

An MSSC-approach to Hungarian classifiers

Brigitta R. Schvarcz & Kata Wohlmuth¹

bridget.schvarcz@gmail.com & katawohl@gmail.com

Comments are welcome!

February 9, 2020

Keywords: mass/count distinction, classifiers, numerals, atomicity, lexical semantics, formal semantics

Abstract: This paper provides a formal semantic analysis of classifiers in Hungarian. We focus on the puzzle posed by classifier optionality in Hungarian, where most nouns can co-occur with classifier, but do not have to. Here we show that classifiers are not entirely optional, and their presence or absence in an expression has semantic consequences. Evidence in support of our analysis comes from nouns that are polysemous and have a physical object and an informational object sense, such as *könyv*, ‘book’, *festmény*, ‘painting’, *magazin*, ‘magazine’. We argue that the role of count classifiers in Hungarian, such as *darab*, is restricting the counting operation to physically distinct, Maximally Strongly Self-Connected entities (Grimm 2012) in the denotation of the noun they modify.

1 Introduction²

Hungarian allows an apparently optional classifier: many nouns can be directly modified by a numeral (1a), but can also be combined with a classifier before being modified by a numeral (1b). In this paper, we refer to the former expressions as numeral expressions (henceforth: Num-expressions), and the latter ones as classifier expressions (henceforth: CL-expressions).

- (1) a. *három könyv*
three book
‘three books’
b. *három darab könyv*
three CL_{general} book
‘three books’

This paradigm poses a problem for the traditional categorizations which define two major systems for explaining countability in the nominal domain: (i) a system which uses sortal classifiers, such as in Mandarin Chinese; (ii) a system which uses count nouns that can directly be modified by numerals, such as in English (Greenberg 1974). Given this background, classifier optionality

¹Co-first authors, listed alphabetically

²Glosses: 1 – first person; ACC – accusative case; CL – classifier; INE – inessive case; PFV – perfective aspect; PL – plural; SG – singular.

30 displayed in (1) raises questions about the difference between the Num-expressions and the CL-
31 expressions in Hungarian, as well as about the function of classifiers such as *darab* ($CL_{general}$).
32 The aim of this paper is to give a semantic account of the Hungarian classifier phenomenon, in
33 particular, to analyze the interaction among classifiers, numerals, and nouns that have countable
34 entities in their denotation.

35 The structure of this paper is as follows: in Section 2 we discuss the grammatical distinction
36 between mass and count and we review different approaches to the mass/count phenomenon. In
37 Section 3, we give an overview of the mass/count typology, and then we show how Hungarian
38 cannot be easily placed in this typology, partially because of the classifier optionality exemplified
39 in (1). Section 4 presents some novel data that show that classifiers in cases like in (1) are not
40 entirely optional, but their presence/absence has some semantic effects. Then in Section 5 we give
41 a short overview how our data poses a problem for previous approaches to Hungarian classifier
42 optionality. We set out our analysis of Hungarian classifiers in Section 6, and discuss additional
43 cases our analysis accounts for. Finally, Section 7 concludes the paper.

44 2 Mass/count as a grammatical distinction

45 Mass and count predicates have distinct grammatical properties that have been observed cross-
46 linguistically. Some of the tests distinguishing mass from count in English include: direct modifi-
47 cation by numerals, pluralization, and determiner selection.

48 First, count nouns directly combine with numerals, resulting in Num-expressions, mass nouns
49 do not. With numerals, mass nouns require an obligatory mediating element, a classifier or a mea-
50 sure phrase, and form a CL-expression (2). Second, count nouns can be morphologically marked
51 for plural, while mass nouns cannot (3). Third, determiners are sensitive to the mass/count distinc-
52 tion: some select only count nouns, some select only mass nouns (4).

- 53 (2) a. three books
54 b. #three waters
55 c. three buckets/litres of water

- 56 (3) a. books
57 b. #waters

- 58 (4) a. many/*much books
59 b. *many/much water

60 It is widely agreed among linguists that the distinction between nouns such as *book* and *water*
61 is a grammatical distinction (Pelletier 1975). Nevertheless, the basis of grammatical countability
62 and how it is encoded in the grammar of particular languages is subject to ongoing debate. While
63 some emphasize that counting relies on the semantic atomic structure of a predicate (Link 1983,
64 Landman 1991, 2003, 2004), others argue that counting is a measure operation under which pred-
65 icates are evaluated in number of measure units (Krifka 1989, 1995). The former approach treats
66 counting as an operation which applies to sums of atoms and specifies the number of atoms in
67 the denotation of a predicate. The latter approach assumes a covert classifier incorporated into the
68 numeral in order to achieve individuation (Krifka 1995).

69 Rothstein (2010, 2017) offers an account that integrates elements of both approaches above:

70 counting is an operation on atoms, and countability is lexically encoded in count nouns. She
 71 assumes the basic denotation of a noun to be mass, and a COUNT operation specifies what counts
 72 as a semantic atom in a given context. Rothstein further argues that mass/count is a reflection of the
 73 distinction between the operations of counting and measuring. According to her, counting entities
 74 and measuring stuff are two different operations: “counting is putting atomic entities in one-to-
 75 one relation with natural numbers [...] measuring is giving value to a quantity on a calibrated
 76 dimensional scale” Rothstein (2010, p. 5). The main difference between count and mass nouns is
 77 in their denotation of sets of entities: count nouns denote sets of entities that can be counted, while
 78 mass nouns denote sets of entities that can be measured.

79 These above approaches are concerned with the basis of the mass/count distinction which does
 80 not play a crucial role in our analysis. We consider the mass/count distinction as a grammatical
 81 phenomenon.

82 There is considerable cross-linguistic variation on how counting and measuring is encoded in
 83 the grammar of particular languages. The next section describes the two major types of languages
 84 discussed in the mass/count literature.

85 3 Mass/count typologies

86 Greenberg (1974) describes two main types of languages in terms of encoding counting. Some
 87 languages use a grammatical distinction between count nouns and mass nouns, while others use
 88 sortal classifiers. In languages such as English, there are count nouns which can be directly modi-
 89 fied by numerals (5a). In contrast, in languages such as Mandarin Chinese, all nouns are mass and
 90 use sortal classifiers upon combining nouns with numerals (5b).

- 91 (5) a. five books
 92 b. *wu *(ben) shu*
 93 five CL_{volume} book
 ‘five books’

94 Chierchia (1998, 2010) argues that the two systems are in complementary distribution, and that
 95 disparate grammatical properties are associated with each system. In typical mass/count languages,
 96 direct modification of count nouns by numerals is possible, as shown in (2), and a morphological
 97 singular/plural distinction on the noun is marked, see (3); neither is the case in typical classifier
 98 languages, see Mandarin Chinese examples in (6)–(7).

- 99 (6) a. *san *(ge) pingguo*
 100 three CL_{general} apple
 ‘three apples’
 101 b. *san *(ping) shui*
 102 three CL_{bottle} water
 ‘three bottles of water’
 103 (7) a. *yi *(ben) shu*
 104 one CL_{volume} book
 ‘one book’

105 b. *wu *(ben) shu*
 five CL_{volume} book
 106 ‘five books’

107 In mass/count languages, sortal classifiers are incompatible with count nouns (8), and bare sin-
 108 gulars cannot appear in argument position (9); both are possible in classifier languages, see the
 109 Mandarin Chinese (7a) and (10).

110 (8) *I have one unit/piece of book.

111 (9) *I bought book.

112 (10) *wo mai le shu*
 I buy PFV book

113 ‘I bought a/the book/books.’ (Schvarcz and Rothstein 2017, p. XX, ex. (2d))

114 The following table sums up the differences between typical mass/count languages on the one
 115 hand, and classifier languages on the other.

	Direct modification by numerals	Plural nominal morphology	Sortal individuating classifiers	Bare singular arguments
Mass/count	✓	✓	✗	✗
Classifier	✗	✗	✓	✓

Table 1: The traits of the two types of languages according to Chierchia (1998, 2010)

116 Hungarian poses a problem for such categorization, as it exhibits both genuine mass/count lan-
 117 guage traits (Schvarcz 2014) and classifier language traits simultaneously. Hungarian allows for
 118 the direct modification of nouns by numerals (11) and exhibits a morphological distinction be-
 119 tween singular and plural nouns (12) – although nouns are not morphologically marked plural
 120 when combined with numerals.

121 (11) *három könyv*
 three book
 122 ‘three books’

123 (12) *könyv/könyv-ek*
 book/book-PL
 124 ‘book/books’

125 Moreover, determiners such as *minden* ‘every’ are restricted to singular count nouns (13a); while
 126 other determiners such as *egy kis* ‘some’ are restricted to mass nouns (13b). These traits make
 127 Hungarian mass/count language-like.

128 (13) a. *minden könyv/#víz*
 every book/water

129 b. *egy kis #könyv/víz*
 one little book/water

(Schvarcz 2014, p. 20)

130

131 However, Hungarian allows for bare singular arguments (14), and a CL-expression upon com-
 132 bining nouns with numerals, as already shown in (1b) (repeated as (15) below). These traits make
 133 the language similar to a classifier language.

134 (14) *Könyv-et vettem.*
 135 book-ACC bought.1SG
 136 ‘I bought a book/books’

136 (15) *három darab könyv* = (1b)
 137 three CL_{general} book
 138 ‘three books’

138 This combination of grammatical properties is unexpected in the Chierchian mass/count typology.
 139 To account for the optional classifier phenomenon in Hungarian, two approaches have been pro-
 140 posed in the literature: one that is based on the rich classifier phenomena (Csirmaz and Dékány
 141 2014); and another which identifies two possible readings – a mass and a count interpretation – of
 142 most nominals in the language (Schvarcz and Rothstein 2017).

143 Following the view that count nouns and sortal classifiers are mutually exclusive, Csirmaz and
 144 Dékány (2014) explain the occurrence of sortal classifiers in Hungarian by treating it as a classifier
 145 language. This approach suggests that in Hungarian all nouns are mass, while counting requires
 146 either an explicit lexicalized classifier, such as *darab*, or a null sortal classifier (both glossed as
 147 CL_{general}).

148 (16) a. *három darab könyv*
 149 three CL_{general} book
 150 b. *három \emptyset_{CL} könyv*
 151 three CL_{general} book

150 Even though the Hungarian-as-a-classifier-language theory in Csirmaz and Dékány (2014) explains
 151 the usage of sortal classifiers, it does not explain other facts about Hungarian, such as the meaning
 152 of the plural marker, or other mass/count-language-like traits. Schvarcz and Rothstein (2017)
 153 show that plural marking is sensitive to the mass/count distinction: with notional count nouns, the
 154 plural suffix yields a predicate denoting a set of pluralities of individuals, while with notional mass
 155 nouns it yields an abundance reading. Furthermore, Schvarcz (2014) discusses numerous ways for
 156 distinguishing between mass and count expressions in Hungarian. (For further discussion on these
 157 issues see the Schvarcz (2014) and Schvarcz and Rothstein (2017) the references therein.) Thus, an
 158 alternative solution has been offered by Schvarcz and Rothstein (2017) according to which nouns
 159 like *könyv*, ‘book’, are ambiguous between a count and mass interpretation. On the mass reading,
 160 the classifier is obligatory (17a), while on the count reading, the classifier is impossible (17b).

161 (17) a. *három *(darab) [könyv]_{MASS}*
 162 three CL_{general} book
 163 ‘three books’
 164 b. *három (*darab) [könyv]_{COUNT}*
 165 three CL_{general} book
 166 ‘three books’

165 Both Csirmaz and Dékány (2014) and Schvarcz and Rothstein (2017) assume that classifiers in

166 Hungarian are required by mass syntax; the role of the classifier is to map mass predicates onto
 167 sets of discrete countable units. They both build on the Chierchian framework in which count
 168 nouns are not expected to take classifiers: Csirmaz and Dékány (2014) equate Hungarian to a true
 169 classifier language; Schvarcz and Rothstein (2017) propose a “mixed system” and two ways of
 170 counting: “a single lexical item can be used in counting contexts either as a count noun [and can
 171 be directly modified by a numeral] or in a sortal classifier construction” (p. 207). Moreover, both
 172 approaches assume that the two possible constructions for expressing the number of some entities
 173 – a Num-expression and a CL-expression – are interpreted in the same way.

174 In this paper we discuss data which suggest a different analysis. We focus on nouns like *könyv*,
 175 ‘book’, that behave both like mass and count nouns based on their co-occurrence with classifiers.
 176 Here we refer to these nouns as *mixed nouns*. We show that mixed nouns can have a different
 177 interpretation depending on the kind of expression (Num or CL) they occur in. Based on this data,
 178 we claim that the classifier optionality in Hungarian is not sensitive to the grammatical distinction
 179 between mass and count nouns. Instead, we argue that classifiers are elements with their own
 180 lexical semantic content, and their role in CL-expressions is to determine what type of objects in
 181 the denotation of the noun are counted by the numeral.

182 4 The key observation

183 The main observation of the paper is that in some contexts, the interpretation of Num-expressions
 184 and that of CL-expressions may be different from each other despite what the previous approaches
 185 suggest. This difference can be observed with mixed nouns that are polysemous having a so-called
 186 *physical object* and an *informational object* sense (Pustejovsky 1995, Asher 2011, Gotham 2017,
 187 a.o.), e.g. *könyv*, ‘book’, *festmény*, ‘painting’, *magazin*, ‘magazine’, etc. If such nouns appear in
 188 Num-expressions, the numeral can count the objects of either type in the denotation of the noun –
 189 given that the context itself does not specify the interpretation of the noun. However, if the noun
 190 occurs in a CL-expressions in the same context, the numeral can only count physical objects. This
 191 meaning-shift between Num-expressions and CL-expressions is illustrated in (18).

- 192 (18) a. SCENARIO: Amelia is an adventurous librarian whose main hobby is to look for
 193 ancient and medieval books that got lost over the centuries. Last month, she had a
 194 particularly successful mission in Alexandria: she found a codex containing three
 195 works by Valentina. (Valentina is a medieval author whose work is mostly lost, but
 196 well-documented, so it is known that she only wrote books.) There was nothing else
 197 in the codex, and Amelia found nothing else last month.
- 198 b. (A *múlt hónap-ban*) Amé^{lia} egy (*darab*) *könyv-et* talált.
 the last month-INE Amelia one (CL_{general}) book-ACC found
 199 ‘Last month, Amelia found one book’
- 200 c. (A *múlt hónap-ban*) Amé^{lia} három (*#darab*) *könyv-et* talált.
 the last month-INE Amelia three (CL_{general}) book-ACC found
 201 ‘Last month, Amelia found three books’

202 In the scenario described in (18a), there is one book in the *physical object* sense (the codex Amelia
 203 found), and there are three books in the *informational object* sense (the works by Valentina in the
 204 codex). That is, the noun *könyv*, ‘book’ can be used to refer to books both in the physical and

205 the informational object sense. In this context, both the Num-expression and the CL-expression are
 206 felicitous if the numeral in the expression corresponds to the number of physical books (see (18b)),
 207 however, only the Num-expression is felicitous if the numeral in the expression corresponds to the
 208 number of informational books (see (18c)).

209 For further examples of this meaning-shift see (19) with the noun *festmény*, ‘painting’, and (20)
 210 with the noun *magazin*, ‘magazine’.

- 211 (19) a. SCENARIO: In the art museum, in Room 303 there is only one item on exhibit: a two-
 212 sided painting depicting people’s front on the one side, and their backs on the other.
 213 The two paintings were painted by two different painters and cannot be separated
 214 from each other.
- 215 b. *A 303-as terem-ben egy (darab) festmény van kiállítva.*
 the 303 room-INE one (CL_{general}) painting is exhibited
 216 ‘In Room 303, there is one painting on exhibit’
- 217 c. *A 303-as terem-ben két (#darab) festmény van kiállítva.*
 the 303 room-INE two (CL_{general}) painting is exhibited
 218 ‘In Room 303, there are two paintings on exhibit’
- 219 (20) a. SCENARIO: Amelia went to the news stand this morning to buy her monthly maga-
 220 zines. She usually buys 3 different ones, but this morning she could only find two of
 221 her usual magazines, published as a special joint issue, where one of the magazines
 222 is published on one side, and the other magazine on the other side, upside down.
 223 Amelia bought nothing else but this special joint issue this morning.
- 224 b. *(Ma reggel) Amélia egy (darab) magazin-t vett.*
 today morning Amelia one (CL_{general}) magazine-ACC bought
 225 ‘This morning, Amelia bought one magazine’
- 226 c. *(Ma reggel) Amélia két (#darab) magazin-t vett.*
 today morning Amelia two (CL_{general}) magazine-ACC bought
 227 ‘This morning, Amelia bought two magazines’

228 As the examples in (18)–(20) demonstrate, Num-expressions and CL-constructions do not neces-
 229 sarily have similar interpretations. We summarized the main difference in Table 2.

Expression	What is counted by the Num
Num	physical or informational objects
CL	physical objects

Table 2: The types of individuals counted in different expressions

230 For this meaning-shift between the two kinds of expressions to occur, the following conditions
 231 have to be met: 1. the noun has to be polysemous and have both *physical object* and *informational*
 232 *object* senses, and 2. the context must not restrict the interpretation of the noun to one of its possible
 233 senses.³

³By *context*, we mean both the linguistic and the extra-linguistic environment of the whole nominal expression. While the extra-linguistic context may specify the interpretation of the noun by providing potential referents for the nominal expression, the linguistic context might involve selectional restrictions imposed by some other expression in the sentence. For example, the verb *write* or *read* select for informational objects in the denotation of *book* when the

234 5 Problems with the previous approaches

235 The meaning-shift discussed in the above section poses a problem for previous approaches to
236 mixed nouns and the classifier-phenomenon in Hungarian. As mentioned before, both Csirmaz
237 and Dékány (2014) and Schvarcz and Rothstein (2017) assume that Num-expressions and CL-
238 expressions result in the same interpretation, but they base this assumption on different grounds.

239 According to Csirmaz and Dékány (2014), the null classifier that is assumed to be present in
240 the structure of Num-expressions, and the overt classifiers are merely different realizations of the
241 same grammatical element – in our examples *darab*, glossed as $CL_{general}$. That is, Csirmaz and
242 Dékány (2014) propose the same syntactic structure for Num-expressions and CL-expressions. If
243 Num-expressions and CL-expressions had the same underlying structure, there would not be any
244 difference in the interpretation of these constructions, as one construction can only be assigned to
245 one meaning. However, as our data indicate, the two expressions do not have the same interpreta-
246 tion, thus they cannot have the same underlying structure, as opposed to what Csirmaz and Dékány
247 (2014) suggests.

248 Since the difference in the interpretation of the two expressions correlates with the presence or
249 absence of the classifier, to maintain the assumption about Hungarian being a classifier language,
250 we could slightly modify the approach in Csirmaz and Dékány (2014) and assign the null classi-
251 fier and the overt classifier *darab*, $CL_{general}$ different interpretations. A clear advantage of such an
252 amendment is that this way, we can account for the fact that some nouns in Hungarian cannot occur
253 in Num-expressions, but can co-occur with *darab* (see an example in (22)), which is unexpected
254 if the null classifier and *darab* are merely different realizations of the same grammatical element.
255 However, this modified version of Csirmaz and Dékány (2014) raises a more general question:
256 if Hungarian is a system where classifiers do not have to be realized overtly, why is it that only
257 certain nouns can co-occur with covert classifiers, while others cannot? In other words, given the
258 assumption that classifiers must modify all nouns in order to form grammatically count construc-
259 tions, there is no obvious explanation why there is a covert classifier available only for some nouns,
260 and there are no other covert classifiers that could potentially modify others.⁴ Moreover, even if
261 we assign different semantics for the null classifier and *darab*, by maintaining the assumption that
262 Hungarian is a classifier language, we still cannot account for the facts about Hungarian suggesting
263 that there is a distinction between mass and count nouns, as discussed in Section 3.

264 (22) *három #(darab) márvány*
three $CL_{general}$ marble

noun appears as (part of) the direct object of the verb. Thus, the meaning-shift described above cannot be observed in sentences with these verbs, as shown in (21).

(21) *Valentina írt/olvasott három (darab) könyv-et.*
Valentia wrote/read three $CL_{general}$ book-ACC
'Valentina wrote/read three books'

⁴An additional conceptual argument against amending Csirmaz and Dékány (2014) as described above is that by assuming that the null classifier is different from all overt classifiers semantically, we assign unique semantic content for a grammatical element that by definition cannot be realized phonologically. It is difficult, if not impossible, to test such a theory empirically, as a grammatical structure with the null classifier and an ungrammatical structure without the null classifier are spelled out exactly the same way.

265 ‘three pieces of marble’

266 Schvarcz and Rothstein (2017) presume that in Hungarian the presence or absence of the classifier
267 correlates with the ‘massness’ or ‘countness’ of the noun in the construction. In their analysis,
268 mixed nouns are ambiguous between being mass and count and that is why they can occur both
269 in Num-expressions and CL-expressions. Both expressions are counting expressions, but while in
270 Num-expressions, the countable entities are denoted by the noun itself, in CL-expressions, these
271 countable entities are in the denotation of the noun modified by the classifier.

272 Even though Schvarcz and Rothstein (2017) could potentially account for differences in the
273 interpretations of the two constructions, the observed pattern we summarized in Table 2 is not a
274 pattern of ambiguity – two mutually exclusive interpretations – but rather one of underspecifica-
275 tion, when one interpretation involves the other as an alternative interpretation. Still, one could
276 assume that the difference in the interpretation of Num-expressions and that of CL-expressions is
277 due to the difference in the interpretations of count and the mass versions of the noun. To account
278 for the observed pattern following Schvarcz and Rothstein (2017), we must assume that the dif-
279 ference in the two versions of the noun is as follows: the count version denotes both physical and
280 informational objects, whereas the mass version denotes physical objects only. While this way we
281 could account for the meaning-shift between Num-expressions and CL-expressions, we could not
282 account for the fact that nouns in CL-expressions can be modified by adjectives that apply to the
283 informational objects in the denotation of the noun, as shown in (23). In (23), the adjective *unal-*
284 *mas*, ‘boring’ can only apply to informational books, hence the denotation of *könyv*, ‘book’ cannot
285 be restricted to physical books. This shows that the meaning-shift between Num-expressions and
286 CL-expressions cannot be reduced to noun ambiguity, thus, the approach in Schvarcz and Rothstein
287 (2017) is not suited to account for the data in (18)–(20).

288 (23) *három darab unalmas könyv*
289 three CL_{general} boring book
‘three boring books’

290 We take that the case of mixed nouns and the potential difference in the interpretation of Num-
291 expressions and CL-expressions show that the classifier optionality phenomenon in Hungarian is
292 not sensitive to the grammatical distinction between mass and count nouns. If classifiers in ex-
293 pressions displaying the phenomenon were purely an indicator of some grammatical property of
294 the noun they modify, there should not be any semantic effect associated with their presence or
295 absence in a construction. However, this is not borne out from the data: the interpretation of CL-
296 expressions is more restricted to that of Num-expressions. Therefore, classifiers are not entirely
297 optional with mixed nouns, but their presence or absence is affected by semantic factors.

298 6 The analysis of Hungarian classifiers

299 Before we give our analysis of Hungarian classifiers, we set out our assumptions about the seman-
300 tics of cardinal numerals and counting. Numerals in Hungarian are treated as expressions referring
301 to entities of type *n*, i.e. the numerical value expressed by the numeral; see our lexical entry for the
302 numeral *three* in (24). In the prenominal modifier position, numerals are shifted from terms of type
303 *n* into predicate modifiers of type $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$ (see a similar analysis of Hungarian numerals in

304 Schvarcz (under review)).⁵ Numerals in prenominal modifier position are devices for the operation
 305 of counting: they give the cardinality of an entity in the denotation of the predicate they modify,
 306 where the cardinality of an entity is understood as the cardinality of the set containing all the parts
 307 of the entity that are the smallest elements in the denotation of the predicate; see (25).⁶

308 (24) $\llbracket \text{three} \rrbracket = 3$

309 (25) $\llbracket \text{three}_{\text{PRENOM}} \rrbracket = \lambda P_{\langle e,t \rangle} . \lambda x_e . P(x) \wedge |x| = 3$, where $|x| = 3 \leftrightarrow |\{y : y \leq x \wedge P(y) \wedge \neg \exists z . z <$
 310 $y \wedge P(z)\}| = 3$

311 Based on our observations summarized in Table 2 above, we propose a purely semantic analysis
 312 for Hungarian classifiers according to which the role of classifiers is to specify what kind of entities
 313 are counted by the numeral in the construction. In our approach, the counting operation introduced
 314 by classifiers selects parts of the entity denoted by the noun that are physically distinct. Here
 315 we assume that physically distinct entities have the property of being Maximally Strongly Self-
 316 Connected (MSSC, Grimm 2012). MSSC-entities are entities that are whole in the topological
 317 sense with respect to a property P , that is, they are the smallest physically bounded entities in the
 318 denotation of the predicate. If y is a Maximally Strongly Self-Connected entity given the predicate
 319 P , the relation MSSC holds between y and P (formally represented as $\text{MSSC}(y, P)$). We analyze the
 320 general classifier *darab* as incorporating this MSSC-property into the counting operation, see (26).

321 (26) $\llbracket \text{darab} \rrbracket = \llbracket \text{CL}_{\text{general}} \rrbracket = \lambda P_{\langle e,t \rangle} . \lambda n_n \lambda x_e . P(x) \wedge |x|_{\text{MSSC}} = n$, where $|x|_{\text{MSSC}} = n \leftrightarrow |\{y : y \leq$
 322 $x \wedge \text{MSSC}(y, P)\}| = n$

323 Given our assumptions about the semantics of numerals in (25) and that of the classifier *darab* in
 324 (26), we capture the semantic difference between Num-expressions and CL-expressions in Hungar-
 325 ian as follows: Num-expressions are analyzed as predicates that hold of any entity in the denotation
 326 of the noun, and this entity has n number of parts that are atomic with respect to the property ex-
 327 pressed by that noun; see (27). We assume that nouns like *könyv*, ‘book’, denote both physical
 328 and informational objects as atoms (and composite objects made up of physical and informational
 329 parts, as in Gotham (2017)), we straightforwardly account for the fact that in Num-expressions
 330 entities both in the physical object and the informational object sense can be counted. In contrast,
 331 in CL-expressions, the classifier restricts the counting operation to MSSC-entities, thus numeral
 332 provides the number of MSSC-parts of the entity to which the CL-expression applies; see (28).]

333 (27) $\llbracket (1a) \rrbracket = \llbracket \text{három könyv} \rrbracket = \lambda x . * \text{BOOK}(x) \wedge |x| = 3$

334 (28) $\llbracket (1b) \rrbracket = \llbracket \text{három darab könyv} \rrbracket = \lambda x . * \text{BOOK}(x) \wedge |x|_{\text{MSSC}} = 3$

335 As (27) and (28) demonstrate, our analysis of *darab*, $\text{CL}_{\text{general}}$, in (26) can account for the meaning-
 336 shift represented in Table 2. Moreover, the analysis can be directly extended to other classifiers

⁵We treat numerals in a prenominal modifier position as intersective modifiers, just as in Landman (2003) and Rothstein (2017). However, there these kind of modifiers are treated as expressions of type $\langle e, t \rangle$ that combine with a other expressions of type $\langle e, t \rangle$ via a special compositional operation, such as MODIFY in Chung and Ladusaw (2006). Here we assume that intersective modifiers combine with predicates via functional application.

⁶There are more fine-grained theories of atomicity and counting (see Rothstein 2010, Sutton and Filip 2016, a.o.), but the one in (25) is sufficient for the purposes of this paper. Our analysis of Hungarian classifiers should be compatible with these other theories.

337 in Hungarian like *szem*, ‘CL_{small round object}’, *fej*, ‘CL_{big round object}’, etc. We assume that these
 338 classifiers select nominals based on the size and shape of the entities in their denotation. Thus,
 339 they can be analyzed as lexically more specified versions of *darab*, CL_{general}, each of them adding
 340 some additional properties to be fulfilled by the entities that are counted. As we did with *darab*
 341 in (28), we represent these extra properties of these other classifiers as restrictions on the entities
 342 selected by the counting operation; see our lexical entry for *szem* in (29).

$$\begin{aligned}
 (29) \quad & \llbracket \text{szem} \rrbracket = \llbracket \text{CL}_{\text{small round object}} \rrbracket = \\
 & = \lambda P_{\langle e,t \rangle} . \lambda n_n \lambda x_e . P(x) \wedge |x|_{\text{MSSC-SMALL-ROUND}} = n, \text{ where} \\
 & |x|_{\text{MSSC-SMALL-ROUND}} = n \leftrightarrow |\{y : y \leq x \wedge \text{MSSC}(y, P) \wedge \text{SMALL}(y) \wedge \text{ROUND}(y)\}| = n
 \end{aligned}$$

346 Before concluding the paper, we would like to highlight some further characteristics of our anal-
 347 ysis of Hungarian classifiers. First, since we assume that the phenomenon of classifier optionality
 348 in Hungarian is not related to the grammatical mass/count distinction, we maintain that the reason
 349 why some nouns in Num-expressions are ill-formed is also semantic. That is to say, there are no
 350 rules in our grammar that prevent any noun from being directly modified by a numeral. According
 351 to our analysis in (25), numerals as prenominal modifiers are sensitive to non/atomicity of the de-
 352 notation of the noun they combine with, as they provide the number of parts of an entity that are
 353 atomic relative to the property expressed by the noun. Thus, if the noun in a Num-expression is
 354 such that its denotation is non-atomic – which is the case in most of the nouns that are grammati-
 355 cally mass – the Num-expression can be assigned no interpretation, as shown in (30a). However,
 356 if such nouns occur in CL-expressions, the expression is semantically well-formed because the
 357 counting operation targets MSSC-parts of the entity in the denotation of the noun, see (31a). In the
 358 case of nouns whose denotation is lexically non-atomic, MSSC-entities may be physically distinct
 359 pieces of matter (blocks, chunks, nuggets, etc.), and we assume that these entities can only be
 360 provided by the context.

- 361 (30) a. #*három agyag*
 three clay
 362 b. $\llbracket \text{három agyag} \rrbracket = \lambda x . \text{CLAY}(x) \wedge |x| = 3$
- 363 (31) a. *három darab agyag*
 three CL_{general} clay
 ‘three pieces of clay’
 364 b. $\llbracket \text{három darab agyag} \rrbracket = \lambda x . \text{CLAY}(x) \wedge |x|_{\text{MSSC}} = 3$
 365

366 Second, while our analysis correctly predicts the meaning shift described in Section 4, it does not
 367 predict that classifiers can in fact occur with abstract nouns, as shown in (32). Abstract nouns in
 368 CL-expressions pose the following problem for our analysis: they are assumed to have abstract ob-
 369 jects in their denotation, and since abstract objects do not have an extension in physical space, the
 370 MSSC-property imposed by the classifier cannot apply to them. Hence, we predict that classifiers
 371 should not appear with abstract nouns, but in fact, they do. This could be because the meaning of
 372 the classifier is coerced so that the MSSC-property is not understood along physical dimensions,
 373 but along some other relevant dimension(s) specified by the lexical meaning of the noun. In the
 374 case of *történet*, ‘story’, in (32), the classifier can restrict the counting operation to some arbitrary
 375 structural requirement imposed on complete individual stories, such as having a beginning, dis-
 376 cussion and an end. Moreover, both authors share the intuition that the main difference between

377 *három történet* and *három darab történet* in (32) can be captured as follows. While the former can
 378 pick out three partially overlapping stories, the latter can only apply to three separate ones. This
 379 can be attributed to a strong non-overlap requirement imposed by the classifier (where the overlap
 380 relation is understood along the dimension specified by the noun). The analysis for cases as in (32)
 381 (and as in (21) in Footnote 4) is subject to future research.

382 (32) *három (darab) történet*
 three CL_{general} story
 383 ‘three stories’

384 Third, the assumptions in this paper are based on nouns which are polysemous and have both physi-
 385 cal and informational object senses. Our analysis can be extended to concrete nouns – nouns which
 386 only denote physical objects. In such cases, we correctly predict that the classifier has no major se-
 387 mantic contribution: even though the Num-expression, *három fotel*, and the CL-expression, *három*
 388 *darab fotel*, in (33) are assigned to different interpretations (see (33a) and (33b), respectively), in
 389 both expressions, the counting operation can only target entities of the physical object type.⁷

390 (33) *három (darab) fotel*
 three CL_{general} armchair
 391 ‘three armchairs’
 392 a. $\llbracket \textit{három fotel} \rrbracket = \lambda x. \text{ARMCHAIR}(x) \wedge |x| = 3$
 393 b. $\llbracket \textit{három darab fotel} \rrbracket = \lambda x. \text{ARMCHAIR}(x) \wedge |x|_{\text{MSSC}} = 3$

394 7 Conclusions

395 In this paper, we presented a formal semantic analysis of Hungarian classifiers, according to which
 396 classifiers require the counting operation to select physically distinct entities. We argued for such
 397 an analysis based on the case of nouns that display the so-called classifier optionality phenomenon,
 398 and that are also polysemous having both a physical object and an informational object sense. We
 399 showed that with these nouns, the classifier is not, in fact, optional: if such nouns combine with
 400 a numeral directly, the numeral can count both the physical objects and the informational objects
 401 denoted by the noun. In contrast, if such nouns appear with a classifier, the counting operation
 402 is restricted to physical objects. This purely semantic difference between counting expressions
 403 with and without a classifiers shows that classifier optionality in Hungarian is not sensitive to
 404 the grammatical mass/count distinction, and so the presence or absence of the classifier does not
 405 indicate whether the noun is count or mass. Rather, classifiers are elements with their own lexical
 406 content that potentially restricts the type of objects to be counted in the denotation of the noun they
 407 modify.

408 References

409 Asher, N. (2011). *Lexical Meaning in Context. A Web of Words*. Cambridge University Press.

⁷Here we ignore the possible kind-reading of Num-expressions that disappears in CL-expressions, as discussed in Schvarcz and Nemes (2019). However, our analysis is compatible with the observations therein.

- 410 Chierchia, G. (1998). Plurality of mass nouns and the notion of “semantic parameter”. In S. Roth-
411 stein (Ed.), *Events and Grammar* (pp. 53–103). Studies in Linguistics and Philosophy 70.
412 Springer Netherlands.
- 413 Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese*, 174(1), 99–149.
- 414 Chung, S. & Ladusaw, W. A. (2006, December). Chamorro evidence for compositional asymmetry.
415 *Natural Language Semantics*, 14.
- 416 Csirmaz, A. & Dékány, É. (2014). Hungarian is a classifier language. In R. Simone & F. Masini
417 (Eds.), *Word Classes: Nature, Typology and Representations* (pp. 141–160). Amsterdam:
418 John Benjamins.
- 419 Gotham, M. (2017). Composing Criteria of Individuation in Copredication. *Journal of Semantics*,
420 34(2).
- 421 Greenberg, J. (1974). *Language Typology: A Historical and Analytic Overview*. Berlin: de Gruyter.
- 422 Grimm, S. (2012). Degrees of countability: a mereotopological approach to the mass/count distinc-
423 tion. In A. Chierchia (Ed.), *Proceedings of Semantics and Linguistics Theory 22* (pp. 584–
424 603). Retrieved from <http://dx.doi.org/10.3765/salt.v22i0.2633>
- 425 Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event seman-
426 tics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and Contextual*
427 *Expressions*. Groningen-Amsterdam Studies in Semantics. Dordrecht: Foris.
- 428 Krifka, M. (1995). Common nouns: a contrastive analysis of English and Chinese. In G. N. Carlson
429 & F. J. Pelletier (Eds.), *The Generic Book*. Chicago: Chicago University Press.
- 430 Landman, F. (1991). *Structures for Semantics*. Dordrecht: Kluwer.
- 431 Landman, F. (2003). Predicate-argument mismatches and the adjectival theory of indefinites. In
432 M. Coene & Y. D’Hulst (Eds.), *From NP to DP, Vol. 1: The Syntax and Semantics of Noun*
433 *Phrases* (pp. 211–237).
- 434 Landman, F. (2004). *Indefinites and the Type of Sets*. Oxford: Blackwell.
- 435 Link, G. (1983). The logical analysis of plural and mass terms: a lattice-theoretical approach. In R.
436 Bäuerle, C. Schwarze, & A. v. Stechow (Eds.), *Meaning, Use and Interpretation of Language*
437 (pp. 302–323). Berlin: Walter de Gruyter.
- 438 Pelletier, F. J. (1975). Non-singular reference: some preliminaries. *Philosophia*, 5(4), 451–465.
- 439 Pustejovsky, J. (1995). *The Generative Lexicon*. MIT Press.
- 440 Rothstein, S. (2010). Counting and the mass count distinction. *Journal of Semantics*, 27(3).
- 441 Rothstein, S. (2017). *Semantics for Counting and Measuring*. Dordrecht: Cambridge University
442 Press.
- 443 Schvarcz, B. R. (under review). *Hárman, sokan, mindannyian*: nominalized numerals and quanti-
444 fiers in Hungarian. In E. Coppock & T. Halm (Eds.), *Approaches to Hungarian 17*.
- 445 Schvarcz, B. R. (2014). *The Hungarians who say -nyi: Issues in Counting and Measuring in Hun-*
446 *garian*. M.A. Thesis, Bar-Ilan University.
- 447 Schvarcz, B. R. & Nemes, B. (2019). Plurality, classifiers and kind interpretation in Hungarian.
448 Poster presented at SinFonJA 12. 12–14 September 2019, Brno, Czech Republic.
- 449 Schvarcz, B. R. & Rothstein, S. (2017). Hungarian classifier constructions, plurality and the mass-
450 count distinction. In H. van der Hulst & A. Lipták (Eds.), *Approaches to Hungarian 15.*
451 *Papers from the 2015 Leiden Conference* (pp. 157–182). Amsterdam: John Benjamins.
- 452 Sutton, P. R. & Filip, H. (2016). Counting in context: Count/mass variation and restrictions on
453 coercion in collective artifact nouns. In M. Moroney, C. R. Little, J. Collard, & D. Burgdorf

454
455

(Eds.), *Proceedings of Semantics and Linguistics Theory 26* (pp. 350–370). Retrieved from <http://dx.doi.org/10.3765/salt.v26i0.3796>