (Pustejovsky and Bouillon 1994):

(54) a. John began the cheese/*cheese. John began the book/*books

b. Jean a commence’ le fromage/*du fromage. Jean a commence’ le libre/*des libles.
In addition to the observed restriction, numerical quantization as in ‘Mary began a novel’ and ‘Mary got into a piece of cheese’ is possible. And, therefore, the condition must be generalized as ‘quantized’ rather than definite and it is my further claim that if an ‘indefinite’ bare plural or mass object occurs in a proper context the event associated becomes generic or habitual for a due interpretation. Consider:

(55) a. The baby began solids at six months. The child began milk at two.
   b. The youngster began alcohol/cigarettes early enough.

Then, the question arises as regards to whether the same constraint applies to article-less languages with numeral classifier constructions and definite NPs with a zero-form definiteness. Let us consider:

(56) Joe -nun pap -ul manhi mok -ki/ sul -ul manhi masi-ki sicakhae-ss -ta
    Top rice Acc much eat Comp wine Acc much drink Comp began
    ‘Joe began to eat rice/to drink wine in a large quantity.’
(57) Joe -nun sul/tambae/sosol -ul sicakhae-ss -ta
    Top wine/cigarette/novel Acc began
    ‘Joe began the wine/wine//the(a) cigarette/cigarettes//the(a) novel/novels.’
(58) haksaeng-tul -un sul han/tu pyong -ul sicakhae-ss -ta
    student Pl Top wine one/two bottle Acc began
    ‘The students began the one/the two//a bottle/two bottles of wine.’

In (56), the ‘begin to V’ expression takes an indefinite NP as the object of the verb, with a modifying vague quantity expression. In this case, the interpretation of the whole VP is generic or habitual. The VP cannot represent a particular event. In (57), the object NPs can be either zero-form definite or indefinite generic and in their latter interpretation the Agent begins the respective activity for the first time. In (58), we can notice the numeral classifier NP constructions as objects of an aspectual verb, which must be quantized in interpretation, being definite, whether specific or non-specific. A non-Topic numeral expression denotes a particular event and the Agents of (58) began a plural number (‘two’) of bottles at the same time in one event. The same holds in English.

If so, what to do about underspecification (or ambiguity) with respect to quantization in articleless languages? In English, definiteness quantization and numerical quantization are marked by the definite article and numerals including ‘a’ (=one), respectively. But in articleless languages, bare nominals can be interpreted as definite, singular (as one), or indefinite plural/group. Quantization is effectively marked by numeral classifier expressions. The
semantics of the expressions involved in ‘begin a novel’ can be shown as follows, as previously done (cf. Mineur & Buitelaar 1996):

\[
\begin{align*}
(59) & \textbf{a. begin/sicakha-}: (\lambda Rx.R, y, \exists e. \text{begin}(e, x, y)) \\
& \textbf{b. a novel/sosol (han kwon)}: (\lambda P, \exists z. \text{novel}(z) \land P(z)) \\
& \textbf{c. begin a novel/sosol-ul sicakha-}: (\lambda x. (\exists z. \text{novel}(z) \land (\exists e. \text{begin}(e, x, z)))
\end{align*}
\]

The telic (purpose) sub-sort event of a novel can be read/write (so, read, write \( \leq \) event). The interpretation is supposed to be ‘begin to read/write a novel.’ However, by the existential quantifier alone, the crucial aspectual distinction between ‘a novel’/sosol han kwon ‘a volume of novel’ and ‘novels’ is not clearly made. The quantized expression ‘a novel’ must include a numerical concept as in (40): something like \( \text{Card}(x)=\text{one} \land \text{novel}(x) \) or \( \text{one}(x) \land \text{novel}(x) \),’ with a numeral as a predicate and it must be associated with a sortal unit in a classifier language, and further, it must be associated with a telic event \( (\text{e}_{\text{tel}}) \) but not with an atelic event. The kind-like generic/habitual reading, with an ‘indefinite’ bare plural in English and a bare common noun singular in Korean, ‘novels’/sosol, must be associated with the latter telic. Existential bares occur differently. Even fuzzy numerals like \( \text{tuo}(-kae) \) ‘two or three,’ \( \text{sono}(-kae) \) ‘three or four’ behave as quantized arguments. It is distinct from vague quantity adverbials/adjectivals such as \( \text{manhi/much} \) and \( \text{manhun/much/many} \), as follows:

\[
\begin{align*}
(60) & \text{I drank much water/I ate many apples for/} \text{in ten minutes.} \\
(61) & \text{I Top 10 minute in/ for much water/apple Acc ate} \\
& \text{‘(Lit.) I ate much water/apple in/ for ten minutes.’}
\end{align*}
\]

7. Concluding Remarks

We have discussed numeral classifier patterns in terms of their syntactic, semantic and pragmatic properties. We could see particularly how their quantization function is crucial in event, aspect and argument structure. Ontologically, mass/kind seems basic and measure/count is possible via a measure/cardinality function. Aspectually, the ‘delimiting,’ ‘strictly bounding’ or ‘measuring out’ function of quantization is crucial and is correlated with telicity in change of state. On the other hand, informationally, information on number is rather novel and tends to lean toward the verbal part rather than toward the Topic part. The notions of specificity and Topic must be considered for possible changes of information status and even semantic type. Further elaboration on semantic and pragmatic/cognitive constraints on numeral classifier constructions and their relations to quantization is required.
References


Li, Audrey (1996) ‘Number Projection,’ ms. USC.


-