

The puzzle

Hungarian allows apparently optional classifiers (henceforth: CLs); see (1).

(1) *három (darab) könyv*
 three (CL_{general}) book

- Why can certain nouns in Hungarian appear with and without a CL?
- What is the difference between constructions with and without the CL?
- What is the function of CLs like *darab*, ‘CL_{general}’ in constructions like (1)?

Our proposal:

- Hungarian CLs have lexical content specifying the type of objects counted by NUM.
- This phenomenon can be explained in an MSSC-framework (Grimm 2012).

Hungarian in mass/count typologies

Chierchia (1998), (2010): the two systems for making nouns countable (systems which use count nouns, and systems which use CLs, Greenberg 1974) are in complementary distribution and each system is associated with a different set of properties.

Schvarcz (2014): Hungarian exhibits genuine count language-like properties.

- Numerals can directly modify nouns; see (2).
- There is a singular/plural distinction in the language; see (3).

(2) *három könyv*
 three book
 ‘three books’

(3) *könyv/könyv-ek*
 book/book-PL
 ‘book/books’

At the same time, Hungarian exhibits some genuine CL language traits.

- Hungarian allows optional CLs; see (1).
- Hungarian also allows bare singular arguments; see (4).

(4) *Könyv-et vettem.*
 book-ACC bought.1SG
 ‘I bought a book/books’

There are **two accounts** for the phenomena:

1. **Hungarian as CL language** (Csirmaz and Dékány 2014, Cs&D14).
 - All nouns are mass – nouns cannot combine with numerals directly.
 - Counting requires an explicit lexical CL (e.g. *darab*) or a null sortal CL (CL₀).
2. **Noun flexibility** (Schvarcz and Rothstein 2017, S&R17).
 - Nouns like *könyv* are ambiguous between a count and mass interpretation.
 - On their mass reading, they must appear in CL-constructions.
 - On their count reading, they must appear in NUM+NP constructions.

Both approaches assume that CLs in Hungarian are required by mass syntax, the role of the CL is to map mass predicates onto sets of discrete countable units.

Our assumption

The appearance of CLs with so called ‘mixed nouns’ (like *könyv*, ‘book’, that can appear in CL-constructions and can be directly modified by NUM) is not related to the mass denotation of a noun.

The function of CLs: to determine the type of objects counted in the denotation of N.

The main observation

There is an unexpected shift in the interpretation of the whole NP where a mixed noun is modified by a NUM with and without the CL.

The data

There are nouns that are polysemous and have the so-called PHYSICAL OBJECT (PHYS) and the INFORMATION OBJECT (INF) senses, e.g. *könyv*, ‘book’, *festmény*, ‘painting’, *magazin*, ‘magazine’, etc. (Pustejovsky 1995, Asher 2011).

- If such nouns appear in contexts that does not specify which sense is used,
 - in Num+NP constructions, the numeral can count the objects of either type in the denotation of the noun; see (5)
 - in CL-constructions, the numeral can only count the physical objects in the denotation of the noun; see (6).

(5) *három könyv*
 three book
 ‘three books_{INF,PHYS}’

(6) *három darab könyv*
 three CL_{general} book
 ‘three books_{PHYS}}’

- (7) a. CONTEXT: Amelia is an adventurous librarian whose main hobby is to dig up ancient and medieval books that got lost over the centuries. Last month, she had a particularly successful mission in Alexandria: she found a **codex containing three works** by Valentina, an ancient author whose work is mostly lost, but well-documented (hence we know she only wrote books). There was nothing else in the codex, and Amelia found nothing else last month.
- b. (*A múlt hónap-ban*) *Amélia egy (darab) könyv-et talált.*
 the last month-INE Amelia one (CL_{general}) book-ACC found
 ‘Last month, Amelia found one book’

- c. (*A múlt hónap-ban*) *Amélia három (#darab) könyv-et talált.*
 the last month-INE Amelia three (CL_{general}) book-ACC found
 ‘Last month, Amelia found three books’

Problems for previous approaches

- Cs&D14: *darab* and the null-CL are different realizations of the CL_{general}, hence supposed to have the same semantic content, but (7) shows it is not the case.
- S&R17: mixed nouns are ambiguous between mass/count interpretations.
 - The pattern we see based on (7) does not suggest ambiguity, rather underspecification.

Construction	What is counted by the NUM
NUM+NP	INF or PHYS
CL-construction	PHYS

Hypotheses

1. Mixed nouns are count nouns in that they have atomic reference (in the sense of Krifka 1989) in both NUM+NP and CL-constructions.
2. The meaning of bare count nouns might be underspecified as to what type of objects are in their denotation as atoms.
3. In NUM+NP constructions, it is not specified what type of objects are counted in the denotation of the NP.
4. In CL-constructions, the CL imposes a restriction as to what types of objects can be counted by the numeral: they have to be physically distinct.

The analysis

Cardinal numerals in Hungarian are interpreted as predicates of type $\langle e, t \rangle$, and they give the cardinality of an entity (Landman 2003; Rothstein 2017); see (8).

$$(8) \quad \llbracket n \rrbracket = \lambda x. |x| = n$$

The **cardinality of an entity** is the cardinality of the set containing all the parts of the entity that are the smallest elements in the denotation of a given predicate; see (9).

$$(9) \quad |x| = n \leftrightarrow |\{y : y \leq x \wedge \exists P.P(x) \wedge P(y) \wedge \neg \exists z.z < y \wedge P(z)\}| = n$$

The **general CL *darab* is a predicate modifier** which takes a predicate over entities, and returns a predicate that holds of any entity that is the sum of MSSC entities with respect to the predicate; see (10).

$$(10) \quad \llbracket \text{DARAB} \rrbracket = \llbracket \text{CL}_{\text{unif}} \rrbracket = \lambda P \lambda x. P(x) \wedge x = \bigoplus \{y : y \leq x \wedge \text{MSSC}(y, P)\}$$

$$(11) \quad \llbracket (6) \rrbracket = \llbracket \text{három darab könyv} \rrbracket = \lambda x. * \text{BOOK}(x) \wedge x = \bigoplus \{y : y \leq x \wedge \text{MSSC}(y, * \text{BOOK})\} \wedge |x| = 3$$

The analysis in (10) can be straightforwardly extended to **other CLs** in Hungarian like *szem*, ‘CL_{small round object}’, *fej*, ‘CL_{big round object}’, etc.

- These CLs select nominals based on size and shape of the entities in their denotation.
- These CLs can be analyzed as lexically more specified versions of *darab*, each of them adding some extra property to be fulfilled by the parts of the entity that is counted; see the lexical entry for *szem* in (12).

$$(12) \quad \llbracket \text{CL}_{\text{small round object}} \rrbracket = \lambda P \lambda x. P(x) \wedge x = \bigoplus \{y : y \leq x \wedge \text{MSSC}(y, P) \wedge \text{SMALL}(y) \wedge \text{ROUND}(y)\}$$

There are no rules in our grammar that prevents **nouns with mass denotation** to occur in either NUM+NP or CL-constructions.

- (13) a. *három #(darab) agyag*
 three (CL_{general}) clay
- b. $\llbracket \text{három agyag} \rrbracket = \lambda x. \text{CLAY}(x) \wedge |x| = 3$
- c. $\llbracket \text{három darab agyag} \rrbracket = \lambda x. \text{CLAY}(x) \wedge x = \bigoplus \{y : y \leq x \wedge \text{MSSC}(y, \text{CLAY})\} \wedge |x| = 3$

Conclusions

- In Hungarian, count nouns can occur in CL-constructions because the appearance of CLs is not related to the mass denotation of a N.
- CLs is to determine what are the countable atoms in the denotation of the N.
- NUM+NP: the NUM counts any type of atoms in the denotation of the N.
- NUM+CL+NP: only physically distinct entities can be counted by the NUM.

References

- Asher, N. (2011). *Lexical Meaning in Context: A Web of Words*. Cambridge University Press.
- Chierchia, G. (1998). Plurality of mass nouns and the notion of ‘semantic parameter’. In S. Rothstein (Ed.), *Events and Grammar* (pp. 53–103). Studies in Linguistics and Philosophy 70. Springer Netherlands.
- Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese*, 174(1), 99–149.
- Csirmaz, A. & Dékány, É. (2014). Hungarian is a classifier language. In R. Simone & F. Masini (Eds.), *Word Classes: Nature, Typology and Representations* (pp. 141–160). Amsterdam: John Benjamins.
- Greenberg, J. (1974). *Language Typology: A Historical and Analytic Overview*. Berlin: de Gruyter.
- Grimm, S. (2012). Degrees of countability: a mereotopological approach to the mass/count distinction. In A. Chierchia (Ed.), *Proceedings of Semantics and Linguistics Theory 22* (pp. 584–603). Retrieved from <http://dx.doi.org/10.3765/salt.v22i0.2633>
- Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and Contextual Expressions*. Groningen-Amsterdam Studies in Semantics. Dordrecht: Foris.
- Landman, F. (2003). Predicate-argument mismatches and the adjectival theory of indefinites. In M. Coene & Y. D’Hulst (Eds.), *From NP to DP, Vol. 1: The Syntax and Semantics of Noun Phrases* (pp. 211–237).
- Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press.
- Rothstein, S. (2017). *Semantics for Counting and Measuring*. Dordrecht: Cambridge University Press.
- Schvarcz, B. (2014). *The Hungarians who say -nyi: Issues in Counting and Measuring in Hungarian*. M.A. Thesis, Bar-Ilan University.
- Schvarcz, B. & Rothstein, S. (2017). Hungarian classifier constructions, plurality and the mass-count distinction. In H. van der Hulst & A. Lipták (Eds.), *Approaches to Hungarian 15. Papers from the 2015 Leiden Conference* (pp. 157–182). Amsterdam: John Benjamins.