

COMPARATIVE STUDY ON KOREAN AND INDONESIAN NOUN PHRASE INCLUDING NUMERAL CLASSIFIER

Prihantoro* **, Ae-Lim Ahn*, Jee-Sun Nam*

*DICORA/ Hankuk University of Foreign Studies, Korea, **Universitas Diponegoro, Indonesia
prihantoro2001@yahoo.com, aelimahn@gmail.com, namjs@hufs.ac.kr

1. Introduction

Speakers from different language might attend reality in different way. Lucy (1992) experimented on English and Yucatec speakers. One of the experiment questions is how they classify a cardboard. Different answers from different speakers were obtained. While the classification of English speakers for cupboard is material based, Yucatec speakers opted for shape-based classification. It indicates that speakers from different language might classify referent to different class of nouns. This can be seen clearly in Yucatec for the categorization device (something like 'piece') of noun referent is shown on surface level.

Some languages have particular linguistic devices to classify the entities represented by nouns according to the nature, the number, the shape, the location or other inherent semantic features these entities. The range of noun categorization frameworks and the degree of semantic notion vary from one language to another. We refer to this device as classifier.

In linguistics, there are several goals of classifier studies. It ranges from universal and typological study of classifier¹ to language-specific² ones. Another type of research is carried out in more computational perspective. However, in this paper we aim at comparing the characteristics of the numeral classifier between Korean and Indonesian languages. We here discuss the similarity and the difference of the main features of noun phrases including classifiers in both languages.

This paper is organized in the following order. First chapter described background, type of research and subject languages of our research. In chapter two, we highlight types of classifier, and found out that Korean and Indonesian classifier fall to the same category. The category might be the same, but there are differences on the lexical properties and syntactic pattern of classifier constructions. We discuss properties of lexical units that compose classifier construction: numeral, classifier itself, and noun in chapter three. Chapter 4 focuses on the pattern of classifier construction. Summary and perspective for further research is available on conclusion, the last chapter of our research.

2. Type of Classifiers

In the first chapter, I have described briefly how Lucy observed how speaker from two different languages who exactly classify same noun into different category. Noun category itself is an interesting discussion in classifier studies for different noun category might trigger the selection of different classifier³. The degree of semantic motivation might be different from one language to another. That is why each language might have different number of noun category, and different number of classifier. We might consider that study of classifier is a first step to research more on noun category.

Before we go further, let us consider a definition of classifier proposed by Allan. He conducted survey on more than fifty languages and come with these two requirements for a categorization device to be called classifier (Allan 1997:285)

¹ Some of typological studies of classifiers include Allan (1997), Kiyomi (1992), Croft (1994), Jenks (2010)

² Some of language specific studies of classifiers include Marnita (1996) on Minang Language, IIDA (1999) on Japanese, Oh (1994) Korean

³ Spitulnik (1989: 207), in Brown (2006) observed that in Bantu, a language in Africa, there are more than ten classes of nouns. Some of them are mentioned here: human, animal, masses, plants, fruits, long objects, small objects, abstract qualities etc.

- (a) They occur as morphemes in surface structures under specifiable conditions;
 (b) They have meaning, in the sense that a classifier denotes some salient perceived or imputed characteristic of the entity to which an associated noun refers (or may refer).

The requirements indicate that a classifier must explicitly appear in surface form as a morpheme to specify the semantic properties of noun referent. These are the foundation to call a categorization device classifier in all languages. However, the morphological and syntactic realization of classifier might vary from one language to another. Consider the following examples:

Example 1. Various morphological and syntactic construction of classifier in different languages
 Bantu (Collins, 1962)

- (1) *Vi-su vi-dogo vi-wili hi-vi* 'vi+ knife vi+ small vi+two vi+this' '=these two small knives'

Thai (Allan, 1977: 287)

- (2) *ma' si tua* 'dog four body' '=four dogs'

Navajo (Hoijer, 1945)

- (3) *beeso si Pq* 'money perfect lie (of round entity)' '=A coin is lying (there).'
beesb si nil 'money perfect lie (of collection)' '=Some money (small change) is lying (there)'

Dew (Martins, 1994: 53) in Aikhenvald (2003)

- (4) *xoo-ked* canoe-IN:HOLLOW '=in a canoe'
naax-pis-mi 'water-small-IN:LIQUID' 'i=n a small river'

We can observe that the morphological and syntactic construction of classifier varies from one language to another. In Bantu, *vi-* classifier is realized as bound form as well as Navajo and Dew. While in Thai, the classifier is realized as free morpheme inside numeral phrase. Another interesting phenomenon is observed In Navajo and Dew. The classifiers are not part of modifier in noun phrase. Instead, classifiers are attached on different constituent: verb (Navajo) and locative expressions (Dew). This indicates that classifiers come in different guises. For this reason, attempts have been made to make the typology of classifiers. One of them is by Allan (1977: 286-288) who categorized classifier into four different classes⁴. According to his study, Korean and Indonesian are both specified as a numeral classifier language⁵, similar to Thai. This type is called numeral for it usually quantifies number of nouns specified by the classifier.

Example 2. Korean and Indonesian Classifier

Indonesian

- (1) *Dua orang presiden* two CL[HUM] president '= two presidents'
 (2) *Tiga ekor sapi* three CL[ANM] cow '= three cows'

Korean

- (3) *학생 두 명 (haksaeng tu myong⁶)* student two CL[HUM] '=two students'
 (4) *돼지 다섯 마리 (doi-ji tu mari)* pig Five CL[ANM] '=five cows'

From the examples, we can see that both Indonesian and Korean have specific classifiers to mark the amount of the specified noun referents. In Indonesian, *orang*, which is a classifier of a living human noun, is not appropriate to classify animals, like *myung* in Korean, which is not appropriate for classifying animals. However, we notice the lexical and syntactic differences of the constituents in these noun phrases including classifiers in both languages.

⁴ Classifiers Typology from Allan (1997:286-288): Concordial (example 1.1), numeral (example 1.2), predicate (example 1.3), locative (example 1.4)

⁵ Numeral classifier is known as one of the most common classifier type. There are many studies about typological perspective of classifier. Some of them can be mentioned here: Allan (1997), Kiyomi (1992), Croft (1994), Jenks (2010). There are some modifications of Allan's classifier typology from the rest of the authors, but all of them recognize numeral classifier as the most common classifier type.

⁶ In this paper, romanization of Korean Hangeul is automatically obtained from <http://www.tckerrigan.com/Misc/Korean+romanizer>

3. Lexical Properties

There are three linguistic units related to the discussion of classifier. First is classifier itself. Second is numeral. Third is noun. This tripartite relation is well known by the linguists⁷ working on language with numeral classifier.

3.1 Comparison of Numeral Systems

3.1.1 Obligatoriness of Numeral

Numeral⁸ is contingent property of numeral classifier. A numeral classifier cannot appear to the surface with just noun, excluding numeral. This applies for both Korean and Indonesian. Please consider the following examples:

Example 3. Obligatoriness of Numeral

(1)	<i>dua carik kertas</i>	two CL:Sheet paper	'=two sheets of paper'	(Indonesian)
(2)	* <i>carik kertas</i>	CL:Sheet paper	'=paper'	
(3)	<i>종이 한 장 jongi du jang</i>	paper two CL:Sheet	'two sheets of paper'	(Korean)
(4)	* <i>종이 장</i>	paper CL:Sheet	'=paper'	

We observed from the above examples that 'paper' as a noun referent is specified by classifier *carik* and *jang* as classifier, and numeral 'two'. However, when the numeral is omitted, the noun phrases become ungrammatical. This once again attests numeral's obligatoriness to appear in phrases with numeral classifier.

3.1.2. Mono and Multi-Numeral Sets

Numeral set is one important aspect in the discussion of numeral classifiers. Some languages have only one numeral set, but some other languages might have more. There is only one numeral system in Indonesian, which is used for all classifiers and related nouns. In contrast, there are two different numeral systems in Korean (Ihm, 2001: 89-95): Native Korean (NK) numeral system and Sino-Korean (SK) numeral system, which is adopted from Chinese. Native Korean numeral set is limited to 99, and Sino-Korean can reach *췌* (*hae*) 10²⁰. Most nouns co-occur with only one numeral system, although in very restricted context it can take both. Please consider the following examples:

Example 4. Representation of Two Numeral Sets in Korean Noun Phrase with Classifier

(1)	<i>한국인 삼 인 (hangugin sam in)</i>	korean three[SK] CL:Person	'=three Koreans (human)'
(2)	<i>선생님 세 명 (seonsaengnim se myeong)</i>	teacher three[NK] CL:Person	'=three teachers'

We can see from the above examples that two types of numerals can be used to specify nouns with the same category [+Human]. This phenomenon does not exist in Indonesian. In natural conversation or legal document, only one numeral system is used. There are other limited numerals in Indonesian, but the contexts of use are very limited. Some of them originally come from Old Sanskrit⁹, but these numerals are used only in slogans, poet, old sayings, proverbs or named entities. Please refer to the examples below:

Example 5. Use of Numeral in Natural and Specific Context of Indonesian Language

(1)	<i>Hasta brata</i>	eight (loanword) wisdom	'=eight wisdom'
(2)	<i>Delapan buah kebijaksanaan</i>	eight CL:Gen wisdom	'=eight wisdom'
(3)	<i>Tunggal putra</i>	one(loanword) man	'=men's single'
(4)	<i>Seorang lelaki</i>	one-CL:Human male	'=one man'

Example (1) is an old saying in Indonesian, where it originally comes from Old Sanskrit. It indicates eight kinds of wisdoms that a leader must have. The word *brata* in example (1) is from Old Sanskrit as well. However, when we

⁷ Greenberg (1975:28) includes quantifiers, for it includes also numerals. However, we need to consider that not all non-numeral quantifier can co-occur with classifier.

⁸ The discussion in this chapter refer to numeral not in Arabian number, but in word form

⁹ Few examples are *eka* 'one', *dwi* 'two', *tri* 'three', *sapta* 'four', *panca* 'five' and etc.

just want to refer to eight wisdoms in natural conversation, example (2) is more preferred. Numeral *tunggal*, which means ‘one’ in Example (3) is very restricted in use: badminton, tennis, squash (all of them are sports with racquet). When we want to refer to ‘one man’ in natural conversation, example (4) is more preferred.

3.1.3. Numeral as Head Noun and Noun Modifier

Numeral, either as head noun or noun modifier in Indonesian is similar in its form. There is no distinction between them. Please refer to the example below from Indonesian:

Example 6. Identical Form of Numeral as Noun and Modifier in Indonesian

- (1) *Satu tambah satu sama dengan dua* one plus one same with two ‘=one plus one is two’
 (2) *Satu sisir pisang* one CL:Banana Cluster banana ‘=one cluster of banana’

In example (6), we observed two representations of numeral. Example (6.1) is representation of numeral as pure noun, and example (6.2) as modifier of noun referent banana. We can see here that they are not distinct. Both of them use *satu* to represent ‘one’ as noun or modifier. However, this is not the case in Korean

Native Korean numerals take different forms when they appear as pure noun and modifier of noun. It does not apply for all number, but just on particular number involving 1, 2, 3, 4 and 20. In classifier construction, Korean numeral takes form as modifier. Therefore, this distinction must be taken into consideration. Please consider the following examples:

Example 7. Distinct Form of Numeral as Noun and Modifier in Korean

- (1) *Hana-neun sutja ida* one number is ‘one is a number’
 (2) 책 한 권 (caeg han gweon) book one CL:Book ‘one book’
 (3) *책 하나 권

On the above examples, *hana* can only be used as a pure noun. When it takes role as modifier, it must be changed into modifier *han*, which is used in classifier construction. Maintaining noun form of numeral in classifier construction will make the construction become ungrammatical as in example (3). However, when numeral conveys anaphoric function, noun form can be taken into consideration. Please refer to the following example:

Example 7. Anaphoric Function of Numeral

- A: *Seonsengnimi myeot myeong-I isseoyo* teacher how many CL:person EXT ‘how many teachers are there?’
 B: *hana* one ‘one (person)’
 C: *anio, tu myeongi isseoyo* no, two CL:person EXT ‘no, there are two’
 D: *Taseot myeong-I seonsengnim-I isseoyo* five CL:person teacher EXT ‘there are five teachers’

We observe that in (7B), the respond from A’s question is only *hana*, which means ‘one’. The word numeral in this context carries anaphoric function. In the above discourse, *hana* as B’s response is preceded by A’s question which already carries classifier construction. Therefore, in the reply, B drops both the classifier and head noun. The problem is, a modifier cannot be independent. It must co-occur with head noun. Therefore, the surface form is changed from modifier to noun.

3.1.4 Morphological Realization of Numeral

There are two morphological realizations of numerals. One as free morpheme and another as bound morpheme. A free morpheme can be independent, and this is the morphological realization of Korean numeral set both native and Sino Korean. Unlike Korean numerals, which are all in free form, there is bound form in Indonesian numerals. The bound numeral is expressed by prefix *se-* for numbers one. For example, the bound form *se-* is used for the number one, it must be attached to classifier like in “*se-orang guru*(one-CL:person teacher =a teacher). The discussion of numeral lexical properties can be summarized as follow:

Table 1a. Comparison of Numeral System, Function and Type in Korean and Indonesian

	System	Function	Type	
		Noun & Modifier	Free	Bound
Indonesian	One	indistinguishable	V	V
Korean	Native Korean System	Some are distinguishable	V	-
	Sino Korean	Indistinguishable	V	-

Table 1b. Numerals from 1-10 in Korean and Indonesian

		1	2	3	4	5	6	7	8	9	10
Native Korean	Noun	<i>hana</i>	<i>dul</i>	<i>ses</i>	<i>nes</i>	<i>taseot</i>	<i>yeosot</i>	<i>ilgop</i>	<i>yeotol</i>	<i>ahop</i>	<i>yeol</i>
	Modifier	<i>han</i>	<i>du</i>	<i>se</i>	<i>ne</i>	SAME					
Sino Korean		<i>il</i>	<i>i</i>	<i>sam</i>	<i>sa</i>	<i>o</i>	<i>yuk</i>	<i>sip</i>	<i>pal</i>	<i>gu</i>	<i>sip</i>
Indonesian		<i>satu</i>	<i>dua</i>	<i>tiga</i>	<i>empat</i>	<i>lima</i>	<i>enam</i>	<i>tujuh</i>	<i>delapan</i>	<i>sembilan</i>	<i>sepuluh</i>

3.2 Comparison of Classifiers

3.2.1 Distinguishing Classifier and Non-Classifier

3.2.1.1 Classifier and Measure Terms

A problem in language with numeral classifier is usually on how to distinguish classifier and measure terms. Some linguists such as Becker (1975) and T'sou (1973) use the term 'numeral classifier' more restrictedly to refer to items like 'head' in the following example:

Example 8. Numeral Classifier for Animal on Burmese
pe voq so 3 head dog 'three dogs'

Classifiers may be distinguished from measure terms in terms of the individualizing function (Unterbeck, 1994:368). This individualizing function belongs only to classifiers and not to measure terms. Here, the term numeral classifier is not used for other items such as measure words such as 'litre'. Only lexemes semantically functioning to individuate the noun being counted are considered numeral classifiers. However, some other linguists, such as Burling (1965) and Nguyen Din Hoa (1957), use the term classifier to refer to all lexical items that occur in the classifier slot or which are adjacent to numerals. It means that measure words such as 'litre' are included.

I agree with Unterbeck, that classifiers and measure terms must be distinguished. Because, measure terms do not specify inherent features of a referent. A book can be measured by using various measure terms such as length, weight, volume and etc, but the quantity can be individuated only by classifier. We can test them by using two examples below from Indonesian and Korean.

- Example 9. Distinguishing Numeral Classifier and Measure Terms in Korean and Indonesian
- (1) *Tiga buah apel* three CL:Gen apple 'three apples' (Indonesian)
 - (2) *tiga kilo apel* three MT:Kilogram apple 'three kilograms of apple'
 - (3) *sagwa sam gae* three CL:Gen apple 'three apples' (Korean)
 - (4) *sagwa sam killo* three MT:Kilogram apple 'three kilograms of apple'

In example (9.1 & 9.3), we observe that apples are individuated into THREE apple fruits by classifier for general things [Gen], meaning that there are really three apples. However, measure terms *kilo* (9.2) and *killo* (9.4) does not individuate apples. There can be three, four, or more apples in 'three kilogram of apples'

3.2.1.2 Classifier and Compound Noun

The problem of distinguishing classifier and compound is identified in Minang (Marnita, 1996). What is meant by compound noun is two nouns compounded into one, such as 'glass house', 'pencil case', 'plastic toy' etc.

Compound nouns are basically noun phrase composed of nouns. In [N N] compound noun, one noun functions as head, and another functions as modifier [N_{MODIFIER} N_{HEAD}] or [N_{HEAD} N_{MODIFIER}]. First structure, can lead to confusion, because, surface representation N_{MODIFIER} can sometimes fill classifier slot. However, I believe that they are distinguishable. Let us consider the example below from Indonesian:

Example 10. Distinguishing Compound Noun

1. Dua <i>tangkai</i> mawar	two stem rose	‘two rose stems’
2. Dua <i>tangkai</i> mawar	two CL:Flower rose	‘two roses’
3. <i>tangkai</i> mawar-nya di-potong	stem rose-POSS PASS-cut	‘have the rose stem cut’

Observe *tangkai* in Indonesian from above examples. Although the lexical item and syntactical position of *tangkai* in (10.1) is similar to classifier (10.2), they are very different. *Tangkai* in example (10.1) is not classifier, but it is a modifier for noun *mawar*. The meaning is also different. Example (3) indicates two roses (flower and the stem altogether) but example (4) refers only to the stem of roses, without the flower. Example (10.3) might make it clearer. The sentence is in the passive form, but it is actually a causative sentence requiring someone to cut (only) the stem of the rose.

3.2.2 Morphological Realization of Classifier

In both Korean and Indonesian, morphological realizations¹⁰ of numeral classifiers are free morpheme. None of the numeral classifiers that we found in two languages are in bound morpheme (affix). However, it is interesting to notice that some numeral classifiers either in Indonesian or in Korean are dependent on head noun, and some of them can be independent. Let us consider the examples below:

Example 11. Numeral Classifier Dependency of Noun

1. <i>Secarik kertas</i>	one-CL:Sheet letter	‘one letter’	(Indonesian)
2. <i>carik</i>	vice of village chief	‘vice of village chief’	
3. <i>koyangi han mari</i>	cat one CL:Anm	‘one cat’	(Korean)
4. <i>*mari-ka itta</i>	animal	‘animal’	
5. <i>Mekju han pyeong</i>	beer one bottle	‘one bottle of beer’	
6. <i>Pyeong-I itta</i>	bottle-TOP EXT	‘there is one bottle’	

Some classifiers in free form are bound to noun (not very independent), while other can be very independent. In example (11.1 & 5), Indonesian *carik* and Korean *pyeong* numeral classifiers are free morphologically. When they are not in classifier construction, they can take grammatical role as pure noun.

If we notice, Indonesian numeral classifier *carik* in Indonesian example (11.2) lost the classifier meaning and function, while in example (11.5) *pyeong* lost classifier function, but still carries same meaning as a noun. This does not apply to all numeral classifiers. Some numeral classifiers are bound word class. *Mari* in Korean example (11.4) is one of them. Outside classifier construction, it becomes ungrammatical.

3.2.3 Broad Categorization of Numeral Classifier

In terms of individuating function, both Korean classifiers can be sub-grouped into three broad classes (Oh, 1994: 101). Numeral classifier in Indonesian can also be treated in the same way. First class individualizes noun as a whole, like when classifying a person. We refer to this as ‘singular classifier’. Second class of classifier individuates just part of whole noun. An example of this is a classifier for parts of wood, bamboo, food etc. We refer to this classifier as ‘partial classifier’. Third class of classifier is called ‘multiple classifiers’ that classify several nouns into one group, like ‘pair’. Please consider the following table:

¹⁰ Other than free and bound morpheme, some languages have fusion classifiers. This classifier is fused with numerals (mostly). In result, different noun category, require different numeral set as in Kusaeian (Lee, 1975). In Kusaeian language, there are two sets of numeral. Set A is used in counting fishes, insects, four-legged animals, plants, means of transportation, and long, pointed objects. Set B is used for everything else. These numerals are morphologically unanalysable and can be considered suppletive (like ‘go’ ‘went’ and ‘gone’ in English)

Table 2. Individuating Function of Numeral Classifier in Korean and Indonesian

	Single	Partial	multiple
Korean	<i>Myeong</i> [+person] <i>Mari</i> [+animal]	<i>Seethe</i> [+set] <i>Kyeolle</i> [+pair]	<i>Tapal</i> [+bundle] <i>Ttey</i> [+group]
Indonesian	<i>Orang</i> [+person] <i>Ekor</i> [+animal]	<i>Potong</i> [+cut things] <i>Iris</i> [+sliced things]	<i>Ikut</i> [+bunch] <i>Kelompok</i> [+group]

Even though these two languages share the above similar properties, we observe there are some differences between them. We list some of them in the following section.

3.2.3.1 Social Status Sensitivity

Some of the numeral classifiers in Korean are sensitive to social status. This is quite different from Indonesian where all noun[+Human] take *orang* as numeral classifier. Please consider the following example from Korean and Indonesian.

Example 12. Numeral Classifier Sensitivity to Social Status

1. <i>enam orang presiden</i>	six CL:Hum president	‘six presidents’ (Indonesian)
2. <i>enam orang siswa</i>	six CL:Hum student	‘six students’
3. <i>enam orang penjahat</i>	six CL:Hum criminal	‘six criminals’
4. <i>daetonglyeong yeoseos bun</i>	president six CL:Hum, High	‘six presidents’ (Korean)
5. <i>hakseng yeosot myong</i>	student six CL:Hum, Mid	‘six students’
6. <i>beomin yeoseos nom</i>	six criminal CL:Hum, Low	‘six criminals’

In Indonesian, regardless of social status of human referent, *orang* can be used as in examples (12.1, 2, 3). However, in Korean, there are several classifiers for living human, depending on the social status of the human referent. We observe from Korean examples (12.4, 5, and 6), referent from high, mid and low social status take different classifiers: *bun* (high), *myeong* (mid), *nom* (low)¹¹. I believe that this sensitivity only applies for noun referent in which the inherent property is [+Human, +Animate], and does not apply for [+Animal] and [-Animate]. So far, we have not found any evidence of social sensitive classifiers in Korean or Indonesian.

3.2.3.2 Individual Properties of Classifier

Numeral classifiers in Korean and Indonesian are available in very vast number. Discussing them one by one is regarded not enough in this paper. Therefore, we only took some samples from the existing classifiers in the field of translation.

In the previous chapter we have briefly discussed that classes of noun might be different from one language and another. What classified to be, let us say, class X in one language might be class Y in another language. As consequence, in languages with classifiers, classifier selected by noun referent is also different. This also happens when we try to translate numeral classifier in noun phrase from Korean to Indonesian or from Indonesian to Korean. Consider the following table

Table 3. Classifier Equivalences

	Korean	Indonesian	Referent
1	마리 [+Animal]	Ekor [+Animal]	Chicken, cow, bird
2	명 [+Human, neutral] 분 [+Human, high] 사람 [+Human, +Neutral] 놈 [+Human,+Low]	Orang [+Human]	Teacher, student, president, thief

¹¹ From Sociolinguistic perspective, this is very interesting phenomenon to research, whether the use of these classifiers is influenced only by social status, or other factors as well like: social distance, degree of formality, the existence of third person etc.

3	개 [+GEN]	Sisir [+Banana Cluster]	Banana cluster
4	대 [+vehicles, +machinery]	Buah [+GEN]	Cars, buses
5	개 [+GEN]	Buah [+GEN]	Computer files

As we see, each classifier might have different individual properties. For *다/즈* and *ekor* both of them happen to be classifier for every animal. Either Korean to Indonesian or Indonesian to Korean, the equivalence is perfect. However, if we observe classifier for human, there are several options for Korean. This will pose problem when the source language is Indonesian. The reason is there is only one classifier for human in Indonesian, which is used regardless of who the referent is. In row (3.3) and (3.4) there are also problems for general classifier. Because of the inexistence of classifier, referent must select general classifier[GEN] for all inanimate.

The most challenging part is to translate [GEN] from source to target language. This cannot be carried out instantly without considering the referent for we are not sure what the classifiers must specify. Real classifier might exist, or might not. It is also possible that the translation of [+GEN] in source language is also +GEN, in target language. ‘Computer files’, as an example, takes general classifier both in Korean and Indonesian. This still needs further research, but I notice that in Indonesian, noun referent that are loan words are specified by [+GEN], such as : computer files, car, hardware, software, computer, home run, smash (badminton), goal (soccer) etc.

3.2.3.3 Interactivity of Classifiers’ Properties

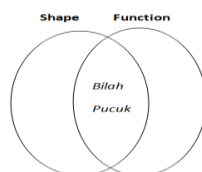
Allan (1997) has proposed seven¹² general properties of classifiers in his previous study, but not all languages use them. It is possible that only few are used. Some languages even take properties outside of what Allan suggested. Adams and Conklin (1973) simplified this into just three: Shape, function and animateness. For Indonesian, I use Adams and Conklin’s properties, plus one property from Allan, which is material.

There are four basic properties of numeral classifiers in Indonesian: Animacy, Shape, Material and Function. For human and animal (Animacy based), these properties are quite clear: *orang* for human, *ekor* for animal, *bujur* for non-living human. Let us turn to another classifier. As an example, *batang*, which is *shape-based numeral classifier*, is opted by different nouns: trees, pencils, cigarette, lollipop candy, ballpoint, cane and etc. although there are various nouns specified by this numeral classifier, the property is clear. *Batang* can be used for everything, which shapes like a stick or bough with different dimension and size. But, there are also some other classifiers, where the properties are overlapping one and each other. It is quite difficult to determine their properties. Consider the following examples and illustration in Indonesian.

Example 13. Interactivity of Numeral Classifiers’ Properties

- | | | | |
|----|-------------------------|--|------------------|
| 1. | <i>Se-bilah bambu</i> | one-CL: [sharp weapon, cutlery] bamboo | ‘one cut bamboo’ |
| 2. | <i>Se-pucuk surat</i> | one-CL: [Letters, long weapon] letter | ‘one letter’ |
| 3. | <i>Se-bilah pedang</i> | one-CL: [sharp weapon, cutlery] sword | ‘one sword’ |
| 4. | <i>Se-pucuk senapan</i> | one-CL: [Letters, fire weapon] letter | ‘one gun’ |

Illustration 1. Interactivity of Numeral Classifiers’ Properties of *Pucuk* and *Bilah* in Indonesian



It is difficult to decide whether the property of *bilah* and *pucuk* are shape-based or function based. *Bilah* in example (13.1) seemed to be shape-based, for one of the literal translations of *bilah* is a long thin object. However, *bilah* in example (11.3) is not only shape based, but also function based. It can be a classifier for everything that can slice

¹² Material, shape, consistency, size, location, arrangement, quanta (Allan, 1977:297)

things, or hurt anybody by stroking, stabbing or other function of sharp weapon. ‘*Bambu runcing*’ which used to be a hand-on-hand combat weapon for Indonesian freedom fighters is also classified by *bilah*.

Pucuk in example (13.2) is more on shape based. *Pucuk* literally mean sharp tip, usually used to refer to the tip of tree leaves. In terms of ‘letter’ *pucuk* is more shape-based. If we look at the shape of letter, there are four sharp tips. But in terms of *pucuk* classifier for gun, we cannot say that they are only shape-based. *Pucuk* there, beside shape-based, is also function-based for it specifies referents firearm.

Korean numeral classifier basic properties and examples of properties interaction are required

3.3 Comparison of Nouns

Aikhenvald (2003:98) mentioned three properties of numeral classifier with regards to noun. First, the choice of classifier by noun is predominantly semantic. We have discussed above that the selection of CL by noun is based on the semantic property on the previous part.

3.3.1 Nouns that has no Classifier

Aikhenvald also mentioned that in language where the CL system is numeral classifier, there is a tendency that not all nouns have CL. It is proven to be right since some nouns has CL in Indonesian, but not in Korea. The opposite case also happens. As an example is ‘book’. This noun has CL in Korean, which is *cheyk*, but not in Indonesian. As an opposite case, ‘banana cluster’ selects *sisir* as CL in Indonesian, but in Korean CL does not exist.

One of the solutions, although does not completely solve the problem, is by using general classifier (CL:Gen). For some nouns that have no CL, they can take default classifier. This default classifier can be limitedly used to co-occur with nouns, which do not have classifier. In Indonesian and Korean, CL:Gen can only replace inanimate objects. And we need to remember that not all CL can be replaced by CL:Gen. This is also the case for Indonesian and Korean. Often, when the real CL is replaced by CL:Gen, the expression becomes unnatural even somehow acceptable.

Example 14. Generic Classifier in Korean and Indonesian

(1) <i>satu carik kertas</i>	one CL:Sheet paper	‘a sheet of paper’
(2) <i>?satu buah kertas</i>	one CL:Gen paper	‘a sheet of paper’
(3) <i>jongi han cang</i>	paper one CL:Sheet	‘a sheet of paper’
(4) <i>?jongi han gae</i>	paper one CL:Gen	‘a sheet of paper’

Examples (13.2 & 4) showed that the preference of CL:Gen caused the expression to be less natural, when appropriate classifier is available (14.1&3). Therefore, we must be really careful about when to select this classifier. There are misperception so far that CL:Gen in Indonesian, which is *buah* can be used to specify all inanimate nouns. By using this example, we have shown that the perception is not completely correct for it cause expressions to be less natural. Another limitation is that, some nouns really do not have any CL. ‘God’ as an example, totally has no CL either in Korean or in Indonesian.

3.3.2 Development of Classifiers from Content Words

Aikhenvald also mentioned that numeral classifiers are different in term of how they are grammaticalized. In Korean and Indonesian, we observed that some nouns select CL, which originated from dependent noun. However, some of the classifiers are realized in independent noun.

In Indonesian, some of CL originates from nouns or verbs such as: *orang* which literally means human (n), or *potong* which literally means ‘to cut’(v). They are free form. But if we look at *carik* (classifier for paper), it must appear with noun (independent). When it appears alone, it signifies another meaning instead of indicating sheet of paper (n). One of the categorizations of Korean nouns we used as a reference is from DECO Database of Korean Lexicon. We found that some of the entries in dependent noun are also found in independent noun database as

classifier. It means that some independent nouns can be functioned as classifier like : *pyeong* bottle, *ken* can, *khop* cup and etc. Most of them are shape-based container. These classifiers also exist in Indonesian and they can stand alone as independent noun.

3.3.3 Mass, Count and Plural Noun Distinction

Noun has inherent property to be mass (non-count) or count. Count noun might take plural form. There are some means to make a count noun to be plural in Indonesian. It might take quantifier, or reduplication. In Korean, besides quantifier a noun might be inflected by plural marker suffix. But when it comes to constructions that involve classifier, reduplication in Indonesian, or plural suffix attachment is avoided. Please refer to the examples below:

Example. Plural Nouns in Construction with Classifier

- | | | | |
|----|---------------------|---------------------|---------------------|
| 1. | Ada pulau-pulau | EXT island-island | 'there are islands' |
| 2. | Delapan buah pulau | eight CL:Gen island | 'one island' |
| 3. | Hakseng-del-ege | student-PLU-to | 'to the students' |
| 4. | Hakseng yeol myeong | student ten CL:Hum | 'ten students' |

We can observe the above examples from Indonesian (1&2), where reduplication (1) of island does not take place again in classifier construction (2). Same method applies for Korean (3&4) where plural suffix attachment (3) is no longer opted in example (4). This proves that numeral classifier construction already indicate a noun to be count noun without having to undergo reduplication or plural suffix attachment. But what about mass noun?

Numeral Classifier for mass noun is considered pseudo-classifiers¹³ for it is not purely classifier. Think about the previous definition that numeral classifier already indicate count noun. In classifier for mass noun, beer for example, numeral classifier tries to individuate mass noun. What it does is it selects content words that can contain, wrap, cover or anything that has individuating function. Consider the examples below

Example. Pseudo-Classifier

- | | | | | |
|----|--------------------------|----------------------|-----------------------|--------------|
| 1. | <i>Sepuluh botol bir</i> | ten CL: Bottle beer | 'ten bottles of beer' | 'Indonesian' |
| 2. | <i>Mekju yeol pyeong</i> | mekju ten CL: Bottle | 'ten bottles of beer' | 'Korean' |

In example (1&2) above, each numeral classifier is considered pseudo. The inherent property of referent in the two examples are liquid, and it can select many classifiers such as glass, cup and bottle. *Botol* in Indonesian and *pyeong* in Korean refer to bottle, where it has individuating function. There are some linguists that consider these kind of containers as numeral classifiers, but some others not. The reason for not including containers into classifiers is because containers work like measure term. If we still agree on our previous definition, measure terms are excluded from classifiers and we must exclude classifiers as well. In some ways it is true that containers work like measure terms, but even so containers also have individuating function (where measure terms don't have this function). Therefore, it is better to think of a continuum where measure terms and numeral classifier stand on different ends, and pseudo-classifiers or containers stand in the middle.

4. Syntactic Constructions

Greenberg (1975:28) mentioned that there are three significant units in the construction of noun phrase with numeral classifier: numeral (Num), classifier (C) and noun (N). We have discussed the lexical properties of each in the previous section. According to Greenberg, there are four possible patterns of these three linguistic units in classifier construction:

1. Num+CL+N
2. N +Num+C
3. CL+Num+N
4. N+CL+Num

¹³ Following Croft (1993)

However, there are some problems with this approach. Allan (1979) argues that there is a universal syntactic constraint in the Classifier Phrase (CP). He points out that a head noun cannot be placed between Num and CL. In this way, Num and CL will always co-occur as a pair, not letting anything come inside. Therefore, Num and CL must be protected by brackets to prevent N (or other linguistic unit) to come inside [Num CL]. If you accept this definition, then simple combinations such as [Num N CL] and [CL N Num] are syntactically not recommended. Obeying this restriction, we only have two basic patterns of numeral classifiers constructions:

1. [[Num CL] N]
2. [N [Num CL]]

4.1 Floating Construction in Korean

Jenks (2010:1) mapped numeral classifiers in East and Southeast Asian languages into some sub-classes in terms of their syntactic diversity. One of the findings is that there are two syntactic structures QN (Quantifier-Noun) and NQ(Noun-Quantifier). They can further be described as:

- a. If a phrase has QN structure then a language has [[Num Clif] N] surface order.
- b. If a phrase has structure [N [Num CL]] surface order, then it has the structure in NQ
- c. Languages with multiple classifiers¹⁴ vary between QN and NQ
- d. Floating construction only appear in (b) structure

He categorized Korean into NQ language, where a floating construction is allowed. This floating construction does not exist in Indonesian. Please refer to the examples below:

Example 15. Floating Construction

- | | |
|--|---|
| <ol style="list-style-type: none"> 1. 말이 어제 3 마리가 다쳤다
<i>mal(horse)-i(nmtf) oje(yesterday) se(3) mari(CL)-ka(nmtf) dachi(to hurt)-ess(Past)-da(Dec)</i>
'=Yesterday, three horses hurt' 2. *tiga kemarin ekor kuda terluk
three yesterday CL:Animal horse hurt
'=Yesterday, three horses hurt' | <p style="text-align: right;">Korean</p> <p style="text-align: right;">Indonesian</p> |
|--|---|

The adverb 'yesterday' comes between noun and numeral (15.1). We tried to insert adverb in Indonesian, but it failed (15.1). It supported Jenks' analysis that floating construction only available in NQ language, not QN. Whereas Indonesian is categorized in QN languages in his system, research that is more refined is required for this topic.

4.2 Permutation and Drop

4.2.1 Permutation

Jenks (2010:2) considers among ten languages in East and Southeast Asia that he surveyed, Korean and Japanese are the exception of the NQ languages, which also allow QN construction¹⁵. For example, it is possible to swap numeral and noun in Korean noun phrases. The phenomenon is called permutation (Hockett, 1996:289). Please consider the following examples in Korean:

- | | | |
|--|------------------|------------|
| Example 16. Permutation in Korean and Indonesian Classifier Construction | | |
| (1) 나무 한 그루 | tree one CL:Tree | 'one tree' |
| (2) 한 그루의 나무 | one CL:Tree tree | 'one tree' |

¹⁴ Most languages has only one type of classifiers. However, In Baniwa and Tariana, there are more than one type of classifiers in use, as it was observed by Aikhenvald (2006)

¹⁵ Oh (1994:28) believed that QN construction is an influence of Chinese. I tested this with native Chinese, and the primary construction is Chinese is QN, where from a test with native Koreans, primary form is NQ

- | | | |
|----------------------------|------------------------|------------|
| (3) <i>sebatang pohon</i> | one-CL:Stick-like | ‘one tree’ |
| (4) <i>?pohon sebatang</i> | noun Num-CL:Stick-like | ‘one tree’ |

We can observe from the above examples (16.1&2) that constructions of noun phrases with numeral classifier in Korean have more flexibility rather than Indonesian. As we can see together, NQ construction in Indonesian is unnatural (Example 16.4). This does not mean NQ construction in Indonesian is not totally prohibited. It is allowed in very limited context.

Example 17. NQ construction in Indonesian

- | | | |
|---|---|---------------------------------|
| (1) <i>Disana ada tiga orang musuh</i> | at-there EXT three CL:Human enemy | ‘there are three enemies there’ |
| | <Emphasis on referent’s existence and its quantity> | |
| (2) <i>Disana ada musuh, tiga orang</i> | at-there EXT enemy three CL:Human | ‘there are three enemies there’ |
| | <Emphasis on referent’s existence> | |

From those two examples, we can observe two functions: emphasis on existence and quantity. NQ construction in example 18.2 covers existence function only. A comma separator must also be used after noun. It is highly recommended to use QN construction as in example 18.1 instead of NQ. The reason is, this construction covers those two functions (Emphasis on noun existence and quantity) altogether. In Indonesian, preference will be weighed more on QN construction.

4.2.2 Drop

Classifier drop is a linguistic phenomenon that is available in both Korean and Indonesian. However, it must be noted that not all nouns allow classifier drops. Consider the following examples.

Example 18. Classifier Drop in Indonesian

- | | | |
|------------------------------------|-----------------------|--------------------------|
| 1. <i>Tiga puluh lembar kertas</i> | thirty CL:Sheet paper | ‘thirty sheets of paper’ |
| 2. <i>*tiga puluh kertas</i> | thirty paper | ‘thirty sheets of paper’ |
| 3. <i>Tiga buah buku</i> | three CL:Gen book | ‘three books’ |
| 4. <i>Tiga buku</i> | three book | ‘three books’ |

In the above Indonesian examples, we can see that classifier drop is only allowable for ‘book’ (example 18.3), but not for ‘paper’ (example 18.1). When specifying quantity of paper, the presence of numeral classifier *lembar* is required. Same cases are also observed in Korean. Consider the following examples:

Example 19. Classifier Drop in Korean

- | | | |
|----------------|----------------------------|-----------------------|
| 1. 집 한 채가 있다 | house one CL:House-TOP EXT | ‘there is one house’ |
| 2. 집이 하나 있다 | house one EXT | ‘there is one house’ |
| 3. 소 다섯 마리가 있다 | cow five CL:Anm-TOP EXT | ‘there are five cows’ |
| 4. ???소 다섯 있다 | cow five EXT | ‘there are five cows’ |

Classifier omission or dropping is allowable, BUT not for all nouns. Dropping classifier for noun ‘house’ is OK (Example 19.2), but it is prohibited for noun ‘cow’ (example 19.4).

Considering the permutation and omission phenomena in Korean there are four possible syntactic structures (or five if floating construction is counted). The syntactic structure in Indonesia is quite strict, by allowing only the structure [[Num CL] N]. If we agree that a classifier can be omitted (or optional), then there will be only two possible syntactic structures (i.e. [[Num CL] N] and [Num N]). However, we should consider that most nouns require an obligatory classifier in both languages. We summarized the possible construction in the table below:

Tabel 4. Possible Pattern of Noun Phrase with Numeral and Classifier

	Floating Construction	Non-Floating Construction	General Pattern of NP with Numeral and Classifier
Korean	N (floating) Num CL	Num(mod) CL N N Num(mod) CL Num(mod) N N Num(N)	NP → N Q Q N Q → Num (CL)
Indonesian	X	Num CL N Num N	NP → QN Q → Num (CL)

5. Conclusion

In this paper, we study the similarity and the difference between the noun phrases including classifiers in Korean and those in Indonesian: we discuss their lexical properties (i.e. numerals, classifiers, and nouns) and their syntactic characteristics. The result of this study can be used as reference in various research areas such as language learning, machine translation or information extraction.

References

- Asher, R., Simpson, J. 1995. *Encyclopedia of Language and Linguistics*. Oxford: Pergamon.
- Akhenvald, A. 2003. *Classifiers: A Typology of Noun Categorization Device*. Oxford University Press: USA.
- Allan, K. 1977. *Classifiers*. *Language* . 53. 285-311.
- Alwi, H. Dardjowidjojo, S. Lapoliwa, H. Moeliono, A. 2003. *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Balai Pustaka.
- Ihm, Ho-Bin. 2001. *Korean Grammar for International Learner*. Yonsei University Press: Seoul.
- Adams, K., Conklin N. 1973. *Towards a Theory of Natural Classification*. *Chicago Linguistic Society* 9. 1-10.
- Chierchia, G. 1998. *Plurality on Mass Nouns and the Notion of ‘Semantic Parameter’*. In S.Rothstein (ed), *Events and Grammar*. Kluwer Academic Publishers, Boston, p 53-103.
- Chung, S. 2000. *On Reference to Kinds in Indonesian*, *Natural Language Semantics*. Vol. 8. Netherlands: Springer.
- Dardjowidjojo, S. 1978. *Sentence Patterns of Indonesian*. Honolulu: University of Hawaii Press.
- Dalrymple, M., Mofu, S. 2008. *Plural semantics, reduplication, and numeral modification in Indonesian*. Handout for the talk presented at AFLA 15. July 2, 2008. University of Sydney.
- Iida, A. 1999. *Descriptive Study of Japanese Classifier (Ph.D Thesis)*. Tokyo: University of Tokyo.
- Nam, J.-S. 2009. *Finite Local Grammars of Korean Time Expressions*, *Journal of Korean Linguistics* 42, 155-191.
- Nam, J.-S. 2010. *Korean Electronic Dictionary DECO-V02. DICORA-TR-2010-01*. DICORA, Hankuk University of Foreign Studies.
- Jenks, P. 2010. *Evidence for the syntactic diversity of numeral classifiers*. Poster in *The 84th Annual Meeting of the Linguistics Society of America*. Baltimore, MD.
- Kridalaksana, H. 2005. *Kelas Kata dalam Bahasa Indonesia*. Gramedia: Jakarta.
- Lyons, J. 1977. *Semantics 2*. Cambridge: Cambridge University Press.
- Marnita, R. 1996. *Classifiers in Miningkabau*. MA thesis. Australian National University, Canberra.
- Oh, S-R. 1994. *Korean Numeral Classifier: Semantic and Universal*. Seoul: Seoul National University Press.
- Sohn, H.-M. 1999. *The Korean Language*. Cambridge: Cambridge University Press.
- Sneddon, J. 2004. *The Indonesian Language: Its History and Role in Modern Society*. UNSW Press. Sydney.
- Hockett, Ch. F. 1966. *Language, Mathematics and Linguistics, Current Trend in Linguistics*. Vol 8. New York: Walter De Gruyter Inc.
- Unterbeck, B. 1994. *Korean Classifier, Theoretical Issues of Korean Linguistics*. Stanford University Press: USA 367-385.