Singular generic ‘the’ and uniqueness

Luca Gasparri[a, c] and Gerhard Schaden[b, c]

[a] Centre National de la Recherche Scientifique, [b] Université de Lille, [c] UMR 8163, STL

AAM, February 2026

D. Bekki et al. (Eds.), *Proceedings of LENLS 21*

https://doi.org/10.1007/978-3-032-31949-4_6

Abstract. Generic sentences with singular *the* (GSTs) are standardly assumed to obey the uniqueness presupposition of the ι-operator, extended to kinds. We argue that when GSTs are embedded in discourse contexts featuring competing taxonomic satisfiers of the NP, the presuppositional demands of the definite article are not enforced as expected by prominent analyses. We begin by reviewing two major accounts of definite singular generics: Chierchia’s (1998) determiner account and Dayal’s (2004) NP account. Next, we document the pattern and argue that it is not immediately predicted by either analysis. Finally, we block some natural repair strategies and sketch two possible ways forward: equipping the definite article with a shifter that guards against violations of uniqueness, and predicating GSTs upon a taxonomic property that is by construction uniqueness-compliant.

Keywords: Genericity · Kinds · Definiteness · Uniqueness · Anaphora

1. Introduction

This note is about generic sentences whose subject is a singular DP with the definite singular determiner *the*. We refer to these as GSTs. A representative example is (1) under the reading that Iberian wolves usually hunt at night.

(1) The Iberian wolf hunts at night.

GSTs are a subtype of the broader category of “definite singular generics”, whose DPs can take a determiner other than singular *the*, as illustrated in (2).

(2) Each zoo has a different kind of Iberian wolf: our Iberian wolf is voracious.

We further distinguish GSTs from definite-singular-*the* sentences with kind-selecting predicates

like (3), which do not distribute over instances of the DP category,[1] and which, we assume, form their own compositional type.

(3) The Iberian wolf is widespread.

GSTs have a puzzling distribution, and the precise contours of the factors that govern the productivity of singular generic *the* are not well understood [9]. Notwithstanding the open debate over the conditions that make GSTs possible, several accounts of the semantics of attested GSTs are available. These can be divided into two families: accounts that derive GSTs by building an ambiguity into the definite article, and accounts that derive GSTs by building an ambiguity into the NP. Common to both is the idea that GSTs are governed by an extension of the presuppositional requirements imposed by the ι-operator on ordinary objects. Just as (4a) requires the referent of the subject DP to uniquely satisfy the NP *Coke bottle* among the token entities in the domain, so the GST reading of (4b) requires the grammatical subject to uniquely satisfy λx.*coke-bottle*(x) among the kinds in the domain.

(4) a. The Coke bottle fell from the table.
b. The Coke bottle has a narrow neck.

In this note, we offer a critical discussion of some GSTs whose discourse behavior does not fit these expectations, and we reflect on how a theory of GSTs should treat such cases. We begin (§2) by reviewing two major accounts of GSTs: Chierchia (1998) [6] and Dayal (2004) [11]. We then (§3) describe examples of GSTs whose behavior challenges both analyses. Finally (§4), we sketch two ways forward: equipping the definite article with a shifter that guards against violations of uniqueness, and predicating GSTs upon a taxonomic property that is by construction uniqueness-compliant. Beyond these specific points, our discussion carries a broader methodological message. Inquiry into GSTs tends to focus on null-context examples or on GSTs in miniature environments: it should pay closer attention to their behavior in more realistic discourse settings.

2. Chierchia (1998) and Dayal (2004)

We begin with [6]. Chierchia's account of GSTs combines three basic components: NPs consistently

[1] (1) can convey that the VP property is generally true of individual instances of the kind IBERIAN WOLF. In (3), only the kind IBERIAN WOLF itself can be widespread.

denoting properties of token entities (sets of ordinary objects); a mass shifter, which we abbreviate as μ; and a definite article ambiguous between two meanings: ι[·] in combination with a count denotation, and g(ι[·]) in combination with a mass denotation, where 'g' is a function from pluralities into atomic groups, or as we shall call them, "totalities".[2] In essence, GSTs are derived by applying g(ι[·]) to count NPs massified by μ. The derivation can be summarized as follows. First, μ is applied to the NP. The shifter outputs the union of the singular NP denotation and the plural NP denotation: ⟦NP⟧ ∪ *PL*(⟦NP⟧). This is the union of all atoms and all pluralities of atoms that fall under NP. Next, ⟦NP⟧ ∪ *PL*(⟦NP⟧) is forwarded to the determiner g(ι[·]). The ι-operator applies first. According to Chierchia, the presuppositional requirements of the ι-operator are consistent across g(ι[·]) and ι[·],[3] and can be modeled in Frege-Russellian terms as spelled out in (5).

(5) ι[x] = the largest member of x

In plural definite reference to ordinary tokens (e.g., (6a)), the ι-operator applies to a set of pluralities and returns the largest plurality in that set. In singular definite reference to tokens (e.g., (6b)), the ι-operator seeks the largest atomic individual in that set. But atoms have no relative magnitude; no atom is larger than another. So *the* NP.SG is defined only if the domain contains exactly one individual satisfying the NP.

(6) a. The Iberian wolves marched in the valley.
b. The Iberian wolf was wounded, desperately looking for shelter.

Applied to ⟦NP⟧ ∪ *PL*(⟦NP⟧), the ι-operator also seeks the largest element in the union. The largest element in ⟦NP⟧ ∪ *PL*(⟦NP⟧) is the maximal plurality of all NP atoms. This maximal plurality is retrieved and passed to g, which converts it into an atomic totality. For concreteness, the denotation of the DP *the Iberian wolf* under the GST interpretation of (1) is (7): the totality whose content is the largest set of token Iberian wolves in the domain.

[2] The term is introduced purely to distinguish this specific construct from other group-theoretic notions invoked later in the discussion. As usual, a plurality {*a*, *b*, *c*} has plural content and takes a plural denoting phrase, whereas the totality g({*a*, *b*, *c*}) has plural content but takes singular morphology.

[3] The definite article "has a uniqueness presupposition" [6, p. 397] and "the only thing which is special about generic uses of the singular definite article is the fact that it induces a mass reading. […] Other than that, 'the' has its ordinary meaning" [6, p. 383]. In other words, *the* NP.SG is ambiguous between ι[⟦NP⟧] and g(ι[μ(⟦NP⟧)]); the meaning of ι is constant.

(7) g(ι[μ(λx.*Iberian-wolf*(x))])

Anaphoric GSTs such as (8) can be accounted for in a similar way.

(8) The Iberian wolf_1 lives in the mountains, whereas the Cantabrian fox_2 dwells in swamps. (Of the two,) Only the wolf_1 hunts at night.

In (8), the GST with *the wolf* denotes the totality of atoms that satisfy the NP *Iberian wolf*; it does not concern any and all kinds of wolf. This is achieved by restricting the atoms NP-accessible at the evaluation of the GST to the denotations of the initial DPs. The first sentence introduces the predicative meanings in (9a). λx.*Iberian-wolf*(x) is intersected with λx.*wolf*(x), yielding λx.*Iberian-wolf*(x), as illustrated in (9b). Applying the mass operator μ to it produces the union in (9c), which in the context of (8) is extensionally equivalent to the pluralities denoted by *PL*(λx.*Iberian-wolf*(x)). Finally, in (9d), ι and g apply as before.[4]

(9) a. λx.*Iberian-wolf*(x), λx.*Cantabrian-fox*(x)

b. λx.*Iberian-wolf*(x) ∩ λx.*wolf*(x) ≡ λx.*Iberian-wolf*(x)

c. μ(λx.*Iberian-wolf*(x)) ≡ λx.*Iberian-wolf*(x) ∪ *PL*(λx.*Iberian-wolf*(x))

d. g(ι[λx.*Iberian-wolf*(x) ∪ *PL*(λx.*Iberian-wolf*(x))])

We now turn to [11]. In essence, Dayal's strategy is to outsource the work carried out by μ and g in [6] to the ambiguity of the NP.[5] Singular *the* consistently denotes ι[·]. NPs are instead ambiguous between the two meaning types in (10), with subscripts 'o' and 'k' indicating ordinary tokens and kinds, respectively.

(10) a. $\lambda x.P_o(x)$ $\langle e_o, t\rangle$

b. $\lambda x.P_k(x)$ $\langle e_k, t\rangle$

For example, *lion* is ambiguous between the $\langle e_o, t\rangle$ denotation λx.*lion*(x) and the $\langle e_k, t\rangle$ denotation

[4] For the record, [6] does not discuss anaphoric GSTs in detail, but the analysis just provided is a charitable extension of the base account.

[5] For a recent cross-linguistic application of Dayal's account, see [18] on Turkish.

λx.LION(x).[6] λx.*lion*(x) is the property of being a token lion. In contrast, λx.LION(x) is the property of being a kind of lion. GSTs are derived by combining ι[·] with the ⟨e_k, t⟩ meaning of the NP. In simple GSTs like (11), *lion* ranges over the taxonomic group illustrated in Figure 1: the generic species LION itself (the kind dominating the hierarchy under the ⟨e_k, t⟩ interpretation of *lion*) together with its relevant subkinds.[7]

(11) The lion has a mane around its neck.

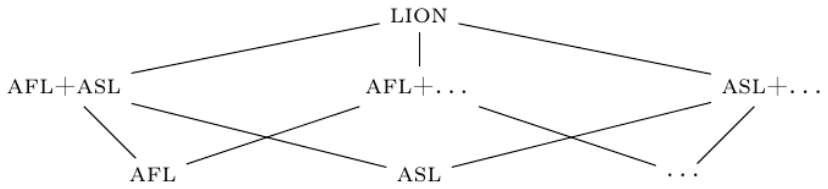

Fig. 1. A sample taxonomic group

The requirements of the ι-operator apply as usual. Singular reference to tokens requires that the context features exactly one atom that satisfies the ⟨e_o, t⟩ meaning of the NP. Likewise, in GSTs "the uniqueness requirement of the ι-operator in combination with a singular common noun will be satisfied if the domain of quantification does not include subkinds of the relevant type" [11, p. 426]. That is: in GSTs, the ι-operator seeks the largest taxonomic object that satisfies the taxonomic reading of the NP, and returns it provided that the domain does not feature other (non-maximal) taxonomic objects of which the kind reading of the NP is true. By analogy with collective NPs such as *committee*, if a GST picks out the taxonomic group [LION, [AFL+ASL, … [AFL, ASL, …]]], the group functions as a structured singular object whose members determine group size but are not themselves categorized as objects in the domain of quantification.[8] Crucially, for Dayal "context determines what level of the hierarchy will be relevant to the interpretation in a particular case, allowing us to include one or more individuals in the denotation of the singular predicative [kind

[6] Henceforward, we use italics for predicates applying to tokens, and small caps for predicates applying to kinds.

[7] AFL and ASL abbreviate AFRICAN LION and ASIATIC LION, respectively. Henceforward, we will notate these taxonomic groups in a flat notation: e.g., [LION [AFL+ASL, … [AFL, ASL, …]]].

[8] Otherwise a *lion*-GST denoting the taxonomic group [LION, [AFL+ASL, … [AFL, ASL, …]]] would defeat Dayal's ban on subkinds trivially. Contrast this to a situation where the domain has two taxonomic discourse referents, the maximal group [LION, [AFL+ASL, … [AFL, ASL, …]]], and the *non*-maximal group [AFL+ASL, … [AFL, ASL, …]]. In such a case the domain would feature two objects of which λx.LION(x) is true.

term]" [11, p. 426]. For instance, in the context of (12a), the denotation of *the lion* includes the generic kind LION, whereas in (12b) *lion* takes a subkind interpretation, which excludes the generic kind LION from the denotation of the indefinite.

(12) a. Unlike other mammals, the lion φs.
b. Not all mammals φ. In fact, a lion has been found not to φ.

The mechanism allows for a straightforward account of anaphoric GSTs. (13) repeats (8).

(13) The Iberian $wolf_1$ lives in the mountains, whereas the Cantabrian fox_2 dwells in swamps. (Of the two,) Only the $wolf_1$ hunts at night.

Normally, ι[λx.WOLF(x)] would be [WOLF, [IBW+AFW, … [IBW, AFW, …]]].[9] But in (13), the portion of $wolf_k$ accessible at the evaluation of *the wolf* is restricted to IBW. Since IBW is a wolf, it also returns true in combination with $wolf_k$, and IBW is the largest taxonomic object accessible to *wolf* in the domain. It is also the only taxonomic object in the domain that satisfies λx.WOLF(x). The requirements of the ι-operator are met, and *the wolf* denotes IBW.

Summing up, Chierchia's account bans GSTs from denoting totalities (whose content is a maximal plurality of atoms) that non-uniquely satisfy the token predicative meaning of the NP. GSTs take totalities with largest sets of ordinary atoms, where a set *s* is the largest NP-set *iff* the NP is true of members of *s* and in the domain there is no *s*' such that the NP is true of members of *s*' and $|s'| > |s|$. In contrast, Dayal's account bans GSTs from denoting a taxonomic entity that is not both maximal and unique in the extension of the $\langle e_k, t\rangle$ meaning of the NP. GSTs take +unique and +maximal taxonomic entities, where a taxonomic entity *k* is maximal, and unique relative to a taxonomic predicate $\lambda x.P_k(x)$, *iff k* is the largest taxonomic object of which $\lambda x.P_k(x)$ is true, and in the domain there is no $k' \neq k$ such that $\lambda x.P_k(x)$ is true of *k*'.

3. Non-uniqueness in GSTs

Chierchia's and Dayal's analyses offer converging predictions about the cases where GSTs should be barred because of a violation of the demands of the ι-operator, and both expect GSTs to pattern closely with definite reference to ordinary atoms. Consider (14) (after [17]).[10]

[9] IBW and AFW abbreviate IBERIAN WOLF and AFRICAN WOLF, respectively.

[10] Subscripts [o] and [k] mark token and kind readings, respectively.

(14) There is a thief$_{[o]}$ down the road, and another one$_{[o]}$ downtown. This situation worries me.
   a. ?? The thief$_{[o]}$ may attack us.
   b. ✓ The downtown thief$_{[o]}$ may attack us.

GSTs are subject to similar effects, as shown by (15) (after [10]).[11]

(15) We manufacture three kinds of bottle in this plant: a tall green bottle, a round green bottle, and a tall blue bottle.
   a. ✓ The round bottle has a narrow neck.
   b. ?? The tall bottle has a narrow neck.

Now consider (16).

(16) There is a lion on the mountains, and another one in the valley. Be careful: the lion preys on people.

There are two possible readings of (16). On one reading, (16) features two indefinites for token lions and a GST.

(17) There is a lion$_{[o]}$ on the mountains, and another one$_{[o]}$ in the valley. Be careful: the lion$_{[k]}$ preys on people.

On another reading, (16) combines two indefinites for subkinds of LION and a GST.

(18) There is a lion$_{[k]}$ on the mountains, and another one$_{[k]}$ in the valley. Be careful: the lion$_{[k]}$ preys on people.

(17) is unproblematic. On Chierchia's account, *the lion*$_{[k]}$ picks out the group comprising the unique largest set of lions accessible to the NP. On Dayal's, *the lion*$_{[k]}$ is $\iota[\lambda x.\text{LION}(x)]$ and requires a maximal taxonomic object uniquely satisfying $\lambda x.\text{LION}(x)$; neither of the preceding indefinites

[11] One might worry that on [6]'s account (15b) should license an anaphoric GST for the fusion of the kinds denoted by *tall green bottle* and *tall blue bottle*. We return to the point below.

supplies a taxonomic object of which λx.LION(x) is true; the GST is therefore safe. Now turn to (18). Chierchia's framework accommodates the example. ι[μ(λx.*lion*(x))] takes the largest set of lions in the domain of quantification. Since the preceding indefinites take subkind interpretations, neither can possibly denote the largest set of lions. As a result, the presuppositional demands of the ι-operator are met. The situation is less clear with Dayal's theory. The indefinites introduce two subkinds of lion. λx.LION(x) is true of these taxonomic referents. The GST at the end of the example seeks the largest unique taxonomic entity that satisfies λx.LION(x) and successfully returns a semantic value. The problem is that this should not be possible under the semantics assumed for the NP, assuming, reasonably, that the two kind-referring indefinites deposit taxonomic referents that persist in the domain at the point where the GST is evaluated. At the derivation of the GST, the group [LION, [AFL+ASL, … [AFL, ASL, …]]] is maximal, but it is not the only taxonomic object in the domain of which λx.LION(x) is true; the categories supplied by the indefinites also land in the extension of the $\langle e_k, t \rangle$ meaning of the NP. The GST should then be filtered out by the presuppositional demands of the definite, but it is not. Why?

First, note that the GST in (18) cannot be explained away by the hypothesis that the DP picks out the conjunction of the kinds introduced by the indefinites. Suppose, with [19], that English has the two definite articles described in (19), with (19a) requiring true contextual uniqueness and (19b) (formalized for deixis and anaphoricity) a weaker property in the range of familiarity.

(19) a. $[\![\textit{the}_{\text{weak}}]\!] = \lambda s_r \lambda P_{\langle e, st \rangle}.\iota x[P(x)(s_r)]$

b. $[\![\textit{the}_{\text{strong}}]\!] = \lambda s_r \lambda P_{\langle e, st \rangle} \lambda y.\iota x[P(x)(s_r) \,\&\, x = y]$

Suppose, further, that (18) passes because the definite article involved in the GST projects the presuppositional demands of $\textit{the}_{\text{strong}}$, and returns the anaphoric conjunction of the lion kind inhabiting the valley and the lion kind inhabiting the mountains.

(20) There is a $\text{lion}_{1[k]}$ on the mountains, and another $\text{one}_{2[k]}$ in the valley. Be careful: the $\text{lion}_{1+2[k]}$ preys on people.

This does not seem right. The reading imparted by (18) is clearly one on which *all* taxonomic entities dominated by λx.LION(x) are such that their members generally prey on people, and on which the addressee should be careful in the area because the propensity to prey on people generalizes across

the species under the generic kind. The correct distribution of indices for (18) is (21).[12]

(21) There is a lion$_{1[k]}$ on the mountains and another one$_{2[k]}$ in the valley. Be careful: the lion$_{3[k]}$ preys on people.

(18) is also fundamentally different from cases involving situational uniqueness obtained through bridging.

(22) Ten species of mammals have just arrived at our zoo; we received fifty exemplars of each species. The lion is friendly with people.

Again, in (22) *the lion* takes a contextually restricted interpretation; the GST asserts that the lions that arrived at the zoo are friendly with people but remains agnostic about whether lions in general (including those in completely different contexts) are friendly with people. In contrast, (18) *does* carry a commitment to the proposition that lions in general (including those in completely different contexts) prey on people. See the continuation test in (23).

(23) a. Ten species of mammals have just arrived at our zoo; we received fifty exemplars of each species. The lion is friendly with people, ✓ though lions generally aren't.

b. There is a lion$_{[k]}$ on the mountains, and another one$_{[k]}$ in the valley. Be careful: the lion$_{[k]}$ preys on people, ?? though lions generally don't.

Recall instead [11, p. 426]: "context determines what level of the hierarchy will be relevant to the interpretation in a particular case, allowing us to include one or more individuals in the denotation of the singular predicative [kind term]". We see two ways to leverage the point as part of a response to (18).

The first is *intensional*: context fixes the taxonomic property expressed by *lion*. Suppose the intension of the taxonomic meaning of *lion* can be coerced based on the QUD. Depending on the

[12] We do not rule out that (18)'s capacity to pick out the largest taxonomic group under λx.LION(x) could be downstream from a raising-to-salience mechanism triggered by the two indefinites, and that [LION, [AFL+ASL, … [AFL, ASL, …]]] can be targeted by definite reference because the indefinites have made it accessible. That, however, would not dissolve the problem: the GST of (18) can be felicitously produced despite the apparent coexistence in the domain of a plurality of taxonomic objects of which the NP is true.

context, *lion*$_{k}$ can contribute λx.LION(x) (the property of being a lion kind), λx.LION$_{TOP}$(x) (the singleton property of being the generic kind topping the taxonomic hierarchy in the range of λx.LION(x), i.e., LION), or λx.LION$_{SUB}$(x) (the property of being a kind scoping taxonomically under LION). Whatever satisfies λx.LION$_{SUB}$(x) or λx.LION$_{TOP}$(x) also satisfies λx.LION(x), though not vice versa. Suppose, further, that in (24a) the NP *lion* contributes λx.LION$_{TOP}$(x), because the partition of the domain that matters to the QUD is restricted to immediate subkinds of MAMMAL, whereas in (24b) the NP contributes λx.LION$_{SUB}$(x), because the QUD is ostensibly centered on the properties generally true of subkinds of LION.

(24) a. (*QUD: the properties of various kinds of mammals*)
The whale does not have a mane; the lion does.

b. (*QUD: the properties generally true of subkinds of lion*)
The lion has a beautiful mane.

Now extend the treatment to (18): the indefinites are predicated on λx.LION(x); in the GST, in contrast, *lion* is the singleton property λx.LION$_{TOP}$(x). Crucially, λx.LION$_{TOP}$(x) is not satisfied by subkinds. The subkinds of LION are not covered *truth-conditionally* by λx.LION$_{TOP}$(x); they are covered only *conceptually* by it, relative to extra-semantic knowledge that LION dominates the taxonomic group in the extension of λx.LION(x). But if λx.LION$_{TOP}$(x) does not cover the subkinds truth-conditionally, the property expressed by the NP of the GST is uniquely satisfied in the taxonomic domain.

We grant that on this hypothesis the problem posed by (18) would be resolved, but there are some problematic aspects. First, there is a cost to arguing that in (18), *lion* shifts predicative content between the indefinites and the GST. Second, admitting a treatment of this kind for (18) may make it harder to predict the oddity of (25).

(25) ? The lion$_{[k]}$ preys on people; the African lion$_{[k]}$ doesn't.

That is: if *the lion*$_{[k]}$ can be intensionally coerced to λx.LION$_{TOP}$(x) to preserve felicity, it should be straightforward to conjoin it with another GST asserting that it is generally false of AFL that its instances prey on people. There is no logical inconsistency in claiming that members of the generic kind LION generally prey on people while members of the subkind AFL generally do not. But (25) feels odd. Objection: the oddity of (25) is no proof that the example feels degraded because at *the*

lion the predicative content of the NP is not $\lambda x.\text{LION}_{\text{TOP}}(x)$. The effect could stem from a general pragmatic aversion to conjoined reference to objects sitting at uneven taxonomic heights. In response, we concede the point. But the generality of the pragmatic stricture is unclear. For instance, (26) feels unproblematic (with or without the contrastive stress indicated with italics).[13]

(26) Teachers are well paid; *philosophy* teachers aren't.

The third point is the comparison to cases featuring adverbs that introduce, intuitively, quantification over regions of the taxonomic hierarchy. See (27) on the reading that it is generally false of members of AFL that they prey on people, though it is generally true of the taxonomic individuals dominated by LION that their members prey on people.

(27) (*QUD: the properties of various subspecies of animals compared to the properties generally true of the species under the general animal concept*)
The African lion does not prey on people, (though) the lion generally does.

At least *prima facie*, it is difficult to see a way to account for the second GST if the predicative meaning of *lion* does not encompass the subkinds. If so, the problem persists.

The second way to leverage Dayal's comment on context sensitivity is *extensional*. Suppose in (18) the NP *lion* is intensionally uniform and that its kind interpretation consistently supplies $\lambda x.\text{LION}(x)$. However, suppose that the state of the domain filters the objects available to $\lambda x.\text{LION}(x)$ during reference assignment to the DP. Now, (18) is problematic only if the taxa introduced by the indefinites are still present in the domain at the point where the GST is evaluated. However, note that the GST is preceded by a transition: *Be careful*. The intervener clearly plays a role here, as its omission causes a slight loss of quality.

(28) There is a $\text{lion}_{[k]}$ on the mountains, and another $\text{one}_{[k]}$ in the valley. ✓/? The $\text{lion}_{[k]}$ preys on people.

Rhetorical markers and bridging conditionals are known to affect the availability of discourse

[13] We recognize that these judgments are subtle and should be confirmed through more extensive empirical testing. However, in this context, they constitute one component of a larger argument, so we believe some residual uncertainty on this point does not undermine the overall case. We thank a reviewer for input on this.

referents in sentences about token objects, as described by classical theories of discourse coherence (e.g., [1]).

(29) There is a lion$_{1[o]}$ on the mountains and another one$_{2[o]}$ in the valley. ??(If you go up the mountains,) Watch out: the lion$_{1[o]}$ could attack you.

Suppose then that in (18) *be careful* licenses a uniqueness-compliant GST by eliminating from the domain of quantification the subkinds supplied by the indefinites.[14] Thanks to the intervener, to echo [11, p. 426], the domain no longer includes "subkinds of the relevant type", and the GST passes.

Again some observations. First, we can generate versions of (18) where the subkinds are mentioned right before the GST without markers of discourse transition or rhetorical structure.

(30) The white Siberian tiger$_{1[k]}$ has been seen in the valley, and the spotted Siberian tiger$_{2[k]}$ has been seen on the mountains. The Siberian tiger$_{3[k]}$ hunts humans.

The subkinds can also be anaphorically recalled right next to the definite.

(31) The white Siberian tiger$_{1[k]}$ has been seen in the valley, and the spotted Siberian tiger$_{2[k]}$ has been seen on the mountains. Be very careful with both$_{1,2[k]}$: the Siberian tiger$_{3[k]}$ hunts humans.

Finally, even with the intervener in place, the subkinds of (18) remain accessible enough to support a personal pronoun. See (32). Personal pronouns are generally assumed to require salient antecedents (e.g., [12]); if the subkinds had left the domain of quantification at this point, pronoun retrieval targeting them should be semantically undefined.

(32) There is a lion$_{1[k]}$ on the mountains and another one$_{2[k]}$ in the valley. If you go up the mountains, be careful: they$_{1,2[k]}$ prey on people.

All in all, pursuing the idea that in (18) (and analogous cases) the subkinds have left the domain of

[14] We distinguish elimination from the domain of quantification (via, e.g., contextual domain restriction) from backgrounding in the attentional sense of, e.g., [1]. Backgrounding demotes a referent in the attentional hierarchy but does not remove it from the domain; elimination does. We take the response to require the stronger operation.

quantification appears to require that their elimination has been enforced by the GST itself. The hypothesis can be defended by appeal to the idea that the interpretive system operates under a pressure to maximize felicity and accommodates whenever necessary. The following is a concise reconstruction. In (18), *lion* is λx.LION(x); at the evaluation of the candidate GST, there are subkinds of LION in the domain; upon checking the domain, the ι-operator finds $n \neq 1$ satisfiers of λx.LION(x) and returns ∅; the repair strategy kicks in; the subkinds introduced by the indefinites are purged; the domain is repopulated with [LION, [AFL+ASL, … [AFL, ASL, …]]]; the group is now a maximal and unique satisfier of λx.LION(x); and the definite article selects that.[15] We do not question the consistency of the proposal. The issue is its general empirical plausibility. One potential issue of empirical coverage comes from discourse fragments that switch repeatedly between different levels of the taxonomic hierarchy. If, as seems reasonable, the content of the domain of quantification is stable across the predications of (33) and the subkinds are not repeatedly purged and reinstated across each predication, we reencounter the initial issue.

(33) The lion φs. In fact, the African lion has been observed to φ. The lion also ψs, since the Asiatic lion has been observed to ψ. […]

We now return to [6]. We have said that Chierchia's account is not threatened by (18) because in his framework GSTs denote totalities of maximal sets of tokens without the mediation of a taxonomic NP. Consider (15) again, repeated as (34).

(34) We manufacture three kinds of bottle in this plant: a tall green bottle$_{1[k]}$, a round green bottle$_{2[k]}$, and a tall blue bottle$_{3[k]}$.
   a. ✓ The round bottle has a narrow neck.
   b. ?? The tall bottle has a narrow neck.

Dayal has a straightforward explanation for the contrast. The subject DP of (34b) requires a taxonomic entity that uniquely satisfies λx.TALL-BOTTLE(x), but there are two kinds of "tall bottle" in the context, so the GST fails. The opposite holds for (34a). For Chierchia, a GST reading of (34a) and (34b) should pick out g(ι[μ(λx.*round-bottle*(x))]) and g(ι[μ(λx.*tall-bottle*(x))]), respectively. The NP can be either anaphorically confined to the bottles introduced by the indefinites or free to

[15] The reasoning can be replicated on the assumption that the domain is repopulated with the generic kind LION alone and the definite article selects that.

range over tall bottles unrestrictedly. Suppose the former is the case.

(35) We manufacture three kinds of bottle in this plant: a tall green bottle$_{1[k]}$, a round green bottle$_{2[k]}$, and a tall blue bottle$_{3[k]}$.
   a. ✓ The round bottle$_{2[k]}$ has a narrow neck.
   b. ?? The tall bottle$_{1+3[k]}$ has a narrow neck.

Why the pattern? As far as we can tell, nothing should prevent *the tall bottle* from returning the largest set of tall bottles in μ(λx.*tall-bottle*(x)) under a contextual restriction of λx.*tall-bottle*(x) to the tall bottles introduced by the kind-referring NPs of the set-up sentence. In which case, (35b) should warrant a GST interpretation denoting the group with the largest set of tall bottles manufactured in the plant. Setting aside the internal consistency of the mechanism, note that something in its vicinity is probably needed to account for the interpretation of the mass N *furniture* in (36).

(36) So, I see a chair$_1$, a table$_2$, a French horn$_3$, and a baseball bat$_4$ here.
   ✓ Come on, pick up the furniture$_{1+2}$ and leave the rest.

One possibility is that we are dealing with an extraneous stricture on a viable compositional pathway. It is often claimed that GSTs require *well-established* kinds (see [4] and others). The basic idea is: (37a) can be read generically because COKE BOTTLE is a well-established kind; (37b) cannot because green bottles do not form a kind; they are just bottles that happen to be green.

(37) a. The Coke bottle has a narrow neck.
   b. The green bottle has a narrow neck.

Now, suppose that for a kind *k* to be well-established at the interpretation of a candidate GST, it is necessary either that speakers occasion-insensitively take *k* to individuate a robust category of things, or that *k* has been overtly introduced in the context. For instance, the NP *round bottle* does not denote a well-established kind and does not license GSTs in out-of-the-blue scenarios. Yet, in (35a) the category counts as well-established in the context; the GST therefore passes. That two kinds *k* and *k*' have been explicitly introduced entails that their plurality is available for definite plural reference, as shown by (38). It does not, however, also entail that their sum is well-established.

Hence the failure of (35b).

(38) We manufacture three kinds of bottle in this plant: a tall green bottle$_{1[k]}$, a round green bottle$_{2[k]}$, and a tall blue bottle$_{3[k]}$. The tall bottles$_{1,3}$ are sturdy and reliable.

In response, two observations. The first is that the well-establishedness condition is its own can of worms.[16] Consider (39).

(39) a. The Boltzmann bottle has a narrow neck.
b. We have a few alien species in this lab. The bluxer is really smart.
c. In the wild, the snarg is aggressive.

Boltzmann bottles do not exist; thus, they should fail any commonsense checklist for well-establishedness. Yet, we can easily interpret (39a) as a GST about a made-up category. In (39b), the GST with the nonce N *bluxer* passes even if all we have introduced in the context is an unspecified set of alien species. The same goes for (39c), which we can interpret as a GST about an unknown kind, possibly an animal, even if all we have are the quantificational incentives provided by the PP *in the wild*. If the response is that well-establishedness is accommodated in these cases, then the question is why the conjunctive kind required for (35b) is not accommodated in the same way. The second observation, orthogonal to our reservations about well-establishedness, is that it seems that GSTs *à la* (35b), with a definite picking out an anaphoric sum, are more liminal than one would think. See (40).

(40) We regularly host four types of guests at this B&B: the philosophy professor from Rutgers$_{1[k]}$; the physics professor from Princeton$_{2[k]}$; the biology professor again from Rutgers$_{3[k]}$; and the chemistry professor from Cornell$_{4[k]}$.
✓/? The Rutgers professor$_{1+3[k]}$ tips much better than the other two academic types$_{2,4[k]}$.

Note that the generic reading supported by (40) does not convey that all kinds of Rutgers professor tip better than "the other two academic types"; the example appears to impart a local generalization about Rutgers philosophy and biology professors, a generalization consistent with a world where all Rutgers professors with the exception of those in philosophy and biology are bad tippers. If our

[16] For another argument questioning the role of well-establishedness in GSTs, see [7, p. 48].

impressions about this example are correct, the puzzle deepens. On the one hand, (40) offers a defeasible reason to think that the failure of (34b) is not due to well-establishedness. But (40) is problematic for Dayal for the same reason (18) was: its GST has an NP with an $\langle e_k, t\rangle$ meaning non-uniquely true of a plurality of taxonomic denizens of the domain.

4. Two Ways Forward

We have argued that there are aspects of the discourse behavior of GSTs not immediately accounted for by two prominent treatments of their semantics: [6] and [11]. In this section, we sketch two possible developments: i) a revised determiner strategy combining $\iota[\cdot]$ with a definite article supplying a shifter that guards against violations of uniqueness; ii) a revised NP strategy on which GSTs are predicated upon a taxonomic property that circumvents the difficulties at play. We do not exclude other (more or less conciliatory) moves,[17] nor do we necessarily assume that GSTs must form a "monistic" derivational kind rather than a family of constructions exploiting different derivational pathways. However, these two approaches should warrant further scrutiny.

We begin with the revised determiner strategy. We mentioned [19]'s distinction between a weak and a strong definite determiner. There are other anti-uniformity views of definite determiners on the market. For instance, [20] claims that the French singular definite article is ambiguous between an ι-operator meaning, and a semantically null one which lets the property denotation of the NP percolate up to the DP. Now, suppose NPs are ambiguous as per Dayal's model. Suppose, additionally, that our determiner inventory is organized as in (41), with ${}^{\text{nom*}}(\cdot)$ defined only for taxonomic properties.

(41) a. $[\![\textit{the}_{\text{weak}}]\!] = (19a)$

b. $[\![\textit{the}_{\text{strong}}]\!] = (19b)$

c. $[\![\textit{the}_{\text{nom*}}]\!] = \lambda s_r \lambda P_{\langle k, st\rangle}.\iota x_k[({}^{\text{nom*}}P)(x_k)(s_r)]$

The standard *nom* function maps properties of ordinary objects onto kinds [5, 15]. Suppose *nom** takes instead a taxonomic property, and returns the singleton property of being the entity at the apex of the taxonomic hierarchy under the taxonomic property. Thus, if $\textit{lion}_k = \lambda x.\text{LION}(x)$, ${}^{\text{nom*}}\textit{lion}_k = \lambda x.\text{LION}_{\text{TOP}}(x)$. The ι-operator is still needed to shift the singleton property to *e*, but now $\lambda x.\text{LION}_{\text{TOP}}(x)$

[17] For instance, renegotiating from the ground up the claim that singular *the* projects requirements in the ballpark of situational uniqueness and having it impose weaker, epistemically flavored constraints like unique identifiability [13]. We will not address this option here, as that would require shifting to another debate altogether.

is guaranteed to be uniquely satisfied in the domain, and uniqueness is settled at the level of *nom**. Next, we mix the ingredients as follows: the_{strong} combining with a taxonomic property to account for anaphoric and deictic GSTs; and $the_{\text{nom*}}$ combining with a taxonomic property to account for free GSTs. On paper, the results are encouraging: (18) is no longer problematic, since $^{\text{nom*}}lion_{\text{k}}$ is not true of subkinds of LION. Other cases that raised concerns for Dayal's original account, like (31), are also accounted for. Granted, (40) remains an oddball; also, assuming that (24b) and (27) are not anaphoric, it remains to be seen whether an NP expressing $^{\text{nom*}}lion_{\text{k}}$ is consistent with the readings we have identified for these examples. (27) appears to require that *the lion* simultaneously express the singleton property of the apex kind and carry a generalization whose contrastive scope covers the full taxonomic hierarchy. But there is room to think that an approach of this sort could extend the empirical coverage of a theory using taxonomic NP meanings.

As for the second strategy, suppose that the definite article is consistently $\iota[\cdot]$, that the ι-operator is uniqueness requiring, and NPs are Dayal-style ambiguous between sets of ordinary individuals and sets of subkinds. However, suppose that the $\langle e_k, t\rangle$ meanings of the NP further split into two subkinds of taxonomic properties: the taxonomic properties invoked by Dayal and name-bearing properties analogous to the metalinguistic predicates that [14] associates with proper names. [14] offers an argument that proper names should be treated as type $\langle e, t\rangle$ expressions lexicalizing "being called" properties. For instance, while on traditional referentialist treatments *John* semantically denotes an individual, for [14] *John* should be treated as a predicative expression denoting $\lambda x.$*called-*$\ulcorner$John$\urcorner(x)$. Now, the idea that GSTs exploit metalinguistic relations is not new. For instance: "[t]he definite DP in subject position in [*La ballena está en peligro de extinción*, 'the whale is in danger of extinction'] corresponds to a definite kind. It denotes […] the name of a kind, the name of the property denoted by the noun *ballena* 'whale' " ([3, p. 136]).[18] Suppose therefore that *lion* is ambiguous between (42a), (42b) and (42c), and that GSTs with *the lion* are predicated upon (42c), whereas generic singular indefinites and plurals use (42b).

(42) a. $\lambda x.$*lion*(x)

b. $\lambda x.$LION(x)

c. $\lambda x.$CALLED-$\ulcorner$lion$\urcorner(x)$

[18] Note that the quoted passage concerns an example with a kind-selecting predicate; we set this aside here. Also, we do not attribute to [3] the proposal described below; the citation is provided purely as evidence of independent sympathy for a metalinguistic approach to definite kind reference. For a similar outlook on GSTs in French, see [2].

(42c) is still a taxonomic property, but does not project structured groups of kinds the way (42b) did. λx.CALLED-⌜lion⌝(x) is the set of kinds whose name is ⌜lion⌝. Interpreted strictly, the property applies exclusively to the taxonomic entity topping the tree of λx.LION(x). LION is in fact called ⌜lion⌝; in contrast, AFL is *not* called ⌜lion⌝; it is called ⌜*African* lion⌝. An analysis along these lines would provide additional resources to address the cases at issue. (43) repeats (18).

(43) There is a $\text{lion}_{1[k]}$ on the mountains, and another $\text{one}_{2[k]}$ in the valley. Be careful: the $\text{lion}_{3[k]}$ preys on people.

Suppose in (43) the two indefinites combine with λx.LION(x) and introduce unspecified subkinds of lion, whereas the GST relies on λx.CALLED-⌜lion⌝(x). If only the generic kind at the top of λx.LION(x) satisfies λx.CALLED-⌜lion⌝(x), the threat of a violation of uniqueness is lifted. Setting aside how the move would improve the analysis of other cases, the reasoning does face a concern when applied to GSTs whose NPs receive a compositional interpretation. See (44) (after [8]).

(44) The successful politician is ambitious, manipulative and well-connected.

Intuitively, (44) does not pick out a category individuated by the property of being called ⌜successful politician⌝ (at least not in the same sense in which ⌜lion⌝ can be taken to "name" LION), and seems difficult to treat as involving a (42c)-like predicate. Compare with (45), after [16].

(45) a. The lion ✓is so called because of *x*, *y*, *z*.
b. The successful politician ??is so called because of *x*, *y*, *z*.

As we see it, the task for a generalized metalinguistic approach to GSTs would be to clarify how a metalinguistic treatment of NPs *à la* (44) could be made to work. One way to do that would be to weaken the relational content of the *being-called* condition and allow it to accommodate mixed labels combining aspects of full names and the compositional transparency of usual descriptions.[19] The formal details of such a model would of course need to be developed in full, but we are not aware of any account of GSTs that has pursued this specific angle of analysis. Note that the treatment would again have to account for (24b) and (27), and the associated cases of GSTs predicating properties of lower members of the taxonomic hierarchy. A mechanism allowing the

[19] Thanks to Dorothy Ahn for discussion on this point.

quantificational adverb to restore the subkinds by ranging over parts of the taxonomic group labeled by the NP could make the conjecture a viable contender.

5. Conclusion

We have examined the behavior of definite singular generic sentences in discourse contexts containing competing taxonomic satisfiers of the NP. We began by reviewing Chierchia's (1998) and Dayal's (2004) accounts of GSTs. We then offered evidence that the discourse behavior of GSTs challenges both analyses. Finally, we sketched two ways forward: allowing the definite article to supply a shifter that guards against violations of uniqueness, and predicating GSTs upon a taxonomic property that circumvents the difficulties at issue. Further work should examine the remaining derivational options, develop the available alternatives at the required level of formal detail, and determine which fares best against the full range of GST data. Beyond the specific points we have covered, we also hope to have shown that inquiry into GSTs stands to gain from closer attention to their behavior in more extended discourse contexts.

Acknowledgments

We thank the organizers and the participants in *LENLS 21* for stimulating discussion, and an anonymous reviewer for thorough and constructive feedback. All errors are ours.

References

1. Asher, N., Lascarides, A.: Logics of Conversation. Cambridge University Press, Cambridge (2003)
2. Beyssade, C.: Les définis génériques en français: noms d'espèces ou sommes maximales. In: Dobrovie-Sorin, C. (ed.) Noms nus et généricité, pp. 33–63. Presses Universitaires de Vincennes, Saint-Denis (2005)
3. Borik, O., Espinal, M.T.: On definite kinds. Recherches linguistiques de Vincennes 41, 123–146 (2012). https://doi.org/10.4000/rlv.2104
4. Carlson, G.N.: Reference to Kinds in English. Outstanding dissertations in linguistics, Garland, New York (1980)
5. Chierchia, G.: Topics in the Syntax and Semantics of Infinitives and Gerunds. Ph.D. thesis, University of Massachusetts, Amherst (1984)
6. Chierchia, G.: Reference to kinds across languages. Natural Language Semantics 6(4), 339–405 (1998). https://doi.org/10.1023/a:1008324218506

7. Cohen, A.: Think Generic! The Meaning and Use of Generic Sentences. CSLI Publications, Stanford, CA (1999)
8. Cohen, A.: Something Out Of Nothing. Brill, Boston (2020)
9. Cohen, A.: Genericity. In: The Oxford Research Encyclopedia of Linguistics. Oxford University Press (2022). https://doi.org/10.1093/acrefore/9780199384655.013.326
10. Dayal, V.: The singular-plural distinction in Hindi generics. In: Proceedings of Semantics and Linguistic Theory II. vol. 40, pp. 39–58. OSU Working Papers in Linguistics (1992). https://doi.org/10.3765/salt.v2i0.3041
11. Dayal, V.: Number marking and (in)definiteness in kind terms. Linguistics and Philosophy 27(4), 393–450 (2004). https://doi.org/10.1023/B:LING.0000024420.80324.67
12. Gundel, J.: Reference and accessibility from a Givenness Hierarchy perspective. International Review of Pragmatics 2(2), 148–168 (2010). https://doi.org/10.1163/187731010X528322
13. Lyons, C.: Definiteness. Cambridge University Press, Cambridge (1999)
14. Matushansky, O.: On the linguistic complexity of proper names. Linguistics and Philosophy 31, 573–627 (2008). https://doi.org/10.1007/s10988-008-9050-1
15. Partee, B.H.: Noun phrase interpretation and type-shifting principles. In: Groenendijk, J., de Jongh, D., Stokhof, M. (eds.) Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers, pp. 115–143. Foris, Dordrecht (1986). https://doi.org/10.1515/9783112420027-006
16. Quine, W.V.O.: Notes on existence and necessity. The Journal of Philosophy 40, 113–127 (1943). https://doi.org/10.2307/2017458
17. Roberts, C.: Uniqueness in definite noun phrases. Linguistics and Philosophy 26, 287–350 (2003). https://doi.org/10.1023/a:1024157132393
18. Sağ, Y.: Bare singulars and singularity in Turkish. Linguistics and Philosophy 45, 741–793 (2022). https://doi.org/10.1007/s10988-021-09323-0
19. Schwarz, F.: Two kinds of definites cross-linguistically. Language and Linguistics Compass 7, 534–559 (2013). https://doi.org/10.1111/lnc3.12048
20. Vergnaud, J.R., Zubizarreta, M.L.: The definite determiner and the inalienable constructions in French and in English. Linguistic Inquiry 23(4), 595–652 (1992). https://doi.org/10.2307/4178791