

# (In)definiteness-driven typology of arbitrary items.<sup>i</sup>

## 1. Introduction

This paper explores a typology of a number of impersonal [1] and passive [2] constructions in several European languages based on their semantic characterisation and discourse-pragmatic properties.

- (1) They speak English in America.
- (2) The enemy ship was sunk.

I will refer to all of the constructions considered here by a descriptive cover-term *constructions with arbitrary interpretations (arbs)*<sup>ii</sup>. Various authors have described as *arbitrary* the interpretations of pronouns and null syntactic elements (traces and PRO) that do not involve antecedents or bound-variable interpretations (Jaeggli 1986, Lebeaux 1984, Cabredo-Hofherr 2002, inter alia). These items then yield sentence interpretations that have a generic or impersonal flavour to them. Another property that all of the constructions considered here have in common is the demotion (in the several senses to be defined below) of the agent. While these constructions have been extensively explored over the last several decades, the proposed analyses remain extremely diverse, and semantics of arbitrariness is far from being thoroughly understood. At the same time discourse-related properties of these constructions have received very little attention in the literature (Koenig and Mauner 1999 and Prince 2003 being two exceptions I am aware of).

The goal of this work is thus twofold: first, I want to introduce a semantically-driven typology into the diverse realm of impersonals and passives. I will argue that arbs are either semantically indefinite, or semantically definite, drawing attention to previously unobserved behaviour of different arbs with respect to adverbial quantification. Secondly, I will point out further differences between the two types of arbs (and non-arbitrary NPs) that emerge in their interaction with topic structure and discourse anaphora.

The semantics of arbitrariness has received significant attention in the literature (Jaeggli 1986, Cinque 1988, Condoravdi 1989, Kim 1991, Chierchia 1995, Koenig and Mauner 1999, Alonso-Ovalle 2002, Cabredo-Hofherr 2002, inter alia), with proposals for the interpretation of arbs ranging from uniformly indefinite analyses (e.g., Chierchia 1995), to work treating arbs as special kinds of definite pronouns (e.g. Alonso-Ovalle 2002), to accounts arguing that arbs are ambiguous between several formal translations (Cabredo-Hofherr 2002). The chief focus here will be on definiteness and indefiniteness in arbs, and I will present evidence that both types of arbs exist in European languages. The empirical data comes from 3<sup>rd</sup>-person plural constructions with antecedentless pronouns in Russian (where the subject is obligatorily null, unlike the anaphoric pronouns), English (where both anaphoric and antecedentless 3<sup>rd</sup>-person plural pronouns are overt), and Italian (where both kinds of pronouns are null) [3], short verbal passives in Russian and English [4], Russian *SJA*-passives [5] and morphosyntactically somewhat similar Italian *si*-impersonals [6], and the German impersonal pronoun *man* [7].

- (3) Russian
  - a. *V Amerike govornjat po-anglijski.*  
In America speak.3PL in-English.

‘They speak English in America’

- b. *Zdes’ govorjat po-ispanski.*  
Here speak.3PL in-Spanish  
‘They speak Spanish here’=‘Spanish is spoken here’

Italian

- c. *In America parlano inglese.*  
In America speak.3PL English  
‘They speak English in America’
- d. *Qui parlano spagnolo.*  
Here speak.3PL Spanish  
‘They speak Spanish here’

English

- e. They speak English in America  
f. They speak Spanish here.

(4) Russian

- a. *Vchera byl potopen vrazheskij korabl’.*  
Yesterday was sunk enemy.NOM ship.NOM  
‘Yesterday, an enemy ship was sunk’
- b. *V Amerike vchera byl radostno otmechen Den’ Blagodarenija.*  
In America yesterday was joyfully celebrated Day.NOM of.Thanks  
‘Thanksgiving Day was joyfully celebrated in America yesterday.’

English

- c. Yesterday, an enemy ship was sunk  
d. Yesterday in America, Independence Day was celebrated with joy

(5) Russian

- Novyj god prazdnovalsja dolgo i radostno v etot raz.*  
New.NOM year.NOM celebrated.SJA long and joyfully in this time  
‘New Year was celebrated long and joyfully this time around’

(6) Italian

- a. *In Italia, si beve molto vino* (from Chierchia 1995: 107)  
In Italy SI drinks much wine  
‘They drink lots of wine in Italy’
- b. *In Italia ieri si è giocato male* (from Chierchia 1995: 108)  
In Italy yesterday SI is played badly  
‘In Italy yesterday they played badly.’

- (7) German  
*Man wäscht die Hände vor dem Essen.*  
 One washes the hands before the meal.  
 ‘One washes one’s hands before meals’

Morphosyntactically, this is an extremely diverse group of items. 3<sup>rd</sup>-person plural forms and *man* are special pronouns that are subjects of sentences they occur in. On the other hand, in passives and *SJA*-passives, the patients and not the implicit agents serve as subjects. By way of background information, it must be noted that non-arbitrary 3<sup>rd</sup>-person pronouns in Russian are overt – Russian is not a pro-drop language. This provides a convenient way of distinguishing the obligatorily null 3<sup>rd</sup>-person plural arbs from the discourse-anaphoric 3<sup>rd</sup>-person plural pronouns – a formal distinction unavailable in English, where personal and impersonal uses of *they* are both overt, or in Italian, which drops its personal subject pronouns most of the time. In Russian and Italian, 3<sup>rd</sup>-person plural arbs are in fact null subjects, with verbs exhibiting overt 3<sup>rd</sup>-person plural agreement. Verbs occurring with *man* have 3<sup>rd</sup>-person singular agreement marking.

While the form of Russian short verbal passives is a mirror of its English counterpart, involving a copula and a passive participle of the (formerly) main verb, verbal morphology in Russian *SJA*-passives is not passive, but reflexive. In all these constructions, the agent is not overtly expressed, and the patient occupies the subject position and agrees with the verb. The agent, however, is present in the syntactic representation. For instance, it can control the PRO in purpose clauses [8].

- (8) Russian
- a. *Komnaty zdes’ chasto provetrivajutsja, chtoby uluchshit’ kachestvo vozduha.*  
 Rooms.NOM here often air.*SJA* to improve quality.ACC of.air  
 ‘Rooms here are frequently aired to improve the air quality’
- b. *Gaz byl ispol’zovan chtoby usypit’ terroristov.*  
 Gas.NOM was used to put.to.sleep terrorists.ACC  
 ‘Gas was used to put the terrorists to sleep’

Italian *si*-impersonals are structurally very similar to Russian *SJA*-passives, with *si* (usually a reflexive morpheme) occupying a head position in the verbal domain (Infl, in Cinque 1988). *Si* in Italian occurs in a wider range of sentence types than Russian *sja*, since the former can occur with intransitive verbs [6b] – a context in which a *sja*-passive would be impossible in Russian.

All of the items under investigation have been claimed to have arbitrary or impersonal interpretations, in that they are “used when the intention of the speaker is to remain vague about the exact identity of the subject” (D’Alessandro 2004 on *si*). Passive and impersonal constructions have been treated by a number of researchers in the last several decades (Jaeggli 1986, Cinque 1988, Condoravdi 1989, Kim 1991, Chierchia 1995, Koenig and Mauner 1999, Alonso-Ovalle 2002, D’Alessandro 2004), with different authors focussing on their syntactic and morphosyntactic properties, as well as various aspects of their semantics and pragmatics. In this work I will be focussing on the interpretive properties of these constructions, first exploring their truth-conditional semantics, and then their interaction with linguistic and extra-linguistic context.

The range of interpretations available for each of the constructions introduced above is extremely wide. Sentences with 3<sup>rd</sup>-person plural arbs are compatible with generic or almost-universal interpretation for the subject [3a], as well as with (seemingly) existential interpretations [3b]. The sentences in [4a] and [4b] are short passives in which the implicit agents have preferred existential-feeling and almost-universal interpretations, respectively; the Russian *sja*-passive in [5] could be interpreted in both ways, depending on the context. While the Italian *si*-impersonal in [6a] is likely to be interpreted existentially, the one in [6b] could be interpreted in both ways (Chierchia 1995). The constructions illustrated in [3-6] vary with respect to their exclusion or inclusion of the speaker and hearer<sup>iii</sup>. Sentences involving *man* can have existential, universal, speaker-exclusive and speaker-inclusive readings, which are extensively discussed in Kratzer (1997).

While the ranges of interpretation available for these items may seem very similar (and very wide), a closer examination of their semantic behaviour reveals important differences. The chief claim I will be defending here is given in [9] below:

- (9) Arbs fall into two sets that differ in their semantics and their effects on subsequent discourse. Set A consists of the 3<sup>rd</sup>-person plural arbs, and the implicit agents in passives and Russian *sja*-passives; Italian *si*-impersonals and German impersonal pronoun *man* comprise Set B.
- a. Set A arbs behave like (plural) definites, while Set B arbs are semantically indefinite.
  - b. Set A arbs are possible, but unlikely (rare) topics and antecedents for discourse anaphora, while Set B arbs do not participate in topic structure or discourse anaphora at all.

The effects of linguistic and extra-linguistic context on the interpretation of arbs have received relatively little attention in previous literature. The discussion of the role of arbs in discourse has been more or less limited to their referential properties. Authors concerned with syntactic and semantic properties of Italian *si*-impersonals, for instance, noted that while it can support reflexive anaphors, it is unable to provide antecedents for intersentential anaphora (Cinque 1988, Chierchia 1995). In fact, the only item that can be used to refer to the agent denotation in a *si*-impersonal construction (outside of very local contexts licensing reflexive anaphors) is *si* itself. The same behaviour emerges when we examine referential properties of the German *man* or its Yiddish equivalent (Kratzer 1997, Prince 2003).

In contrast, 3<sup>rd</sup>-person plural arbs can support intersentential anaphora – a fact that emerges most clearly in Russian, where personal and arbitrary 3<sup>rd</sup>-person plural pronouns have different realisations. Koenig and Mauner (1999) argue that implicit agents in short verbal passives satisfy the argument slot of the predicate, but do not participate in the referential structure. That is, they claim that these items fail to introduce a new discourse referent or to pick up a discourse referent introduced earlier. I shall argue against this claim, showing that in important respects implicit agents (like 3<sup>rd</sup>-person plural arbs) behave like definite noun phrases in discourse, and that they are not referentially impotent.

The claim in [9] has consequences for the interaction of arbitrary items with context. The interpretation of definite plurals is intrinsically context-dependent, a fact captured in the framework of Schwarzschild (1991) adopted here by including a contextually-determined variable in the formal translation for definite plurals. Moreover, the interpretation of plural

indefinites is also subject to (weaker) contextual constraints, mostly due to the relevance-governed domain restrictions for the existential quantifier that (eventually) binds the indefinite variable.

The rest of this paper is organised as follows. In the next section I will present evidence for the typology in [9a], and a semantic analysis capturing the range of interpretation available to both types of arbs. The ways in which the context influences the interpretation of arbs, and the different effects that items with arbitrary interpretations have on subsequent discourse will be discussed in section 3.

## 2. Semantics of arbs

### *Definiteness in arbs*

Many researchers have claimed that the 3<sup>rd</sup>-person plural arb is translated as an indefinite (Jaeggli 1986, Condoravdi 1989, Kim 1991, Chierchia 1995:547 inter alia). The variable introduced by the indefinite is then bound by the sentence-level operator, e.g., in episodic sentences, the existential quantifier over events unselectively binds the free variable introduced by the indefinite; in generic sentences, the unselective binding is done by the generic operator.

Cabredo-Hofherr (2002) notes that arbs cannot be uniformly analysed as indefinites. In particular, examples in [10] show that the 3<sup>rd</sup>-plural arbs cannot be analysed simply as free variables that take on generic or existential readings depending on the sentential context they are in. That is because the arb in the habitual sentence [10a] can be understood existentially, and the one in the episodic sentence [10b] can be interpreted as generic.

(10) Russian

a. *Zdes' igrajut v karty na urokax.*  
Here play.3PL in cards on classes  
'They play cards here during classes.'

b. *Russkij voennyj lager' schitali shedevrom taktiki.*  
Russian military camp.NOM considered.3PL chef-d'oevr.INSTR tactics.GEN  
'Russian military camp was considered a chef-d'oevr of tactics.'

The same reasoning applies to the implicit agents in the habitual sentence [11a] and episodic sentence [11b] – the implicit agent in the former can be understood existentially, while the one in the latter is most naturally interpreted as almost-universal among relevant individuals.

(11) English

a. Spanish is spoken here.

Russian

b. *Russkij voennyj lager' schitalsja shedevrom taktiki.*  
Russian military camp.NOM considered.SJA chef-d'oeuvre.INSTR tactics.GEN  
'Russian military camp was considered to be a chef d'oeuvre of tactics.'

Cabredo-Hofherr (2002) proposes an ambiguity analysis for these arbs, claiming that they can be translated as definites or as indefinites. When the context provides an individual-level restriction for the arbs (such as a locative expression “in America” or our encyclopaedic knowledge that it’s people “in the government” that raise or lower taxes), then the arbs can be translated as definite plurals, giving rise to the quasi-universal or generic readings in the sentences containing them. In all cases, arbs can also be translated as free variables, giving rise to the various existential readings.

In Malamud (2004), I present evidence that an indefinite translation for these items is unavailable, arguing that the ambiguity analysis of Cabredo-Hofherr (2002), which allows an indefinite translation for the arbs, is inadequate. The evidence comes from contexts that normally cause quantificational variability effects (QVE) in sentences containing indefinites. These contexts contain quantificational adverbs (Q-adverbs) like *always* or *usually*, which denote quantifiers that at least sometimes target situation variables: *always* (*for every situation*), *usually* (*for most situations*), etc. (see Lewis 1975 for an influential analysis of Q-adverbs). In sentences with indefinite subjects, Q-adverbs yield the effect of quantification over the variable introduced by the indefinite (QVE) [12], whether singular, as in the classical QVE sentence in [12a], or plural [12b] (Lewis 1975).

- (12) English (QVE available: *Most students = QVE on students*)  
 a. A Penn student is usually smart  
 b. Students at Penn are usually smart.

In contrast, QVE over a plural definite is achieved with operators like *for the most part* [13], not with Q-adverbs [14] (Nakanishi and Romero 2003).

- (13) The students sitting over there now are for the most part smart.  
 (QVE available: *Most students = QVE on students*)#  
 (14) The students sitting over there now are usually smart.  
 (QVE not available: the only reading is ‘*Now they are smart, now they are not*’)

The same pattern holds in Russian [15], Italian [16]<sup>iv</sup>, and German [17].

- (15) Russian  
 a. *Student MGU obychno horoshij chelovek / ljubit dekana.*  
 Student MGU.GEN usually good.SING person / loves dean.ACC  
 ‘A student of Moscow University usually is a good person / loves the dean’  
 (QVE available: *Most students = QVE on students*)  
 b. *V MGU studenty obychno horoshie rebjata / ljubjat dekana.*  
 In MGU students usually good.PL guys / love.PL dean.ACC  
 ‘Students of Moscow University usually are good people / love the dean’  
 (QVE available: *Most students = QVE on students*)  
 c. *Eti studenty bol’shej chast’ju ljubjat dekana.*  
 These.NOM students.NOM most.INSTR part.INSTR love.3PL dean.ACC

‘These students for the most part love the dean.’  
(QVE available: *Most students = QVE on students*)

- d. *Eti studenty obychno ljubjat dekana.*  
These.NOM students.NOM usually love.3PL dean.ACC  
‘These students usually love the dean.’  
(no QVE: the only reading is ‘*Now they love the dean, now they don’t*’)
- (16) Italian<sup>v</sup>
- a. *In questa città uno studente universitario di solito sa parlare inglese.*  
In this city a student university of usual knows to.speak English  
‘In this city, a college student usually knows English.’  
(QVE available: *Most students = QVE on students*)
- b. *Questi studenti in gran parte sanno parlare inglese / sono tifosi dell’Atalanta.*  
These students in most part know.3PL to.speak English/are fans of.the.Atalanta  
‘These students for the most part speak English / are fans of the Atalanta’  
(QVE available: *Most students = QVE on students*)
- c. *Questi studenti di solito sanno parlare inglese / sono tifosi dell’Atalanta.*  
These students of usual know.3PL to.speak English / are fans of.the.Atalanta  
‘These students usually speak English / are fans of the Atalanta.’  
(QVE not available: the only reading is ‘*Now they are, now they are not*’)
- (17) German
- a. *An dieser Fakultät sind Studenten gewöhnlich klug.*  
On this department are students usually smart  
‘Students in this department are usually smart’  
(QVE available: *Most students = QVE on students*)
- b. *Diese Studenten sind gewöhnlich klug.*  
These students are usually smart  
‘These students are usually smart.’  
(QVE not available; the only reading is ‘*Now they’re smart, now they’re not*’)
- c. *Diese Studenten sind grossenteils klug.*  
These students are for.most.part smart  
‘These students are for the most part smart.’  
(QVE available: *Most students = QVE on students*)

Notably, QVE readings with 3pl arbs in Russian, English, and Italian are achieved with *for the most part*, but not with Q-adverbs [18].

- (18) English

- a. In this department, they're for the most part proud of Maribel.  
(QVE available: *Most people in this department are proud of her*)
- b. In this department, they're usually proud of Maribel.  
(QVE not available: the only reading is 'Now they're proud, now they're not')

Russian

- c. *Na etom fakul'tete bol'shej chast'ju ljubjat dekana.*  
On this department most.INSTR part.INSTR love.3PL dean.ACC  
'In this department, they for the most part love the dean.'  
(QVE available: *Most people in the department = QVE on people in the department*)
- d. *Na etom fakul'tete obychno ljubjat dekana.*  
On this department usually love.3PL dean.ACC  
'In this department, they usually love the dean.'  
(no QVE: the only reading is 'Now they love the dean, now they don't')

Italian

- e. *In questa città in gran parte sanno parlare tedesco.*  
In this city in most part know.3PL to. speak German  
'In this city for the most part they speak German'  
(QVE available: *Most people in this city = QVE on people in this city*)
- f. *In questa città di solito sanno parlare tedesco.*  
In this city of usual know.3PL to. speak German  
'In this city usually they speak German'  
(no QVE: the only reading is 'In most situations, people here speak German')

As this pattern of QVE readings indicates, 3pl arbs are never translated as indefinites, but are strikingly similar to overt definite plurals. Moreover, while discourse conditions for using them are quite different, the truth-conditional semantics for 3<sup>rd</sup>-person plural arbs can be extended to implicit agents as well. Indeed, contrary to the common assumptions that implicit agents in short verbal passives are indefinites (Dowty 1978), placing them in QVE contexts shows that they pattern rather with overt definite plurals [19].

(19) English

- a. In Spain, Michael Jackson is for the most part admired.  
(QVE available: *Most Spaniards = QVE on Spaniards*)
- b. In Spain, Michael Jackson is usually admired.  
(QVE not available: the only reading is 'Now he's admired, now he's not.')

Similarly, the implicit agents in Russian *sja*-passives pattern with overt definites in QVE contexts [20]:

(20) Russian

- a. *V Rossii, "Pravda" bol'shej chast'ju preziraetsja.*  
In Russia, "Pravda.NOM" most.INSTR part.INSTR despises.SJA.  
'In Russia, "Pravda" is for the most part despised.'  
(QVE available: *Most Russians = QVE on Russians*)
- b. *V Rossii, "Pravda" obychno preziraetsja.*  
In Russia, "Pravda.NOM" usually despises.SJA.  
'In Russia, "Pravda" is usually despised.'  
(QVE not available: the only reading is '*Now it's despised, now it's not.*')

It is important to note that this variability is truly dependent on the presence of the implicit agent. This is illustrated by the absence of any such effect in adjectival passives [21], where neither the quantificational adverb, nor "*for the most part*" have the agent denotation to quantify over.

(21) English

- a. The door was usually/for the most part (un)opened.  
\*Intended reading: *Most people did(n't) open the door.*

Russian

- b. *Dver' byla obychno otkryta.*  
Door was usually open(ed)  
'The/a door was usually open(ed)'  
\*Intended reading: *Most people opened the door.*

In Malamud (2004) I offer an analysis of 3<sup>rd</sup>-person plural arbs and implicit agents in short verbal passives and Russian *sja*-passives that translates them uniformly as definite plurals (due to the humanness restriction, the readings are similar to those for overt definite NP '*the people*'). This straightforwardly accounts for the generic/quasi-universal readings for these arbs: 'They speak English in America' is treated as truth-conditionally equivalent to 'The people in America speak English,' with the universal feeling resulting from the maximal (within the discourse-relevant limits) interpretation of the definite subject. Existential-like readings are then derived using covers and team credit – mechanisms developed to account for distributivity and pragmatic weakening in overt plural definites by Schwarzschild 1991 and Brisson 1998, respectively.

Examples like [22] illustrate that distributivity possibilities for 3<sup>rd</sup>-person plural arbs (or, in fact, for indefinite plurals) are like those available to definite plurals. A predicate distributing over its definite plural subject does not have to go all the way to atoms (Schwarzschild 1991). Instead, this subject denotation can be broken into intermediate pieces, and then the predicate distributes up to these pieces. For example, both sentences in [22] (on the antecedentless reading of [22a]) are true in a situation [23], where it's neither the case that each boy individually lifted the piano, nor that all of them collectively did so.

- (22) a. On the fourth floor, they lifted the piano.  
b. The boys lifted the piano.

- (23) Residents of the fourth floor are the boys Ed, John, Bill, and Harry.  
Ed and John lifted the piano together, then Bill and Harry did.

The different distributivity possibilities are achieved by using covers over the domain of discourse. A cover is defined as follows:

DEFINITION. Cover - a collection of sets of entities, such that the union of the sets in the cover gives you the entire universe of discourse.

Assuming the universe of discourse  $D = \{\text{Ed, John, Bill, Harry}\}$ , the cover in (24) corresponds to the situation in (23):

- (24) Cover =  $\{\{\text{Ed, John}\}, \{\text{Bill, Harry}\}\}$

A (pragmatically determined) variable over covers is an inherent part of each VP denotation, allowing predicates to distribute up to each cover-cell in the NP denotation, as illustrated in [25].

- (25)  $(\text{Part}(\text{Cov}_i)(\text{lifted-the-piano}'))(\text{the boys}) =$   
 $= \forall x [x \in [[\text{Cov}_i]] \ \& \ x \subseteq [[\text{the boys}']] \rightarrow x \in [[\text{lifted-the-piano}']]]$

The distributive operator applies to the ‘covered’ denotation: every cell  $x$  in the cover, such that  $x$  is a subset of ‘the boys’, lifted the piano. Assuming definite denotations for arbs, we can thus account for the different distributivity levels available for [22a].

Brisson (1998) builds on the ideas of Schwarzschild (1991) to account for pragmatic weakening of sentences with plural definites. For example, [22] can be true even if only a few of the boys participated as in [26].

- (26) *Scenario*: only Ed and John lifted the piano.

This approach allows sentences with definites to be compatible with essentially existential scenarios. This is equivalent to assigning team credit to the boys for lifting the piano. In Brisson (1998), pragmatic weakening is achieved by putting the exceptional members of the NP denotation in a cover-cell with non-members (‘pragmatic junkpile’). Then, the boys in [19b], interpreted with respect to the cover in [27] will essentially mean ‘the boys except Bill and Harry’.

- (27) Universe =  $\{\text{Ed, John, Bill, Harry, Mary, Sue}\}$   
 Cover =  $\{\{\text{Ed, John}\}, \{\text{Bill, Harry, Mary, Sue}\}\}$

Whatever the implementation, team credit assignment allows existential readings for the definite-plural translation of the arbs, reducing the ambiguity analysis to a uniform treatment of these items as definite descriptions.

*Indefiniteness in arbs*

An investigation taking into account other types of arbs shows that not all of them pattern the same way with respect to (in)definiteness. In particular, QVE contexts were used above to rule out the indefinite translation for the 3<sup>rd</sup>-person plural arbs and implicit agents. However, applying this crucial test to German *man*-impersonals and Italian *si*-impersonals gives an entirely different result. Quantificational adverbs produce the quantificational variability effects on the denotation of *man* [28a] and subjects of *si*-impersonals [28b], just as they do with overt indefinites [compare 28c]<sup>vi</sup>.

(28) German

- a. *An diese Fakultät ist man gewöhnlich klug/stolz auf den Dekan.*  
 On this department is one usually smart/proud of the dean  
 ‘In this department one is usually proud of the dean’  
 (QVE available: *Most people in this department...*)

Italian<sup>ii</sup>

- b. *Qui si è di solito tifosi dell’Atalanta.*  
 Here SI is of usual fans of.the.Atalanta  
 ‘Here one is usually a fan of the Atalanta’  
 QVE available: *Most people here are fans of the Atalanta*)

English

- c. Penn students are usually smart/proud of the dean  
 (QVE available: *Most Penn students...*)

Even more strikingly, the variable introduced with *man* and *si* in the if-clause of a donkey sentence can be bound by quantificational adverbs [29a-b], a hallmark behaviour of indefinites [compare 29c-d]<sup>vii</sup>.

(29) German

- a. *Wenn man klug ist, ist man gewöhnlich stolz.*  
 If one smart is, is one usually proud  
 ‘If a person is smart, he/she is usually proud’  
 (QVE available: *Most smart people are proud*)

Italian

- b. *Se si è intelligenti, si è di solito fieri.*  
 If SI is intelligent.PL, SI is usually proud.PL  
 ‘If a person is intelligent, he/she is usually proud’  
 (QVE available: *Most smart people are proud*)

English

- c. If people are intelligent, they are usually proud  
 d. If a man is intelligent, he is usually proud

This pattern shows that constructions with *man* and *si* behave like indefinites, both in the scope of quantification adverbs and in the if-clauses of donkey sentences.

In her 1997 presentation, Kratzer argues that *man* is actually definite, based on its unacceptability in *there*-sentences [30]:

- (30) German
- a. \**Es war man gekommen.* (Kratzer 1997, example 1)  
\*There was one come  
Intended reading: ‘Someone had come’
- b. *Es war wer/jemand gekommen.* (Kratzer 1997, examples 2-3)  
There was someone come  
‘Someone had come’

However, given the entirely indefinite-like behaviour of *man* in QVE contexts, we must conclude that something else is at work in these examples. As Prince (2003) points out for the impersonal pronoun *me(n)* - the Yiddish counterpart of *man* – the pronoun is not focusable, in the sense that it cannot receive intonational prominence, or be placed in syntactic configurations in which it receives contrastive focus (Rooth 1992) [31].

- (31) Yiddish
- a. *Emitser / #Men, zog ikh dir, iz do geven.* (Prince 2003: example 1e)  
Someone / #One, say I you, is here been  
‘Someone / #One, I’m telling you, was here.’
- German
- b. *Jemand / #Man, sage ich, war hier.*  
Someone / #One, say I, was here  
‘Someone / #One, I’m telling you, was here.’

However, the post-*es-war* position is exactly the one in focus, both in the sense of being hearer-new information (Ward and Birner 1995), and in having intonational prominence. I therefore conclude that it is the unfocusability of *man* (however explained) that is responsible for ruling it out in *there*-sentences, and not its definiteness or indefiniteness. I address the issue of explaining the unfocusability of *man* in section 3 of this paper, essentially following the account given in Prince (2003).

Chierchia (1995) analyses *si* straightforwardly as an indefinite. In his implementation, *si* introduces an existentially quantified variable, which can either scopally interact with quantificational adverbs or generic operators, or if need be, be “disclosed” and (unselectively) bound by them. This account correctly derives existential readings for the agents in episodic [32] and generic sentences [33], and also generic/quasi-universal readings in generic/habitual contexts [34].

- (32) Italian
- a. *Ieri ad un partito si ha ballato molto.*  
Yesterday at a party SI has danced much  
‘Yesterday at a party, people danced much.’

German

- b. *Man tanzte auf der Party.*  
One danced at the party  
'People danced at the party.'

(33) Italian

- a. *Si parla tedesco qui.*  
SI speaks German here  
'German is spoken here.'

German

- b. *Hier spricht man Spanisch.*  
Here speaks one Spanish  
'Spanish is spoken here.'

(34) Italian

- a. *Si parla inglese in America.*  
SI speaks English in America  
'People speak English in America.'

German

- b. *Man spricht Englisch in Amerika.*  
One speaks English in America  
'People speak English in America.'

In her 1997 presentation, Angelika Kratzer claims that *man* is in fact plural, because it can support plurality-taking reflexives, like *each other* [35]

(35) German

- a. *Man redete miteinander.* (Kratzer 1997, example 5)  
One talked with.each.other  
'People talked with each other.'
- b. *In diese Familie ist man aufeinander stolz.*  
In this family is one of each.other proud  
'Members of this family are proud of each other.'

Similarly, *si* also supports anaphoric reference with '*each other*', at least in certain usages [36].

(36) Italian

- Si era parlato l'uno con l'altro.* (Cinque 1988, example 39)  
Si was talked the.one with the.other  
'People talked with each other.'

It must be noted that the plural-ness of *man* and *si* cannot be purely semantic, since '*each other*' cannot refer to singular NPs that denote pluralities, as is shown in [37].

(37) German

- a. \**Diese Familie ist auf einander stolz.*  
This family is of each.other proud  
Intended reading: ‘Members of this family are proud of each other.’

Italian

- b. \**Questa famiglia parla l’un l’altro.*  
This family talks the.one the.other  
Intended reading: ‘Members of this family talk to each other.’

In two corpus studies of *man* and *si* (with total of just under 100 occurrences of each in a corpus), and in previous linguistic literature, I was unable to find examples of generic/universal use of *man/si* that would not involve quantification over situations, times, or worlds [38a-d]. In fact, once the elements carrying this quantification are removed from the examples, the sentences become infelicitous/anomalous [38e].

(38) Italian

- a. *Ieri in Italia si è celebrato il Natale.*  
Yesterday in Italy SI is celebrated the Christmas  
‘Yesterday in Italy people celebrated Christmas’

German

- b. *Gestern feierte man Weihnachten in Deutschland.*  
Yesterday celebrated one Christmas in Germany  
‘Yesterday, people in Germany celebrated Christmas’

- c. *Man ist ein bißchen einsam in der Wüste...*  
One is a bit lonely in the desert  
‘One is a bit lonely in the desert...’

- d. *Ich wußte gut, daß man ihn nicht fragen durfte.*  
I knew well, that one him not ask should  
I knew well that one shouldn’t ask him

- e. #*Man ist ein bißchen einsam.*  
One is a bit lonely  
‘One is a bit lonely’ (infelicitous unless context like *in the desert* is supplied)

I thus conclude that generic/universal use of *man/si* in unassailably episodic sentences is impossible, and that both *man* and *si* are plural indefinites, translated as variables ranging over groups of people (possibly existentially quantified as in Chierchia (1995), depending on the framework for deriving their behaviour in generic contexts).

### 3. Pragmatics of arbs

*Effects of arbs on discourse*

Truth-conditionally, I have claimed that arbs are parallel to other definite and indefinite plural noun phrases. A natural question that arises is whether their discourse properties are those of non-arbitrary definites and indefinites. In fact, definite and indefinite arbs affect discourse quite differently from non-arbitrary noun phrases, and from each other.

Koenig and Mauner (1999) show that implicit agents in English short verbal passives can, under certain circumstances, be subsequently followed by intersentential anaphora (which, they argue, is achieved by an inferential coercive process, rather than normal coreference). This is true for 3<sup>rd</sup>-person plural arbs [39a] as well as for implicit agents in Russian *sja*-passives [39c], as the following constructed follow-ups to naturally occurring examples show (Note that in Russian, conveniently, 3pl arbs must be null, while anaphoric pronouns must be overt).

(39) Russian

- a. *Ivan uzhe privyk, chto ego rasskazu ne verjat ili  
Ivan already got.used, that his story.ACC not believe.3PL or  
ponimajut ego kak-to izvrashchenno. Oni ved' ne  
understand.3PL it.ACC somehow perversely. They indeed not  
vstrechalis' s d'javolom.  
have.met with devil.*

'Ivan already got used to the fact that his story was not believed or understood somehow perversely. Unlike him, they haven't met the devil.'

English

- b. The ship, made of iron, was cleaned and made environmentally safe before it was sunk to the seabed. They wanted to create an artificial reef.

Russian

- c. *Na trave mojetsja krasnaja loshad'. Oni skrebut jej golovu...  
On grass washes.SJA red.NOM horse.NOM They scrub.3PL it head.ACC  
'On the grass a red horse is being washed. They're scrubbing its head...'*

This referential possibility, however, is not evidence enough for declaring that the definite arbs are identical to non-arbitrary NPs in their effect on subsequent discourse. I have conducted a corpus study of their status on a corpus of Russian literary texts collected from the online library of Maxim Moshkow ([www.lib.ru](http://www.lib.ru)). A Centering analysis of the corpus investigated the topicality and pronominalisation patterns for 3<sup>rd</sup>-person plural arbs.

Centering theory (Joshi and Kuhn 1979, Grosz, Joshi and Weinstein 1995, Walker, Joshi and Prince 1998) was proposed as a model of local discourse coherence. The connection between salience and coherence made in the theory was later the basis for pronoun-resolution algorithms (e.g., Brennan, Friedman, and Pollard 1987). In the framework, the local discourse segment is broken up into utterances (roughly equivalent to sentences). Several basic notions of the theory are defined as follows.

DEFINITION 1. *Cf-list* or *the list of forward-looking centers* is the set of discourse entities evoked explicitly (and sometimes implicitly) in each utterance.

The entities on the Cf-list are ranked according to their salience, with language-specific principles determining the ranking. In English, Italian, and Russian, previous studies have determined the ranking as being roughly SUBJECT >> OBJECT >> OTHER (Walker and Prince 1996, DiEugenio 1998, Malamud 2001).

DEFINITION 2. *Cp* or *the preferred centre* is the highest-ranked entity in the current utterance, which is the predicted preferred topic for subsequent discourse, most likely to be pronominalised in future utterances.

DEFINITION 3. *Cb* or *the backward-looking centre* is the most salient (highest-ranked) entity from preceding utterance realised in the current one. This is the centre of the hearer's attention, the actual topic of an utterance, linking this utterance to preceding discourse

The backward-looking centre is the item most likely to be realised as a pronoun in the current utterance. Based on the change and retention of backward-looking centres and preferred centres, transitions of different levels of coherence are defined between adjacent utterances. For example, in [40a] below, the same backward-looking centre (John) is retained in utterance 2 and 3, and the *Cp* of utterance 3 is also the same entity. This is the smoothest type of transition, called *Continue*. On the other hand, in [40b], while the *Cb* and *Cp* of the third utterance are the same, indicating a probability that the speaker will continue talking about Mary, there is a change of backward-looking centre between utterance 2 and utterance 3, making this transition a *Smooth-Shift*. The roughest type of transition is illustrated in [40c], where the topic (i.e., the backward-looking centre) changes from utterance 2 to utterance 3, and the projected topic for subsequent discourse (i.e., the preferred centre) is also different from the *Cb*.

- (40) a. 1. John went to school this morning. 2. He met Mary on his way. 3. He was in a good mood.  
b. 1. John went to school this morning. 2. He met Mary on his way. 3 She carried a pink schoolbag.  
c. 1. John went to school this morning. 2. He met Mary on his way. 3. Bill had been annoying her.

Previous Centering studies (Malamud 2001) showed that non-sentence-final subjects in Russian are the preferred centres in their utterances – they are most likely to be pronominalised in subsequent discourse. However, the analysis of corpus data showed that 3<sup>rd</sup>-person plural arbs are almost never actually followed with intersentential anaphora (5 cases out of 967 items). Moreover, an analysis that considered the phonologically null arbs to be the backward-looking centres, like the personal null pronouns in pro-drop languages, predicted that the discourse in which they occurred had an unusual number of incoherent (rough-shift) transitions between utterances. However, both the intuitive perception of the discourse and an analysis that only counted non-arbitrary pronouns as especially salient showed that the discourse was perfectly coherent. This suggests that while agents of sentences involving arbs *can* provide antecedents for future anaphora (at least when this is the only way to make sense of a discourse), they are not preferred topics nor highly salient items.

For comparison, I conducted a smaller corpus study for overt definite noun phrases in an English corpus (comprised of several Wall Street Journal articles). Previous Centering studies (Brennan, Friedman, and Pollard 1987, Prince 1992, Walker and Prince 1996, inter alia) of English showed that subjects are preferred topics in this language. However, when only the definite noun phrases in the subject position are considered, we find that they are pronominalised quite rarely in subsequent discourse (3 cases out of 63 items). This observation is confirmed in the studies conducted by Francis and Michaelis<sup>viii</sup> on a corpus of spoken English, where definite noun phrases are pronominalised in subsequent discourse a third less than indefinites. While some of the definite descriptions in my corpus were the most salient entities of their sentences, most were neither the centre of attention of their utterance, nor the preferred topics for future discourse. While a much larger corpus is necessary to achieve statistical significance for these results, the two corpus studies strongly suggest that while English overt definites are not strongly salient or topical, Russian 3<sup>rd</sup>-person plural arbs are even less so.

This becomes apparent in sentences containing a competing antecedent for plural anaphora. Even nonsensical overt antecedents provide strong competition for arbs [41].

- (41) Russian  
*Zdes' pekut horoshije pirogi. Oni starajutsja ugodit' klientam.*  
 Here bake.3PL good.ACC pies.ACC. They try.3PL to.please clients.DAT  
 'They bake good pies here. They try to please the clients.'

A very salient (and for 6 out of 10 native speakers asked, the preferred) reading of [40] is one where the pies are trying to please the clients. Of course, overt and sensible antecedents are almost unbeatable competition [42].

- (42) Russian  
*Zdes' vybirajut horoshih artistov. Oni starajutsja ugodit' klientam.*  
 Here choose.3PL good.ACC performers.ACC. They try.3PL to.please clients.DAT  
 'They choose good performers here. They try to please the clients.'

The only reading in [42] is the one where the performers, and not their employers, try to please the clients. Thus, the pronoun (which refers to the most central, the highest-ranked salient entity from the preceding utterance) is preferentially (and in [42], categorically) resolved to *something other* than the arbitrary subject.

Given that, in general, subjects in Russian are typically the likeliest antecedents for future anaphora (and likeliest topics for subsequent discourse), we can conclude that definite-plural arbs are closer to definite plurals in English than to typical Russian subjects. Whatever their chief discourse function(s), utilising 3<sup>rd</sup>-person plural impersonals, short passives and SJA-passives has the effect of decreasing the topicality or salience of the agent. In Centering terms, we say that these arbs succeed in placing a discourse entity on the ranked list, thus making future anaphora possible, but that the entities are ranked extremely low on the list, making future anaphora very much dispreferred.

In contrast, it has been noted by many researchers that *man* and *si*-impersonals cannot provide antecedents for discourse pronouns of any kind (Chierchia 1995, Kratzer 1997, inter alia). Prince (2003) reports on a Centering study of a Yiddish corpus containing *me(n)*, the Yiddish correlate of German *man*. The patterns of pro-drop and pronominalisation force a

conclusion that *me(n)* fails to place an entity onto the ranked list; that is, the agent denotation in sentences with *me(n)* is never considered to be a potential topic for subsequent discourse, or a potential antecedent for intersentential anaphora. I conducted a replication of Prince's study for small corpora of German and Italian. In each case, evidence from pronominalisation and pro-drop (zero instances for both *man* and agents of *si*-sentences; rather, the entity that would have been ranked second after the agent was the most likely to be dropped or pronominalised in subsequent discourse) showed that an entity other than the agent denotation in sentences with *man* or *si* was the preferred centre. This provides a striking contrast with overt indefinites, which are exactly the providers of new discourse entities to the ranked list of potential topics. Given these Centering results, I conclude that the chief discourse function of *man* and *si*-impersonals is to remove the subject/agent denotation from the topic-structure of its sentence (the list of forward-looking centres), allowing the next-highest-ranked entity to become the topic in the Centering sense.

This Centering-based analysis naturally accounts for the unfocusability of *man* – if its denotation does not participate in the computation of topichood and salience, specifically focussing the hearer's attention on it would violate its *raison d'être*.

The corpus studies (in particular, the rareness of subsequent pronominalisation of definite arbs, and the absence of pronominalisation or pro-drop for indefinite ones) support the analysis in which the definite arbs are placed on the bottom of the salience hierarchy (the Cf list), making subsequent discourse anaphora possible yet extremely dispreferred; at the same time, the indefinite arbs would not be placed on the list of potential antecedents at all, making subsequent discourse anaphora entirely impossible, as suggested in Prince (2003) for Yiddish *me(n)*.

### *Effects of discourse on arbs*

In this section, I will discuss the effects of linguistic context on the interpretation of definite and indefinite arbs, focussing on the way different contextual influences can be traced to speaker's communicative goals. Two factors I will consider especially are the entailingness of the immediate linguistic context in which the arbs occur (i.e., on whether the sentences are upward- or downward-entailing), and the situational influences. While definites are traditionally analysed as context-dependent in contrast to indefinites, both definites and indefinites are (in different ways) affected by linguistic and extra-linguistic context in which they occur. In this section, I will first address the phenomenon of domain narrowing and selection, applicable to both definite and indefinite nominals, before proceeding to the main portion, which discusses the effects of situational and discourse context on the structure rather than the size of the domain, and concentrates on definite plurals for that purpose. Finally, I address the effects of entailingness of sentences containing definites and indefinites on their interpretation.

Domain narrowing and selection is one aspect that makes the interpretation of nominals, whether arbitrary or not, context-dependent. All definites, existentials, and universals are subject to what has been called in the literature "contextual domain restrictions" (von Stechow 1999). This is an operation narrowing the restrictor-set of the quantifiers or the maximality operator to just the relevant individuals. In general, the domain restriction that applies in interpreting definites, indefinites, and non-existential QPs is the one most relevant to the speaker or hearer communicative goals, and a salient part of the common ground. In the case of definites, this has been partially analysed as anaphoric links to previous discourse. Note, however, that a property serving as the restrictor set can be made salient by extra-linguistic context [43].

- (43) a. Everyone was at the party. – everyone of the set of beings whose presence at the party makes sense (i.e., excluding people from faraway geographical locations, dead politicians, imaginary animals, etc.).
- b. The students are here. – when spoken by one TA to another in the TA office, this usually includes the students who are taking the class currently, as opposed to all those who've taken it over the past ten years, and all the other students in the world.
- c. I talk to students. – when spoken by one TA to another, this includes students from the appropriate class, and most certainly has no bearing on such groups as elementary-school students.

The same operation applies most noticeably during the interpretation of definite arbs: a salient domain restriction applies to the 'base' meaning of the definite. If a locative expression is present, it provides a very salient property that serves as the domain restriction [44a,b]; otherwise, the definite is restricted to mean "the relevant individuals" [44c] in the same way as are overt NPs in [43].

(44) Russian

- a. *Na galerke zataili dyhanie.*  
 On gallery held.3PL breath.ACC  
 'In the gallery, they held their breath.' = 'The people in the gallery held their breath.'

English

- b. They speak English in America = 'The people in America speak English.'

Russian

- c. [from 'The Gulag Archipelago', on various kinds of arrests]  
*Chernorabocheho vyzyvajut v kontoru.*  
 Blue.collar.worker.ACC call.3PL into office.ACC  
 'They call a blue-collar worker into the office.' = 'The perpetrators of GULAG call a blue-collar worker into the office (in order to arrest him).'

Another way in which definites are context-dependent is the determination of their level of distributivity, modelled in our framework as the choice of cover. A particular cover can be made particularly preferable (and thus a certain reading for a definite particularly salient) by explicit mention in the immediately preceding discourse [45a], or by extra-linguistic considerations [45b] (Schwarzschild 1991). It should be noted that in both sentences in [45], but most obviously in [45b], speaker goals are a factor in the choice of cover, in addition to the salience of the cover provided by the context.

(45) Examples adopted from Schwarzschild (1991)

- a. [out-of-the-blue] Farmer Giles keeps his sheep and his cows in the same barn. The animals fight each other.
- b. [Scenario: Apples arrive to the grocery pre-packaged in baskets. Grocery has only a tiny scale for weighing a couple of fruits at a time, and a huge scale for small truckloads. The speaker is the grocer]  
The apples are too light for the huge scale and too heavy for the tiny scale!

Pragmatic strengthening effects in overt definites, discussed in Krifka (1996), also point to Gricean-like factors influencing the choice of cover and allowance of exceptions. In particular, Krifka notes that utterance situation influences the interpretation of plural definites. This influence is basically the speaker's and hearer's goals affecting the choice of cover, as illustrated in [46a,b].

- (46) *Scenario a:* Before a thunderstorm, John has to decide whether to go on with daily business (action 1, if all windows are closed) or return home (action 2, if some windows are open). [i] resolves his dilemma (and is true) in case if some windows are open. For [ii] to be relevant (and true) all windows have to be closed.
- i. The windows are open.      ii. The windows are closed.
- Scenario b:* A bank vault can be reached through three successive sets of doors. John, a worker in the bank, wants to reach the vault either by going there (action 1) or by asking the manager for keys first (action 2). [i] is resolves the dilemma (and is true) if all the doors are open. For [ii] to be resolving (and true) it's enough if only some of the doors are closed.
- i. The doors are open.      ii. The doors are closed.

In [46a-i], the cover is allowed to be ill-fitting, so that even if one window is left open, the cell containing it (i.e., the cell with *all the relevant windows*, the ones that will influence John's actions) will be a subset of [[open]], making the sentence true while pointing to action 2 as the right solution.

On the other hand, in [46a-ii], the cover has to be both well-fitting and fine, so that the sentence is true while making action 1 the way to go. An ill-fitting cover, e.g. {{window1},{window2}, {window3, door1, door2}}, making *the windows* read "all the windows except window3" would not provide John with an assurance that he needn't go back, failing to resolve his dilemma, and so would not be relevant or appropriate in this situation. In [46b], again, it is the choice of action that determines the choice of cover: John will pick that interpretation of the plural, which will uniquely direct him to one or the other action. As a general principle, we see that what is relevant is basically is what helps the speaker uniquely determine his or her course of action, and that definites are sensitive to exactly this notion of relevance (see van Rooy (2003a,b) for a formal implementation of this notion of relevance).

A further factor has been noted for Krifka to influence the interpretation of definite plurals, and we shall claim, of indefinite plurals as well. When situational and grammatical factors do not enforce a particular reading, upward-entailing contexts favour a universal/exhaustive interpretation for the definite, while downward-entailing contexts trigger an existential interpretation for the definite [47], an effect termed *pragmatic strengthening*:

(47)

- a. (I believe) The windows are made of security glass.  
→ I believe that all of the windows are made of security glass
- b. I doubt the windows are made of security glass.  
→ I believe that none of the windows are made of security glass

Krifka proposes a two-rule principle that derives these pragmatic strengthening effects in plural predications, similar to Horn's (1984) R-based implicatures:

*Rule 1.* Grammar does not fix whether predication P over a sum individual is universal or existential, except if there is explicit information enforcing one or the other.

Examples of explicit information include overt lexical items such as *All the windows...*, *The men each...*, or extra-linguistic information about the predicate and speaker goals (do all subparts need to fulfil the predicate in order for the plural to fulfil it, or is one/some enough). This latter type of explicit information is exemplified in [46] above.

*Rule 2.* When multiple interpretations are available, pick the strongest one that's consistent with general background assumptions.

Krifka's observations hold for arbs as well. Here is a naturally-occurring example involving a 3pl arb, with two simpler variants given below.

(48) Russian

- a. *Odnako vcherashnij opyt pokazal, chto etomu rasskazu ne verjat ili ponimajut jego kak-to izvraschenno.*  
But yesterday's experience showed, that this.DAT story.DAT not believe.3PL or understand.3PL it.ACC somehow perversely  
'But yesterday's experience showed that this story is not believed, or is understood somehow perversely.'
- b. *Opyt pokazal, chto etot rasskaz ponimajut izvraschenno.*  
Experience showed, that this.ACC story.ACC understand.3PL perversely  
'Experience showed that this story is understood perversely.'  
→ I think everyone misunderstands this story
- c. *Ja somnevajus', chto etot rasskaz ponimajut izvraschenno.*  
I doubt, that this.ACC story.ACC understand.3PL perversely  
'I doubt that this story is understood perversely.'  
→ I think nobody misunderstands this story

The entailingness of the context affects implicit agents as well:

(49) Russian

- a. *V etom godu, Rozhdestvo bylo otprazdnovano bujnym vesel'jem.*  
In this year, Christmas.NOM was celebrated wild.INSTR joy.INSTR  
'This year, Christmas was celebrated by wild merry-making'
- b. *Ja somnevajus', chto v etom godu Rozhdestvo bylo otprazdnovano*

I doubt,                    that in this year Christmas.NOM was celebrated  
*bujnym vesel'jem.*  
 wild.INSTR joy.INSTR  
 'I doubt that this year , Christmas was celebrated by wild merry-making'

These changes of interpretation do not depend on the choice of cover, but have a flavour of almost a scope-related phenomenon. That is, the resulting reading is as if the definite plural, interpreted with respect to a well-fitting and finest cover, “out-scopes” the downward-entailing operator, so that the negative property distributes to the individuals in the domain of the plural. This pattern is captured when the sentence interpretation depends on Gricean principles. That is, in the absence of information about speaker goals, in interpreting the definite plural (whether regular or arbitrary), the hearer will go for the strongest relevant interpretation<sup>ix</sup>.

It is important to note that, while Krifka’s rules (and Rooth’s Generalisation for which the rules account) were formulated for expressions denoting sum individuals (i.e., for plural definites), it also applies to plural indefinites, or rather, to bare plurals in English and Russian.

(50) English

- a. (I think that) Children are smart.  
     → I believe that almost all children are smart
- b. I doubt that children are smart.  
     → I believe that pretty much no children are smart

Russian

- c. (*Ja dumaju, chto*) *deti pravil'no pojmut etot rasskaz.*  
 (I think that) children.NOM correctly will.understand this story.ACC  
 '(I think that) children will understand this story correctly.'  
     → I believe that almost all children will get the story
- d. *Ja somnevajus', chto deti pravil'no pojmut etot rasskaz.*  
 I doubt that children.NOM correctly will.understand this story.ACC  
 'I doubt that children will understand this story correctly.'  
     → I believe that pretty much no children will get the story

Just as in the case of definite plurals, the interpretation of each sentence in [50b,d] is not a negation of an existential claim, but rather a negative generalisation; that is, [50b] can be paraphrased as in [51] below.

(51) I think that children are stupid.

Unsurprisingly, Rooth’s Generalisation also applies to indefinite arbs discussed here. In particular, upward-entailing contexts favour universal-like readings for agents of the sentences in [52] below, and downward-entailing contexts favour existential interpretations [53].

(52) Italian

- a. *Questa storia si capisce correttamente.*

This story SI understands correctly  
'One (generally) understands this story correctly.'

German

- b. *Man versteht die Geschichte richtig.*  
One understands the story right  
'One (generally) understands this story correctly.'

(53) Italian

- a. *Dubito che si capisca questa storia corretamente.*  
I.doubt that SI understands this story correctly  
'I doubt that one understands this story correctly.'  
→ I believe that almost everyone misunderstands this story

German

- b. *Ich bezweifle, dass man die Geschichte richtig versteht.*  
I doubt that one the story correctly understands  
'I doubt that one understands this story correctly.'  
→ I believe that almost everyone misunderstands this story

As currently formulated, Krifka's rule does not account for this phenomenon in plural indefinites. However, this further supports our claim that indefinites, too, are subject to contextual influences. Noting that the entailingness of the immediate context has an effect on the preferred interpretation for the sentences with both definite and indefinite plurals, I will leave an account of this effect to future research.

#### 4. Conclusion

This paper explores a typology of impersonal and passive constructions in several European languages that have been claimed to involve 'arbitrary interpretations'. The typology of arbitrary items is based on their status in the definiteness hierarchy. I claim, in particular, that certain types of arbitrary items should be translated as plural definites, arguing that the ambiguity analysis proposed for 3<sup>rd</sup>-plural arbs by Cabredo-Hofherr (2002) is both unneeded and empirically wrong. The existential-like readings for these arbs are derived using a mechanism proposed by Schwarzschild (1991) and Brisson (1998) to account for distributivity facts in overt definite plurals: cover-sets and team credit.

At the same time, I adopt Chierchia (1995) approach to *si*-impersonals to account for the interpretation of both *si*-impersonals and *man*, arguing that these items are semantically indefinite. The behaviour of the different kinds of arbs in the scope of quantificational adverbs and in donkey-sentences provides a test for the availability of an indefinite translation for these items.

The effects of context and speaker goals on the interpretation of arbs suggest further, non truth-conditional similarities between arbs and overt definite and indefinite plurals. At the same time, the properties of arbs and overt plurals with respect to discourse anaphora are shown to be quite different. Corpus studies showing that arbs are not preferred antecedents for subsequent anaphora suggest that one discourse function of arbs is precisely to 'demote' the agent

pragmatically. For the indefinite arbs, this non-topicality becomes categorical rather than preferential – they do not participate in the topic-structure of their discourse segments and thus never provide antecedents for discourse anaphora. A Centering analysis provides a natural explanation for this behaviour.

A further investigation is under way to explore the differences in preferred readings for 3pl arbs and implicit agents: a preliminary corpus study suggests that implicit agents tend to existential-like readings more than 3pl arbs. The two types of sentences have quite different focus structures, a very strong factor influencing the preferred interpretation. Similarly, a cross-linguistic comparison of the constructions with arbs (short verbal passive, 3pl arbs, *se/sja/si* constructions) suggests that finer semantic and pragmatic distinctions are necessary to account for their distribution, since syntactically identical constructions can have different truth-conditions and felicity conditions in different languages [54]. An investigation into the nature and possible explanation for these differences is being conducted.

- (54) *Ruki mojut pered jedoj.*  
Hands.ACC wash.3PL before eating  
'They wash hands before eating.' = 'Hands are washed before eating.'  
(impossible reading in English)

Finally, cross-linguistic coverage of the typology presented here is being expanded in ongoing work, with the inclusion of 3<sup>rd</sup>-person plural arbs in several other languages, impersonal pronouns in French and Finnish, as well as the generic uses of the (singular) second-person pronoun (*you*) in several languages.

## References

- Alonso-Ovalle, L. 2002. "Arbitrary pronouns are not that indefinite". In *Proceedings of Going Romance 2000*, C. Beyssade, R. Bok-Bennema, F. Drijkoningen and P. Monachesi (eds), 1-14. Amsterdam. John Benjamins.
- Brennan, S., Friedman, M., and Pollard, C. 1987. "A Centering approach to pronouns". In *Proceedings of the 25<sup>th</sup> Annual Meeting of the ACL*, 155-162. Stanford, CA.
- Brisson, C. 1998. *Distributivity, Maximality, and Floating Quantifiers*. Ph.D. thesis, Rutgers University.
- Cabredo-Hofherr, P. 2002. "Arbitrary readings of 3pl pronominals". Talk at *Sinn und Bedeutung VII*. University of Konstanz, Germany.
- Chierchia, G. 1995. "The Variability of Impersonal Subjects". In *Quantification in Natural Languages*, E. Bach, E. Jelinek, A. Kratzer and B.H. Partee (eds), 107-143. Kluwer, Boston.
- Cinque, G. 1988. "On *Si* Constructions and the Theory of Arb". *Linguistic Inquiry*, 19:521–582.
- Condoravdi, C. 1989. "Indefinite and generic pronouns". In *Proceedings of the Eight Annual West Coast Conference on Formal Linguistics*, E.J. Fee and K. Hunt (eds), 71–85. Stanford: CSLI.
- DiEugenio, B. 1998. "Centering in Italian". In *Centering in discourse*, M. Walker, A. Joshi and E. Prince (eds), 115-138. Oxford: Oxford University Press.
- Francis, H. S. and Michaelis, L. A. *Lexical Subjects Project*.  
<http://ucsu.colorado.edu/~francish/substart.html>

- von Stechow, P. 1999. "Quantifier domains and pseudo-scope". Talk at *Cornell Context-Dependence Conference*.
- Grosz, B., Joshi, A. and Weinstein, S. 1995. "Centering: a framework for modelling local coherence of discourse". *Computational Linguistics* 21 (2): 203-225.
- Grosz, B. and Sidner, C. 1986. "Attention, intentions, and the structure of discourse". *Computational Linguistics* 12: 175-204.
- Horn, L. R. 1984. "Towards a new taxonomy for pragmatic inferences: Q-based vs. R-based implicatures". In *Meaning, Form and Use in Context: Linguistic Applications*, D. Schiffrin (ed), 11-42. Washington: Georgetown University Press.
- Jaeggli, O. 1986. "Arbitrary plural nominals". *Natural Language and Linguistic Theory*, 4:43-76.
- Joshi, A., and Kuhn, S. 1979. "Centered logic: the role of entity centered sentence representation in natural language inferencing". In *Proceedings of the 6<sup>th</sup> International Joint Conference on Artificial Intelligence*, 435-439.
- Kamp, H. and Reyle, U. 1993. *From Discourse to Logic. Introduction to Model Theoretic Semantics of Natural Language, Formal Logic, and Discourse Representation Theory*. Dordrecht: Kluwer.
- Kim, B. 1991. "An indefinite analysis of pro<sub>arb</sub>". *MIT Working Papers in Linguistics* 14:147-163.
- Koenig, J.-P. and Mauner, G. 1999. "A-definites and the semantics of implicit arguments". *Journal of Semantics*, 16:207-236.
- Kratzer, A. 1997. "German impersonal pronouns and logophoricity". Talk at *Sinn und Bedeutung II*. Berlin, Germany.
- Krifka, M. 1996. "Pragmatic strengthening in plural predications and donkey sentences". In *Proceedings of Semantics and Linguistic Theory VI*, T. Galloway and J. Spence (eds), 136-153. Cornell University: CLC Publications.
- Lebeaux, D. 1984. "Anaphoric Binding and the Definition of PRO". In *Proceedings of NELS 14*. University of Massachusetts, Amherst.
- Lewis, D. 1975. "Adverbs of Quantification". In *Formal Semantics of Natural Language*. E. L. Keenan (ed), 3-15. Cambridge University Press, Cambridge.
- Malamud, S. 2001. "Centering in Russian". In *Penn Working Papers in Linguistics (Volume 7.2: Current Work in Linguistics)*, E. Kaiser (ed), 95-108. Philadelphia, PA.
- Malamud, S. 2004. "Arbitrariness: a definite account". In *WCCFL 23 Proceedings*. V. Chand, A. Kelleher, A. J. Rodríguez, and B. Schmeiser (eds), 541-554. Somerville, MA: Cascadilla Press.
- Moshkow, Maxim. On-line Russian Library. <http://lib.ru>
- Nakanishi, K. and Romero, M. 2003. "Two constructions with *Most* and their semantic properties". Talk at *NELS 34*. SUNY Stony Brook.
- Prince, E. 1992. "The ZPG letter: subjects, definiteness, and information-status". In *Discourse description: diverse analyses of a fund raising text*, S. Thompson and W. Mann (eds), 295-325. Philadelphia/Amsterdam: John Benjamins.
- Prince, E. 2003. "The Yiddish impersonal pronoun *me(n)* 'one': a Centering analysis". Presented at the *Annual Meeting of the LSA*. Atlanta.
- Rooth, M. 1987. "Noun Phrase Interpretation in Montague Grammar, File Change Semantics, and Situation Semantics". In *Generalized Quantifiers. Linguistic and Logical Approaches*, P. Gärdenfors (ed), 237-268. Dordrecht: Reidel.
- Rooth, M. 1992. "A theory of focus interpretation". *Natural Language Semantics* 1: 75-116.

- van Rooy, Robert (2003a) “Questioning to Resolve Decision Problems”, *Linguistics and Philosophy* 2003: 727-763
- van Rooy, Robert (2003b) “Negative Polarity Items in Questions: Strength as Relevance”, *Journal of Semantics*, 2003, 20: 239-273
- Schwarzschild, R. 1991. *On the Meaning of Definite Plural Noun Phrases*. Ph.D. thesis, University of Massachusetts, Amherst.
- Suñer, M. 1983. “pro<sub>arb</sub>”. *Linguistic Inquiry* 14: 188-191.
- Walker, M. and Prince, E. 1996. “A bilateral approach to givenness: a Hearer-Status Algorithm and a Centering Algorithm”. In *Reference and referent accessibility*, T. Fretheim and J. Gundel (eds), 291-306. Philadelphia: John Benjamins.
- Walker, M., Joshi, A. and Prince, E. (eds). 1998. *Centering Theory in Discourse*. Oxford: Oxford University Press.
- Ward, G. and Birner, B. J. 1995. “Definiteness and the English Existential”. *Language* 71:722-742

---

<sup>i</sup> I would like to thank Maribel Romero for prolonged and productive discussions and criticism she has provided of this work. Special thanks are also due to Tatjana Scheffler, Augustin Speyer, and Beatrice Santorini for help with German judgements, to Ivano Caponigro, Roberta D’Alessandro, and Roberto Zamparelli for help with Italian examples and judgements, and to Vita Markman, Zoya Romm, and Alexandra Malamud for help with Russian data. Special thanks to Vita Markman for proofreading and discussions of earlier versions of this paper. I am very grateful to Gennaro Chierchia for pointing out to me the indefinite behaviour of *si* in donkey sentences. I would also like to thank the audience at WCCFL 23 for thoughtful questions and comments regarding the semantics of definite arbs.

<sup>ii</sup> I am extending here the usage of the term *arbitrary* from its application to the null subjects in the Spanish 3<sup>rd</sup>-person plural null pronouns [i] (compare with [1]) (Suñer 1983), which in turn derives from its usage for the agents of certain infinitival clauses [ii] (e.g., Lebeaux 1984).

- i. Spanish (example and translation quoted from Cabredo-Hofherr 2002)
  - Tocan a la puerta
  - knock.3pl on the door
  - ‘Someone’s knocking on the door’ (lit. ‘They’re knocking on the door’)
- ii. [CP PRO To write a dissertation] is hard.

<sup>iii</sup> Factors influencing this variation present an intriguing field of inquiry which, however, lies outside the scope of this paper.

<sup>iv</sup> Italian plurals are quite a bit more complicated. For example, definites are the normal way to express kind terms (the function reserved in English for bare plurals, usually analysed as a kind of indefinite). This makes it difficult to construct Italian QVE examples with plural indefinites. However, to illustrate the point that Q-adverbs bind the variable introduced by the indefinite, a singular example suffices, and definite plurals constructed with demonstrative determiners pattern just like their Russian and English counterparts [16].

<sup>v</sup> Special thanks to Roberto Zamparelli for providing all the examples involving fans of Atalanta, Bergamo’s football team (p.c.).

<sup>vi</sup> Operators like “for the most part”, while often sounding weird, could still be used with indefinites and with these arbitrary items to produce the QVE. However, I use the examples in [25] to show that Q-adverbs, which are powerless to produce QVE in definites, work perfectly to create these effects in sentences with *man* and *si*.

<sup>vii</sup> These examples are complicated by the fact that *man/si* are the only forms available to refer back to *man/si*. So, they are used for this purpose instead of regular pronouns in the consequent clause of the donkey sentences. Since the variables denoted by the two *man/si* occurrences are co-varying under quantification, their co-construal cannot be driven by inference or other pragmatic considerations, as argued for the French impersonal pronoun *on* in Koenig and Mauner (1999), but has to be present in the linguistic representation.

<sup>viii</sup> Thanks to an anonymous reviewer for pointing out this work (and its online location) to me.

---

<sup>ix</sup> In fact, Cabredo-Hofherr's (2002) ambiguity analysis for 3<sup>rd</sup>-person plural arbs in Spanish, Italian, and French, among other languages predicts that in the absence of a corporate-reading-inducing predicate<sup>ix</sup> [iiia] or an individual-level locative predicate [iiib], the existential reading is the only one available (at the very least, strongly preferred).

iii. Spanish

- a. *Planean convocar elecciones.*  
Plan.3PL to.call elections  
'They plan to call elections.'
  
- b. *En España hablan español.*  
In Spain speak.3PL Spanish  
'In Spain, they speak Spanish.'

This is because in her approach, the non-existential readings become available when there is a special means of content identification for the 3<sup>rd</sup>-person plural arb, either provided by the locative expression or the predicate. Thus, in the absence of the methods of content identification, upward-entailing context should not induce universal-like preferences. This is not supported by the Russian data. In fact, the behaviour of these arbs with respect to pragmatic strengthening is very much the same as that of overt definites, as expected under our approach.