Hungarian is a classifier language

Aniko Csirmaz and Éva Dékány
University of Utah and RIL HAS
acsirmaz@linguistics.utah.edu and dekany.eva@nytud.mta.hu

We present an overview of classifiers, a subgroup of what Raffaele and Masini (this volume) call light nouns. We distinguish three major types: group, sortal and mensural classifiers. Focusing on group and sortal classifiers, we establish a battery of tests which diagnose the membership in the appropriate classifier subgroup. We argue that some of the tests established have universal validity, while the applicability of others depends on language-specific factors. We appeal to these tests to support the claim that Hungarian is a classifier language. We show that Hungarian has the hallmarks of a classifier language indeed, which warrants a treatment similar to the more familiar Southeast Asian classifier languages. As for the category of sortal and group classifiers, we suggest that while the former represent a functional category in the extended projection of the noun, the latter are nouns themselves that take an optional nominal complement. We finally show how the distributional differences between sortal and group classifiers fall out from this proposal.

1. Introduction

The name 'classifier' (CL) is an umbrella term that covers various kinds of lexemes which categorize (classify) nouns into subgroups. This categorization is standardly based on semantic features or properties of the classified items, as opposed to their syntactic or morphological properties. Aikhenvald (2000), a seminal survey of noun categorization devices, distinguishes noun, numeral, verbal, possessive and locative/deictic classifiers. As the names suggest, the different types of classifiers have diverse syntactic-semantic functions and occur in multifarious syntactic environments.

In this article we focus on numeral classifiers; these are classifiers that occur in the context of counting devices such as numerals or quantifiers. The empirical domain of the investigation is the range of numeral classifiers of Hungarian, a Finno-Ugric language. Even though this language has various numeral classifier constructions, they have not yet been subject to detailed investigations.

We will show that Hungarian has three types of numeral classifiers: sortal, group and mensural classifiers. As distinguishing mensural classifiers from sortal ones has a vast literature (see, for instance, Aikhenvald 2000, Grinevald 2000, Borer 2005, Beckwith 2007, Zhang 2009a,b and

references cited therein), in this article we will have very little to say about mensural classifiers, and we will focus on the differences between sortal and group classifiers, instead.

Setting apart sortal and group classifiers based on formal criteria is far from being a trivial task for a language such as Hungarian, as we argue below. It is necessary then to define a battery of tests that reliably distinguish these classifier types; we present these tests in section 4. The conclusion we draw from applying these tests is that sortal classifiers and group classifiers belong to different word classes: sortal classifiers are functional elements that act as satellites of the noun, while group classifiers are nouns themselves. We will also show that the structural position occupied by sortal and group classifiers is distinct, and the position correlates with the word class membership of the classifiers. We will suggest that some of our tests are universally applicable, while others are subject to certain conditions holding in a language, and thus cannot be applied in languages across the board. The classifiers established for Hungarian will also be contrasted with the classifiers identified in standardly accepted `classifier' languages, including South-East Asian languages.

The discussion will proceed as follows. In Section 2 we introduce the three types of numeral classifiers. Motivating the existence of Hungarian sortal classifiers in detail will be a major concern of ours in Section 3. Section 4 addresses the issue of distinguishing sortal and group classifiers, and Section 5 concludes the paper.

2. Three types of classifiers

2.1 Bare nouns denote an undifferentiated mass

Bare nominals in Hungarian are non-atomic; they denote an undifferentiated mass (see Farkas and De Swart 2003). This means that individuals, which are required for the formation of plurals, among others, must be derived in the syntactic component. The addition of a classifier to the nominal lexical item yields an individual. In absence of a classifier the bare noun denotes a mass, as expected:

(1) János level-et írt.
John letter-ACC wrote
"John wrote a letter / letters." (either complete or partial letters)

(2) János bélyeg-et gyűjtött
John stamp-ACC collected
"John collected stamps."

(3) János szendvics-et evett
John sandwich-ACC ate
"John ate a sandwich / sandwiches." (either whole sandwiches or parts)

A bare, lexical noun object of write and eat is multiply ambiguous, as the translation shows. The affected object can be a single letter or sandwich, multiple letters/ sandwiches, or parts of either a single or multiple letters/ sandwiches. These readings are all expected if the noun denotes a mass; the resulting interpretation is vague and it can result in either of the readings described above. Collective predicates such as *collect* (as in 2) require a plural argument. The morphologically singular *bélyeg* `stamp' can appear with this predicate, because the bare noun can refer to a plurality of individuals of the relevant type, i.e. multiple stamps.

For constructions where units are required (including plurals), there must be a way to establish units for the undifferentiated mass that the lexical noun denotes. We assume (following Borer 2005) that classifiers can fulfill this function. It is not surprising then that Hungarian has a range of classifiers, as we show below.

2.2 Classifiers in Hungarian and elsewhere

In Hungarian, three types of classifiers are licensed in the context of a numeral or quantifier. Sortal classifiers combine with count nouns and typically categorize the noun according to shape and size (e.g. as small spherical, extended rigid, long flexible, cf. (Grinevald, 2000)).[1] Group classifiers also occur with count nouns, but while sortal classifiers refer to individual units, group classifiers refer to an assembly of individuals that function together as a unit in some sense (cf. *a flock of sheep, a deck of cards*). Mensural classifiers combine with both mass and count nouns, and they may be independent of the shape and size of the noun they occur with. Aikhenvald (2000) defines them as classifiers "used for measuring units of countable and mass nouns (pg. 115). Typical mensural classifiers name containers (*a box of*) or other canonical measure units (*a kilo of*).[2]

In (4), (6) and (8) we give examples of each type of classifier from various Southeast Asian languages, which are widely recognized to be 'classifier languages'. In (5), (7) and (9) we provide

[1]In English, item and piece could be considered to function as sortal classifiers in phrases such as a piece of furniture/silverware, and item of news, because they combine with ontologically count but syntactically mass nouns, and make them countable (Wiltschko 2012).

[2] More on the the sortal vs mensural CL distinction can be found in Cheng and Sybesma (1998, 1999); Aikhenvald (2000); Borer (2005); on sortal versus group CLs, see Beckwith (1992, 2007); Zhang (2009a).

corresponding Hungarian examples.[3] The different meaning contributions of the different types of classifiers are also illustrated in the minimal triplet from Hungarian given in (10), in which the same noun, *gyógyszer* 'medicine, pill', appears with classifiers of each type.

Sortal CL

(4) yi ke tang
one CL candy
"one candy"
Mandarin Chinese (Zhang, 2007, pg. 50)

(5) egy szem cukor
one CL_{eye} candy
"one piece of candy"

Group CL

(6) yi bao xiangyan
one CL cigarette
"a pack of cigarettes"
Mandarin Chinese (Zhang, 2007, pg. 48)

(7) egy csomó zöldhagyma
a CL_{bunch} green.onion
"a bunch of green onions"
(usually 5, sold as one unit, tied together)

Mensural CL

(8) yāt dihkhyut
one CL blood
"a drop of blood"
Cantonese (Matthews and Yip, 1994, pg. 98)

(9) egy csepp vér
one CL_{drop} blood
"one drop of blood"

(10) a hét szem gyógyszer — sortal Cl
seven CL_{eye} medicine
"seven pills"

b hét levél gyógyszer — group Cl
seven CL_{strip} medicine
"seven strips of pills"

[3] All Hungarian sortal classifiers are homophonous with a noun. Eg. fej means 'head' as a noun and 'big spherical' as a classifier, szál means 'thread' as a noun and 'long thin' as a classifier. We gloss sortal classifiers as CL and give the nominal interpretation in subscripts, eg. CL_{eye}, CL_{thread}.

c hét kanál gyógyszer mensural Cl
seven CL_{spoon} medicine
"seven spoons of medicine/pills"

The three types of numeral classifiers are alike in terms of their external distribution. In Hungarian, every classifier is licensed by numerals (11)-(13), quantifiers (14)-(16) as well as demonstratives (17)-(19).[4] Classifiers cannot co-occur with bare nouns (20)-(22) and they are not licensed when only a definite determiner is present (23)-(25).[5]

Numeral: ✓

(11) három fej saláta sortal CL
three CL_{head}lettuce
"three (heads of) lettuce"[6]

(12) három falka farkas group CL
three CL_{pack}wolf
"three pack of wolves"

(13) három tepsi süti mensural CL
three CL_{pan} pastry
"three pans of pastries"

[4] While numeral classifiers are normally restricted to quantificational contexts indeed, the Hungarian pattern is by no means unique. We cite Greenberg (1972, pg. 36): "The synchronic universal seems to hold that whenever a numeral classifier construction is also used in non-quantifier constructions, the construction with demonstratives is one of these, often the only one". Mandarin and Hungarian are examples of languages that may use a classifier with demonstratives even in the absence of a numeral.

[5] Demonstratives require an overt definite article in Hungarian. Clearly, in these examples it is the presence of the demonstrative that licenses the CL, not the definite article, cf. (23) - (25).

[6] The English translations of some examples may suggest that English is a classifier language, too. The unit word use of English *head, piece*, etc. has been grievously neglected in the literature, and due to space considerations we cannot do justice to this topic in this article. We note that sortal CLs in classifier languages generally take nouns without a linker such as English *of*. Mandarin, for instance, disallows the modification marker *de* between a noun and a sortal Cl but allows it with mensural CLs (Cheng and Sybesma 1999).

Quantifier: ✓

(14) sok fej saláta — sortal CL
many CL_{head}lettuce
"many lettuces"

(15) sok falka farkas — group CL
many CL_{pack}wolf
"many packs of wolves"

(16) sok tepsi süti — mensural CL
many CL_{pan} pastry
"many pans of pastries"

Demonstrative: ✓

(17) az a fej saláta — sortal CL
that the CL_{head}lettuce
"that (head of) lettuce"

(18) az a falka farkas — group CL
that the CL_{pack}wolf
"that pack of wolves"

(19) az a tepsi süti — mensural CL
that the CL_{pan} pastry
"that pan of pastries"

Bare N: ✗

(20) *fej saláta-(k) — sortal CL
CL_{head}lettuce-PL

(21) *falka farkas-(ok) — group CL
CL_{pack} wolf-PL

(22) *tepsi süti-(k) mensural CL
CL_{pan} pastry-PL

Definite article: ×

(23) *a fej saláta-(k) sortal CL
the CL_{head}lettuce-PL
"the (head of) lettuce"

(24) *a falka farkas group CL
the CL_{pack}wolf
"the pack of wolves"

(25) *a tepsi süti mensural Cl
the CL_{pan} pastry
"the pan of pastries"

Hungarian sortal and group classifiers have some properties in common that they don't share with mensural classifiers. Frequently, a selectional restriction is operative between the sortal or group classifier and the lexical noun (cf. *egy szál/*fej/*karika gyertya*, 'one CL_{thread}/*CL_{head}/*CL_{ring} candle', *egy pakli/*falka/*levél kártya* 'one CL_{deck}/*CL_{pack}/*CL_{book} card' for sortal and group classifiers, respectively). Furthermore, both sortal and group CLs combine only with count nouns, and they easily form compounds with the noun they modify (*gyertya-szál* 'candle-CL_{thread}' meaning 'candle', *kártya-pakli* 'card-deck' meaning 'a deck of cards'). Mensural classifiers, in contrast, can combine with mass nouns *két kancsó bor* "two jugs of wine" (lit. two jug wine) and compound formation is often impossible **bor-kancsó* "wine-jug" (cf. *boros kancsó* "wine-ADJ jug").

In light of the similarities between sortal and group CLs, one needs reliable tests to distinguish them on formal grounds, as opposed to the rather vague and semantics-based characterization in terms of denoting an individual unit a vs. set of units. It should be noted that overt plural marking is of no help in distinguishing between these classifiers. As shown in the previous examples, e.g. in (18), the lexical noun has no plural marking in group classifier structures in Hungarian; this is unlike English, where the sortal *a <u>piece</u> of furniture* and the group *a <u>pack</u> of wolves* are easily distinguished by the plural marking on the lexical noun).

The formal criteria that can be used to distinguish sortal CLs and group CLs rely on the internal makeup of phrases containing these classifiers. Before we turn to these formal criteria (section 4), we justify the identification of certain lexemes of Hungarian as sortal CLs in section 3. This is necessary because Hungarian is generally not considered to be a language with sortal classifiers. As we will show, however, the Hungarian sortal classifier system shares crucial similarities with sortal classifier systems of Southeast Asian classifier languages, and potential counter-arguments against positing sortal CLs in Hungarian are weak at best.

3 Motivating the word class of sortal classifiers in Hungarian

An illustrative list of Hungarian sortal classifiers is given in (26). The items in (26 a) have already been identified as classifiers in Beckwith (1992, 2007).

(26) a. fő, kötet, szál, szem, fej, tő gerezd,
Cl_{head} CL_{volume} Cl_{thread} Cl_{eye} Cl_{head} Cl_{stem} CL_{clove}

b. karika, cső, cikk, rúd, bokor, vekni, cserép, csík, darab, rózsa, ív
Cl_{ring} Cl_{tube} $Cl_{article}$ Cl_{rod} Cl_{bush} Cl_{loaf} CL_{pot} Cl_{strip} CL_{piece} Cl_{rose} CL_{sheet}

In this section we motivate treating the items in (26) - and possibly even more - as sortal classifiers and we discuss the similarities and differences between Hungarian and Southeast Asian classifier systems.

The best known examples of classifier languages that make extensive use of sortal classifiers are Southeast Asian (SEA) languages such as Mandarin Chinese, Japanese, Thai, Vietnamese and Burmese (Gil 2008). Not only do these languages share the property of having sortal classifiers, but their classifiers also have a number of important properties in common, as discussed below. We now show which properties of these classifier systems Hungarian shares and which ones are not.

3.1 Similarities in sortal classifier systems

The sortal classifier systems of Hungarian and SEA classifier languages show strong typological parallels. First, in classifier languages nouns may be compatible with more than one classifier, depending on what sort of unit we want to obtain. Thus the Hungarian noun *kukorica* 'sweetcorn' may be classified either by *szem* `eye' or *cső* `tube', giving rise to different types of sweetcorn units, grains

vs. ears of corn (27). A similar example from Cantonese is reproduced from Matthews and Yip (1994: 106) in (28).

(27) a. két szem kukorica
two CL_{eye} sweetcorn
"two grains of sweetcorn"

b. két cső kukorica
two CL_{tube} sweetcorn
"two ears of sweetcorn"

(28) a. nībouh dihnlóuh[7]
this computer
(classified as model)

b. nīga dihnlóuh
this computer
(as a machine)

Second, several CL languages have a generic or general classifier that can appear with nouns not associated with a specific shape-based sortal classifier. Examples include Mandarin Chinese *ge* (as in (30), taken from Cheng and Sybesma, 1998), Japanese *tsu* (for inanimates, Downing, 1996), Korean *kay* (for inanimates, Lee and Ramsey, 2000) and Vietnamese *cái* (Greenberg, 1972). Hungarian also possesses such a classifier: *darab* "piece" is exemplified in (29).

(29) hét darab szó
seven $CL_{generic}$ word
"seven words"

(30) san ge ren
three $CL_{generic}$ people
"three people"

[7] Our source does not parse the determiner and the classifier separately in the glosses. Under the most plausible parse, nī is the demonstrative, while bouh and ga correspond to different classifiers.

Generic classifiers can frequently replace a more specific classifier both in SEA languages and Hungarian. This is illustrated for Hungarian in (31) and for Mandarin Chinese in (32).

(31) két szál / darab rózsa
two CL_{thread} / $CL_{generic}$ rose
"two roses"

(32) 3 zhāng/ ge zhuōzi
3 CL / $CL_{generic}$ table
"three tables"
(Zhang 2009b:8)

Third, sortal classifiers typically classify nouns according to animacy, shape, size and structure. The sortal classifiers of Hungarian also tend to express shape and size, and the interpretation is tied to the meaning of the classifier when used as a lexical noun. For instance, *szem* "eye" is used for small spherical objects (33), *fej* "head" for big spherical objects (34), *szál* "string" for long and thin objects (35), and *karika* "circle" for flat roundish objects. Animacy plays a role in distinguishing the general classifier *darab* "piece" from the more specific *fő* "head": the former can be used with any noun, the latter is specialized for human animates, esp. in regimented situations (36).

(33) egy szem rizs / alma / homok
one CL_{eye} rice / apple / sand
one grain of rice, one apple, one grain of sand

(34) öt fej káposzta / saláta / karfiol
five CL_{head} cabbage / lettuce / cauliflower
five heads of cabbage/lettuce /cauliflower

(35) egy szál gyufa / gyertya / virág / kolbász
one CL_{thread} match / candle / flower / sausage
one match/candle/flower/sausage

(36) a. öt darab könyv / orvos
five $CL_{general}$ book / doctor
five books/doctors

b. öt fő legénység / *könyv
five CL_{human} crew / book
five crew-members/books

Fourth, lexical nouns for body parts and objects with canonical shapes are often used as classifiers for inanimate objects. Typical nouns that become grammaticalized as classifiers include head for big round objects, eye for small spherical objects and thread for long thin objects (Aikhenvald, 2000). Hungarian *fej* "head", *szem* "eye" and *szál* "thread" have become grammaticalized in exactly this way, as shown above.

Finally, classifiers can occur in anaphoric phrases in both Hungarian and SEA languages. Hungarian and Thai examples are given in (37) and (38) respectively (the latter is from Allan, 1977, pg. 286).

(37) a. az a szem gyöngy
that the CL_{eye} pearl
that pearl

b. az a szem
that the CL_{eye}
that one (reference contextually determined; possibly refers to a pearl)

(38) a. maˇ• tua nán
dog CL that
that dog

b. tua nán
CL that
that one

3.2 Contrasts between SEA and Hungarian sortal classifier systems

Still, Hungarian does not show all features of SEA classifier systems. This may cast doubt on the claim that Hungarian has sortal classifiers, and it may warrant an attempt to reanalyze Hungarian classifiers as a member of some other word class. We will show, however, that the use of classifiers exhibits considerable variability in the languages of the world, the Southeast Asian pattern being just one of the possible patterns. While it is true that Hungarian uses sortal classifiers differently from e.g. Mandarin Chinese, Hungarian is neither special nor unique in this respect.

The first major difference between Hungarian sortal classifiers and SEA classifiers concerns the optionality of the classifier. While in SEA languages the use of the classifier is obligatory in quantificational contexts, Hungarian classifiers are optional (we will elaborate on the notion of optionality presently). Only Hungarian examples are provided below.

(39) a. hét (fej) saláta
seven CL_{head}lettuce
"seven lettuces"

b. hét (szem) cukor
seven CL_{eye} candy
"seven pieces of candy"

c. hét (szál) gyertya
seven CL_{thread} candle
"seven candles"

The obligatory nature of the classifier, however, is not nearly as central a feature of classifier languages. In the sample of 400 languages shown in Gil (2008), classifiers are absent from 260 languages, optional in 62 languages, and obligatory in 78 languages. Among the languages with optional classifiers, we find Akatek Mayan (Zavala, 2000), Minangkabau (Aikhenvald, 2000), informal Khmer (Greenberg, 1972 and Allan 1977) and Cambodian (Goral, 1979), to mention just a few. Thus the optionality of *fej*, *szem* and *szál* in (39) does not warrant the conclusion that these words are not classifiers.

Above, 'optional' is used in a purely descriptive sense, meaning that the classifier may or may not have overt phonological realization. Depending on the particular theory, this can be interpreted in a number of ways. For example, optionality can be described as classifiers being optionally overt or covert; as arising from the simple absence of a classifier; or from the existence of a specific,

phonologically empty classifier. As Hungarian nouns have a mass denotation (Section 2), and the very definition of masses is that they need to be partitioned before they can be counted, it follows from compositionality that some element must perform the partitioning function in the absence of an overt classifier. We take this as evidence for a phonologically zero classifier in the language. (We also assume without argument here that the zero classifier is an empty counterpart of the general sortal classifier darab `piece'.) Whether a language possesses a phonologically zero classifier or not depends on the lexicon of the particular language. Hungarian has a zero classifier, therefore its overt classifiers appear to be optional. As Southeast Asian languages require the overt classifier in all count noun phrases, it follows that they cannot have a phonologically zero classifier.

The second point where the Hungarian and SEA sortal classifier systems diverge is the number of classifiers in the language. SEA languages have a large number of classifiers, while the number of sortal classifiers in Hungarian is just a handful. However, the sheer number of classifiers does not show anything significant about the language in question. Languages show great variation in the number of sortal classifiers the utilize. Cebuano has only one classifier (Rijkhoff, 2002), Nung has four, Iwam and Chambri have five each. On the other end of the scale, Vietnamese has approximately 140 and Burmese has around 200 classifiers (Aikhenvald, 2000). Yet all of these languages are characterized as being classifier languages. Given these counts of classifier items in classifier languages, Hungarian - with about twenty sortal classifiers - is far from being at the low end of the scale.

Finally, Hungarian differs from SEA languages significantly in the number of nouns that can be classified with a specific sortal classifier, as opposed to the all-purpose general classifier. Only a fraction of Hungarian nouns can occur with a specific, selected sortal classifier. This is not so in SEA languages, where a noun will more often take a specific classifier than not.

(40) két (*$CL_{specific}$) ceruza
two CL pencil
"two pencils"

It may appear at first blush that most Hungarian nouns are unclassifiable. This is not a correct description of the facts, however. Every Hungarian (non-human) `count' noun can occur with the general classifier *darab* "piece".

(41) két darab ceruza / saláta / cukor / gyerta
two $CL_{generic}$ pencil / lettuce / candy/ candle
"two pencils / heads of lettuce / pieces of candy / candles"

Even in well-established classifier languages, there are nouns that do not occur with specific, selected classifiers. The Mandarin *ren* 'people', for instance, takes only the general classifier *ge* (see example (30)). Furthermore, even SEA sortal classifier languages may have entirely unclassifiable nouns. Allan (1977:286, fn. 2) claims that this holds for a 'large number of nouns' in Burmese and Vietnamese (fn. 2). Additional classifier languages with unclassifiable nouns include Bengali, Omani Arabic and Kana (Aikhenvald, 2000) and Akatek Mayan (Zavala, 2000).

Let us summarize the discussion in this section. We have seen that a number of parallels can be detected between the classifiers of Hungarian and SEA languages. These include the following: a) nouns are compatible with more than one specific classifier, b) the existence and range of use of the general classifier, c) the role of shape, size and animacy in the choice of the classifier, d) the grammaticalization of body parts as classifiers and e) the ability of the classifier to occur in anaphoric phrases. These similarities lend further support to categorizing the lexical items in (26) as sortal classifiers.

While it is true that Hungarian classifiers differ from the classifiers of SEA languages in a number of ways, the differences remain within the normal range of variation attested among classifier languages. Neither the relatively small number and the optionality of classifiers, nor the small number of nouns taking a specific classifier can be taken as evidence against the word class of classifiers in Hungarian. The language remains entirely within the independently attested range of diversity of classifier languages.

4. Distinguishing sortal and group classifiers

We have stated above the group classifiers name groups of individuated units, and that these classifiers occur with morphologically singular lexical nouns. The phrase *három falka farkas* "three pack wolf" (three packs of wolves) served as an illustration in (12). Group classifiers clearly differ from sortal ones in requiring a plurality of individuals, but this difference may be too vague to constitute a reliable diagnostic. We enumerate some diagnostic properties in this section. We first show several contrasts between sortal and group classifiers, and then offer principled reasons for those contrasts.

4.1 Distinguishing classifiers

At first sight, distinguishing the two classifiers is easy: sortal CLs appear with singular, and group CLs with plural nouns. This difference, while transparent in some languages such as English, does not reliably distinguish these classifiers in all languages, including Mandarin or Hungarian. We suggest that universally applicable tests rely on the distinct functions of these classifiers. Following Borer (2005), we assume that in and of themselves all bare nouns denote only a mass, or 'stuff'. 'Stuff' is neither singular nor plural, and as it is not unitized, it is not countable either. Sortal classifiers partition out the lexically given `stuff' and the resulting partitions are countable units. This function requires sortal classifiers to appear between lexical nouns and counting elements (numerals and quantifiers). Group classifiers, in contrast, do not partition inherent `stuff' in the same sense. Rather, they resemble counters because they require a plurality of partitioned units. As a consequence, group classifiers appear structurally higher than the Classifier phrase headed by sortal classifiers (the CLP), which perform the dividing function. In essence, sortal classifiers divide, and group classifiers require a complement that has already been divided (cf. also Beckwith 2007). We will argue that this distinction correlates with the syntactic category of the classifier, and determines its behavior with respect to the diagnostics shown below.

Adjectives and classifiers. Dékány and Csirmaz (2010) show that in Hungarian, the range of adjectives that may precede sortal classifiers is restricted to adjectives located high on the adjective hierarchy of Sproat and Shih (1991), Cinque (1994) and Scott (2002), and that these adjectives are marked in a position after the classifier.

(42) $Adj_{quantification} > Adj_{quantity} > Adj_{size} > Adj_{shape} > Adj_{color} > Adj_{nationality}$ (Cinque 1994)

The property shared by these adjectives is that they require a partitioned element they modify, i.e they are not able to modify an undifferentiated mass.[8] At the same time, adjectives which are low on the hierarchy, such as color adjectives, can only follow sortal classifiers.

(43)

két	nagy	(*zöld)	szem	(*nagy)	zöld	gyöngy
two	large	green	CL_{eye}	big	green	pearl

"two large green pearls"

[8] If they follow a sortal classifier, the noun is coerced into a type or kind reading, and a token reading is not possible.

Group classifiers, on the other hand, can be either preceded or followed by any adjective. Furthermore, adjectives which precede classifiers differ in the element they modify: a size-denoting adjective which precedes a sortal classifier modifies the head noun, while if it precedes a group classifier, the adjective modifies the classifier itself.

(44) két nagy szem gyöngy — sortal CL
two large CL_{eye} pearl
"two large pearls"

(45) két nagy falka kutya — group CL
two large CL_{pack} dog
"two large packs of dogs"

Accordingly, given a size-denoting adjective that precedes the classifier, a contradictory or synonymous adjective after the classifier is only possible with group classifiers (47):

(46) ?? két nagy szem nagy / kis gyöngy — sortal CL
two large CL_{eye} big / small pearl
"two large pearls (of the big / small type)"

(47) két nagy falka nagy / kis kutya — group CL
two large CL_{pack} large / small dog
"two large packs of large / small dogs"

No lexical noun. The second consistent difference between group and sortal classifiers arises in elliptical constructions. If the lexical noun is elided, the meaning of group classifier structures is not significantly affected (48).

(48) két falka kutya ≈ két falka — group CL
two CL_{pack} dog ≈ two Cl_{pack}
"two packs of dogs ≈ two packs"

This does not hold for sortal classifiers, however. With these classifiers, a systematic ambiguity arises in noun ellipsis. To appreciate the nature of the ambiguity, recall that sortal classifiers in Hungarian are always homophonous with a noun, from which they possibly have grammaticalized. In case the classified noun is omitted, the lexeme following the numeral can be interpreted as a sortal classifier that classifies the elided noun (b, first interpretation), or it can receive a literal nominal reading (b, second interpretation).

(49) a. két szem gyöngy sortal CL
two CL_{eye} pearl
"two pearls"

b. két szem
"two ones"(small spherical object, possibly refers to pearl, reference defined by context)
"two eyes"

Classifiers and interpretation. Finally, the lack of an overt classifier does not significantly affect the meaning of a structure with sortal classifiers, but the semantic difference is significant if a group classifier is absent.

(50) a. két szem gyöngy = két gyöngy sortal CL
two CL_{eye} pearl = two pearl
"two pearls = two pearls"

b. két falka kutya ≠ két kutya group CL
two CL_{pack} dog ≠ two dog
"two packs of dogs ≠ two dogs"

4.2 Explaining the contrast

We suggest that the differences noted above follow from the category of the classifiers and from their position within the nominal projections. Specifically, we propose that sortal classifiers are functional elements acting as noun satellites, hence they are not the head of the noun phrase they occur in. Group classifiers, on the other hand, are lexical nouns that act as the head of the noun phrase they occur in, and they embed a noun phrase complement, the head of which is the classified noun. The structure of noun phrases with sortal and group classifiers is schematized in (51), where N stands for noun, F stands

for functional element and ... shows the position of potential adjectival modifiers. This amounts to saying that there is a full range of 'extended nominal projection' between the classified noun and the classifier in (b), or in other words, (b) contains two noun phrases, but (a) contains only one.

(51) a. Sortal: [D [... [F(Cl) [... [N]]]]]
b. Group: [D [... [N(Cl) [... [N]]]]]

The difference in adjectival modification arises because an adjective always modifies the head noun. If group classifiers are nouns - in contrast with sortal classifiers - then the difference follows. At the same time, we assume that for every noun, the full range of adjectival modifiers is possible; these can be seen as parts of the extended nominal projection. Accordingly, the full range of adjectives is predicted to appear between a group classifier and a lexical noun, but sortal classifiers cut the extended nominal projection in two. Given that sortal classifiers yield partitioned units, it is expected that adjectives which modify individual units - including size - are restricted to a position above the classifier. These predictions are borne out for Hungarian, as shown in (43)-(44).

The adjectival modification facts are expected to be universal, because sortal classifiers invariably have a partitioning function and group classifiers universally require elements that are already partitioned. On the assumption that only functional elements can appear within the extended nominal domain, and that only a noun can take a complement that has been partitioned, all languages are predicted to conform to the Hungarian pattern. Thai, for instance, corroborates this expectation (Hundius and Kölver 1983, pp 169-171):

(52) a. nók tua jàj — sortal CL
bird Cl big
"the big bird"

b. nók fŭuN jàj — group CL
bird swarm big
"a large swarm of birds"

c. nók jàj fŭuN jáj — group CL
bird big swarm big
"a large swarm of big birds"

Let us turn to elliptical structures. The constructions in (48) and (49), where the full classifier structure is compared to the construction with an elided noun, corroborates the view that the syntactic category of classifiers is not uniform. A pack of dogs can be described as a *pack*, revealing a similarity in the interpretation of the group classifier and the same word when used as a noun. The content, and arguably the function, of *pack* is the same in both structures. This state of affairs contrasts with sortal classifiers. A *head of lettuce* cannot be described as a *head*; the semantic import of the lexical, contentful noun and the classifier are rather different. This interpretive difference follows from the proposal in (51) that group classifiers are nouns. At the same time, we claim that sortal classifiers are functional elements rather than full nouns. The functional nature is supported by the impoverished, bleached interpretation when compared to the homonymous nouns; the contrast is reminiscent of the difference between restructuring verbs which can also appear as full-blown lexical verbs (see Cinque 2006).

The ambiguity which arises from the ellipsis of sortal classifiers is predicted to be language specific. The fact that the test works reliably in Hungarian stems from the combination of two independent factors: i) sortal classifiers are optional in Hungarian (more precisely, there is a phonologically null classifier) and ii) sortal classifiers are always homophonous with nouns. Thus the output string of noun ellipsis can be parsed in two ways. If the lexeme following the numeral is parsed as a member of the classifier word class, then the whole string is interpreted with the head of the noun phrase unpronounced ('two ones' interpretation of (48b)). On the other hand, if the lexeme following the numeral in the output string is parsed as a member of the noun word class, then it is interpreted as the head of the noun phrase that takes a null classifier ('two eyes' interpretation of (48b)). Thus the test is predicted to be applicable beyond Hungarian if and only if the language in question has a null classifier, and the sortal classifier used in the given example is homophonous with a lexical noun.

Example (50a), where the sortal classifier is omitted, is expected to be language-specific as well. In our account, the omission of the sortal classifier fails to affect interpretation because Hungarian has a null sortal classifier (cf. footnote 7). The lack of equivalence with a group classifier omitted in (50b) is predicted to be universal: the numeral either specifies the cardinality of the groups (if the group classifier is present) or that of partitioned units of the lexical noun.[9] These interpretations are clearly distinct.

[9] This test relies on the assumption that cross-linguistically, there is no null group classifier. We suggest that this has to do with group classifiers being nouns, hence lexical instead of functional elements. Lexical material in general is expected to have overt phonological exponence: we are not aware of examples of null lexical nouns, lexical verbs and adjectives in languages.

5. Conclusion

This paper makes two major claims. First, it was argued that the classifier languages (where classifiers fulfill some kind of individuating function) include Hungaian, a language where classifiers were not assumed to play a major role earlier. We showed that Hungarian shares a number of properties with SEA (sortal) classifier systems, and the contrasts between the two classifier systems are within the range of attested variation for classifier languages. We suggested that the existence of a phonologically null general classifier in Hungarian may be the reason why classifiers were not considered to play a crucial role in Hungarian nominal interpretation earlier. The second major point was identifying a range of test which distinguish sortal and group classifiers. We noted that such tests are necessary because number marking on the head noun does not universally distinguish the two types of classifiers. We noted a number of universal and specific tests, and for the specific tests, we stated the properties of those languages where they are applicable. The diagnostic tests make use of essential, inherent differences between sortal and group classifiers, which we treat as universally valid properties of these items. The wider range of classifier languages and universal properties of classifiers paves the way to a general description and a better understanding of members of this word class.

References

Aikhenvald, Alexandra Y. 2003. *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Allan, Keith. 1977. "Classifiers". *Language* 53.285–311.

Beckwith, Christopher I. 1992. "Classifiers in Hungarian". *Approaches to Hungarian 4: The Structure of Hungarian* ed. by István Kenesei & Csaba Pléh, 197–206. Szeged: JATE.

Beckwith, Christopher I. 2007. *Phoronyms: Classifiers, class nouns and the pseudopartitive construction*. New York: Peter Lang.

Borer, Hagit. 2005. *In Name Only*. (= Structuring Sense, 1.) New York: Oxford University Press.

Cheng, Lisa L.-S., and Rint Sybesma. 1998. "Yi-wan Tang, Yi-ge Tang: Classifiers and Massifiers". *Tsing-Hua Journal of Chinese Studies* 28.385–412.

Cheng, Lisa L.-S., and Rint Sybesma. 1999. "Bare and Not-so-bare Nouns and the Structure of NP". *Linguistic Inquiry* 30.509–542.

Cinque, Guglielmo. 1994. "On the Evidence for Partial N-movement in the Romance DP". *Paths towards Universal Grammar: Studies in Honor of Richard S. Kayne* ed. by Guglielmo Cinque & Jan Koster & Jean-Yves Pollock & Luigi Rizzi & Rafaella Zanuttini, 85-110. Washington, D.C.:

Georgetown University Press.

Cinque, Guglielmo. 2006. *Restructuring and Functional Heads*. (= *The Cartography of Syntactic Structures*, 4.) New York: Oxford University Press.

Dékány, Éva and Aniko Csirmaz. 2010. "Classifiers and the Functional Structure in DPs". Paper presented at FU11: Universal Grammar and the Syntax of Finno-Ugric Languages Workshop, Piliscsaba, Hungary, August 2010.

Downing, Pamela. 1996. *Numeral Classifier Systems: the case of Japanese*. (= *Studies in Discourse and Grammar*, 4.) Amsterdam and Philadelphia: John Benjamins.

Farkas, Donka and Henriëtte de Swart. 2003. *The Semantics of Incorporation*. Stanford, Calif.: CSLI Publications.

Gil, David. 2008. "Numeral Classifiers'. *The World Atlas of Language Structures Online* ed. by Martin Haspelmath & Matthew S. Dryer & David Gil, &Bernard Comrie, chapter 55. Max Planck Digital Library.

Goral, Donald R. 1979. "Numeral Classifier Systems: a Southeast Asian Cross-linguistic Analysis". *Linguistics of the Tibeto-Burman Area* 4.1–72.

Greenberg, Joseph H. 1972. "Numeral Classifiers and Substantival Number: Problems in the Genesis of a Linguistic Type". *Working Papers on Language Universals* 9.1–39.

Grinevald, Colette. 2000. "A Morphosyntactic Typology of Classifiers". *Systems of Nominal Classification* ed. by Gunter Senft, 50–92. Cambridge: Cambridge University Press.

Lee, Iskop, and Robert Ramsey. 2000. *The Korean Language*. Albany: State University of New York Press.

Matthews, Stephen, and Virginia Yip. 1994. *Cantonese: a comprehensive grammar*. London & New York: Routledge.

Simone, Raffaele and Francesca Masini. this volume. "Light Nouns". *Proceedings of Word Classes 2010*, ed. by Simone, Raffaele & Francesca Masini. Amsterdam & Philadelphia: John Benjamins.

Rijkhoff, Jan. 2002. *The Noun Phrase*. Oxford: Oxford University Press.

Scott, Gary J. 2002. "Stacked Adjectival Modification and the Structure of Nominal Phrases". *Functional Structure in DP and IP: The Cartography of Syntactic Structrures* ed. by Guglielmo Cinque, 91-120. New York: Oxford University Press.

Sproat, Richard and Chilin Shih. 1991. "The Cross-linguistic Distribution of Adjective Ordering Restrictions". *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda* ed. by Carol Georgopoulos & Roberta Ishihara, 565—593. Dordrecht: Kulwer Academic Press.

Wiltschko, Martina. 2012. “Decomposing the Mass/Count Distinction: Evidence from Languages that Lack it”. *Count and Mass Across Languages* ed. by Diane Massam, 120-146. Oxford: Oxford University Press.
Available at http://www.linguistics.ubc.ca/sites/default/files/mass_count_wiltschko.pdf

Zavala, Roberto. 2000. “Multiple Classifier Systems in Akatek Mayan”. *Systems of Nominal Classification* ed. by Gunter Senft, 114–146. Cambridge: Cambridge University Press.

Zhang, Hong. 2007. “Numeral Classifiers in Mandarin Chinese”. *Journal of East Asian Linguistics* 16.43– 59.

Zhang, Niina Ning. 2009a. “Counting Unit Words as Nominal Auxiliaries”. Ms., National Chung Cheng Univesity. Available from http://ling.auf.net/lingbuzz/000936

Zhang, Niina Ning. 2009b. “Syntactic Properties of Numeral Classifiers in Mandarin Chinese”. *Proceedings of the 2009. Summer Conference of the International Society for Chomskyan Studies*, 93-108. Seoul: The International Society for Chomskyan Studies.