Hungarian classifiers∗

Aniko Csirmaz and Éva Dékány
University of Utah and University of Tromsø
acsirmaz@linguistics.utah.edu, eva.dekany@uit.no

March 24-26, 2010

1. Introduction

South-East Asian languages are well-known for being classifier languages:

Chinese:

(1)  Wo zhege xueqi shangle san men ke.
    I this semester took three Cl courses
    I took three courses this semester.
    (Zhang, 2007, pg. 52, ex. 23 a)

Hungarian is generally considered to be a non-classifier language. However, Hungarian has numerous classifiers that show striking typological parallels with classifiers in South-East Asian languages:

(2)  Vett-em három szál rózsá-t
    buy.PAST-1sg three CL_thread rose-ACC
    I have bought three roses.

Goals

• Describe the range of classifier constructions and their distribution in Hungarian
• Compare classifier constructions in Hungarian and other, esp. South-East Asian languages
• Provide tests that make a clear-cut between the different types of classifiers
• Give a syntactic analysis of Hungarian classifier constructions

Claims

• Hungarian bare nouns (e.g. könyv ‘book’) are non-atomic, they denote an undifferentiated mass (<e,t>)
• Bare nouns need a sortal Cl to carve them up into individuals (Cl can be overt or covert)
• No sortal Cl in environments where an <e, t> type of predicate is also possible
  (e.g. könyvet olvas ‘read book’ (= read a book / books))

∗This talk reports on work in progress. We would like to thank Tarald Taraldsen, Antonio Fábregas and Edward J. Rubin for helpful discussions and comments. All errors are our own.
2. Bare nouns

2.1. Borer (2005) on the denotation of NP

• The categories ‘mass’ and ‘count’ are not lexical specifications, they correspond to a piece of structure
  Nouns are neither inherently ‘mass’ nor count in and of themselves, they merely denote ‘stuff’

• Nouns must be divided or portioned out before they can be counted
  This dividing out is done by a classifier system, hosted in CIP

• Derived count / mass distinction
  – Absence of dividing structure (bare nouns) → ‘mass’
  – Presence of dividing structure → ‘count’

2.2. Hungarian bare nouns

• Hungarian is fully consistent with bare Ns denoting undivided ‘stuff’

• Bare nouns are restricted to the predicate operator position immediately preceding the verb\(^1\)
  In this position morphologically singular count nouns can have a non-atomic interpretation

\[(3)\] János level-et ír
  John letter-ACC write
  John is writing a letter

\[(4)\] *János ír level-et
  John write letter-ACC
  John is writing a letter
  (OK as a non-neutral sentence, with John in focus)

• Collective predicates require a semantically plural argument
  They can occur with morphologically singular bare Ns:

\[(5)\] János bőlyeg-et gyűjt.
  John stamp-ACC collect
  John collects stamps

• Other predicates are multiply ambiguous between a singular, plural or partitive reading

\[(6)\] Feri szendvics-et eszik
  F-NOM sandwich-ACC eats
  Feri eats (a) sandwich
  ((a) an entire sandwich; (b) part(s) of one or more sandwiches; (c) more than one entire sandwich)

\(^1\)The data and argumentation under the first three points in this section are directly taken from Csirmaz and Szabolcsi (in progress), also cf. Farkas and de Swart (2003).
2.3. Partitioning non-individuated denotation

- Individuation (atoms) required for counting and specific reference
  - Numerals / quantifiers: only atoms can be counted
  - Specificity must be established for atoms; no specific interpretation for bare nouns
- Classifiers required for atomic interpretation
  - *Sortal* classifiers (with ‘count’ nouns only)
    * General, productive classifier: *darab* ‘piece’
      (a productive classifier, some classifiers select a specific noun or a specific range of meaning)
    * Selected classifiers (*borsó* ‘pea’ → *szem* ‘small.spherical’, *gyertya* ‘candle’ → *szál* ‘long.thin’,
      *gyalogos* ‘infantryman’ → *fő* ‘person’, etc)
    * Overt and covert classifiers
  - *Mensural* classifiers (measure unit)
    *(a bucket of water, a pound of meat, a meter of ribbon)*
- Lexical contrasts
  - Lexically non-atomic, ‘count’ (individuation by sortal Cl)
  - Lexically non-atomic, ‘mass’ (individuation by mensural Cl)
  - as a tentative hypothesis, ‘count’ and ‘mass’ nouns differ in how their ‘stuff’ denotation can be partitioned
    * there exist unique conventionalized units for ‘count’ nouns, (e.g. for ‘book’, ‘cup’, all animate individuals)
    * for ‘mass’ nouns, either no conventionalized unit exists (e.g. sand’), or there is no unique conventionalized unit (e.g. coffee’, ‘wine’)
3. Classifiers: General considerations

- Classifier distinction: (a) sortal CL, (b) group CL, (c) mensural CL
  (on the storal CL vs mensural distinction, see Cheng and Sybesma (1998, 1999); Borer (2005); on sortal
  versus group Cls, see Beckwith (1992, 2007); Zhang (2009a))

<table>
<thead>
<tr>
<th>Hungarian</th>
<th>English</th>
<th>South-East Asian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sortal CL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) egy szem cukor</td>
<td>(10) —</td>
<td>(11) yi ke tang one CL candy one candy</td>
</tr>
<tr>
<td>one CL szem candy one candy</td>
<td></td>
<td>Chinese (Zhang, 2007, pg. 50)</td>
</tr>
<tr>
<td>Group CL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12) egy csomó zöldhagyma</td>
<td>(13) a book of needles</td>
<td>(14) yi bao xiangyan one CL cigarette</td>
</tr>
<tr>
<td>a bunch green onion</td>
<td></td>
<td>a pack of cigarettes</td>
</tr>
<tr>
<td>a bunch of green onions</td>
<td></td>
<td>Chinese (Zhang, 2007, pg. 48)</td>
</tr>
<tr>
<td>(usually 5, tied together</td>
<td>and sold as one unit)</td>
<td></td>
</tr>
<tr>
<td>Mensural CL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(15) egy csépp vér</td>
<td>(16) a drop of blood</td>
<td>(17) yāt dīh khyut</td>
</tr>
<tr>
<td>one drop blood one drop of blood</td>
<td></td>
<td>one drop blood</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a drop of blood</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cantonese (Matthews and Yip, 1994, pg. 98)</td>
</tr>
</tbody>
</table>

- Classifiers are licensed by numerals, quantifiers (including how many) or a singular demonstrative pro-
  noun\(^2\)

\(^2\)"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" (Greenberg, 1972, pg. 36). Mandarin
Chinese and Hungarian are examples of languages that require (Mandarin) or allow (Hungarian) a classifier with demonstratives
even in the absence of a numeral."
Licensing classifiers in Hungarian

<table>
<thead>
<tr>
<th>Sortal CL</th>
<th>Group CL</th>
<th>Mensural CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL + bare noun: *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(18) *fej saláta(k)</td>
<td>(19) *falka farkas</td>
<td>(20) *tepsi süti</td>
</tr>
<tr>
<td>CL_head lettuce-pl</td>
<td>pack wolf</td>
<td>pan pastry</td>
</tr>
<tr>
<td>(head of) lettuce</td>
<td>pack of wolves</td>
<td></td>
</tr>
<tr>
<td>Definite article + CL + N: * (definiteness in itself does not license CL)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(21) *a fej saláta(k)</td>
<td>(22) *a falka farkas</td>
<td>(23) *a tepsi süti</td>
</tr>
<tr>
<td>the CL_head lettuce-pl</td>
<td>the pack wolf</td>
<td>the pan pastry</td>
</tr>
<tr>
<td>the (head of) lettuce</td>
<td>the pack of wolves</td>
<td>the pan of pastries</td>
</tr>
<tr>
<td>Numeral + CL + N: ✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(24) három fej saláta</td>
<td>(25) három falka farkas</td>
<td>(26) három tepsi süti</td>
</tr>
<tr>
<td>three CL_head lettuce</td>
<td>three pack wolf</td>
<td>three pan pastry</td>
</tr>
<tr>
<td>the (heads of) lettuce</td>
<td>three pack of wolves</td>
<td>three pans of pastries</td>
</tr>
<tr>
<td>Singular demonstrative + def. article + CL + N(^3): ✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(27) az a fej saláta</td>
<td>(28) az a falka farkas</td>
<td>(29) az a tepsi süti</td>
</tr>
<tr>
<td>that the CL_head lettuce</td>
<td>that the pack wolf</td>
<td>that the pan pastry</td>
</tr>
<tr>
<td>that (head of) lettuce</td>
<td>that pack of wolves</td>
<td>that pan of pastries</td>
</tr>
<tr>
<td>Plural demonstrative + def. article + CL + N(_\text{PL}): ??</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(30) ??az-ok a fej saláta(k)</td>
<td>(31) ??az-ok a falka farkas-ok</td>
<td>(32) ??az-ok a tepsi süti-k</td>
</tr>
<tr>
<td>that-PL the CL_head lettuce-pl</td>
<td>that-PL the pack farkas-ok-PL</td>
<td>three-PL the pan sütí-PL</td>
</tr>
<tr>
<td>salátá-k</td>
<td>farkas-ok</td>
<td>sütí-k</td>
</tr>
<tr>
<td>lettuce-PL</td>
<td>wolf-PL</td>
<td>pastry-PL</td>
</tr>
<tr>
<td>those (heads of) lettuce</td>
<td>those packs of wolves</td>
<td>those packs of pastries</td>
</tr>
<tr>
<td>Quantifier + CL + N: ✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(33) sok fej saláta</td>
<td>(34) sok falka farkas</td>
<td>(35) sok tepsi süti</td>
</tr>
<tr>
<td>many CL_head lettuce</td>
<td>many pack wolf</td>
<td>many pan pastry</td>
</tr>
<tr>
<td>many (heads of) lettuce(s)</td>
<td>many packs of wolves</td>
<td>many pans of pastries</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
- We confine our attention to sortal and group CLs in this talk
- Sortal and group CLs group together; distinguishing them is not trivial (especially in light of the lexical specification of count' nouns in Hungarian and Mandarin)
4. Sortal versus group classifiers

4.1. A bird’s eye view

4.1.1. Tests to tease apart the two types

Differences between the two Cl types in terms of syntactic position and distribution are expected, as sortal and group Cls have a different function:

- Sortal Cls portion out ‘stuff’ so that ‘stuff’ becomes countable; thus theses Cls must be between the projection of ‘stuff’ (N) and counters (numerals, quantifiers)

- Group Cls do not portion out ‘stuff’ they are more similar to counters instead; as they require a (plural) set consisting of already divided stuff; thus these Cls must come on top of the division-performing CIP

Such differences are indeed attested and can be used to tell whether a given Cl is of the sortal or group type. Partially building on Zhang (2009a,b), we present the following tests to distinguish sortal and group Cls:

- Adjectival modification
  - Adjectives predecing the Cl modify (a) the noun with sortal Cls but (b) the Cl itself with group Cls

\[
\begin{align*}
(36) & \quad 2 \text{nagy szem gyöngy} \\
& \quad 2 \text{big CL}_{\text{eye}} \text{pearl} \\
& \quad 2 \text{big pearls}
\end{align*}
\]

\[
\begin{align*}
(37) & \quad 2 \text{nagy falka kutya} \\
& \quad 2 \text{big pack dog} \\
& \quad 2 \text{big packs of dogs}
\end{align*}
\]

- The same adjective or adjectives with contradictory meaning (a) cannot appear with sortal Cls, but (b) can appear with group Cls

\[
\begin{align*}
(38) & \quad *2 \text{nagy szem nagy/kis gyöngy} \\
& \quad 2 \text{big CL}_{\text{eye}} \text{big/small pearl} \\
& \quad 2 \text{big pearls}
\end{align*}
\]

\[
\begin{align*}
(39) & \quad 2 \text{nagy falka nagy/kis kutya} \\
& \quad 2 \text{big pack big/small dog} \\
& \quad 2 \text{big packs of big/small dogs}
\end{align*}
\]

- Numerals and ellipsis\(^4\)

  - The presence or absence of the classified noun makes a significant meaning difference only with (a) sortal Cls, but not with (b) group Cls

\[
\begin{align*}
(40) & \quad 2 \text{szem gyöngy} \neq 2 \text{szem} \\
& \quad 2 \text{CL}_{\text{eye}} \text{pearl} \neq 2 \text{CL}_{\text{eye}} \\
& \quad 2 \text{pearls} \neq 2 \text{eyes}
\end{align*}
\]

\[
\begin{align*}
(41) & \quad 2 \text{falka kutya} \approx 2 \text{falka} \\
& \quad 2 \text{pack dog} \approx 2 \text{pack} \\
& \quad 2 \text{packs of dogs} \approx 2 \text{packs}
\end{align*}
\]

\(^4\)All examples in (40) - (43) have an elliptical interpretation. In these elliptical readings all examples are grammatical and have the same reading as the non-ellipted, full phrase. In this section we are interested only in the non-elliptical interpretations.
4.1.2. Accounting for the facts

- The presence of absence of the Cl (a) does not affect the interpretation with sortal Cls, but (b) it yields a significant meaning change with group Cls

(42) 2 szem gyöngy = 2 gyöngy  
2 Cl_{eye} pearl = 2 pearl  
2 pearls = 2 pearls

(43) 2 falka kutya ≠ 2 kutya  
2 pack dog ≠ 2 dog  
2 packs of dogs ≠ 2 dogs

- Sortal Cls
  - Are not nouns, they belong to the functional category Cl
  - Phrases with a sortal Cl involve one extended NP, that of the classified noun
    \[ NP \ Cl \ ... \ N \ ]

- Group Cls
  - Are nouns
  - Phrases with a group Cl involve two extended NPs: one projected by the group Cl, the other by the noun
    \[ NP \ N_{Cl} [NP \ N] \]

- Adjectives and Cls
  - The adjective facts follow because
    * adjectives always modify the head noun
    * every noun has its own range of adjectives that modify it
  - Sortal and group Cls (subscripts on AP indicate the N modified)
    * Sortal Cl: \[ NP \ (AP_{N}) \ Cl \ ... \ N \ ]
    * Group Cl: \[ NP \ (AP_{N_{cl}}) \ N_{Cl} [NP \ (AP_{N}) \ N] \]
  - We expect that the pattern of adjectival modification holds cross-linguistically
    c.f Thai examples mirror the Hungarian order and show the same pattern with adjective modification (Hundius and Kölver, 1983, pg. 169–171):

(44) nők tua jáj
bird CL big
the big bird

(45) *nők fűnų sű-khiaw
bird swarm green
(swarms cannot be green)

(46) nők fűnų jáj
bird swarm big
a large swarm of birds

(47) nők jáj fűnų jáj
bird big swarm big
a large swarm of big birds

- No overt N head
  - The facts with omission of the classified noun follow because
    * Every phrase must have a head,
    * In (40) szem will be interpreted as the nominal (N) head (and not as a sortal classifier Cl),
* In (41) the group Cl is the (N) head, taking the classified noun as a complement (c.f. eat ≈ eat pizza)

- The facts with omission of the classifier follow because
  * (40) is possible because overt sortal classifiers are optional in Hungarian
    (there exists a covert, phonologically empty sortal classifier)
  * In (41) the classified noun is an optional complement of the group Cl

- We expect the classified noun and classifier omission facts to be language-specific;
  in languages with only overt (sortal) classifiers, the omission in (40) - (43) will lead to ungrammaticality

4.2. Sortal classifiers

Dixon (1982, 1986): isolating languages are likely to have classifiers, inflectional and strongly agglutinative languages tend to have noun classes. This is admittedly only a tendency — one which Hungarian does not follow (agglutinative but has classifiers).

Beckwith (1992, 2007) notes that there are sortal Cls in Hungarian and lists some sortal Cls, but his lists are not exhaustive (only 7 Cls) and he does not discuss the distribution of sortal Cls

\[(48) \text{ fő, kötet, szál, szem, fej, tő, gerezd, karika, cső, cikk, rúd, }\]
\[\text{ CL\_head, CL\_volume, CL\_thread, CL\_eye, CL\_head, CL\_stem, CL\_clove, CL\_ring, CL\_tube, CL\_article, CL\_rod, }\]
\[\text{ bokor, vekni, cserép, csák, darab, rózsa, űv, }\]
\[\text{ CL\_bush, CL\_loaf, CL\_pot, CL\_strip, CL\_piece, CL\_rose, CL\_sheet}\]

4.2.1. The Hungarian sortal Cl system is just like others

- In Cl languages, nouns may be associated with more than one classifier, depending on which property of the noun is in focus

\[(49) \text{ két szem kukorica } \quad \text{ (50) két cső kukorica}\]
\[\text{ two CL\_eye sweetcorn } \quad \text{ two CL\_tube sweetcorn}\]
\[\text{ two grains of sweetcorn } \quad \text{ two ears of sweetcorn}\]

Cantonese (Matthews and Yip, 1994, pg. 106)

\[(51) \text{ nàbou-dînhûh } \quad \text{ (52) nûg diânhûh}\]
\[\text{ this computer (classified as model) } \quad \text{ this computer (as a machine)}\]

Japanese (Beckwith (2007), Naoyuki Yamato, p.c.)

\[(53) \text{ taoru-o go-nai } \quad \text{ (54) taoru-o go-hon}\]
\[\text{ towel-ACC five-CL } \quad \text{ towel-ACC 5-CL(extended cylindrical)}\]
\[\text{ five towels } \quad \text{ five towels (rolled up)}\]

Akatek Mayan (Zavala, 2000, pg. 127)

\[(55) 'ox-eb’ kûpan 'ixim paat \quad \text{ (56) 'ox-eb’ xoûan 'ixim}\]
\[\text{ 3-inanim CL NOUNCL tortilla } \quad \text{ 3-inanim CL\_circleshaped NOUNCL}\]
\[\text{ three half-folded tortillas } \quad \text{ tortilla}\]
\[\text{ three tortillas}\]
• Several CL languages have a generic/general CL that can appear with nouns not associated with a specific shape-based sortal CL. Examples include Chinese ge, Japanese tsu for inanimates (Downing, 1996), Korean kay for inanimates (Lee and Ramsey, 2000), Vietnamese cái (Greenberg, 1972), Thai tsu and tan (Deepadung, 1997). The generic sortal CL in Hungarian is darab ‘piece’.

Hungarian

(57) hét darab szó
seven CLgeneric word
seven words

Chinese (Cheng and Sybesma, 1998)

(58) san ge ren
three CL people
three people

• The generic classifier darab frequently replaces a more specific classifier in Hungarian:

(59) két szál gyertya
two CLthread candle
two candles

(60) két darab gyertya
two CLgeneric candle
two candles

and in Chinese: (Zhang, 2009b, pg.8)

(61) 3 zhăng zhūōzī
3 CLtable
3 tables

(62) 3 ge zhūōzī
3 CLgeneric table
3 tables

• The specific semantic content expressed by selected sortal classifiers is most often tied to animacy, shape, size and structure. In Hungarian sortal C's tend to express shape and size:

(63) szem fej, szál, karika
small.spherical, big.spherical, long.thin, flat.circular

Hungarian

(64) egy szál gyufa
one CLthread match
one match

Thai (Beckwith, 2007, pg.151.)

(65) máj-khiit nǐŋ kāan
wood-match one CL(stick-like)
one match

Mandarin (Borer, 2005, pg.86.)

(66) egy szem rīzs
one CLeye rice
one grain of rice

Hungarian

(67) yi lī mi
one CLrice
one grain of rice

Taskhent Uzbek (Aikhenvald, 2000, pg.102.)

(68) egy fej káposzta
one CLhead cabbage
one head of cabbage

Languages that have a large class of numerals often have sortal classifiers

Hungarian has a large class of numerals (see (70-b))

---

5 The Cebuano generic CL bu?uk also means ‘piece’ (Allan, 1977); it can appear with all N heads, including those without a specific associated sortal CL.
(70) a. öt szál virág
five CL_{thread} flower
five flowers

b. két-millió szál virág
two-million CL_{thread} flower
two million flowers

- Classifiers typically evolve from nouns
  All Cs in Hungarian have a nominal use as well\(^6\)

- Nouns for body parts and objects with canonical shapes are often used as classifiers for inanimate objects
  Some typical nouns that can function as classifiers:

  - head for big round objects, c.f. fő ‘head’ and fej ‘head’ in Hungarian

    (71) öt fej káposzta
    five CL_{head} cabbage
    five heads of cabbage

    (72) öt fő személyzet
    five CL_{head} personnel
    five personnel

  - eye for small spherical objects, c.f. szem ‘eye’ in Hungarian

    (73) hét szem gyöngy
    seven CL_{eye} pearl
    seven pearls

  - thread for long thin objects, c.f szál ‘thread’ in Hungarian

    (74) tíz szál kolbász
    ten CL_{thread} sausage
    ten sausages

  - leaf for flat flexible objects, c.f. levél ‘leaf’ in Hungarian

    (75) húsz levél gyógyszer
    twenty CL_{leaf} pill
    twenty strips of medicine

- classifiers occur in anaphoric phrases in all CL languages (Allan, 1977) and in Hungarian, too

  Hungarian

    (76) 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)

  Thai (Allan, 1977, pg. 286)

    (77) a. mā• tua nān
dog CL that
that dog

    b. tua nān
CL that
that one

\(^6\)The use of a lexical entry as a noun or as a classifier are clearly distinct, of course, as an N it is the head of the extended nominal projection, as a C it it not.
4.2.2. Frequently raised objections, and why they can be dismissed

"These lexemes in Hungarian are probably not sortal CLs because . . ."

- . . . because a large number of nouns cannot take a sortal Cl
  BUT:
  - a noun can always take the generic Cl darab (if non-human),
  - unclassifiable nouns are not uncommon in Cl languages
    examples include several nouns in Burmese, and Thai (Allan, 1977), Vietnamese (Greenberg, 1972), Bengali, Omani Arabic and Kana (Kegboid, Cross-River), (Aikhenvald, 2000), Akatek Mayan (Zavala, 2000)

\[
\begin{align*}
(78) & \quad \text{'két \((\ast\text{CL}_{specific})\) ceruza} \\
& \quad \text{two CL} \quad \text{pencil} \\
& \quad \text{two pencils}
\end{align*}
\]

\[
\begin{align*}
(79) & \quad \text{két darab ceruza} \\
& \quad \text{two CL}_{generic} \quad \text{pencil} \\
& \quad \text{two pencils}
\end{align*}
\]

- . . . because the number of sortal Cls is fairly small. We may expect to find more Cls in a "Cl-language"
  BUT: the number of classifiers can range from just a handful to around two hundred (data from Aikhenvald (2000))
  - Nung, Thai family: 4
  - Iwam and Chambri: 5
  - Vietnamese: \(~140\)
  - Thai and Burmese: \(~200\)
  - Hungarian: \(~20\)

- . . . because sortal CLs are obligatory in Chinese but in Hungarian they are optional\(^7\)
  BUT: classifiers are optional in several CL languages, e.g. Akatek Mayan, (Zavala, 2000), Minangkabau (Aikhenvald, 2000), informal Khmer (Greenberg, 1972), Cambodian (Goral, 1979), to mention a few

\[
\begin{align*}
(80) & \quad \text{a. két fej hagyma} \\
& \quad \text{two CL}_{head} \quad \text{onion} \\
& \quad \text{two onions}
\end{align*}
\]

\[
\begin{align*}
(81) & \quad \text{b. két hagyma} \\
& \quad \text{two CL}_{head} \quad \text{onion} \\
& \quad \text{two onions}
\end{align*}
\]

4.2.3. The position of sortal Cls

- decomposition of the DP: \[DP \ D [NumP \ Num [CLP \ Cl] [NP \ N]]] \]
  (c.f. Cheng and Sybesma, 1999; den Dikken, 2003; Borer, 2005; Cinque and Krapova, 2007; Svenonius, 2008)

- we propose that Hungarian has a phonologically zero classifier, which results in the perceived optionality of Cls\(^8\)
  - condition on interpretation: only canonical units are possible

---

\(^7\)‘Optional’ is used in a purely descriptive sense here, meaning that the Cl may or may not appear on the surface in the DP. This can be interpreted in a number of ways: true syntactic optionality of ClP; overtness vs. covertness, a phonologically zero Cl, etc. For concreteness, we assume that if an overt sortal Cl is absent in Hungarian, the nominal expression contains a covert counterpart of the generic darab.

\(^8\)See Borer (2005) for a different proposal.
– condition on distribution: same as specific Cls and *_darab_

• Some researchers claim cross-linguistic complementary distribution between plural marking and classifiers (T’ou, 1976; Chierchia, 1998)

• Suppose that this claim is true, then

• Are plurals and Cls are in complementary distribution because they would like to fill the same slot in the tree, and having both violates the one lexical item per terminal rule?

No, the following data show that plurals and Cls occupy different positions in the tree:

– the plural and sortal Cls appear on different sides of the head N

(81) rágó-k
   chewing gum-
   chewing gums

(82) két csík rágó
   two Cl_{strip} chewing gum
   two chewing gums

– the plural must copy onto a demonstrative, but a sortal CL cannot\footnote{One may try to explain the data in (81) - (86) by appealing to a morphological distinction between sortal Cls and the plural. Then it could be maintained that sortal Cls and the plural occupy the same position in the tree, and their distributional differences in (81) - (86) could be explained by this morphological property. Note that this property, crucially, could not be the suffixhood of the plural. Some phonologically independent nominal modifiers, such as postpositions, do copy onto the demonstrative just like the suffixal plural marker:}

(83) Ez-ek a ház-ak
   this-PL the house-scpl
   These houses

(84) *Ez a ház-ak
   this the house-scpl
   These houses

(85) *Ez szem a szem mogyoró
   This Cl_{eye} the Cl_{eye} hazelnut
   This hazelnut

(86) Ez a szem mogyoró
   This the Cl_{eye} hazelnut
   This hazelnut

In addition, some nominal suffixes do not copy onto the demonstrative. This is the case with the possessive suffix \*a/e:

(iii) *ez-e a ház-a
   this-poss the house-poss
   this house of his

(iv) ez a ház-a
   this the house-poss
   this house of his

(87-b) is a knock-down analysis against an analysis in which sortal Cls and the plural compete for the same structural position.
• Further, plural and sortal Cl are not in complementary distribution in Hungarian
  An anaphoric sortal Cl co-occurs with the plural suffix in (87-b)\textsuperscript{10}

\begin{align*}
(87) \quad & a. \text{Vett-em t\textsuperscript{\text{-}}z szem} & b. \text{Ez-\text{-}ek a szem-ek rohadt-ak.} \\
& \text{Buy.PAST-\text{-}1\text{SG ten Cl\text{\text{-}ge} mogyor\text{\text{-}t.}} & \text{this-PL the Cl\text{\text{-}ge-PL rotten-PL} These (ones) are rotten.} \\
& \text{hazelnut-ACC} & \text{I have bought ten hazelnuts.}
\end{align*}

**Conclusion:**

i) plural and sortal CLs are not realizations of the same position

ii) according to common wisdom, both Pl and CL are heads

→ they are in different projections

• Is plural in the same projection as numerals (NumP), or does it have its own projection?

  – Numerals don’t co-occur with plural-marked NPs in Hungarian
    (compatible with numerals and Pl competing for the same position, but also with numerals selecting for a bare NP)

\begin{align*}
(88) \quad & \text{három takaró-(\text{*}k)} & \text{this three \ the three \ blanket} \\
& \text{three blanket-PL} & \text{three blankets}
\end{align*}

  – Numerals don’t copy onto demonstratives like the plural does
    (not compatible with numerals and Pl competing for the same position, but compatible with numerals sitting in the spec of the projection of which the head is Pl)
    assumption: copying is a property of certain heads;
    also for English: assume that structure of cl and non-cl languages is essentially identical

\begin{align*}
(89) \quad & \text{*ez három a három takaró} & \text{this three \ the three \ blanket} \\
& \text{this three \ blanket} & \text{these three blankets}
\end{align*}

  – English also shows that numerals and the plural marker are not in the same position: *three nuts
  – Evidence is compatible with (90) or (91)

\begin{align*}
(90) \\
\text{DP} \\
\text{D} \quad \text{NumP} \\
\text{numerals} \quad \text{Num'} \\
\text{Num \ } \text{plural} \quad \text{CLP} \\
\text{Div} \quad \text{NP} \\
\text{sortal CLs}
\end{align*}

\textsuperscript{10}The plural marker on the demonstrative and the predicative adjective are agreement morphemes (demonstratives and predicative adjectives always agree in number with the head N), but the plural on the CL is not agreement, as Cls never agree with the head N in number (see all other examples on this handout).
– unless there is good evidence for (91), parsimony favors (90)

### 4.3. Group classifiers

<table>
<thead>
<tr>
<th>Language</th>
<th>Example</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungarian</td>
<td>(92) az a csapat lány</td>
<td>(93) gó bāan léuihái</td>
</tr>
<tr>
<td></td>
<td>that the bunch girl</td>
<td>that bunch girl</td>
</tr>
<tr>
<td></td>
<td>that bunch of girls</td>
<td>that bunch of girls</td>
</tr>
<tr>
<td>Chinese</td>
<td>(94) egy raj méh</td>
<td>(95) yī qún mìfēng</td>
</tr>
<tr>
<td></td>
<td>one swarm bee</td>
<td>one collection bee</td>
</tr>
<tr>
<td></td>
<td>a swarm of bees</td>
<td>a swarm of bees</td>
</tr>
</tbody>
</table>

- We concluded in Section 4 that group Cls are nouns that take a small nominal category as a complement. A sortal Cl is required is the embedded nominal category, because having plural individuals (required by group Cls) is only possible if there is individuation. Since sortal Cls individuate the mass-like denotation, a sortal Cl is necessary.

(96) Group Cl structure
• Crucially, in (96) the embedded nominal category is smaller than a full-fledged DP
• Support for this claim:
  – The nominal embedded category below a group Cl cannot contain an article or a demonstrative

(97) *egy raj ez a méh
    one swarm this the bee

(98) *egy raj a méh
    one swarm the bee

(97) and (101) may be out for semantic reasons (definite article has a uniqueness presupposition, group Cl denotes a plurality of individuals), but no such conflict arises with (99). Using a partitive construction, (100) approximates the meaning that (99) tries to express but cannot.

(99) *egy raj ez-ek a méh-ek
    one swarm this-PL the bee-PL

(100) egy raj ez-ek-ből a méh-ek-ből
      one swarm this-PL-FROM the bee-PL-FROM
      a swarm of these bees

We do not find evidence that D and Dem are structurally represented in the NP embedded by the group Cl → we conclude that these categories are not projected in the embedded NP

– The nominal embedded category below a group Cl cannot contain a numeral (101) (yet this cannot be due to semantic reasons, see example (102)) or a plural morpheme

(101) *egy raj 20 méh
      one swarm 20 bee

(102) egy 20 méh-ből álló
      a 20 bee-from comprising
      (méh)raj
      bee-swarm
      a swarm of 20 bees

(103) *egy raj méh-ek
      one swarm bee-PL
      a swarm of bees

This receives a natural explanation if the embedded nominal category is smaller than NumP

• If a sortal Cls is necessary for individuation in the embedded NP, then why are overt sortal Cls ungrammatical in the embedded NP?

(104) egy levél gyógyszer
      one strip medicine

(105) *egy levél szem gyógyszer
      one strip Cl\textsubscript{eye} medicine

We are led to conclude that group Cls subcategorize for the covert sortal Cl

5. Conclusion

• Classifier systems may be more widespread than usually considered
• Hungarian is a classifier language with various overt classifiers and a phonologically null classifier
• Distinction between ‘standard classifier’ vs. other languages
  (would English be considered to be similar to Hungarian in being more like a classifier language? C.f.: a blade of grass, an ear of corn, a stick of wood, a stick of candy)

• Test battery shows that sortal and group CIs have a different category and they are in a different structural position

• Adjective tests are expected to show the same pattern cross-linguistically, omission tests are not

• Plural marking is not in complementary distribution with classifiers

6. Appendix

(106) lists the sortal CIs of Hungarian along with their meaning when used as a noun, and a non-comprehensive inventory of the nouns they classify. In (107) - (124) we illustrate their use.

(106) Sortal CIs and the nouns they classify

<table>
<thead>
<tr>
<th>classifier</th>
<th>nominal gloss</th>
<th>nouns classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>bokor</td>
<td>bush</td>
<td>potato, raspberry, rose</td>
</tr>
<tr>
<td>cikk</td>
<td>article</td>
<td>garlic, orange, mandarin</td>
</tr>
<tr>
<td>cserép</td>
<td>pot</td>
<td>pot plants</td>
</tr>
<tr>
<td>csik</td>
<td>strip</td>
<td>chewing gum</td>
</tr>
<tr>
<td>cső</td>
<td>tube</td>
<td>sweet corn, red and green pepper</td>
</tr>
<tr>
<td>darab</td>
<td>piece</td>
<td>any noun</td>
</tr>
<tr>
<td>fej</td>
<td>head</td>
<td>cabbage, onion, lettuce, kohlrabi, cauliflower, broccoli</td>
</tr>
<tr>
<td>fő</td>
<td>head</td>
<td>people in regimented situations, e.g. staff</td>
</tr>
<tr>
<td>gerezd</td>
<td>clove</td>
<td>garlic, orange, mandarin</td>
</tr>
<tr>
<td>ív</td>
<td>sheet</td>
<td>paper</td>
</tr>
<tr>
<td>karika</td>
<td>ring</td>
<td>sausage</td>
</tr>
<tr>
<td>kötet</td>
<td>volume</td>
<td>books and other bound volumes</td>
</tr>
<tr>
<td>rózsa</td>
<td>rose</td>
<td>cauliflower, broccoli (the natural units into which a head of caulifower can be)</td>
</tr>
<tr>
<td>rúd</td>
<td>rod</td>
<td>chitterlings and salami</td>
</tr>
<tr>
<td>szál</td>
<td>thread</td>
<td>hair, match, sausage, salami, flower, candle, green onion, cigarette, plank, carrot, welt</td>
</tr>
<tr>
<td>szem</td>
<td>eye</td>
<td>all types of nuts and berries, tomato, pepper, biscuit, pearl, sand, potatoe, cucumber</td>
</tr>
<tr>
<td>tő</td>
<td>stem</td>
<td>grape, rose, nursling, any specific type of plant that has a nursling</td>
</tr>
<tr>
<td>vekni</td>
<td>loaf</td>
<td>bread</td>
</tr>
</tbody>
</table>

Examples

(107) hét bokor krumpli seven CL_{bush} potato seven potato plants
(108) hét cikk narancs seven CL_{clove} orange seven cloves of orange
(109) hét cserép virág seven CL_{pot} flower seven potted plants

(110) hét csik seven CL_{strip} régő chewing gum seven chewing gums (long, thin, not pellets)
(111) hét cső kukorical seven CL_{tube} sweet corn seven ears of sweetcorn
(112) hét darab kabát seven CL_{generic} coat seven coats
| (113)  | hét fej | seven CL<em>head</em> |
|        | karfiol | cauliflower        |
|        | seven clauflowers |

| (114)  | hét fő | seven CL<em>head</em> |
|        | légénység | seven people belonging to |
|        | a crew |

| (115)  | hét | seven CL<em>close</em> |
|        | gerezd narancs | orange |
|        | seven cloves of orange |

| (116)  | hét ív | seven CL<em>arc</em> |
|        | papír | paper |
|        | seven sheets of paper |

| (117)  | hét | seven CL<em>circle</em> |
|        | karika kolbász | sausage |
|        | seven slices of sausage | (ring-shaped) |

| (118)  | hét kötet | seven CL<em>volume</em> |
|        | könyv | book |

| (119)  | hét rózsa | seven CL<em>rose</em> |
|        | karfiol | cauliflower |
|        | seven pieces of cauliflower |

| (120)  | hét | seven CL<em>rod</em> |
|        | rúd | pieces |
|        | szalámi | sausage |
|        | seven |

| (121)  | hét | seven CL<em>thread</em> |
|        | szál | cigarette |
|        | seven cigarettes |

| (122)  | hét szem | seven CL<em>eye</em> |
|        | kukorica | sweet corn |
|        | seven grains of sweetcorn |

| (123)  | hét | seven CL<em>stem</em> |
|        | tő | pieces |
|        | palánta | salami |
|        | seven |

| (124)  | hét | seven CL<em>loaf</em> |
|        | vekni kenyér | bread |
|        | seven loaves of bread |

**References**


