Counting Aggregates, Groups and Kinds: Countability from the Perspective of a Morphologically Complex Language

Scott Grimm and Mojmír Dočekal

1 Introduction: The Components of Countability and Their Interrelation

Theories of countability face the task of explaining how various nouns' different participation in grammatical number constructions corresponds to meaning contrasts among those nouns.[1][2] What, if anything, in a noun's meaning impinges on its ability to appear in different morphosyntactic contexts related to counting and/or measuring? Or in the other direction, how do morphosyntactic contexts impinge on the possible interpretations of a noun? An explanation of countability must show how and why countability distinctions arise, from morphosyntax or lexical meaning or from a combination thereof, and be predictive of grammatical number systems both generally and in particular languages.

Different proposals have staked out different positions along the spectrum of possible ways morphosyntax, lexical meaning and countability are related—from arguing that lexical meaning simply does not influence countability patterns (see Borer 2005 i.a.) to arguing that lexical meaning fully determines countability patterns (Wierzbicka, 1988; Wisniewski et al.,

[2]   Our use of the term Countability subsumes various phenomena broadly related to grammatical number, including the count/non-count contrast, collectives, duals, pluralia tantum and so forth, following the usage in Payne and Huddleston (2002), Joosten (2003), and Grimm (2012).

2003). The evidence for which analysis is best comes from natural language data, of course, and consequently the choice of which languages to discuss will predetermine the range of tenable conclusions about the nature of countability.

This paper adds to this discussion data from Czech and, as a consequence, directs the focus towards languages that have a morphologically rich nominal system. Importantly, more expressive nominal morphology may specifically target entity types which in less rich nominal systems remain morphologically undifferentiated. In the case of Czech, several complex cardinal expressions along with derivational morphology specifically target entity types beyond the object-level: namely aggregates, groups and (taxonomic) kinds, none of which, for instance, in English are morphologically distinguished in such a clear fashion. While the data presented here come mostly from Czech, we note that much of these phenomena appear in Slavic languages more generally: In all three sub-groups of the Slavic language group (West Slavic, East Slavic and South Slavic), there are numerical expressions and derivational morphology processes which restrict the domain of quantification to aggregates, groups and (sub)-kinds. Polish possess all three types of these expressions (only the taxonomic numerals are less productive and more archaic than in Czech, see Wągiel (2015) for details), in Russian there are at least group and aggregate denoting numerical expressions (see Khrizman 2016, to appear); finally Bosnian/Croatian/Serbian (BCS) seems to have in its inventory both group numerals and aggregate denoting expressions, too. Thus, while Czech is far from being exceptional among Slavic languages as far as possessing a rich derivational morphology system for nouns and numerals, investigating these constructions in Czech is advantageous as all three types of expressions can be found in contemporary spoken Czech, whereas only a subset of the three types of expressions can be found in other Slavic languages.

The greater complexity of the system of cardinal expressions has implications even at the level of the very diagnostics of countability. One of the clearest diagnostics of countability

is the combination of nouns with cardinal expressions, as in *two cats* (the impossibility of which is named the "signature property" of non-countable nouns by Chierchia 2010). In Czech, certain nouns may combine with some cardinal expressions yet not others, and accordingly be "countable" in different ways.

In this paper, we relate Czech's nominal system to three questions that are central to theories of countability. First, what must theories of countability assume as primitive? There are two sub-questions here. What elements serve as the ontological foundations of countability? And is one countability category, e.g. non-countable, more basic than another? Second, as Czech provides morphology for explicitly counting aggregates, what can be determined about the nature of aggregates? Finally, to what extent is countability lexically-specified as opposed to dependent on pragmatic context? Put in other terms, how intrinsic is nominal flexibility to countability? We now provide initial discussion of each question in turn.

Primitives of Countability The early theory of Link (1983) took the count-mass distinction to be a lexical contrast grounded in a basic ontological distinction between atomic and non-atomic entities, from which the syntactic patterns of countability followed. From this starting point, researchers have taken more and less parsimonious positions. Chierchia (1998a, 2010), for instance, takes a more parsimonious position wherein all nouns derive from a common ontological sort: All nouns have an atomic domain, but mass nouns, due to their inherent vagueness in Chierchia (2010), impede the ability to refer to the atomic parts. Alternately, many researchers account for three (grammatical) types of nouns — substances, individuals and nouns such as *furniture* that are non-countable yet refer to individuals — while retaining two ontological sorts, namely atomic entities and non-atomic entities, where nouns like *furniture* are comprised of atomic entities but are non-countable for some reason, e.g. a lexical feature is stipulated as in Barner and Snedeker (2005) or Bale and Barner (2009). On the other hand, Grimm (2012) accepts an ontological contrast between atomic and non-atomic entities, and

further argues that even more ontological contrasts are needed than just the one between atomic and non-atomic (substance) reference, namely reference to aggregates, for example, for nouns such as 'hair' or 'foliage', while <u>Grimm and Levin</u> (<u>2017</u>) argue for recognizing yet other ontological types in the artifactual domain, distinguishing between individual artifact nouns (*hammer*, *chair*) and artifactual aggregates (*furniture*, *laundry*).

Czech provides more evidence that nominal roots may be ontologically richer than the individual/substance contrast which has garnered the most attention in the literature. This is in part due to the complex cardinals mentioned, which directly target several ontological sorts (aggregates, groups and kinds), but also due to derivational morphology which observably derives aggregates, which form a sub-class of non-countable nouns in Czech with particular properties.

A related, yet slightly different question facing theories of countability is what to take as basic and what to derive. Countable and non-countable nouns could have equal status, or one class could be derived from the other. For <u>Borer</u> (<u>2005</u>), again an example of a parsimonious theory, all nouns begin as mass or "stuff", then are given further structure in syntax through the application of functional heads. To anticipate, the Czech system provides a complicated, and perhaps fatal, data point for theories where one class of nouns is derived from the other.

The Nature of Aggregates The discussion of aggregates as a separate ontological sort leads to the more specific question of what, if anything, can be determined about the nature of aggregates, or the various different sorts of things which have been labeled "aggregates". While the recent literature has often proposed a three-way contrast in nominal semantic types, substance (non-countable), aggregate (non-countable) and individuals (countable), where aggregate nouns are complete atomic join semi-lattices, that is, including atoms and sums (<u>Bale</u>

and Barner, 2009; Deal, 2017), many questions remain, most critically, why these particular nouns rather than others have aggregate denotations. Other questions include what the nature of any specific lexical semantics in play is and the cross-linguistically commensurability of different classes of aggregates. Czech's morphology isolates a reasonably coherent lexical class of aggregate nouns, prototypically based on spatial proximity of elements, which in turn sheds light on some of the preconditions that associate with aggregate nouns.

Nominal Flexibility In English, famous shifts in meaning have been attributed to operations known as 'grinding', 'packaging', and 'sorting' (see Pelletier 1979; Bunt 1985, i.a.), demonstrated with examples (1), (2) and (3) from Bach (1986). For these cases in English, erstwhile countable nouns appear as non-countable, as in (1), and vice versa, as in (2) and (3), given suitable context. The core question is how does the meaning of a word impinge on its ability to appear in different contexts, if at all?

(1) a. There was **dog** splattered all over the road.

     b. **Much missionary** was eaten at the festival

(2) ice-creams = 'portions of ice-cream'

(3) muds = 'kinds of mud'

While the existence of contextual shifts is uncontroversial, their importance for theories of countability is a matter of ongoing debate. Some researchers have taken nominal flexibility to be a foundational property of countability (Bale and Barner, 2009; Borer, 2005; Chierchia, 2010; Pelletier and Schubert, 2004), which is compatible with the data in English. On the other hand, there are serious questions as to how well the flexibility data in English carries over to other languages.

A valuable set of recent studies, including Dalrymple and Mofu (2012) on Indonesian, Lima (2014) on Yudja and Deal (2017) on Nez Perce, have reported on languages with restricted morphological means to grammatically code number distinctions and which simultaneously permit a large degree of nominal flexibility. In these languages, canonical mass nouns, like *mud* or *blood*, may be counted when appearing in contexts which license such interpretations. Although the authors have not settled whether in these languages there is a lexically encoded distinction between countable and non-countable nouns or not (see discussion in Deal 2017), these studies have greatly added to our knowledge about how such languages function.

Examining Czech, which uses rich inflectional and derivational morphology to code countability, provides a counterpoint study. As will become clear, Czech presents the opposite behavior than theories that emphasize flexibility would predict: While Czech possesses rich number morphology, it manifests limited nominal flexibility. More generally, we seek to provide evidence bearing on a larger question: As the syntax and morphology varies from language to language, how does this affect the distribution of labor between pragmatically available senses and morpho-syntactically specified senses?

This paper is organized as follows. In section 2, we introduce the core countability data from Czech, including the basic diagnostics as well as derivational morphology and complex cardinals. To the best of our knowledge, even this basic data has not received any substantial attention in the countability literature. Further, as we have indicated, the data itself already provide directions towards answering some of the questions brought up here, which we return to in section 3. Sections 4 and 5 develop a formal analysis of Czech's nominal system, providing an extended investigation of how the different semantic distinctions are parceled out to different parts of the morphosyntax. Altogether, this exploration of countability in a morphologically rich language advances the understanding of the interplay between morphosyntax, lexical meaning and countability.

2 Countability in Czech: The Core Data

As with the majority of Slavic languages, Czech is an SVO language possessing rich inflectional and derivational morphology. Czech nominals lack overt articles and those that appear bare can be interpreted in various ways, including definite, indefinite or generic interpretations.[3]

Czech manifests the familiar distributional contrasts for notionally countable and non-countable nouns. For instance, *pes* 'dog', a prototypical count noun, intuitively denotes individual dogs. Such nouns can be pluralized, as shown in (4-b), and also may combine both with cardinal numerals, as shown in (4-c) and (4-d). Cardinal numerals fall into two syntactic categories: adjectives for cardinal numbers 1, 2, 3 and 4, as in (4-c), and nouns for cardinal numerals 5 and greater, as in (4-d). There are two reasons for classifying numerals 5 and greater as nouns: (i) they (unlike cardinal numerals for 1, 2, 3 and 4) assign genitive case to their nominal complements and (ii) if the cardinal numeral for numbers 5 and above is the subject of a sentence, the verb does not agree with the nominal complement (of the numeral). Countable nouns also combine with determiners such as *mnozí* 'many', a vague adjectival determiner agreeing with the noun it modifies in both case and number.

(4) a. pes-∅

      dog-SG

      'dog'

---

[3]   See Dayal (2004), Filip (1999), and Krifka (1992), among others, for discussions of the interpretation of bare nouns in Czech in connection with lexical and grammatical aspect.

b. ps-i

   dog-PL

   'dogs'

c. dva/tři   ps-i

   two/three dog-PL

   'two/three dogs'

d. pět ps-ů

   five dogs-GEN.PL

   'five dogs'

e. mnoz-í   ps-i

   DET-NOM.PL dog-NOM.PL

   'many dogs'

We take *bláto* 'mud' as an instance of a prototypical non-countable noun, denoting a substance. As shown in (5), these nouns cannot be pluralized, nor do they combine with cardinal numerals or with determiners such as 'many'. However, these nouns do combine with certain determiners, such as *všechno*, roughly equivalent to English 'all', which requires singular agreement on the noun, shown in (6).

(5) a. blát-o

      mud-SG

      'mud'

   b. *blát-a

      mud-PL

      'muds'

c. *dvě/tři   blát-a

   two/three mud-PL

   'two/three mud'

d. *mnoh-á   blát-a

   DET-NOM.PL mud-NOM.PL

   'many mud'

(6) a. všechn-o blát-o

   DET-SG mud-SG

   'all mud'

b. všechn-a vod-a

   DET-SG water-SG

   'all water'

Another diagnostic from English, that only non-countable nouns may appear as bare singulars, does not succeed in Czech, as countable nouns may also appear as bare singulars. We also note that Czech bare singulars are able to refer to objects (7-a) as well as to kinds (7-b) (at least when the bare singulars are arguments of kind-level predicates as in (7-b)).

(7) a. Blesk   způsobil požár   a   zabil mamut-a.

   thunderbolt caused   wildfire and killed mammoth-ACC.SG

   'A thunderbolt caused a wildfire and killed a mammoth.'

b. Mamut-Ø   vymřel i   v oblastech, kde   bylo málo lidí.

   mammoth-NOM.SG died-out even in areas   where was few   people

   'Mammoths died out even in areas where there were only few people.'

For certain prototypical non-countable nouns, such as *voda* 'water', Czech also manifests an additional interpretation, that is, a "packaging" use, as witnessed in (8-a). Yet, the ability for

non-countable nouns to be used with a packaging interpretation is highly constrained, typically only found in conventionally licensed instances, similar to what has been observed for English. Attempting to establish a parallel packaging interpretation with *písky* 'sand' in (8-b) fails: Even when speakers are given a background scenario in which a quarry sells different packages of sand to builders, Czech native speakers reject (8-b) and deem it unnatural.

(8) a. mnoh-é   vod-y   už   byl-y   vyprod-ané

        many-NOM.PL waters-NOM.PL already were-PL sold-PST.PART

        'many bottles of water were already sold out'

    b. #mnoh-é   písk-y   už   byl-y   vyprod-ané

         many-NOM.PL sands-NOM.PL already were-PL sold-PST.PART

       Intended: 'many packages of sand were already sold out'

So far, the nominal classes that can be isolated through distributional tests match up with expectations derived from English and other Western European languages. Yet, the data in Czech is far more involved. The remainder of this section explores the various types and interpretations of nouns, made visible through complex morphology on both nouns and cardinal numerals, and then turns to consider the issue of nominal flexibility. Through its derivational morphology, Czech identifies entity types critical to understanding countability, but which have garnered less discussion in the literature than the individual/substance distinction, namely derived non-countable aggregates (*list-í* 'foliage'), and complex numerals which count groups (*dvoj-ice* 'a group of two'), connected clusters (*dvoje hranolky* 'two collections of French fries') and taxonomic kinds (*dvojí metr* 'two kinds of measure'). We discuss each in turn.

2.1 Derived Aggregates

In addition to nouns which we consider to be lexically non-countable, Czech possesses derivational morphology, the suffix -*í*[4], which derives non-countable nouns. Table 1 presents the nouns known to be derived by -*í*.[5] While these nouns are restricted to be inanimate, they arise in several semantic domains. Schematically, -*í* applies to a countable noun root *N* and results in an interpretation 'a collection of N', although some of the examples illustrate that an even more specific (lexical) semantics may be necessary.[6]

---

[4]   There are several variant forms of -*í*, such as -*oví*, which as far as we can discern, do not differ semantically.

[5]   The items presented in the table comprise the majority of Czech nouns derived by -*í*. A few other lexical items exist but they are archaic or at the periphery of contemporary Czech.

[6]   For 21 of the 22 nouns given, -*í* clearly applies to purely countable nouns, as we verified through applying the standard tests for distinguishing countable and non-countable nouns. These nouns are often naturally found in counting contexts as shown in the following example for *krajka* 'lace', where the word *krajkoví* 'lacework' designates pieces of lacework.

(i) přida-l-a   ještě další dv-ě   krajk-y   a   vyda-l-a   nový malý
     add-PST-3SG still next two-ACC.PL lace-ACC.PL and publish-PST-3SG new   small
     sešit   se   čtyř-mi   krajka-mi
     notebook with four-INS.PL lace-INS.PL
     'She added two pieces of lacework and published a new small notebook with four pieces of lacework.'

     The only exceptional noun is *dříví*, derived from *dřevo* 'wood', which is ambiguous between countable and non-countable interpretations. We assume that -*í* applies to the countable form.

Table 1: Non-Countable Nouns Derived by *í*

Noun	Gloss	Derivational Source
Trees		
stromoví	'clump of trees'	*strom* 'tree'
boroví	'clump of pine trees'	*borovice* 'pine tree'
habroví	'clump of hornbeam trees'	*habr* 'hornbeam tree'
olšoví	'clump of alder trees'	*olše* 'alder tree'
vrboví	'clump of willow trees'	*vrba* 'willow tree'
Plants		
listí	'foliage'	*list* 'leaf'
rákosí	'rushes'	*rákos* 'reed'
trní	'thorns, brambles'	*trn* 'thorn'
jahodí	'clump of strawberry plants'	*jahoda* 'strawberry'
lískoví	'clump of hazel bushes'	*líska* 'hazel tree'
maliní	'clump of raspberry plants'	*malina* 'raspberry'
ostružiní	'clump of blackberry plants'	*ostružina* 'blackberry'
Complex Objects		
cihloví	'brickwork'	*cihla* 'brick'
krajkoví	'lacework'	*krajka* 'lace'
laťkoví	'fence (made from laths)'	*laťka* 'lath/slat'
nádobí	'dishes'	*nádoba* 'container'
síťoví	'netting/nets'	*síť* 'net'
Nautical Terms:		
lanoví	'rigging/ropes'	*lano* 'rope'
plachtoví	'sails'	*plachta* 'sail'
ráhnoví	'spars'	*ráhno* 'spar'
Other:		
dříví	'firewood'	*dřevo* 'wood'
kamení	'rocks'	*kámen* 'rock'

Nouns derived with *-í* are strongly non-countable. Unlike the root nouns from which they are formed, nouns derived by *-í* do not pluralize nor combine with simple cardinal numerals or vague quantifiers, as shown in (9)-(11)[7], nor do they manifest packaging or otherwise countable interpretations, as shown in (12), unlike lexically non-countable nouns such as *voda* 'water'.

(9) a. list-í-m

      leaf-í-INST.SG

      'foliage'

    b. *list-í-mi

      leaf-í-INST.PL

      'foliages'

(10) a. dva   list-y

      CARD.Masc leaf-Masc.PL

      'two leaves'

    b. *dvě   list-í

      CARD.Neut leaf-í.Neut

      'two foliages'

(11) *mnohé list-í už   spadl-y

    many   leaf-í already fell-PL

    'many foliage already fell'

(12) ??břízy a   smrky   shodil-y mnohé list-í

    birches and spruces shed-PL many   leaf-í

    'birches and spruces shed many foliage'

---

[7]   Since case syncretism often obscure the morphological patterns in Czech, we often present nouns in the instrumental case which is morphologically more transparent.

As the examples in table 1 indicate, the derived nouns implicate multiple elements designated by the root noun, that is, a clump of pine trees implicates multiple pine trees. Further, nouns derived by *-í* pattern in their semantic properties with plural nouns: They are cumulative (*list-í + list-í = list-í*) and non-divisive, since there are atomic parts for which divisiveness does not hold — parts of *list-í* are *list-í* but the atomic parts (*list* 'leaf') are not *list-í*.[8]

Despite these similarities, the meanings of nouns derived by *-í* contrast strongly with ordinary plural meanings. As the glosses given in table 1 express, the resultant meaning of nouns derived by *-í* is not simply a set of, for example, pine trees, but a set in which the members are coherently related. This is most often in terms of spatial proximity, as in the case of clumps of trees or plants, but also may be in terms of functional interdependence, as in the case of *lanoví* where it is not the individual ropes at issue, but their coherent organization as part of a ship's rigging. Similarly, *cihloví*, derived from *cihla* 'brick', does not signify a random collection of bricks, rather they must be related in some manner, for instance, used together in an architectural motif.[9]

Additionally, *-í* is not nearly as productive as one would expect if it were only to signal plurality. As the examples in (13) demonstrate, the ability to derive new nouns with *-í* is limited. Such limited productivity is, however, consistent with the lexical restrictions typically found with collective morphology (see Acquaviva 2008 for discussion of various systems).

---

[8]   For reference, we present definitions of predicates which are cumulative, divisive and atomic (relative to a property) in (i).

    (i)   a.   Cumulative(P) = $[P(x) \land P(y) \rightarrow P(x \sqcup y)]]$
              b.   Divisive(P) = $\forall x[P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$
              c.   Atomic(x,P) $\leftrightarrow P(x) \land \neg \exists y[y < x \land P(y)]$

[9]   We confirmed these intuitions by examining uses of *-í* nouns along with their related images on the internet. For instance, in one illustrative example, the noun *cihloví* designates various clusters of bricks used in an architectural pattern, differentiated by colors. (https://www.novinky.cz/bydleni/tipy-a-trendy/240130-karnevalove-vile-rekonstrukce-vratila-puvodni-rozverny-styl.html)

(13) a. *stol-í

  table-í

  'collection of tables'

 b. *kabát-í

  jacket-í

  'collection of jackets'

 c. *planet-í

  planet-í

  'collection of planets'

Given these observations, the non-countability, at least at the object-level, of derived aggregates is expected, since these nouns do not simply designate entities, but coherently related groupings of entities. As will be seen in detail in section 2.2.2, Czech provides morphological mechanisms for counting precisely these more complex composite entities.

2.2 Complex Numerals

Czech, like other Slavic languages, possesses simple cardinal numerals, such as those shown in (14) as well as morphologically derived complex numerals.

(14) a. tři ps-i

  three dog-NOM.PL

  '3 dogs'

 b. tři hrnk-y

  three cup-NOM.PL

  '3 cups'

This section lays out the basic data from complex numerals in Czech, through which speakers make reference to groups, aggregates, and taxonomic kinds.

2.2.1 Derived group numerals

We first examine the complex numeral derived with the suffix *-ice*, which we term GROUP NUMERALS. These numerals produce an interpretation of "group of *n* Xs", such as "group of three sailors" given in (15-b). The distribution of the *-ice* suffix is restricted, both in the numerals it combines with, limited to basic numerals from 2 to 8, and in the nouns is combines with, restricted to animate nouns.

(15) a. tři   námořníc-i

           three sailor-NOM.PL

           '3 sailors'

       b. troj-ice   námořník-ů

           three-ICE sailor-GEN.PL

           'a group of 3 sailors'

We first make some remarks on the basic syntactic configuration of these numeral phrases.[10] The derivational suffix *-ice* applies to a non-cardinal numeral stem and takes a nominal complement in the genitive case (*námořníků* 'sailors' in the example in (15-b)). The derived complex numeral, rather than the noun, behaves as the head of the construction. First, *-ice* numerals assign genitive case to their nominal complement and there is no further agreement

---

[10]   We present here only properties which will be critical to our analysis of these constructions, while acknowledging that the literature on Slavic numerals and their syntactic properties is intricate. See Veselovská (2001), Ionin and Matushansky (2006), among others, for discussion.

between the derived numeral and the nominal complement. Additionally, if such a numeral is the subject of a sentence, the verb agrees with the numeral, not with the nominal complement. In these respects, the *-ice* complex numerals parallel syntactically the cardinal numerals for numbers 5 and above, since both cardinal and complex numerals are syntactically not modifiers, but rather behave like heads of the extended NP projection. But the two classes differ semantically: The cardinal numerals ≥ 5 permit collective and distributive interpretations, whereas the *-ice* complex numerals only permit a collective interpretation.[11]

The complex numeral formed with *-ice* may occur in both singular and plural, as shown in (16).

---

[11]   Some properties of the Czech group numeral are shared by a similar type of complex cardinal number in Russian, the collective numeral, discussed in Khrizman (2016, to appear). First, both types of numerals favor a "collective" interpretation, and second, Russian collective numerals are also restricted to combine with animate nouns (more specifically, the nouns must denote humans or young animals). Yet, there are many important differences between the Russian collective and Czech group numerals. Most critically, Russian collective numerals allow distributive readings (when they act as antecedents of dependent indefinites) and cumulative readings, while such readings are simply unacceptable for Czech group numerals. The Czech examples in (i-a) and (i-b) witness this fact. In (i-a), the obligatory distributive marker *po* heading the dependent indefinite phrase clashes with the collectivity of the *-ice* numeral construction, resulting in the sentence being unacceptable. In (i-b), the sentence has only collective-distributive interpretation: the group of three sailors has been lost twice (in different harbors). The cumulative interpretation, for example where Sailor A has been lost in a harbor 1 and Sailors B and C have been lost in a harbor 2, is unacceptable. Consequently, Czech group numerals seem to be a clear case of group interpretation unlike Russian collective numerals. In terms of the classification of collective predicates introduced by Winter (2001), Czech group numerals act like collective set predicates as indicated by the differing acceptability of Czech group numerals with quantifiers *all* and *every*, shown in (ii), similarly to other collective set predicates like *meet, gather, be brothers*, or *be similar*. This is formalized in section 5.2 accordingly. More data concerning Czech group numerals can be found in Dočekal (2012).

(i)   a.   *Troj-ice námořník-ů   vyhrál-a   po litru rum-u.
              three-ICE sailor-GEN.PL won-3SG.F DIS liter rum-GEN.SG
              'A group of three sailors won a liter of rum.'
       b.   Troj-ice   námořníků se   ztratil-a   ve dvou přístavech.
              three-ICE sailors   REFL lost-3SG.F in two   harbors
              'A group of three sailors has been lost in two harbors.'
(ii)   a.   ?Všichni podezřelí byl-a   troj-ice   námořník-ů.
              all   suspects AUX-3.SG three-ICE sailor-GEN.PL
              'All suspects were a group of three sailors.'
       b.   *Každý podezřelý byl troj-ice   námořník-ů.
              every suspect   was three-ICE sailor-GEN.PL
              '*Every suspect was a group of three sailors.'

On the other hand, Polish collective numerals like *dwójka* 'group of two' appear to behave exactly like Czech *-ice* numerals, as reported in Wągiel (2015).

(16) a. s   troj-ic-í   námořník-ů

          with three-ICE-INST.SG sailor-GEN.PL

          'with group of three sailors'

     b. s   troj-ice-mi   námořník-ů

          with three-ICE-INST.PL sailors-GEN.PL

          'with groups of three sailors'

Complex numeral phrases formed through the application of *-ice* can be further quantified over with cardinal numerals, as shown in (17). Further, they are compatible with the vague adjectival determiners *mnozí* 'many', as shown in (18). At the same time, unlike simple cardinal numerals, the use of *-ice* derived numerals is incompatible with the singular universal quantifier *všechno* 'all', as shown in (19).

(17) dvě troj-ice   námořník-ů

       two three-ICE sailor-GEN.PL

       'two groups of three sailors'

(18) mnohé troj-ice   námořník-ů

       many   three-ICE sailor-GEN.PL

       'many groups of three sailors'

(19) *všechna troj-ice   námořník-ů

        all   three-ICE sailor-GEN.PL

        'all group of three sailors'

Accordingly, counting groups through complex numerals provides only a partial match for the diagnostics of countability present for simple cardinal numerals, or countable nouns for that

matter. The other complex cardinals, to which we now turn, also show only partial matching, but in different fashions than the group numerals.

2.2.2 Complex Numerals for Aggregates

A second complex numeral, which we call an AGGREGATE NUMERAL, is formed through combining the suffix *-oje* with cardinal numeral roots and designates a number of collections of an entity. For instance, applied to the cardinal numeral root *dv-* 'two' yields *dv-oje*, which roughly translates to 'two collections', as shown in (20-a)-(20-d). This complex numeral adds (morphologically) the derivational suffix *-oje* to a cardinal stem. Unlike the derived group numerals discussed in the Section 2.2.1, *-oje* derived numerals are much more productive and appear to be derivable for any cardinal numeral, though there is allomorphy: the morpheme *-oje* is used with cardinal numerals 2 and 3, but for the cardinal numerals 4 and up the morpheme *-ery* is used. (The distribution of the morphemes *-oje/-ery* is standardly described in the academic and descriptive grammars of Czech as allomorphy: see Komárek, Kořenský ad. 1986, Karlík, Nekula ad. 1995). Second, the noun is the head of the numeral phrase. The numeral must agree with its head noun in case and number, as shown in (20-a), in contrast to the group numerals. (This is often difficult to see since the agreement may be morphologically syncretic — for example, the phonologically zero NOM.PL morpheme on the numeral in (20-a) is syncretic with ACC.PL.) When the numeral is part of a subject of a clause, the verb agrees with its head noun, as in (20-a). The allomorphs *-ery* and *-oje* are identical in terms of morphosyntactic behavior (see again Komárek, Kořenský ad. 1986, Karlík, Nekula ad. 1995 for more details), as shown in (20-e). We will mostly use examples with the morpheme *-oje* (numerals 2-3) since smaller cardinality of aggregates typically yields more comprehensible

and intuitively acceptable examples, yet our analysis and argumentation extends to complex numerals with *-ery* as well.

(20) a. Dv-oje-∅   kart-y   ležel-y na stole.

two-OJE-NOM.PL card-NOM.PL lie-3PL on table

'Two sets of cards were lying on the table.'

b. dv-oje   klíč-e

two-OJE key-PL

'two rings/sets of keys'

c. dv-oje   bot-y

two-OJE shoe-PL

'two pairs of shoes'

d. dv-oje   schod-y

two-OJE stair-PL

'two staircases'

e. Do zámku vedly   sedm-ery-∅   schod-y.

to castle led-3PL   seven-ERY-NOM.PL stair-PL

'Seven staircases led to the castle.'

On the other hand, this complex numeral is quite restricted in the sorts of nouns it may combine with, primarily nouns derived by *-i*, pluralia tantum, and entities that typically come together in multiples. Examples of each type are shown in (21)-(23), respectively. There is clearly a lexical semantic theme concerning which nouns may combine with *-oje* numerals: The nouns must designate entities which are sets of individuals that typically come together in groups or are "connected" to one another in some manner.

(21) NOUNS DERIVED BY -*í*: *nádobí* 'dishes', *dříví* 'firewood', *kamení* 'rocks'

(22) PLURALIA TANTUM: *dveře* 'doors', *housle* 'violins', *brýle* 'glasses'

(23) ENTITIES THAT TYPICALLY COME TOGETHER IN MULTIPLES: *sirky* 'matches', *hranolky* 'french fries', *schody* 'stairs'

In the case of nouns derived by -*í*, native speakers always prefer the use of the complex numeral over the simple cardinal numeral, despite a slight oddity for some combinations. A naturally-occurring example is given in (24-a) where the sequence of counting with cardinal numerals is changed from simple cardinals to an -*oje* numeral for a noun derived by -*í*. The use of a simple cardinal numeral would be unacceptable for native speakers. This is further illustrated in (24-b), a constructed example representing in a typical fast-food ordering scenario where the -*oje* numeral is the only felicitous and natural option. In this context, *dvě hranolky* (feminine), or *dva hranolky* (masculine), is odd, and infelicitous, because it is naturally interpreted as referring to two individual French fries, not to two portions of French fries.

(24) a. Dvě kuchyně, dva nákupy,   dvě lednice,   dv-oje nádobí

         two kitchens two purchases two refrigerators   two-oje dishes …

         'two kitchens, two purchases, two refrigerators, two sets of dishes'   Czech National

Corpus (2018)

      b. [context: fast food order]

         dvě kávy,   dv-oje/#dvě hranolky

         two coffees   two-oje/#two French.fries

         'two coffees, two portions of French fries / #two individual French fries'

Applying *-oje* to nouns other than those denoting entities which are sets of individuals that typically come in groups or are connected to one another in some manner typically result in infelicities, as given (25) (but see footnote 21, further below).

(25) a. ??dv-oje   stol-y

      two-OJE table-PL

      'two sets of tables'

  b. ??dv-oje   kabát-y

      two-OJE jacket-PL

      'two sets of jackets'

  c. ??šest-ery   aut-a

      six-ERY   car-PL

      'six sets of cars'

Combining *-oje* numerals with všechny 'all', as in (26), or with simple cardinals, as in (27), results in unacceptability:

(26) a. ??všechny dv-oje   housl-e

      DET   two-OJE violin-PL

      'all two sets of violins'

  b. ??všechny dv-oje   klíč-e

      DET   two-OJE key-PL

      'all two sets of keys'

  c. ??všechny sedm-ery   schod-y

      DET   seven-ERY stair-PL

      'all seven staircases'

(27) a. *tři   dv-oje   klíč-e

      three two-OJE key-PL

       'three two sets of keys'

   b. *tři   dv-oje   dveř-e

     three   two-OJE key-PL

     'three two sets of doors'

   c. *tři   sedm-ery   schod-y

     three   seven-ERY stair-PL

     'three seven staircases'

Thus, the sort of object delivered by combination with an aggregate numeral is not one that can be further counted, at least not in a standard manner.

2.2.3 Taxonomic numerals

The final type of complex numeral, derived with the suffix *-ojí*, yields taxonomic plurals, namely, a "different kinds" reading, which we call the TAXONOMIC NUMERAL. This complex numeral is morphologically parallel to *-oje*: The derivational suffix combines with a cardinal numeral root and then with a noun. The noun is the head of the numeral phrase and the numeral must agree with its head noun in case and gender. Like *-oje, -ojí* appears to be derivable for any cardinal numeral. Applying *-ojí* to the cardinal numeral root *dv-* 'two' yields *dv-ojí*, roughly translated as 'two kinds'. As in case of *-oje*, there is an allomorphy: for numerals 2-3, the suffix *-ojí* is used, for numerals 4 and up the morpheme *-ero* is the derivational suffix of the taxonomic numerals (and again we follow the standard allomorphy analysis as described in more detail in the academic/descriptive grammars of Czech: see Komárek, Kořenský ad. 1986,

Karlík, Nekula ad. 1995). Examples are given in (28). Note that *sýry* 'cheese' in example (28-b) does not combine with ordinary cardinals, and by that diagnostic is uncountable. Notice that the morpheme *-ero* in (28-d) applied to the numeral becomes the syntactic head and assigns genitive to its complement (analogous to the case of derived group numerals).

(28) a. dv-ojí   život

        two-OJI life

        'two kinds of life' (as in a Jekyll and Hyde scenario)

      b. dv-ojí   sýry

        two-OJI cheese

        'two kinds of cheese'

      c. dv-ojí   tvář

        two-OJI face

        'two faces' (as in a Janus face)

      d. čtv-ero   sýr-ů

        four-ERO cheese-GEN.PL

        'four kinds of cheese'

Unlike the previous two complex numerals, the nouns with which *-ojí* combines are not as restricted. Still, there are some noun types for which usage is more usual than for others, namely liquid and substance nouns, pluralia tantum and abstract nouns. Examples of each type are shown in (29)-(31), respectively.

(29) LIQUID AND SUBSTANCE NOUNS: *krev* 'blood', *víno* 'wine', *sýr* 'cheese'

(30) PLURALIA TANTUM: *housle* 'violin', *šaty* 'dress'

(31) ABSTRACT NOUNS: *život* 'life', *metr* 'measure', *občanství* 'citizenship'

Similarly to what was found with complex aggregate numerals derived by *-oje*, further combination with *všechny* 'all' or simple cardinals is disallowed, as shown in (32) and (33), respectively.

(32) a. *všechny dv-ojí   sýry

      DET   two-OJI cheese

    'all two kinds of cheese'

  b. *všechen dv-ojí   život

    DET   two-OJI life

    'All two kinds of life'

  c. *všechno čtv-ero   sýr-ů

    DET   four-OJI cheese-GEN.PL

    'All four kinds of cheese'

(33) a. *tři   dv-ojí   sýry

    three two-OJI cheese

    'three two kinds of cheese'

  b. *deset dv-ojí-ch   životů

    ten   two-OJI-GEN.PL lives

    'ten two kinds of life'

  c. *devět čtv-ero   sýr-ů

    nine   four-ERO   cheese-GEN.PL

    'nine four kinds of cheese'

Having set out the data for complex numerals in Czech, we now turn to examining the ability of nouns in Czech to shift their interpretation as a function of the context.

2.3 Restricted Nominal Flexibility

In comparison with a language such as English, the grammatical elements impacting countability and nominal interpretation in Czech are far more elaborate. At the same time, and, as we will argue, relatedly, the interpretative possibilities of nouns appear to be more impoverished than in English. At issue is the reduced number of interpretations a noun licenses.

As discussed at the beginning of this paper, packaging is permitted in Czech, if there is sufficient conventional use associated with the entity at issue. Yet, two other operations frequently claimed to be "universal", namely the Universal Grinder (Pelletier, 1979) and the Universal Sorter (Bunt, 1985), are even more restricted. "Grinding" appears to be broadly rejected by Czech native speakers. The examples in (34) and (35) provide standard grinding contexts in Czech, the first providing a "splattered animal" context and the second providing an "animal-as-food-product" context. Both are anomolous in Czech.[12]

(34) #Po celé   silnici byla kráva.

   on whole road   was cow

   Intended: 'There was cow all over the road.'

(35) #V salátu bylo prase.

   in salad   was pig

Intended: 'There was (a) pig in the salad.'

Turning to taxonomic interpretations, Czech's grammatical means to reference taxonomic sub-kinds are more elaborate, and this corresponds to greater intricacy in both distribution and interpretation. First, in Czech, interpreting plural nouns as referring to "different kinds" of the relevant noun is often not possible to the extent it is in English. While the sentence in (36-a) is unremarkable in English, its Czech counterpart distinctly odd.

(36) a. I used two oils in this salad.

    b. # V salátu jsem   použil   dva oleje.

        to   salad   AUX.1SG used   two oils

        Intended: `I used two oils in this salad.'

While taxonomic sub-kind readings may occur for typically non-countable nouns both as bare plurals and with simple cardinals, their distribution is restricted, both in terms of their syntactic context and in terms of their interpretation.

Acceptable examples of taxonomic sub-kind uses with a bare plural noun phrases and with a simple cardinal phrase are given in (37) and (38)[13], respectively.

(37) a. Prodává-me vína   lahvová   i   stáčená

     sell-1PL   wine.PL in-bottles and wine-on-tap

     'We sell wines in bottles and on tap.'

   b. Prodává-me oleje pro osobní, nákladní a užitková vozidla

     sell-1PL   oil.PL for personal cargo and utility   car.PL

---

[13] We thank an anonymous reviewer for providing example (38-a) and several other examples of simple cardinal phrases with taxonomic readings.

'We sell oils for personal cargo and utility cars.'

(38) a. V Brně mají   na čepu další tři   piva

in Brno have.3PL on tap   next three beer.PL

'They have three beers on tap in Brno.'

Countable nouns appear also to receive a taxonomic sub-kind interpretations, although usually manifest post-nominal modification which ensures a taxonomic interpretation, as in (39).

(39) Chováme   psy   (různých   ras)

breed.we.pl dog.PL (different.GEN.PL types.GEN.PL)

'We breed (different types of) dogs.'

A common trait of all the examples of taxonomic uses of non-countable nouns was they were found in generic, existential or otherwise non-episodic contexts. This contrasts with clearly episodic uses of nouns in their taxonomic interpretation in English, as in (36-a). This further contrasts with the use of the taxonomic numeral, which is licensed both in episodic and non-episodic contexts. This is shown in (40), where simple cardinal phrases used in generic and episodic descriptions produce a contrast in acceptability, shown in (40-a) and (40-b), while the use of the taxonomic numeral is acceptable in both contexts, as shown in (40-c) and (40-d).

(40) a. Naše benzínka   prodává   tři   paliva.

our   gas-station sells.IMPERF-HAB three fuel.PL

'Our gas station sells three fuels.'

b. ??Naše benzínka   včera   prodala   tři   paliva

our   gas-station yesterday   sold.PERF   three fuel.PL

'Yesterday our gas station sold three fuels.'

c. Naše benzínka   prodává   trojí   palivo.

our   gas-station sells.IMPERF-HAB three-kind fuel.SG

'Our gas station sells three fuels.'

d. Naše benzínka   včera   prodala   trojí   palivo

our   gas-station yesterday sold.PERF three-kind fuel.SG

'Yesterday our gas station sold three fuels.'

A second difference between the use of simple cardinals or bare plurals to indicate taxonomic sub-kinds and the use of taxonomic numerals is in the type of sub-kind that is identified. With the use of simple cardinals or bare plurals, specific, named, and well-established "sub-specimens" appear to be the most salient, and possibly the only, type of taxonomic entities referenced. For instance, in the example (40-a), the use of the simple cardinal with a taxonomic interpretation would be interpreted as contrasting, for example, specific types of gasoline, such as with different octane levels (e.g., 85, 95, or 98). The use of the taxonomic numeral, however, typically brings about upper-level taxonomic contrasts, such as between diesel, gasoline and natural gas, although the more specific types are also a possible interpretation.

We now take stock of the implications of Czech's nominal system for theories of countability.

3 Interim Discussion

All theories of countability have at their core a claim of what countability is about. By measuring the claims of different theories against the Czech data, we can work towards narrowing down the space of theoretical possibilities. We return to the three issues introduced in section 1: (i) whether nouns are ontologically uniform or multi-sorted, (ii) the nature of aggregates, and (iii) the trade-off between a language's potential for nominal flexibility and its capacity for expression through morphology.

To gain insight into the first issue, we consider what it means within different theories to be non-countable and then compare those conjectures with Czech data from the non-countable nouns derived by -*í*. In brief, we will show that the Czech data poses challenges for theories in which there is one overarching explanation of how (non-)countability arises. In particular, these explanations falter when extended to account for the second form of non-countable nouns in Czech derived by -*í*.

First, we consider the theory of Borer (2005) for which all nouns begin as non-countable, and then through combination with a functional head, e.g. Div^0, may become countable. Two facets of the Czech system in conjunction undermine this claim: the lack of grinding and the nouns in Czech derived by -*í*. Instead, the opposite trajectory is observed. Non-countable aggregate nouns such as *listí* 'foliage' begin as nominal roots for which there is only evidence that they are fully countable nouns, that is, *list* 'leaf' only has a countable interpretation and no effects from 'grinding' may be observed. Through combination with

morphology, these nouns become then non-countable with respect to combination with simple cardinal numerals, which is exactly opposite of the prediction in Borer (2005). This is not to say that the system in Borer (2005) could not be altered to account for the immediate data under consideration, in the way that others have expanded her system to account for different countability phenomena (De Belder, 2013; Mathieu, 2012). Rather, the phenomenon of -*i* nouns in Czech stands in contradistinction to the conceptual predictions of Borer (2005).

Similar unmet expectations are found when considering how the data aligns with the theory of Chierchia (2010), for which non-countable nouns differ from countable ones in that non-countable nouns designate entities for which the atomic parts are vague. Applying this line of reasoning to -*i* nouns would also be counterintuitive, since the parts, designated by the derivational source, e.g. *list*- 'leaf/leaves', are non-vaguely atomic and fully countable. One would have to argue either that the semantic contribution of -*i* brings about vagueness into the derived forms, or that there is a secondary cause for non-countability.[14] In either case, this subclass of non-countable nouns would go against the grain of the explanation of non-countability in Chierchia (2010).

More generally, the subclass of non-countable nouns derived by -*i* provides a strong argument that non-countability may arise from more than one source. Accordingly, theories for which there is a single path of explanation that gives rise to (non-)countable interpretations are simply not expressive enough. In other words, the domain of nominal meaning is not of a single ontological sort, but many-sorted, containing at least substances and what we have termed aggregate nouns, those derived by -*i*, the nature of which we turn to now.

---

[14] An anonymous reviewer suggests that noun meanings derived with -*i* may be analyzed akin to *furniture*-nouns in Chierchia (2010), that is, as singular properties; however, it is difficult to see a path forward for using a denotation which is simply all the instantiations of, say, leaves, as it does not speak to the particular conditions of relatedness, discussed in 2.1, nor explain the restricted distribution. Chierchia (2010, fn. 11) carefully distinguishes *furniture*-nouns from "collectives", noting that single entities may qualify as furniture, as in "That chair is furniture", but this is at odds with -*i* nouns which may not felicitously reference atomic parts, but only coherent collections.

Granting the existence of aggregates nouns as separate from substance nouns, many questions remain about their nature. Again, measuring the theoretical possibilities by the Czech data proves illuminating. The grammatical number system of Czech forces the recognition of (i) aggregate nouns as a distinct class, both through devoted derivational morphology (-*i*) and through a devoted complex numeral, and (ii) a stronger notion of aggregate than is often employed in theories of countability. The first point dovetails with increasing acceptance in the literature of at least three types of nouns, substance (non-countable), aggregate (non-countable) and individuals (countable). For instance, Bale and Barner (2009) and Deal (2017) both explicitly set out to model those three noun types. Despite increasing recognition of two sorts of non-countable nouns, the data from Czech implicates a more specific semantics than is usually given in the analysis of nouns such as *furniture* or *footwear* in English. First, in those models, aggregate nouns are essentially treated as akin to plural nouns but for the inclusion of atoms, technically speaking as atomic join semi-lattices. Second, even though theories such as Barner and Snedeker (2005) or Bale and Barner (2009) do recognize aggregate nouns as a distinct class, they analyze them through lexical fiat, i.e. the non-countable status of such nouns is idiosyncratic. For these accounts, extending such a treatment to -*i* nouns leaves their regularity and lexical semantic cohesion unexplained. Why should these particular nouns, and not others, receive such distinguished morphological treatment?

As shown in 2.2.2, Aggregate Numerals, complex numerals derived with -*oje*, apply to nouns derived by -*i*, pluralia tantum nouns, and also certain morphologically regular plural nouns which designate entities that typically come together in multiples. Assuming that there is some form of semantic selection, then there should be a common denominator among these noun types. Analyzing aggregates as akin to plurals however does not provide enough discriminatory power to bring this to the fore. Since morphologically regular plurals and aggregates are selected for by Aggregate Numerals, the common denominator would be

individuated lattices, yet, this would overgenerate as any plural noun should be able to combine with Aggregate Numerals, contrary to fact. Instead, our semantics must distinguish entity types for whom its members canonically co-occur, as is the case for these nouns in Czech.

Finally, we turn to considering the nature of nominal flexibility, which we argue requires a more nuanced view. The countability literature has often put forth that the possible interpretation of nouns is unrestrained, in that every countable noun can find a non-countable use, and vice versa, as is most clearly articulated in Pelletier and Schubert (2004) or Borer (2005). What then should be made of Czech's limited flexibility? Under such a view, an interpretation as a taxonomic plural should be licensed when conceptually possible, yet, the taxonomic plural is systematically absent as an interpretation of ordinary plural nouns in episodic contexts, even though parallel examples in English demonstrate such uses are conceptually possible. The cause for this discrepancy is obvious: Czech has morphology devoted to expressing taxonomic interpretations which can be employed freely. Accordingly, regular plural noun phrases only receive taxonomic interpretations in grammatical contexts that already foster kind interpretations. Similarly, grinding with bare singular nouns is absent in Czech. Again, there is a plausible reason for this: Unlike in English, Czech does not have articles, so bare singular nouns already serve a role in Czech (see also Cheng et al. 2008; Rothstein 2017).

The broader implication is that the potential for nominal flexibility in a given language is influenced and constrained by distinctions already expressed within the morphosyntax of nominals in that language. That is, the lexical semantics of nouns manifests structure that is coordinated with what is expressed in the grammatical structure.

4 A Formal Treatment of Czech's Grammatical Number System

We now turn to examining how to integrate Czech's nominal system within a formal analysis of countability. We develop an analysis by extending a version of the theory of Krifka (1995). The system in Krifka (1995) provides a useful starting point: It develops explicit representations of several ingredients to the semantics of countability, such as taxonomic reference and a notion of "natural units" (in contrast to several other theories, discussed in sections 1 and 3, committed to parsimonious views of the semantics underpinning countability). As such, it provides a suitably expressive framework to model many of the distinctions present in Czech. Even so, after presenting the basic components of the theory of Krifka (1995), we still must further develop a variety of extensions to the system in section 5 in order to fully account for the data.

4.1 Nominal Semantics in Krifka (1995)

Krifka (1995) integrates two lines of research on nominal semantics. On the one hand, building on Link (1983) and Krifka (1989), among others, the domain of objects is structured according to the basic principles of mereology, and as such models nominal meaning as complete semilattice structures lacking the null element. On the other hand, it builds on the work on generics, in which nominal meaning includes both reference to objects and reference to kinds (Carlson, 1980; Krifka et al., 1995; Zamparelli, 1999; Müller-Reichau, 2006). Krifka (1995) proposes a revision to kind-based nominal semantics: instead of kinds, the broader category of *concepts* is used, of which kinds are a special subset. We will adopt this distinction for a general framework of nominal meaning, yet as it will not be critical for the purposes of the present discussion, we will often use the terms *kind* and *concept* interchangeably. The two levels of

nominal meaning are related by a realization relation R between *concepts* and the instances of the concept at the level of objects. That is, the referential use of *dog* is tied to the realizations of the concept *dog*, that is, the instances of a concept. A second relation discussed in Krifka (1995) is a taxonomic relation T holding between *kinds/concepts* and their subkinds or subconcepts, where the subkind reading of *dogs* would correspond to "different types of dogs", such as *beagle* or *chihuahua*. The basic meaning of a noun in this system is given in (41), where variables ranging over kind-level entities are subscripted and *k*. In prose, (41) describes "the property of being a specimen or subspecies, or an individual sum of specimens or subspecies" (Krifka 1995, p. 399).[15]

(41) $\lambda y_k \lambda i \lambda x[R_i(x, y_k) \lor T(x, y_k)]$

Given (41), an entity may satisfy the predicate *dog* in two ways. First, if it is an individual dog, or a plural individual composed of dogs, which are objects related to the kind *dog* by the realization relation R. Second, the predicate may be satisfied on the taxonomic reading, where the entity must be an individual subkind of dog (*chihuahua*) or sum of subkinds, which are related to the kind *dog* by the taxonomic relation T. The remainder of this section focuses on object-level interpretations, leaving the discussion of taxonomic-level interpretations for section 4.2.

Krifka's account further includes measure functions to model expressions such as *two liters* or *five ounces*, and more pertinently for us, a measure function relevant for countable nouns that counts "natural units" — relative to the kind that the noun names — named the "object unit" operator (OU). The OU operator provides a measure, *n*, of the number of elements which qualify as instances of the kind. Thus, there are two criteria of applicability at work in

---

[15]   In (41), i is a variable of type s ranging over possible worlds—in what follows, we will simplify by extensionalizing the representations.

the semantics of nominals: one which is "qualitative" and represented by the nominal predicate, for example, *gold* or *dog*, and a second which is "quantitative", represented by a "natural unit" measure function.

Countable nouns under this account are two-place relations between numbers and entities ($\langle n, \langle e, t \rangle \rangle$), while non-countable nouns are one-place relations ($\langle e, t \rangle$), as shown the contrasting lexical entries for *dog* and *gold* in (42-a) and (42-b), respectively.

(42) a. $[\![\text{dog}]\!] := \lambda n \lambda x [R(x, \text{DOG}) \wedge \text{OU}(\text{DOG}, x) = n]$

    b. $[\![\text{gold}]\!] := \lambda x [R(x, \text{GOLD})]$

Krifka, partly inspired by classifier languages such as Mandarin, discusses where the natural unit measure function is located in the extended nominal phrase in English. One option is that the OU operator is endemic to the noun, as is the case in (42-a), and cardinal numerals are simply an argument of type *n* as in (43-a). Alternately, the OU operator may be supplied by (cardinal) numerals, that is, they include a built-in classifier as in (43-b).

(43) a. $[\![[three]_{Num}]\!] = 3$

    b. $[\![[three]_{Num}]\!] = \lambda y \lambda x [R(x, y) \wedge \text{OU}(y, x) = 3]$

While Krifka (1995) provides some arguments for including a built-in classifier for numerals in English, these arguments do not transfer over to Czech. One main argument given by Krifka (1995) is that assuming the OU operator as part of nominal meaning does not dispose of the need for a syntactic distinction between countable and non-countable nouns, as there is a ban in English on using singular count nouns as noun phrases. Yet, Czech differs as bare singulars are permitted, as was shown in 2. Second, the very nature of the complex numerals in Czech discussed here indicate that building the OU operator into the numeral would be conceptually difficult for number terms generally: If the semantics of the numeral stem *troj-* `three' included

a built-in classifier as specified in (46-b), then the morphemes *-ice*, *-oje*, and *-oji* would then necessarily have to eliminate or overwrite the built-in classifier material in order to count groups, collections, or kinds.[16]

In sum, we will take kinds as the building block of nominal meaning, as argued for in Carlson (1980) or Krifka (1995) and assumed in much other work.[17] In addition, we assume that for countable nouns the shift from kind-level to object-level reference is accompanied by the lexically specified OU operator. That is, countable common nouns are of type $\langle n, \langle e,t \rangle \rangle$. Cardinal numerals, in their primary function, simply supply a number value.[18]

We make standard assumptions for the nominal syntax. We assume the determiner phrase includes a NUMP layer. The NUM head can be occupied by cardinals 2-4 and 5 and above, which contribute their cardinal value as the number argument and require plural agreement on the noun. Compositionally, an NP of type $\langle n, \langle e,t \rangle \rangle$ has its number argument saturated by a cardinal number of type n, where the plural form is required for agreement. Alternately, NUM can be occupied by non-quantified singular or plural nouns, whereby singular or plural

---

[16] A third argument given by Krifka (1995) is that since many nouns can be used as countable or non-countable nouns in English, as in Czech, building the OU operator into the noun would, erroneously, disallow this flexibility. We are skeptical that the observed flexibility in English or Czech necessitates abandoning specifying countability as part of nominal meaning, since the facts could be accounted for, arguably more successfully, by analyzing flexibility as resulting from either (i) true ambiguity, as in the case of *string* which designates pieces of string or the material, or (ii) pragmatic accommodation. See Grimm (2018) for extended argumentation.

[17] The analysis we articulate below is in principle compatible with taking properties as basic, as in Krifka (2003), with a slightly different architecture guiding the type shifts, although we do not elaborate here.

[18] We note that there is tension in this analysis between the denotation of cardinal numerals as of type $\langle n \rangle$, that is, simply supplying a number value, and the syntactic modifier status of Czech cardinals 1-4. Yet, we assume, following Geurts (2006) and Rothstein (2017), among others, that a type-shift is generally available between the argumental $\langle n \rangle$ and predicative/adjectival type $\langle e,t \rangle$. This is particularly plausible in the case of Slavic numerals, as they have undergone significant grammaticalization whereby mismatches between syntactic and semantic category are not unusual. (See, for instance, Miechowicz-Mathiasen (2012) who discusses the mixed adjectival and nominal properties of lower numerals in Polish). Still, as a reviewer pointed out, it is important to assess how the modifier status of cardinals 1-4 affects the analysis we propose. We provide another, alternate analysis which represents the modifier status of lower cardinals, yet maintains the rest of the semantics that we propose. As this is not central to main thrust of our study, this is discussed in footnotes 23-25 and 27. (We thank Hana Filip and Peter Sutton for discussion and suggesting this line of analysis.)

morphology is semantically valued as the number value 1 or ≥ 2, respectively.[19] In this case, an NP occupies NUM whereby its number argument is saturated by the noun's morphological number value. This view accords with standard current formal syntactic analysis of Slavic numerals (see Miechowicz-Mathiasen 2012, Marušič and Nevins 2009, *inter alia*).[20]

The general theory of nominal reference outlined so far is able to account for the basics of nominal reference in Czech, as well. The representations in (44) present the different types of nouns discussed in increasing complexity: kind-referring bare singular (44-a), non-countable noun (44-b), singular countable noun with number morphology evaluated (44-c), plural countable noun with number morphology evaluated (44-d), a plural genitive NP form, which provides an unsaturated property meaning (44-e), and a (genitive plural) noun combining with a simple cardinal numeral *pet* 'five' (44-f)[21].

(44)   a. $[\![_N pes]\!] = \text{DOG}$

     b. $[\![_{NP} prach]\!] = \lambda x[\text{R}(x,\text{DUST})]$

---

[19] We assume an exclusive view of plurality although nothing in particular hinges on this choice. See Grimm (2013) for discussion.

[20]   At first sight, that numerals ≥ 5 in Slavic languages show singular agreement on the verb when their NP is used in subject position might appear to be an empirical obstacle. Yet, this singular agreement is convincingly argued by Marušič and Nevins (2009) and Marušič et al. (2015), among others, to be a failure of the verb to agree at all, since in Slavic languages, singular (with neuter gender – as in the case of numerals ≥ 5) is the default number for a verb as witnessed by non-argument weather verbs producing singular (neuter) non-agreement as well.

[21]   In an alternate analysis which represents the modifier status of lower cardinals, numerals 1-4 may be analyzed as of type $\langle\langle n, a\rangle, a\rangle$ where a is a variable over types. The relevant type for simple cardinal constructions is $a = \langle e,t\rangle$. The lexical entry for *dva* 'two' is given in (i). This also requires a more specific analysis of the plural morpheme, given in (ii), which derives plural nouns as in (iii). Thus, the derivation of *dva psi* 'two dogs' would proceed as in (iv), delivering a parallel representation to (44-f).

(i) $[\![_{Num'} dva]\!] = \lambda P{:}\langle n,\langle e,t\rangle\rangle[P(2)]{:}\langle e,t\rangle$

(ii) $[\![\text{-PL}]\!] = \lambda k.\lambda n.\lambda x.[\text{R}(x, k) \wedge \text{OU}(k, x) = n \wedge n \geq 2]$

(iii) $[\![_N psi]\!] = [\![\text{-PL}]\!]([\![_N pes]\!]) = \lambda n\lambda x[\text{R}(x,\text{DOG}) \wedge \text{OU}(\text{DOG},x) = n \wedge n \geq 2]$

(iv) $[\![_{NumP} [_{Num'} dva][_N psi]]\!] = [\![_{Num'} dva]\!]([\![_{NP} psi]\!])$

          $= \lambda P[P(2)](\lambda n\lambda x[\text{R}(x,\text{DOG}) \wedge \text{OU}(\text{DOG},x) = n])$

          $= \lambda n\lambda x[\text{R}(x,\text{DOG}) \wedge \text{OU}(\text{DOG},x) = n \wedge n \geq 2](2)$

          $= \lambda x[\text{R}(x,\text{DOG}) \wedge \text{OU}(\text{DOG},x) = 2 \wedge 2 \geq 2]$

c. $[[_{NumP}\,pes]] = \lambda x[\mathrm{R}(x,\mathrm{DOG}) \wedge \mathrm{OU}(\mathrm{DOG},x) = 1]$

d. $[[_{NumP}\,psi]] = \lambda x[\mathrm{R}(x,\mathrm{DOG}) \wedge \mathrm{OU}(\mathrm{DOG},x) \geq 2]$

e. $[[_{NP}\,ps\r{u}]] = \lambda n \lambda x[\mathrm{R}(x,\mathrm{DOG}) \wedge \mathrm{OU}(\mathrm{DOG},x) = n]$

f. $[[_{NumP}\,[_{Num'}\,pet][_{NP}\,ps\r{u}]]] = [[_{NP}\,ps\r{u}]]([[_{Num'}\,pet]]) = \lambda x[\mathrm{R}(x,\mathrm{DOG}) \wedge \mathrm{OU}(\mathrm{DOG},x) = 5]$

Additionally, we will assume, following Krifka (1995), the availability of a shift from a kind to a number-neutral predicate of the kind whereby the OU operator is existentially bound, given in (45). This shift derives a predicate interpretation which can be passed along to the complex numerals for aggregates.

(45) $[[_{NP}\,psi]] = \lambda x \exists n[\mathrm{R}(x,\mathrm{DOG}) \wedge \mathrm{OU}(\mathrm{DOG},x) = n]$

As we now show, this account, with minor modifications, also extends straightforwardly to the taxonomic interpretations of nominals in Czech as well.

4.2 Taxonomic Interpretations and Numerals

To incorporate the taxonomic interpretations found with bare plurals and simple cardinals, discussed in section 2.3, we generalize the nominal semantics to include taxonomic units form the system of Krifka (1995).   Accordingly, the base template for nominal meaning in Czech is (46), disposing of the realization operator R and the taxonomic operator T, which can be generalized into a single operator, RT, given in (47).

(46) $\lambda y_k \lambda x[\mathrm{R}_i(x,y_k) \vee \mathrm{T}\,(x,y_k)]$

(47) $\mathrm{RT}(x,y) \leftrightarrow \mathrm{R}\,(x,y) \vee \mathrm{T}\,(x,y)$

To this we add the analogue for the OU operator for kinds, the KU ('kind unit') operator which, for each possible world, when applied to a kind, delivers the number of subspecies of that kind. Both measure functions, OU and KU, can be joined into one operator, OKU ('object or kind unit'), defined in (48) Krifka 1995 (1995, 406).

(48) $\text{OKU}_i(x)(y) = n \leftrightarrow \text{OU}_i(x)(y) \vee \text{KU}_i(x)(y) = n$

The different types of nominal reference are given in their generalized form in (49).

(49) a. $[[_N pes]] = \text{DOG}$

  b. $[[_{NP} prach]] = \lambda x[\text{RT}(x, \text{DUST})]$

  c. $[[_{NumP} pes]] = \lambda x[\text{RT}(x, \text{DOG}) \wedge \text{OKU}(\text{DOG}, x) = 1]$

  d. $[[_{NumP} psi]] = \lambda x[\text{RT}(x, \text{DOG}) \wedge \text{OKU}(\text{DOG}, x) \geq 2]$

  e. $[[_{NP} ps\mathring{u}]] = \lambda n \lambda x[\text{RT}(x, \text{DOG}) \wedge \text{OKU}(\text{DOG}, x) = n]$

  f. $[[_{NumP} [_{Num'} 5][_{NP} ps\mathring{u}]]] = \lambda x[\text{RT}(x, \text{DOG}) \wedge \text{OKU}(\text{DOG}, x) = 5]$

These representation secure the desired taxonomic interpretations for countable nouns. For instance, *psi* 'dogs' may receive the interpretation, stated in simplified form, $\lambda x[\text{T}(x, \text{DOG}) \wedge \text{KU}(\text{DOG}, x) \geq 2]$, denoting the set of sums of sub-kinds of *dog* greater than equal to 2. It is prudent to retain the representation in (49-b) for non-countable nouns, where no unit operator is specified, as non-countable nouns may be used to refer to assemblages of stuff which do not cohere to any units. At the same time, it is necessary to allow non-countable nouns to shift to the representations in (49-d)-(49-e). We assume a rule which converts non-countable noun templates to countable ones, that is, from $\lambda y \lambda x[\text{RT}(x, y)]$ to $\lambda y \lambda x[\text{RT}(x, y) \wedge \text{OKU}(x, y) = n]$, when contextually licensed, for example, for taxonomic interpretations in generic or otherwise kind-

selecting contexts. This then licenses a reading for, e.g., *tři paliva* 'three fuels', namely $\lambda x[T(x,\text{FUEL}) \wedge \text{KU}(x,\text{FUEL})=3]$. Packaging interpretations are analogously derived.

We now turn to the second method to derive taxonomic interpretations in Czech. We argue that the taxonomic operator T is also found in the meaning of the complex taxonomic numerals derived by *-ojí*. The lexical entry for the derivational suffix *-ojí* is given in (51), designating "different kinds". In prose, *-ojí* takes a numeral, a kind, and a subkind or individual sum of subkinds, checks that all the parts of the (sum of) subkind are a subkind of the kind specified by the nominal head and that their cardinality is equal to that specified by the numeral.

(50) $[\![\text{-ojí}]\!] = \lambda n \lambda k \lambda x [\forall z(z < x \wedge T(k,z)) \wedge |x| = n]]$

This ensures, for example in (51), that when combining with a numeral, here *dv-* 'two', and a kind denoting root noun, here [N *víno*] 'wine', which happens to be non-countable at the object-level, the resulting cardinal phrase designates a sum of subkinds of wine whose cardinality is 2.[22]

(51) $[\![\text{dvojí víno}]\!] = \lambda x [\forall z(z < x \wedge T(\text{Wine},z)) \wedge |x| = 2]$

A further welcome result is an explanation of the oddity when combining taxonomic numerals with uniquely denoting noun phrases, as in (52-a), and with proper names, as shown in (52-b): These nouns do not supply the kind terms which *-ojí* requires.

---

[22] Here too the alternate analysis of lower cardinals as of generalized type $\langle\langle n, a\rangle, a\rangle$ may be pursued to represent their adjectival status as modifiers. In this instance, the type will be $\langle\langle n, \langle k, \langle e,t\rangle\rangle\rangle, \langle k, \langle e,t\rangle\rangle\rangle$, as shown for the lexical entry of *dva* in (i). This combines with the entry for *-ojí,* shown in (ii), ultimately leading to the same result as in (51).

       (i) $[\![_{Num'}\, dva]\!] = \lambda P\!:\!\langle n, \langle k, \langle e,t\rangle\rangle\rangle\ [P(2)]\!:\!\langle k, \langle e,t\rangle\rangle$

       (ii) $[\![dvojí]\!] = [\![_{Num'}\, dva]\!]([\![\text{-ojí}]\!]) = \lambda k \lambda x [\forall z(z < x \wedge T(k,z)) \wedge |x| = 2]]$

(52) a. #dv-ojí   noha   tohoto stolu

      two-OJI leg   this   table

      'two kinds of this table's leg'

   b. #dv-ojí   Petr Novák

      two-OJI Petr Novák

      'two kinds of Petr Novák'

Section 2.3 showed distributional differences between the uses of taxonomic numeral phrases, which are contextually unrestricted, and taxonomic interpretations of bare plurals or simple cardinal phrases, which are restricted to non-episodic contexts. This receives an explanation on the account provided here, since bare plurals or simple cardinal phrases require appropriate context, e.g. a generic context, to shift from non-countable to countable noun meanings, while the taxonomic numerals do not depend on contextual support.

Section 2.3 also demonstrated a difference in interpretation depending on whether a simple cardinal/bare plural was used or a taxonomic numeral was used. Taxonomic uses of simple cardinal and bare plural noun phrases refer to specific, well-establish and named sub-kinds, such as specific types of varieties of wine (Chardonnay versus Riesling). This corresponds to exactly the sort of entities that would be returned by the KU operator, that is, true 'kind units'. On the other hand, taxonomic entities referred to with taxonomic numeral phrases tended to be upper-level taxonomic contrasts, such as red versus white wine. In the analysis of taxonomic numeral phrases given here,   any taxonomic entities standing in sub-kind/concept relation to the general kind/concept are permitted, thus, it is expected that their interpretation may cover upper-level taxonomic contrasts, and further, will have a tendency to do so if specific sub-kinds are already referenced by simple cardinal phrases.

So far, much of Czech's nominal system has been able to be analyzed using the elements of Krifka's (1995) system. We now turn to aspects of Czech's nominal system which require an extension to this basic framework.

5 Extending the Framework: Counting Groups and Aggregates

The last section demonstrated that, with some minor modifications, Krifka (1995) can be successfully used to treat the basic cases of nominal semantics as well as taxonomic numerals. Yet, in order to cope with the data arising from group numerals and from aggregates in Czech, more substantial extensions are required. We first enrich the system with groups in the sense of Landman (1989) to account for group numerals, and then with mereotopology in the sense of Grimm (2012) to account for aggregates.

5.1 Group Numerals

The system of Krifka (1995) provides a method to analyze certain types of group nouns, such as *herd*, making uses of measure functions. As shown in (53), the measure function counts groups in the same way it counts atomic objects.

(53)   [[three herds of cows]] = $\lambda x[R(x, \text{Cow}) \wedge \text{Herd}(\text{Cow}, x) = 3]$

While this analysis is adequate for the data discussed in Krifka (1995), the derived collective numerals in Czech require a different analysis. In particular, counting with the aid of group numerals involves counting both at the level of the number of groups and the number of individuals inside the groups. As was shown in (17), repeated here as (54), *-ice* assigns a

cardinal value to the members of the groups (*troj-ice námořníků* 'a group of three sailors'), and the groups themselves can then also be counted.

(54)   dvě troj-ice   námořník-ů

two three-ice sailors-GEN.PL

'two groups of three sailors'

Accordingly, an analysis of these complex numerals will both need to contribute a generalized way to deliver a semantics which groups objects and counts the number of objects in the group. A straightforward solution is to augment the schema of <u>Krifka</u> (<u>1995</u>) with the group shifting operator ↑ from <u>Landman</u> (<u>1989</u>), which maps sums of individuals to a group.

Unlike the other complex numerals, *-ice* is the head of the noun phrase and takes a genitive argument describing the group members, here, as sailors. We analyze *-ice*, shown in <u>(55-a)</u>, as first combining with a number, which ultimately feeds the OU operator, and then the property provided by the genitive argument. As previously stated, we assume the plural genitive NP provides an unsaturated property meaning. A minimal example is given in <u>(55-c)</u>, where the application of *pět-* 'five' and then *mužů* to *-ice* results in the set of sum individuals which are 'groups of five men'.[23]

(55)   a. $[\![\text{-ice}]\!] = \lambda n \lambda P \lambda x[\uparrow x \wedge P(x)(n)]$

b. $[\![_{NP}\,\text{mužů}]\!] = \lambda n \lambda x[R(x,\text{MAN}) \wedge \text{OU}(\text{MAN},x) = n]$

c. $[\![_{NumP}\,\text{pětice mužů}]\!] = \lambda x[\uparrow x \wedge R(x,\text{MAN}) \wedge \text{OU}(\text{MAN},x) = 5]$

---

[23] Pursuing the alternate analysis of lower cardinals as of generalized type $\langle\langle n, a\rangle, a\rangle$ in this case results in the type $\langle\langle n, \langle\langle e,t\rangle,\langle e,t\rangle\rangle\rangle, \langle\langle e,t\rangle,\langle e,t\rangle\rangle\rangle$, as shown for the lexical entry of *dva* in (i). This combines with the entry for *-ice,* shown in (ii), delivering a parallel result.

(i) $[\![_{Num'}\,dva]\!] = \lambda P{:}\langle n, \langle\langle e,t\rangle,\langle e,t\rangle\rangle\rangle \; [P(2)]{:}\langle\langle e,t\rangle,\langle e,t\rangle\rangle$

(ii) $[\![dvojice]\!] = [\![_{Num'}\,dva]\!]([\![\text{-ice}]\!]) = \lambda n \lambda P \lambda x[\uparrow x \wedge P(x)(2)]$

This meaning is again fully countable, designating a singular individual or atom, but in the domain of groups. In particular, the meaning given in (55-c) can be itself be pluralized, as well as counted by means of standard cardinal numbers, where, for instance, the meaning of *dvě pětice mužů*, as given in (54), consists of the set of two groups of five men. We assume that the simple cardinals which precede the complex cardinals, simply occur as modifiers, that is, the meaning of *dvě* 'two' receives a standard analysis as a modifier as $\lambda P \lambda x[P(x) \wedge |x| = 2]$.[24]

5.2 Deriving and Counting Aggregates

The derived aggregate nouns and complex numerals which count aggregates require further extensions to Krifka (1995), namely the adoption of a theory of aggregate nouns. First, the derived aggregate nouns and complex numerals for aggregates discussed in section 2, unlike the other phenomena discussed here, showed severe restrictions. The derivational suffix *-i* only applied to restricted set of nouns, while the complex numerals for aggregates only applied to *connected entities*, which in turn could be morphologically encoded in the noun, as with *-i* derived nouns or pluralia tantum, or due to a noun's meaning which proffers such a connectedness relation, as in the case of *schody* 'stairs', which typically come in multiples.

We build on the account of Grimm (2012) to provide a semantics for aggregate nouns in Czech. Grimm (2012) provides a topologicial extension of mereology in part to treat non-countable aggregate terms in English such as (*sand* or *foliage*) as well as to treat morphologically recognized collective/singulative classes in languages such as Welsh. The principle tool is to

---

[24]   We also note that this is a point where the analysis of simple cardinals possessing a built-in OU classifier runs aground, since the objects being counted here are not of the type amenable to being counted by a "natural unit" function, that is, there is no kind (or concept) of *group of three sailor* which is natural to measure via the OU function.

add connectedness relations to the standard mereological framework. While topological extension of mereology in Grimm (2012) relates to such core topics in countability, such as how to define atoms and substances, we only discuss the portion of the account relevant for the Czech aggregate data.

Central to extending mereology with topological relations is including a relationship of CONNECTEDNESS. The intuitive definition of being connected is that two entities are connected if they share a common boundary. Including a definition of connectedness then interacts with the different definitions and axioms of standard mereology, as discussed in detail in Casati and Varzi (1999).

We consider first the basic relation C, CONNECTED, which is required to be reflexive and symmetrical, given as axioms in (56) and (57). There are some further intuitive interactions with the mereological relations *part*, $\leq$, and *overlap*, O, that Casati and Varzi (1999) note. First, the axiom in (58) ensures that parthood implies connectedness. From (58) the relation in (59) follows, whereby overlap implies connectedness.

(56) $C(x,x)$ (Reflexivity)

(57) $C(x,y) \rightarrow C(y,x)$ (Symmetry)

(58) $x \leq y \rightarrow \forall z(C(x,z) \rightarrow C(z,y))$

(59) $O(x,y) \rightarrow C(x,y)$

Connectedness can come in a variety of strengths. Two more specific varieties of connectedness relevant for aggregate noun semantics are EXTERNAL CONNECTEDNESS, when two entities touch on their boundaries, notated as $C_{ext}(x,y)$ and PROXIMATE CONNECTEDNESS, when two entities are sufficiently near relative to some distance d, $C_{prox}(x,y)$. We now turn to apply this extended system to aggregates in Czech.

5.2.1 Derived mass nouns

We now give a semantics to the morpheme *-i*, whereby it applies to a root noun and returns a connected set of individuals, which we will term a CLUSTER INDIVIDUAL, described by a root noun. We first define a transitive connection relation in (60). In prose, x and y are transitively connected relative to a property P, a connection relation C, and a set of entities Z, when all members of Z satisfy P and x and y are connected through the sequence of z_is in Z.

(60) TransitiveC(x,y,P,C,Z) $=_{def} \forall z \in Z[P(z) \wedge (x = z_1 \wedge y = z_n) \wedge C_{z1z2} \wedge C_{z2z3} \ldots \wedge C_{zn-1zn}]$ where $Z = \{z_1, z_2, \ldots, z_n\}$ (Transitively Connected)

Cluster individuals, relative to a property and connection type, are defined in (61) as a set of entities of the same type connected to one another by virtue of each being transitively connected to the others (through a set Y relative to the same property and connection relation).

(61) Cluster(x,P,C) $= \exists Z[x = \bigoplus Z \wedge \forall z,z' \in Z \exists Y [TransitiveC(z,z',P,C,Y)]]$

With the additional topological machinery, we are now able to give a semantics to *-i*. We analyze *-i* nouns as aggregate nouns which refer to cluster individuals, that is, *listí* 'foliage' refers to connected clusters of leaves and combinations thereof. For *-i*-derived nouns, the relevant connectedness relation is proximate connectedness, that is, all the individual leaves, or alder trees and other examples from table 1, must be spatially close to one another (to a degree relevant for the noun at hand). Importantly, this is a stronger condition than just being a plural individual, and distinguishes the meaning of *-i*-derived nouns from simple plural meanings. As given in (62), the morpheme *-i* applies to the root (designating the kind) and returns the connected clusters. As no individual objects are at issue, the OU operator is existentially bound.

(62) $[\![-i]\!] = \lambda k \lambda x \exists n[R(x,k) \wedge x \in CLUSTER \wedge OU(k,x) = n]$

The semantics given here delivers on the nouns' properties as discussed in section 2. Since the OU operator is existentially bound, it follows that these nouns resist pluralization and counting with basic cardinal numerals, since composition with such elements would fail. Unlike typical non-countable nouns such as *water* which allow contextual shifts to countable uses by adding a contextually specified measure function which counts units, *-i*-derived nouns resist contextual shifts to countable uses since the OU operator is already part of the denotation, it cannot be supplied from context. Finally, the application of *-i* was seen to have severe lexical restrictions. The given analysis, which results in cluster individuals, is constrained to apply just to those nouns which describe entities which do come in connected sets, which is a much more restricted class of entities than those to which, e.g., plural morphology would legitimately apply.

Although the analysis just given is, we argue, sufficient to account for the core facts concerning *-i* nouns, we note that there is clearly more to be said to give a full account of the special semantics of terms such as *lanoví* 'rigging/ropes' (< *lano* 'rope') or *krajkoví* 'lacework' (< *krajka* 'lace'), similarly for *maliní* 'cluster of raspberry bushes' (<*malina* 'raspberry'). Mere spatial connectedness of a set of ropes is necessary but not sufficient for them to constitute a ship's rigging, since, for example, the ropes must be of the appropriate types and organized in the fashion required as specified with respect to the function of rigging on a ship. We, however, leave such intriguing details to the side on this occasion.

5.2.2 Aggregate Numerals

Armed with a semantics for *-i*-derived nouns, we now return specifying the semantics for the complex numerals for aggregates *-oje*. As the complex numerals count *-i*-derived nouns, they

clearly count cluster individuals; however, complex numerals for aggregates cannot count arbitrary cluster individuals, as there may be many such clusters in a given connected set: Cluster individual of n alder trees will of course contain many smaller cluster individuals of alder trees, $n-1$ and so on.

Instead, the complex cardinals for aggregates count *maximal* clusters, defined in (63), which are then disjoint. We then employ maximal clusters, notated *MaxCluster*, in our semantics for *-oje* given in (64).

(63) $MaxCluster(x,P) = \exists C[Cluster(x,P,C) \wedge \forall y(Cluster(y,P,C) \wedge O(y,x) \leftrightarrow y \leq x)]$

(64) $[\![\text{-oje}]\!] = \lambda n \lambda P \lambda x[P(x) \wedge \exists Y [\forall z(z < x \wedge MaxCluster(z,P) \leftrightarrow z \in Y) \wedge |Y| = n]]$

-oje first combines with a numeral, similarly to the other complex numerals, and then with a noun, whose denotation is filtered to a derived set containing just the maximal clusters of the specified cardinality. As with group numerals, aggregate numerals combine with predicate noun type, as opposed to taxonomic numerals which select for the kind denotation.[25] While in the case of group numerals, we assumed that the genitive NP complement was an unsaturated NP meaning as it entered into the composition with the group numerals, for aggregate numerals there are several cases to consider. They can count derived nouns like *listí* 'foliage', which directly supply clusters, and, in what is no doubt something of a simplification, we will assimilate pluralia tantum nouns under this class, too. Aggregate numerals also count what are in one sense standard countable nouns such as *hranolky* 'French fries', *karty* 'cards' or *klíče* 'keys' but which also regularly occur in connected sets, and thus also supply clusters, although we leave it open here whether this is lexically specified or pragmatically accommodated. The semantics of *listí* 'foliage' and *hranolky* 'French fries' are given as example meanings in (65)

---

[25] The alternate analysis of lower cardinals as of generalized type $\langle\langle n, a\rangle, a\rangle$ for *-oje* is parallel to that of *-ice*, namely analyzing the cardinals as of type $\langle\langle n, \langle\langle e,t\rangle,\langle e,t\rangle\rangle\rangle, \langle\langle e,t\rangle,\langle e,t\rangle\rangle\rangle$, whereby the composition returns the same result.

and (66), respectively. For *hranolky* 'French fries' and other such predicates, we assume they undergo a shift to the number-neutral predicate before entering into composition with the aggregate numeral, which has the effect of binding off the number argument.[26]

(65) $[\![_{NP}\,\text{listí}]\!] = \lambda x \exists n[\text{R}(x, \text{LEAF}) \wedge x \in \mathit{CLUSTER} \wedge \text{OU}(\text{LEAF}, x) = n]$

(66) $[\![_{NP}\,\text{hranolky}]\!] = \lambda x \exists n[\text{R}(x, \text{FRY}) \wedge \text{OU}(\text{FRY}, x) = n]$

We put the different elements together for the case of *hranolky* 'French fries', and abbreviating (66) as FRY(*x*), producing the meaning in (67):

(67) $[\![_{NumP}\,\text{patery hranolky}]\!] = \lambda x \exists n[\text{FRY}(x) \wedge \exists Y\,[\forall z(z < x \wedge \mathit{MaxCluster}(z, \text{FRY}) \leftrightarrow z \in Y) \wedge |Y| = 5]]$

As such, the semantics arrives at a representation of *-oje* numerals already given in traditional grammatical descriptions of Czech, that they count just objects which are somehow connected.

Further, this analysis provides an explanation for the contrast given in (68), where *-oje* is infelicitously used with a noun which does not designate connected entities.

(68)   #Petr viděl na dvorku dv-oje   psy.

Petr saw on yard   two-OJE dogs

'Peter saw two sets of dogs in the yard.'

Other uses of *-oje* numerals indicate that it can be used even when a noun's semantics does not include (maximal) clusters in its denotation. The example in (69) shows an instance where *-oje* counts entities described by a non-countable noun which has been coerced into a packaging

---

[26]   The assumption that Czech nouns are ambiguous between predicates (⟨e,t⟩) and kinds (⟨k⟩) receives some additional support from adjectival modification. Taxonomic numerals are not compatible with stage-level adjectives modifying its head noun – ???*dvojí včerejší chleba* 'two-kinds of yesterday bread' but completely grammatical with kind level adjectives *dvojí tmavý chleba* 'two-kinds of dark bread'. But there is no such restriction with complex numerals for aggregates: *dvoje včerejší/elektronické noviny* 'two-sets of yesterday/digital newspapers'.

reading, where the packages are themselves complex objects for which the number of objects is further specified.[27]

(69) dv-oje   vody po   šesti

two-OJE water DIST six

'two packages of water, each consisting of six bottles'

Thus, -*oje* numerals are at once compatible with nouns which provide the appropriate semantic type, cluster individuals, but may also impose such an interpretation on nouns which do not standardly designate it.

6 Outlook: Countability from the Perspective of Czech

We have made the case that increased attention to languages with complex nominal morphology is valuable for gaining insight into countability. In this section, we conclude by inverting the usual perspective and considering the English nominal system from the perspective of Czech.

Czech demonstrates that aggregates, groups and taxonomic kinds can be counted in their own right through particular morphological means. From the perspective of Czech, the nominal system of English underspecifies the contrasts visible in Czech morphology. More nuanced differences arise as well as concerning the overall properties of the different systems

---

[27]   Under strong pragmatic pressure even examples as (68) can be acceptable but only if the context provides the connected meaning which is required by the -*oje* numerals, for example, in a scenario where the two pluralities of dogs are harnessed to two different sleds. (Thanks to an anonymous reviewer for pressing this point.) Yet, there is a palpable difference between nouns which denote connected entities due to their lexical meaning (and consequently are always compatible with the -*oje* numerals) and other nouns as witnessed by a simple Google search: *dvoje nádobí* 'two sets of dishes' returns 142 results, many of which are in "out of the blue" contexts, while *dvoje psy* 'two sets of dogs' returns 6 results, all of them used in a context providing connected meaning, for instance, two sets of dogs, each set owned separately, or housed in different locations.

in Czech and English affecting how different interpretations are accessed. In both Czech and English, the plural morpheme delivers sums of objects or sums of taxonomic subkinds and therefore, in the terminology of programming languages, is "overloaded" (in the same way that the + operator is often "overloaded" to perform addition on integers and concatenation on strings). Yet, in Czech the presence of overt grammatical means to count taxonomic sub-kinds, namely taxonomic numerals, impinges on the distributional range of the taxonomic plural, which is restricted to non-episodic contexts, and its interpretational range, which is restricted to specific, well-established kind units. In contrast, the taxonomic sub-kind interpretation in English is not grammatically nor interpretationally restricted, but, apparently, only by pragmatic and/or conceptual possibility.

Similarly, we have argued that the restricted interpretative possibilities of bare nouns in Czech with respect to countability are both a function of the richer interpretative possibilities overtly expressed in morphology as well as, and probably more critically, a function of the presence of bare singular count nouns. It remains to be seen how well this generalizes typologically. At present, it is clear that there is no positive evidence for extending to Czech theories which claim that there is a derivational relation between "ground" interpretations of nouns and interpretations as individuals — for example, Bale and Barner (2009) assert that ground interpretations are foundational and that "all count nouns are derived from lexical items that denote non-individuated semi-lattices." Considering English from the perspective of Czech, ground interpretations are a peripheral phenomenon, which only occur in limited contexts, not foundational.

In sum, Czech would appear to be a well-regimented language, where clear constraints determine in which morphosyntactic context different nominal interpretations may arise; in contrast, English would appear to be very permissive and less constrained by morphosyntactic contexts and more open to contextual shifts in meaning. These observations connect to what

has been repeatedly observed in other areas of grammar: what are hard grammatical constraints in one language are soft statistical constraints in another (Givón, 1979; Bresnan et al., 2001). Thus, the oft-noted particular contextual requirements to produce, e.g., ground interpretations could be seen as a soft statistical constraint against secondary interpretations of nouns designating whole individuals.

Clearly much work remains in order to gain a fuller understanding of the relation between morphosyntax and lexical semantics in nominals. Ultimately, this should aid us in understanding the causal foundations of countability: Why do languages make the countability distinctions that they do? Further investigations will undoubtedly discover more distinctions, and even within Slavic languages there is much more variation than needs to be investigated and understood.

References

Acquaviva, P. 2008. *Lexical Plurals: A morphosemantic approach*. Oxford University Press, Oxford.

Bach, E. 1986. The algebra of events. *Linguistics and Philosophy* 15: 5–16.

Bale, A. & D. Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26: 217–252.

Barner, D. & J. Snedeker. 2005. Quantity judgments and individuation: evidence that mass nouns count. *Cognition* 97: 41–66.

Borer, H. 2005. *Structuring Sense I: In Name Only*. Oxford University Press, Oxford.

J. Bresnan, S. Dingare & C. Manning. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In *Proceedings of the LFG01 Conference*, 13–32, CSLI Publications, Stanford, CA.

Bunt, H.C. 1985. *Mass Terms and Model-Theoretic Semantics*. Cambridge University Press, Cambridge.

Carlson, G. 1980. *Reference to Kinds in English*. Garland Publishing, New York and London.

Casati, R. & A. C. Varzi. 1999. *Parts and Places: The Structures of Spatial Representation*. MIT Press, Cambridge, MA.

Chierchia, G. 1998a. Reference to kinds across languages. *Natural Language Semantics*, 6:339–405.

Chierchia, G. 1998b. Plurality of mass nouns and the notion of "semantic parameter". In S. Rothstein (ed.), *Events and Grammar*, 53–104. Kluwer Academic Publishers, Dordrecht.

Chierchia, G. 2010. Language, thought, and reality after Chomsky. In J. Franck & J. Bricmont (eds.), *The Chomsky Notebook*, 142–169. Columbia University Press, New York.

Czech National Corpus, 2018. URL http://www.korpus.cz.

Dalrymple, M. & S. Mofu. 2012. Plural semantics, reduplication, and numeral modification in Indonesian. *Journal of Semantics*, 29(2):229–261.

Dayal, V. 2004. Number marking and (in) definiteness in kind terms. *Linguistics and Philosophy*, 27(4):393–450.

De Belder, M. 2013. Collective mass affixes: When derivation restricts functional structure. *Lingua*, 126:32–50.

Deal, A.M. 2017. Countability distinctions and semantic variation. *Natural Language Semantics*, 25:125–171.

Dočekal, M. 2012. Atoms, groups and kinds in Czech. *Acta Linguistica Hungarica*, 59(1-2):109–126.

Filip, H. 1999. *Aspect, eventuality types and nominal reference*. Taylor & Francis.

Geurts, B. 2006. The meaning and use of a number word. *Non-definiteness and plurality*, 95:311–329.

Givón, T. 1979. *On Understanding Grammar*. Academic Press, New York.

Grimm, S. 2012. *Number and Individuation*. PhD thesis, Stanford University.

Grimm, S. 2018. Grammatical number and the scale of individuation. *Language,* 94:527-574.

Grimm, S. & B. Levin. 2017. Artifact nouns: Reference and countability. In A. Lamont & K. Tetzloff (eds.), *Proceedings from the 47th meeting of the Northeast Linguistic Society*, 55–64, Amherst, MA, GLSA.

Horn, L. R. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (ed.), *Meaning, Form, and Use in Context: Linguistic Applications*, 11–42. Georgetown University Press, Washington, DC.

Ionin, T. & O. Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics*, 23(4):315–360.

Joosten, F. 2003. Accounts of the count-mass distinction: A critical survey. *Linguisticae Investigationes*, 26:159–173.

Karlík, P. & M. Nekula ad. (eds.) 1995. *Příruční mluvnice češtiny*. Nakladatelství Lidové noviny.

Khrizman, K. 2016. *Numerous Issues in the Semantics of Numeral Constructions in Russian*. PhD thesis, Bar-Ilan University.

Khrizman, K. To appear. The Cardinal/Collective Alternation in Russian Numerals. In Bork, T. R (ed.), *Formal Description of Slavic Languages 11*, Peter Lang, Berlin.

Komárek, M. & J. Kořenský ad. (eds.). 1986. *Mluvnice češtiny*. Academia.

M. Krifka, F. Pelletier, G. Carlson, A. ter Meulen, G. Chierchia, & G. Link. 1995. Genericity: An introduction. In G. Carlson and F. J. Pelletier (eds., *The Generic Book*, 1–125. Chicago University Press, Chicago.

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*, 75–115. Foris Publications, Dordrecht.

Krifka, M. 1992. Thematic relations as links between nominal reference and temporal constitution. In I. Sag & A. Szabolcsi (eds.), *Lexical Matters*, 29–53. Center for the Study of Language (CSLI).

Krifka, M. 1995. Common nouns: A contrastive analysis of English and Chinese. In G. N. Carlson & F. J. Pelletier (eds.), *The Generic Book*, 398–411. Chicago University Press.

Krifka, M. 2003. Bare NPs: Kind-referring, indefinites, both, or neither? In R. B. Young & Y. Zhou, (eds.), *Proceedings of the 13th Conference on Semantics and Linguistic Theory*, pages 180–203, Ithaca, NY, Cornell University.

Landman, F. 1989. Groups I. *Linguistics and Philosophy*, 12:559–605.

Lima, S. 2014. *The Grammar of Individuation and Counting*. PhD thesis, University of Massachusetts, Amherst.

Link, G. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bauerle, Ch. Schwarze & A. von Stechow (eds.), *Meaning, use, and interpretation of language*. de Gruyter, Berlin.

Marušič, F., A. I. Nevins & W. Badecker. 2015. The grammars of conjunction agreement in Slovenian. *Syntax*, 18(1):39–77.

Marušič, F. & A. I. Nevins. 2009. Two types of neuter: Closest-conjunct agreement in the presence of '5&ups'. In A. Fisher, E. Kesici, N. Predolac & D. Zec, W. Browne & A. Cooper (eds.), *Annual Workshop on Formal Approaches to Slavic Linguistics, The Second Cornell Meeting*, 301–317, Ann Arbor, Michigan Slavic Publications.

Mathieu, É. 2012. Flavors of division. *Linguistic Inquiry*, 43(4):650–679.

Miechowicz-Mathiasen, K. 2012. Licensing polish higher numerals: An account of the accusative hypothesis. *Current Issues in Generative Linguistics*, 81.

Müller-Reichau, O. 2006. *Sorting the world: On the relevance of the kind-level/object-level distinction to referential semantics*. PhD thesis, Universität Leipzig.

Payne, J. & R. Huddleston. 2002. Nouns and noun phrases. In R. Huddleston & G. Pullum (eds.), *Cambridge Grammar of the English Language*, 323–524. Cambridge University Press, Cambridge.

Pelletier, F. J. 1979. Non-singular reference: Some preliminaries. In F. J. Pelletier (ed.), *Mass Terms: Some Philosophical Problems*, 1–14. Reidel, Dordrecht.

F. J. Pelletier & L. K. Schubert. 2004. Mass expressions. In D. Gabbay & F. Guenthner, (eds.), *Handbook of Philosophical Logic*, volume 10, 249–335. Kluwer, Dordrecht.

Rothstein, S. 2017. *Semantics for Counting and Measuring*. Cambridge University Press.

Veselovská, L. 2001. Agreement patterns of Czech group nouns and quantifiers. In N. Corver and H. van Riemsdijk (eds.)*, Semi-Lexical Categories: The Function of Content Words and the Content of Function Words,* 273–320*.* Mouton de Gruyter, Berlin.

Wągiel, M. 2015. Sums, groups, genders, and Polish numerals. In G. Zybatow, P. Biskup, M. Guhl, C. Hurtig, O. Mueller-Reichau & M. Yastrebova (eds.), *Slavic Grammar from a Formal Perspective. The 10th Anniversary FDSL Conference, Leipzig 2013*, 495–513, Frankfurt am Main, Peter Lang.

Wierzbicka, A. 1988. *The Semantics of Grammar*. John Benjamins, Amsterdam/Philadelphia.

Winter, Y. 2001. *Flexibility Principles in Boolean Semantics*. MIT Press.

E. J. Wisniewski, Ch. A. Lamb & E. L. Middleton. 2003. On the conceptual basis for the count and mass noun distinction. *Language and Cognitive Processes*, 18:583–624.

Roberto Zamparelli. 1999. *Layers in the Determiner Phrase*. Garland, New York, NY.