ATOMS, GROUPS AND KINDS IN CZECH*

MOJMÍR DOČEKAL

Department of Linguistics and Baltic Languages
Masaryk University in Brno
Arna Nováka 1
602 00 Brno
Czech Republic
docekal@phil.muni.cz

Abstract: A semantics for various classes of Czech numerals is presented which is based on Landman's theory but also incorporates Chierchia's Neo-Carlsonian approach to kinds. On the basis of a wide range of empirical data it is argued that Czech overtly lexicalizes at least two covert "sort-shifting" operators that have been stipulated in order to derive certain fine-grained semantic effects.

Keywords: plurality, kinds, numerals, distributivity, collectivity, morphosyntax–semantics interface

1. Introduction

This paper provides evidence that Czech derivational morphology very clearly displays two shifting operators—the group-forming and the kind-forming operator—postulated in the plurality (Landman 2000) and kind-oriented (Chierchia 1998) semantic frameworks. Furthermore, the distributional constraints on Czech numeral/noun combinations emerge neatly

from the interaction of the shifting operators with the lattice-theoretical description of numerals and grammatical number.[1] The crucial piece of evidence comes from derivational morphological processes in Czech, namely from various numeral classes.

2. Data and a puzzle

Czech has a productive system of a semantically driven derivational morphology for numerals. In this paper I examine four classes of Czech numerals:

(i) basic numerals (e.g. *dv-a klíče* 'two keys'—morphological make-up: root-inflectional_marker (case, gender agreement with head noun)),

(ii) numerals with the suffix *-e* (e.g. *dv-oj-e klíče* 'two bunches of keys'—morphological make-up: root-non_cardinal_stem-derivational_suffix),

(iii) numerals with the suffix *-ice* (e.g. *dv-oj-ice mužů* 'a group of two men'—morphologically composed of: root-non_cardinal_stem-derivational_suffix) and

(iv) numerals with the suffix *-í* (e.g. *dv-oj-í housle* 'two kinds of violin'—morphologically composed of: root-non_cardinal_stem-derivational_suffix).

Looking at the last three classes, it should be apparent that different suffixes create meaning distinctions with respect to the things counted: numerals always count objects, but the nature of the entities counted is dictated by the shape of the derivational suffix. So, the suffix *-í* counts objects as instances of a different kind than the suffix *-ice*, for instance, that cleary creates bunch names (similar in certain respects to bunch/group names like *committee* or *team* which, even though they are morphologically singular, denote plural entities). I will remain very vague as to the proper syntax and morphology of the suffixes discussed. However, notice that *-í* and *-é* suffixes derive numerals that necessarily have to agree with their head nouns, as opposed to the class (ii) suffix *-ice* that makes the numerals act more like nouns: such *-ice* numerals assign genitive to their

---

[1] One of the reviewers remarks that any analysis of groups and kinds would be compatible with the data discussed. I have to agree, if only partially: some approaches (beside Landman's work which I use as a framework) are compatible with the data. Nevertheless, plurality theories like Schwarzschild (1996) that discuss the plurality phenomena without group shifting operators are *prima facie* falsified by the data discussed here. I leave the comparison of alternative plurality theories and their relation to the Czech data for future work; the primary goal of the present article is to describe the data using one of the currently most worked-out plurality frameworks.

nominal complement (*mužů* in the example above) and if such a numeral is a sentential subject, the verb agrees with the numeral, not the nominal complement. In that sense, then, *-ice* numerals behave as nominal heads, while *-í* and *-é* numerals behave like adjectives. Nevertheless, as mentioned above, I will largely leave the morphology/syntax–semantics association for future work and will concentrate on the semantics of the numerals from now on.

The goal of the present article is, first, to describe the semantics of these complex numerals and, second, to shed some light on the compositional contribution of various derivational morphemes in the Czech numerals. Before getting there, however, I will introduce (standard) terminology in the plurality frameworks dealing with similar problems (classic references are Link 1983 and Landman 1996).

The basic distinction concerning the interpretation of plural noun phrases is the distinction between collective and distributive interpretations. Consider the example in (1) that can be interpreted in various ways. One interpretation says that six boys cooperated as a group and they together touched the ceiling—this reading is known as the **collective reading.** The other interpretation comprises six boys who each touched the ceiling, so in total the ceiling was touched six times—such a reading is called **distributive**. The readings are schematically depicted in Figure 1.

(1) Šest chlapců se dotklo stropu.
     'Six boys touched the ceiling.'

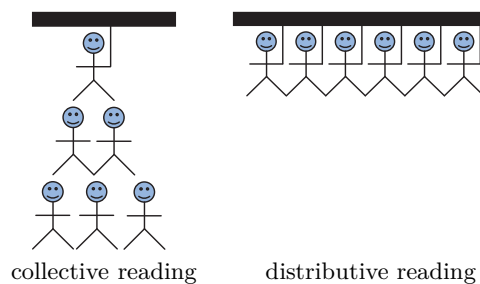collective reading   distributive reading

Fig. 1

The natural question to ask now is whether the distinction between the collective and distributive readings is due to vagueness or whether it represents a real semantic ambiguity. Most of the current semantic approaches opt for the ambiguity treatment. Leaving our intuitions aside, the strongest theoretical argument in favor of semantic ambiguity is found

in Krifka (1999). Let me repeat his basic argumentation in a few steps: a sentence of the form *n boys slept* entails sentences like *n' boys slept*, for any n' smaller then n, but is not entailed by them. Nevertheless, sentences in collective interpretation do not allow for these scalar entailments; assume that Jane and Mary own a dog together, then the sentence *Two girls own a dog* does not entail the sentence *One girl owns a dog*; on the other hand, the same sentence under the distributive interpretation allows for the entailment. This shows that the collective interpretation should be considered as a separate reading: should the distinction between collectivity/distributivity be only due to vagueness, the two different behaviors with respect to scalar entailments would be unexpected.

Let us assume that the distinction between the distributive and the collective reading is real, due to structural ambiguity, as argued above. There is yet another interpretation of the sentences with numerals, though. Consider (2): the sentence has the collective and the distributive meaning, but there is a third interpretation of (2): five boys split the chore of baking cakes: one baked one cake, three other boys baked two cakes together, and the last boy baked the last cake. This reading is known as the **cumulative** reading. All the three readings are in principle possible for simple numerals (their availability depends on the lexical semantics of the predicate in their sentences: predicates vary from the totally distributive ones like *be born in Brno* via the ambiguous ones like *touch the ceiling* to the totally collective predicates like *gather* or *meet*). However, the situation changes if we use a morphologically more complex numeral, as the one in (3). The sentence in (3) lacks the distributive and cumulative readings and I infer it is because the ordinary numeral *pět* was substituted by a complex numeral *pět-ice*. The only interpretation (3) has, then, is the collective interpretation. The lack of the distributive interpretation for the numerals like *pětice* (and generally for all numerals with the suffix *-ice*) is shown by the ungrammaticality of sentences like (4b): such numerals are incompatible with an (obligatorily distributive) adjunct *po praseti.*

(2) Pět chlapců   upeklo   čtyři koláče.
     five boy.gen.pl baked.sg.neut four pies
     'Five boys baked four pies.'

(3) Pětice   chlapců   upekla   koláč.
     five.sg.fem boy.gen.pl baked.sg.fem pie
     '(A group of) five boys baked a pie.'

(4) (a) Pětice   chlapců   vyhrála   prase.
        Five.sg.fem boy.gen.pl won.sg.fem pig
        '(A group of) five boys won a pig.'

  (b) #Pětice chlapců vyhrála po praseti.
        '(A group of) five boys won a pig each.'

The distribution of -*ice* numerals exhibits many morphological subtleties which, for the time being, I need to put asside (but notice in passing that -*ice* numerals can probably combine only with animate nouns and, moreover, their productivity is limited: the suffix can be added to numerals from 2 to 8 only). Nevertheless, from the semantic point of view, the suffix -*ice* is a morphological reflex of the collective interpretation of the -*ice* marked noun phrase. Such a claim can be tested in various contexts:

1. -*ice* numerals cannot cooccur with unambiguously distributive predicates, see (5);

2. -*ice* numerals cannot cooccur with cumulative predicates—see (6);

3. predicates showing ambiguity (collective/distributive/cumulative reading) are disambiguated by the suffix -*ice*: once -*ice* is added, only the collective reading is available, along with the usual collective implications—see (7) and (8).

The fact that the -*ice* numerals cannot cooccur with distributive predicates is not surprising, neither is their ungrammaticality in connection with cumulative predicates like 'lay eggs'. Moreover, this fact supports Landman's claim that cumulative reading is a sort of distributive plurality reading without a scope mechanism.

(5) #Pět-**ice** chlapců má křivý nos.
    '(A group of) five boys have a wry nose.'

(6) #Troj-**ice** slepic snesla pět vajec.
    '(A group of) three hens laid five eggs.'

(7) Šest-**ice** chlapců se dotkla stropu.
    '(A group of) six boys touched the ceiling.'   = only collective interpretation

(8) (a) Troj-**ice** chlapců napsala dopis.
        '(A group of) three boys wrote a letter.'   = only one letter written
    (b) Tři chlapci napsali dopis.
        'Three boys wrote a letter.' = either 1 (collective) or 3 (distributive) letters

To summarize so far: my goal is to show that the behavior of *-ice* numerals (and also the behavior of two other complex Czech numerals) follows from their morphological make-up. Alongside, I will also discuss another puzzle concerning Czech numerals, namely the following: In Czech, pluralia tantum like *kalhoty* 'trousers' or *nůžky* 'scissors' are incompatible with basic numerals, as shown in (9). This is not to say that pluralia tantum in Czech cannot be counted at all; naturally, they can, but not using basic numerals; the numerals with the suffix *-e* or *-í* have to be used. I dedicate the last section of the present paper to this restriction. But first I will present a more general description of the four classes of Czech numerals.

(9)   Petr koupil *dvě/dv-oj-**e**/dv-oj-**í** kalhoty/nůžky.
        'Petr bought *two/two pairs of/two kinds of trousers/scissors.'

3. Proposal

I will model the semantic contribution of the numeral suffixes in the plurality framework of Landman (1998, 2000). Hence I start by presenting the basic machinery. Let us assume a Boolean domain with three individuals in it, as shown in (10). The individuals at the bottom line are singularities, the atoms of the model; the entities above the singularities are plural entities. In the Boolean semi-lattice, the domain is partially ordered by $\subseteq$, the part-of relation, and closed under $\sqcup$, the sum or join operation. The formal axioms of the model can be found in the standard accounts of singular/plural distinction (Link 1983; Landman 1989), where semi-lattices like (10) are used to model denotations of count expressions. I do not take a stand with regards to the lattice/set theoretic treatment of plurality (for elaborate comparisons of the two views, see Landman 1989 and Schwarzschild 1996). For the purposes of this article, both approaches would work; however, I adopt the lattice framework.

        Concerning the denotation of singular and plural nouns, singular count nouns (like *dog*) denote a set of atoms or, as in (10), the elements at the very bottom of the semi-lattice, here a, b, c. The plurals (like *dogs*) denote the set of atoms closed under the sum, that is, the set of elements $\{a,\ b,\ c,\ a \sqcup b,\ a \sqcup c,\ b \sqcup c,\ a \sqcup b \sqcup c\}$ in (10).

        In Link's semantics, a singular predicate like BOY denotes a set of singular individuals only, hence a set of atoms. Pluralization is a closure under sum: $*$BOY adds to the extension of BOY all the plural sums

(10)

	$a \sqcup b \sqcup c$	
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$
a	b	c

that can be formed from elements of BOY, as illustrated in (11). Beside the pluralization operator $*$ (which is a theoretical tool for the description of plural on nouns), let us assume the group-forming operation $\uparrow$, which is an operation that maps a sum onto an atomic (group) individual in its own right. Landman's (2000) definition is shown in (12) and (13) gives mechanics for it to work. The group-forming operation is an operator that packages pluralities into atoms. According to (12), $\uparrow$ operates in the domain of sums of individuals (assembled from individuals, and called SUM-IND in (12)) and its output belongs to the domain of groups (GROUP in (12)). The same process can be observed in the behavior of bunch denoting nouns like *team, committee* or *government*. Bunch nouns, despite their morphological singularity, denote a plurality.

(11) $*\mathrm{BOY} = \{d \in D : \text{ for some non-empty } X \subseteq \mathrm{BOY}: d = \sqcup X\}$

(12) $\uparrow$ is a **one-one** function from SUM into ATOM such that:

    1. $\forall d \in \text{SUM-IND}: \uparrow(d) \in \mathrm{GROUP}$
    2. $\forall d \in \mathrm{IND}: \uparrow(d) = d$

    $\downarrow$ is a function from ATOM onto SUM such that:

    1. $\forall d \in \mathrm{SUM}: \downarrow(\uparrow(d)) = d$
    2. $\forall d \in \mathrm{IND}: \downarrow(d) = d$

(13)

	$a \sqcup b \sqcup c$			$\uparrow(a \sqcup b) \sqcup \uparrow(a \sqcup c) \sqcup \uparrow(b \sqcup c)$		KIND
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$	$\uparrow(a \sqcup b) \sqcup \uparrow(a \sqcup c)$	$\uparrow(a \sqcup b) \sqcup \uparrow(b \sqcup c)$	$\uparrow(a \sqcup c) \sqcup \uparrow(b \sqcup c)$	SUM
a	b	c	$\uparrow(a \sqcup b)$	$\uparrow(a \sqcup c)$	$\uparrow(b \sqcup c)$	ATOM
INDIVIDUAL			GROUP			

As an anonymous reviewer correctly points out, there arise serious problems with the pluralization operator and counting of groups. The data will be discussed in section **3.3**, dedicated to numerals with the suffix *-e*. Intuitively, what a natural language requires is not only counting the number of individuals constituting some group atoms (like with *dv-oj-ice klíčů* 'a bunch of two keys'), but also the counting of groups constituted by an arbitrary number of atoms (like with *dv-oj-e klíče* 'two bunches of keys').

But because the star pluralization operator—as defined in (11)—works only in the individual domain, such counting is not possible. Thus, let me define a minimally different pluralization operator $*^G$ that operates in the domain of GROUPS; this addition will also address the issue raised above. This operator is just a variant of the standard $*$ operator, but allows us to count the cardinality of groups it creates.

(11′) $*^G$ BOY-GROUP $= \{d \in D$: for some non-empty $X \subseteq$ BOY-GROUP: $d = \sqcup X\}$

3.1. Noun phrases

Landman's theory (adopted here) treats non-quantificational noun phrases differently from the quantificational ones. Non-quantificational NPs include proper names, definites and indefinites. For these, let us assume with Landman (1989) that they can shift their interpretation from the plural to the group freely. NPs like *John and Mary* and *three boys* have thus two interpretations: both non-quantificational NPs can be interpreted either as the set of properties that a sum of three boys (or the sum of John and Mary) have, or, alternatively, the NPs can be interpreted as the set of properties that a group of three boys (or the group of John and Mary) has. The first interpretation is called **sum** interpretation and is responsible for the distributive reading of sentences containing such NPs; the second interpretation is called **group** interpretation and is interpretation of non-quantificational NPs in sentences with collective predicates. Unlike non-quantificational NPs, quantifiers get their standard interpretation (as in (15)) and their standard interpretation is obligatorily atomic. That explains the empirical fact that genuine quantifiers like *every* or *no* are obligatorily distributive, as witnessed by the ungrammaticality of sentences like **Every student met in Prague* where the distributivity of the quantifier clashes with the collectivity of the predicate.

(14) (a)   John and Mary $\rightarrow j \sqcup m, \uparrow (j \sqcup m)$

     (b)   three boys $\rightarrow \lambda P.\exists x \in *BOY : |x| = 3 \wedge P(x)$ (sum)
                            $\rightarrow \lambda P.\exists x \in *BOY : |x| = 3 \wedge P(\uparrow (x))$ (group)

(15) every girl $\rightarrow \lambda P.\forall x \in$ GIRL: $P(x)$
     no girl $\rightarrow \lambda P \forall x \in$ GIRL: $\neg P(x)$

3.2. Basic numerals

Now, let us apply the plurality framework to Czech numerals. As predicted, **basic numerals** can have either sum or group interpretation, since they belong to the non-quantificational NPs. The basic numerals can shift their interpretation from the plural to the group one: the plural (sum) interpretation leads to the distributive interpretation of the predicate, the group interpretation leads to the collective interpretation of the predicate. They count atoms in the denotation of the noun they modify and the cardinality of the noun must be equal to the number denoted by the numeral. The atomicity condition on the noun denotation is responsible for the fact that the connections of basic numerals with mass nouns (e.g,. *two snows, *two furnitures*) are ungrammatical. The atomicity of counting is realized independently of the inner complexity of the modified noun—the numerals count ordinary atoms like *keys* as well as the group-atoms like *committee*—see (16a) and (16b).

(16) (a) dva klíče = 2 atoms in the denotation of the noun *klíč* 'key'

   (b) dva výbory = 2 atoms in the denotation of the noun *výbor* 'committee'

To make the distinction between readings easier, I obligatorily quantify-in the sum interpretation of the numerals in (17a′) and leave aside the potential cumulative reading which would be hard to get with the predicate like *bring* and a singular subject. The sentence like (17a), then, has two possible meanings, as expected: the distributive one (the numeral is interpreted as sum) in (17a′) and the collective one (the numeral is interpreted as group) in (17a″). The visualization of the two interpretations is shown in (18): we either count atoms (so in the universe with three keys, we have three possible denotations of the NP *two keys*) or we count group atoms—see the left and the right side of the table, respectively.[2] Czech basic numerals behave the way any other ordinary numerals in natural

---

[2] The purpose of tables like (18), (21) and (24) is purely illustrational. They should help readers unfamiliar with plurality theories to imagine how formulas like (17) work. The tables are here for the benefit of such readers, as presenting the theory in this particular way showed useful for various audiences. An anonymous reviewer remarks on a relative difficulty with the intuitive grasp of the graphics. I have to agree: it is true that in the tables no information is added to what follows from the formulas. Readers familiar with such frameworks are welcome to ignore the tables.

languages do. However, things get more interesting once we consider the complex numerals that we discuss in the following sections.

(17) (a)   Petr přinesl dva klíče. 'Petr brought two keys.'

     (a′) $\exists x \in {*}\text{KEY}: |x| = 2 \land \forall a \in \text{AT}(x): \exists e \in \text{BRING}: \text{Ag}(e) = \text{PETR} \land \text{Th}(e) = x$

     (a″) $\exists e \in \text{BRING}: \text{Ag}(e) = \text{Petr} \land \exists y \in {*}\text{KEY}: |y| = 2 \land \text{Th}(e) = {\uparrow}(y)$

(18)

$a \sqcup b \sqcup c$			$\uparrow (a \sqcup b) \sqcup \uparrow (a \sqcup c) \sqcup \uparrow (b \sqcup c)$			KIND
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$	$\uparrow (a \sqcup b) \sqcup \uparrow (a \sqcup c)$	$\uparrow (a \sqcup b) \sqcup \uparrow (b \sqcup c)$	$\uparrow (a \sqcup c) \sqcup \uparrow (b \sqcup c)$	SUM
a	b	c	$\uparrow (a \sqcup b)$	$\uparrow (a \sqcup c)$	$\uparrow (b \sqcup c)$	ATOM
INDIVIDUAL			GROUP			

3.3. Numerals with the suffix -e

The first type of Czech complex numerals that I discuss here are numerals with the suffix -e. Unlike basic numerals, they count sums or groups, yet they do not care about the number of atoms constituting the sums/groups; the number of atoms in the sums or groups is arbitrary, but the totality of the sums/groups is equal to the number denoted by the numeral. The sums/groups can even be constituted by identical atoms. The compositional contribution of the suffix -e, then, is the shift of granularity of the object counted from atoms to sums/groups. So, for instance, a noun phrase *dva klíče* denotes any sum/group consisting of two atomic keys, but a noun phrase like *dv-oj-e klíče* denotes any two sums/groups consisting of arbitrary cardinality of keys. As the -e numerals can denote both sums and groups, their interpretation is either distributive or collective: see (20a) and its distributive interpretation in (20a′); there is a sum of two pluralities of keys and for each atom in the sum, there was a separate event of its bringing. The collective reading of (20a) shown in (20a″) means that there is a plurality of keys consisting of two group atoms, and this plurality was the theme of the event of being brought by Peter. The table in (21) visualizes the distributive reading. The number of readings (w.r.t. basic numerals) did not change, but the nature of the objects counted is different. As discussed in section **3**, sometimes we count the number of groups instead of the number of atoms constituting the groups and -e numerals present exactly this case. That is the reason why in (20a′) and (20a″) I use the modified pluralization operator $*^{\text{G}}$ instead of the standard $*$ operator.

(19) dvoje klíče = 2 sums/groups (of arbitrary complexity) of keys

(20) (a) Petr přinesl dv-oj-e klíče. 'Petr brought two bunches of keys.'
    (a$'$) $\exists x \in {}^{*\mathrm{G}}\mathrm{KEY}{:}|x| = 2 \wedge \forall a \in \mathrm{AT}(x) : \exists e \in \mathrm{BRING}: \mathrm{Ag}(e) = \mathrm{PETR} \wedge \mathrm{Th}(e) = x$
    (a$''$) $\exists e \in \mathrm{BRING}: \mathrm{Ag}(e) = \mathrm{Petr} \wedge \exists y \in {}^{*\mathrm{G}}\mathrm{KEY}{:}|y| = 2 \wedge \mathrm{Th}(e) = \uparrow (y)$

(21)

	$a \sqcup b \sqcup c$			$\uparrow (a \sqcup b) \sqcup \uparrow (a \sqcup c) \sqcup \uparrow (b \sqcup c)$		KIND
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$	$\uparrow (a \sqcup b) \sqcup \uparrow (a \sqcup c)$	$\uparrow (a \sqcup b) \sqcup \uparrow (b \sqcup c)$	$\uparrow (a \sqcup c) \sqcup \uparrow (b \sqcup c)$	SUM
a	b	c	$\uparrow (a \sqcup b)$	$\uparrow (a \sqcup c)$	$\uparrow (b \sqcup c)$	ATOM
INDIVIDUAL			GROUP			

3.4. Numerals with the suffix *-ice*

The second class of Czech complex numerals, those with the suffix *-ice,* is clearly different from the other types: unlike the *-e* numerals, the *-ice* numerals count atoms in the denotation of the noun phrase they modify, but the result is interpreted as a group-atom (formally through ↑: Landman's group-forming operator). The interpretation of *-ice* numerals is similar to bunch-denoting words like *committee* or *team*: *-ice* numerals (like bunch nouns) carry singular number, as reflected by the agreement of the verb. Look at the examples in (22): (22a) and (22b) with basic and *-e* numerals have plural verbal agreement, but (22c) with an *-ice* numeral shows singular verbal agreement. I assume that this morphological feature of singular on the verb in (22c) is simply a reflex of agreement between the subject NP and the verb, with the *-ice* numeral playing the role of the subject. It also shows that morpho-syntactically *-ice* numerals seem to be more noun-like elements, rather than plain numerals. In general, Czech numerals either agree with their head noun (like *-í* and *-e* numerals or cardinal numerals less than five), or they have a neuter agreement; however, *-ice* numerals show a regular feminine agreement with verbs (or adjuncts which further modify the numeral). The shift to group denotation is, of course, also reflected by the inability of *-ice* numerals to be interpreted in the distributive way—see (23a) and its impossible distributive logical form in (23a$'$) and the only collective interpretation in (23a$''$).

(22) (a) Dva studenti přišli.
      'Two students came.pl'

    (b) Dv-oj-e klíče ležely na stole.
       'Two (bunches of) keys lay.pl on the table'

    (c) Dvoj-ice klíčů ležela na stole.
       'Two (bunches of) keys lay.sg on the table.'

(23) (a)   Petr přivedl dvojici studentů. 'Petr brought along two students.'

(a′) $*\exists x \in *\mathrm{STUDENT}: |x| = 2 \wedge \forall a \in \mathrm{AT}(x) : \exists e \in \mathrm{BRING}: \mathrm{Ag}(e) = \mathrm{PETR} \wedge \mathrm{Th}(e) = x$

(a″) $\exists e \in \mathrm{BRING}: \mathrm{Ag}(e) = \mathrm{Petr} \wedge \exists y \in *\mathrm{STUDENT}: |y| = 2 \wedge \mathrm{Th}(e) = \uparrow (y)$

(24)

	$a \sqcup b \sqcup c$			$\uparrow (a \sqcup b) \sqcup \uparrow (a \sqcup c) \sqcup \uparrow (b \sqcup c)$			KIND
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$	$\uparrow (a \sqcup b) \sqcup \uparrow (a \sqcup c)$	$\uparrow (a \sqcup b) \sqcup \uparrow (b \sqcup c)$	$\uparrow (a \sqcup c) \sqcup \uparrow (b \sqcup c)$	SUM
a	b	c	$\uparrow (a \sqcup b)$	$\uparrow (a \sqcup c)$	$\uparrow (b \sqcup c)$	ATOM
INDIVIDUAL			GROUP			

The obligatory group interpretation of the *-ice* numerals is also linked to their wide scope/independent reading: as groups are interpreted without quantifying-in, they are scopeless; that is, they do not scopally interact with the other plurality operators in a sentence.[3] Consider the following contrast: the basic numerals in (25a) can be interpreted either collectively or distributively. The distributive reading depends on the quantifying-in mechanism. On the other hand, in a sentence like (25b) with two *-ice* numerals as arguments, the object *-ice* numeral must be interpreted scope independently with respect to the subject: the interpretation in which the thieves vary with the policemen is not grammatical. This is expressed by the unavailability of the logical form (25b′) and the only appropriate logical form in (25b″).

(25) (a)   Dva policisté chytili tři zloděje.
          'Two policemen caught three thieves.'

    (b)   Dvojice policistů chytila trojici zlodějů.
          '(A group of) two policemen caught (a group of) three thieves.'

    (b′) $*\exists x \in *\mathrm{POLICEMAN}: |x| = 2 \wedge \forall a \in \mathrm{AT}(x):$
          $\exists e \in \mathrm{CATCH}: \mathrm{Ag}(e) = a \wedge \exists y \in *\mathrm{THIEF}: |y| = 3 \wedge \mathrm{Th}(e) = \uparrow (y)$

    (b″) $\exists e \in \mathrm{CATCH}: \exists x \in *\mathrm{POLICEMAN}:$
          $|x| = 2 \wedge *\mathrm{Ag}(e) = \uparrow (x) \wedge \exists y \in *\mathrm{THIEF}: |y| = 3 \wedge *\mathrm{Th}(e) = \uparrow (y)$

The analysis of *-ice* numerals presented in this section explains the data discussed in section **2**: as the suffix *-ice* is a morphological reflex of the group-forming operator $\uparrow$, we can infer the following predictions confirmed by the data:

---

[3] Of course, in a sentence with one distributively interpreted NP and one *-ice* numeral, the *-ice* numeral can be outscoped or it can be in the scope of the other element. But this scopal interaction is independent of the semantics/morphology of *-ice* numerals.

1. *-ice* numerals cannot cooccur with unambiguously distributive predicates, as shown in (5);

2. *-ice* numerals cannot cooccur with cumulative predicates, as shown in (6). See Landman (2000), where it is convincingly shown that the cumulative readings are non-scopal distributive readings. If *-ice* numerals disallow a distributive reading, then they also (by transitivity) cannot occur as arguments of cumulative predicates.

3. The predicates showing ambiguity in (their) reading are disambiguated by the suffix *-ice*, so that they pick up the collective reading only, along with the usual collective implications—see (7) and (8).

3.5. Numerals with the suffix *-í*

The numerals with the suffix *-í*, which is the last class of numerals discussed in this article, do not count atoms or pluralities: they count sub-kinds in the kind denotation of the noun they modify. So, (26) means that there are two sub-kinds of wine: red wine and white wine, and that the sum of the two sub-kinds is the most widespread among the kinds of wine. Sub-kinds are kinds in the sense that they are spatiotemporal discontinuous entities which share some common property/behavior: e.g. all red wines have some non-trivial properties, which are manifested in all their instantiations (made from grapes, having red color, being an alcoholic beverage . . .). The kind denoting nature of such NPs modified by *-í* numerals shows up in generic/characterizing sentences like (26), but if the kind-denoting NP occurs in the scope of an episodic predicate, it is automatically type-shifted by the operation **R**—see Krifka et al. (1995)—into a realization/instantiation of its kind. In generic sentences like (26), the kind-denoting NPs denote kinds; in episodic sentences like (27a), the kind-denoting NPs denote instantiations of the kind. Sentences like (27a) have two readings: a distributive one (for each sub-kind of keys there was a separate event of bringing an instantiation of the sub-kind), and a collective one (both instantiations of sub-kinds have been brought at once).

If we compare *-í* numerals with *-e* numerals (see example (27b)), we see that *-e* numerals count sums/groups in the denotation of their noun, while *-í* numerals count sub-kinds in the denotation of their noun. On top of it, there is a meaning distinction between (27a) and (27b), let us call it **the identity constraint**: (27a) must mean that the two bunches of keys

are different, which, as we will see, follows from the different sub-kinds interpretation of the *-í* suffix, while the bunches of the keys in (27b) can be identical.

(26) (a) Nejrozšířenější je dvojí víno: červené a bílé.
'The most widespread are two wines: red and white.'

    (a′) MOST_WIDESPREAD(winer^r $\sqcup$ wine^w)

(27) (a) Petr přinesl dv-oj-**í** klíče.
'Petr brought two kinds of keys.'

    (a′) $\exists x \exists x^k \in$ KEY-KIND: $\mathbf{R}(x, x^k) \wedge |x| = 2 \wedge \forall a \in \mathrm{AT}(x) : \exists e \in$ BRING:
$\mathrm{Ag}(e) = \mathrm{PETR} \wedge Th(e) = x$

    (a″) $\exists e \in$ BRING: $\mathrm{Ag}(e) = \mathrm{Petr} \wedge \exists x \exists x^k \in$ KEY-KIND:
$\mathbf{R}(x, x^k \wedge |x| = 2 \wedge \mathrm{Th}(e) = \uparrow(x)$

    (b) Petr přinesl dv-oj-**e** klíče.
'Petr brought two bunches of keys.'

    (b′) $\exists x \in {*}\mathrm{KEY}: |\uparrow(x)| = 2 \wedge \forall a \in \mathrm{AT}(x) : \exists e \in$ BRING:
$\mathrm{Ag}(e) = \mathrm{PETR} \wedge \mathrm{Th}(e) = x$

    (b″) $\exists e \in$ BRING: $\mathrm{Ag}(e) = Petr \wedge \exists y \in {*}\mathrm{KEY}: |\uparrow(y)| = 2 \wedge \mathrm{Th}(e) = \uparrow(y)$

The existence of this non-identity condition for instantiations of sub-kinds follows from the general way kinds work in a natural language. Independently of numerals, it was claimed for sub-kinds that there is a **disjointness condition** (see Carlson 1977), which, intuitively, says that no entity can belong to two (or more) different sub-kinds of the same kind. See the formalization in (28) and its demonstration in (29). All speakers of a natural language see the relation between (29a) and (29b) as an implication: the existence of two different cat sub-kinds in the next room implies the existence of at least two different cats (instantiating at least two different sub-kinds of cats) there, because no cat can belong simultaneously to two different cat sub-kinds. The disjointness condition is also the reason behind the interpretation of (30), where the universal quantifier is obligatorily interpreted with wide scope w.r.t. the kind denoting name *animals*: if we allowed linear scope, the kind referring noun *animals* would outscope the universal quantifier and we would get the implausible reading where the addressee should help animals which belong to all kinds. But as there are no such animals, only reverse scope is attested; that follows immediately from the disjointness condition. The same logic

explains the non-idenity interpretation of *-í* numerals, then. Since *-í* numerals denote different sub-kinds and since no object can belong to two different sub-kinds, the objects counted must be non-identical.

(28) $\neg \exists x^0 \exists y^k \exists z^k [y^k \neq z^k \land S_0(y) \land S_0(z) \land \mathbf{R}'(x,y) \land \mathbf{R}'(x,z)]$

(29) (a) There are two different kinds of cats in the next room.

    (b) $\rightarrow$ There are at least two cats in the next room.

(30) Help animals of all kinds!

If we keep formalizing the numeral suffixes as we did in the preceding sections, we must specify the precise meaning of the kind numerals. For kinds generally, I follow Chierchia's (1998, 348) influential treatment. For him, "kinds are generally seen as regularities that occur in nature. They are similar to individuals like you and me, but their spatiotemporal manifestations are typically 'discontinuous'". Translated into a semi-lattice jargon, the kind denotation of NPs is the maximal entity (spatiotemporally discontinuous), which is the sum of all appropriate atoms in the semi-lattice. Formally, as in (31), we can define this as the ι type shifting of some property P, which, as a result, returns the maximal entity consisting of the sum of all atoms with the property P. In (32), we see that in a semi-lattice with three atoms having the property P, the kind denotation of NP denoting P is the sum $a \sqcup b \sqcup c$. The instantiations of different sub-kinds, then, are any entities constituting the maximal entity.

(31) For any property P and world/situation s,
$^\cap P = \quad \lambda s \iota P_S$, if $\lambda s \iota P_S$ is in K
              undefined otherwise   (Chierchia 1998, ex. 14–16)

P_S is an extension of the predicate P in the world s.

(32)

	$a \sqcup b \sqcup c$			PL	KIND
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$			
a	b	c	SG		

Before moving on to the last section, let us make an interim summary of what has been achieved so far. I have looked at four classes of Czech numerals: basic numerals (like *dva klíče* 'two keys'), numerals with the suffix *-e* (like *dv-oj-e klíče* 'two bunches of keys'), numerals with the suffix *-ice* (like *dv-oj-ice klíčů* 'the group consisting of two keys') and, finally,

numerals with the suffix *-í* (like *dv-oj-í klíče* 'two kinds of keys). I have identified the semantic contribution of each derivational suffix and linked it to the plurality framework of Landman and the kind framework of Chierchia. Specifically, I described the most important meanings of each type of the numerals in the following way:

1. Czech basic numerals count sums or groups; their semantics is the same as basic numerals' in other languages.

2. Czech numerals with the suffix *-e* count sums or groups in the denotation of the noun that the numeral modifies. The counting is made over the number of sums/groups, not over the atoms that constitute the pluralities.

3. Czech *-ice* numerals count atoms in the denotation of the noun and the result is interpreted as a group-atom (formally through ↑—Landman's group-forming operator), the sum interpretation is not possible.

4. Czech *-í* numerals count sub-kinds in the denotation of the noun (formally through ∩—Chierchia's kind-forming operator). In episodic sentences, the denotation is type shifted into the instantiations of the sub-kinds (as with other kind referring noun phrases).

4. The puzzle

Let me end with a little puzzle concerning the distribution of Czech numerals. The description of their various classes done so far will prove handy now. Recall that there is a constraint allowing pluralia tantum to cooccur only with *-e* or *-í* numerals, as shown in (35). Pluralia tantum cannot cooccur with ordinary numerals. I argue that the reason for this restriction lies in the mass-noun algebraic structure of pluralia tantum in Czech. Concretely, Czech pluralia tantum lack singular counterparts that leads to the lack of their obligatory sum interpretation, usually observed with normal nouns. So, the NP *dogs* in most contexts denotes only sums and not atoms, but pluralia tantum like *scissors* denote sums as well as atoms. Pluralia tantum (in any language) can denote atoms and sums while ordinary plural count nouns must—in the majority of contexts—denote sums only. I assume that for ordinary plural count nouns, this ban on atom interpretation is achieved through the 'maximize presupposition' strategy, see Sauerland (2003). Link's semantics of morphological plural allows the atomic denotation of plurals also: recall that ∗ is defined as

closure under sum formation. Schematically, we can draw the distinction between count nouns and pluralia tantum as in (33) and (34).

(33) count nouns

	$a \sqcup b \sqcup c$		PL
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$	
a	b	c	SG

(34) pluralia tantum/mass nouns

	$a \sqcup b \sqcup c$		PL
$a \sqcup b$	$a \sqcup c$	$b \sqcup c$	
a	b	c	

Thus, in fact, I argue for the same type of denotation of pluralia tantum and of mass nouns: mass nouns, like pluralia tantum, are also able to denote atoms and sums (e.g., *snow* or *furniture* can denote one piece, two pieces, a group of the objects falling under the denotation of the noun, etc.). And since ordinary numerals are incompatible with mass nouns in any language (*two furnitures), pluralia tantum, by the same reasoning, are ungrammatical with Czech ordinary numerals. On the other hand, *-e* numerals can modify pluralia tantum because of the inner opacity of the sums/groups that the *-e* numerals count (see the example in (35) repeated from (9) above and its interpretation). The *-í* numerals, finally, can modify pluralia tantum as well, because pluralia tantum are similar to *-í* numerals with respect to their counting domain. Both *-í* numerals and *-e* numerals do not count atoms, but higher level pluralities: in the case of *-e* numerals it is sums/groups and in the case of *-í* numerals it is sub-kinds.

(35) Petr koupil *dvě/dv-oj-**e**/dv-oj-**í** kalhoty/nůžky.
 'Petr bought two/two pairs of/two kinds of trousers/scissors.'

5. Summary

In this article, I have shown the way various derivational suffixes on numerals change their meaning. Four classes of numerals have been described and one puzzle concerning the numeral–noun distribution has

been solved. The overall emerging picture points toward a deep formal elegance of the operations underlying the various classes of Czech complex numerals.[4]

References

Carlson, Greg N. 1977. Reference to kinds in English. Doctoral dissertation, University of Massachussetts at Amherst. Published in 1980 by Garland Press, New York.

Chierchia, Gennaro 1998. Reference to kinds across languages. In: Natural Language Semantics 6 : 339–405.

Krifka, Manfred 1999. At least some determiners aren't determiners. In: Ken Turner (ed.): The semantics/pragmatics interface from different points of view. Current research in the semantics/pragmatics interface 1, 257–91. Elsevier, Oxford.

Krifka, Manfred – Francis Jeffrey Pelletier – Gregory N. Carlson – Alice ter Meulen – Gennaro Chierchia – Godehard Link 1995. Genericity: An introduction. In: Gregory Carlson – Francis Pelletier (eds): The generic book, 1–124. The University of Chicago Press, Chicago.

Landman, Fred 1989. Groups I. In: Linguistics and Philosophy 12 : 559–605.

Landman, Fred 1996. Plurality. In: Shalom Lappin (ed.): The handbook of contemporary semantic theory, 425–59. Cambridge MA & Oxford.

Landman, Fred 2000. Events and plurality: The Jerusalem lectures. Kluwer, Dordrecht.

Link, Godehard 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In: Rainer Bäuerle – Christoph Schwarze – Armin von Stechow (eds): Meaning, use and interpretation, 303–23. Mouton de Gruyter, Berlin & New York.

Sauerland, Uli 2003. A new semantics for number. In: Rob Young – Yuping Zhou (eds): Proceedings of SALT 13, 258–75. CLC Publications, Cornell University, Ithaca NY.

Schwarzschild, Roger 1996. Pluralities. Kluwer, Dordrecht.

---

[4] As always, there remains a lot for future research. I am indebted to one of the reviewers for pointing me in the following direction (to which I hope to dedicate a future article). Notice that I did not discuss syntactic properties of the three numeral suffixes; however, they seem to behave very much like classifiers in languages like Chinese. To bring in this parallel for a full discussion here would take me far beyond the scope of the present article.

Another very important topic for future discussion is the precise truth-conditions for every suffix discussed. There are clear indications of what the compositional semantics of the suffixes should look like but, again, coming from a draft to the full painting is a lengthy process and I leave it for future work.