

The does not encode an anaphoric index: Evidence from kind uses

Sadhwi Srinivas*

Abstract. Two types of semantic theories concerning referring uses of the English definite article *the* have historically held sway: (i) *uniqueness* theories, where *the* is taken to uniquely describe a referent within some contextually restricted domain, and (ii) *familiarity* theories, where *the* picks out a previously mentioned referent. Here, we focus on an observation made in Reed (2024) on the anaphoric potential of the definite article in kind-denoting contexts: namely, that it is limited when compared to occurrences of *the* in anaphoric individual-denoting contexts as well as to occurrences of other referring expressions (e.g., *that*) in anaphoric kind-denoting contexts. Based on these data, we argue for an analysis of the definite article that makes crucial use of domain restriction rather than anaphoric indices.

Keywords. definite article; domain restriction; kind uses; anaphoricity; uniqueness

1. Introduction. Broadly speaking, two types of semantic theories concerning referential uses of the English definite article *the* have figured prominently in the literature. First, the *uniqueness* theories (Russell 2005; Löbner 1985; Strawson 1950; Barker 2004), according to which the definite expression containing *the* identifies a unique referent within a contextually restricted domain. Alternatively, *familiarity* based accounts (Christophersen 1939; Heim 1982; Kamp 1981; Roberts 2003) argue that *the* picks out a previously mentioned entity already represented within the discourse domain – often by way of formal anaphoric indices. In this article, we focus on an empirical observation that originates in Reed (2024) as a means to further illuminate the *uniqueness* vs. *familiarity* debate in the definiteness literature.

The observation in question is that the English definite article, when used in a kind-denoting context such as (1a), exhibits limited capacity to refer back to the previously mentioned kind-level entity. For instance, native speakers report that the highlighted phrase *the mouse* in (1a) is most naturally interpreted not as referring to the already mentioned class of wood mice, but to the entire mouse-kind. This contrasts with occurrences of *the* in anaphoric, individual-denoting contexts such as (1b), as well as with occurrences of other types of referring expressions (e.g., demonstratives like *this* or *that*) in kind-denoting contexts (1c). In both these cases, the highlighted expression is uncontroversially able to receive an anaphoric interpretation – as referring to the mouse living in the speaker’s backyard in (1b) and to the class of wood mice in (1c).

- (1) a. The wood mouse_k is a rodent native to Europe. #**The mouse**_k is a highly resourceful creature.
 b. A wood mouse_k lives in our backyard. **The mouse**_k is a highly resourceful creature.
 c. The wood mouse_k is a rodent native to Europe. **This mouse**_k is a highly resourceful creature.

* Many thanks to Dr. Ann Reed for generously introducing me to the core phenomenon presented here and for engaging me in interesting discussions. Thanks also to the linguistics program and the English department at William & Mary for their feedback and support, as well as the reviewers and audiences at the LSA 2024 meeting. Author: Sadhwi Srinivas, William & Mary (ssrinivas@wm.edu).

In this article, we interpret data such as in (1) to argue for a *uniqueness*-based analysis of the definite article that crucially involves a domain restriction mechanism, over a *familiarity*-based analysis involving formal anaphoric indices. To develop our presentation of the data and related arguments, the upcoming discussion is organized as follows. In the remainder of Section 1, we describe some of the main ideas within *uniqueness* and *familiarity* theories of definiteness. We also lay out our assumptions pertaining to the definite article in kind-denoting contexts, how it relates to its counterpart in individual-denoting contexts. Following this, Section 2 introduces the core data, based on which we describe our theoretical proposal in Section 3. Here, we also relate our proposal to some other recent ideas pertaining to definite and demonstrative determiners in the literature. Section 4 outlines directions for future work, and concludes.

1.1. UNIQUENESS- AND FAMILIARITY-BASED ANALYSES OF *the*. In standard implementations of *uniqueness*-based analyses of definiteness, a nominal phrase containing *the* (henceforth: *the*-phrase) is licensed if and only if there exists within the domain (or situation) under discussion one and only one referent satisfying the descriptive content of this phrase. The highlighted *the*-phrase in (2a), for example, denotes the unique mouse introduced within the scenario from Wednesday that is currently being described. Compare this to (2b) and (2c), which are both cases where the highlighted *the*-phrase sounds odd due to the lack of a unique entity known to satisfy its descriptive content in the relevant scenario.

- (2) a. I saw a mouse in the bathroom on Wednesday. **The mouse** was huge!
 b. I saw two mice in the bathroom on Wednesday. #**The mouse** was huge.
 c. (out of the blue) On Wednesday, #**the mouse** was huge.

Formally, under the *uniqueness*-based view, the definite article receives a denotation along the lines of (3), adopted here from Schwarz (2009). Per (3), the *the*-phrase contributes to the discourse an entity x that is uniquely described by P within the situational domain restriction s , where *situations* are to be formally understood in the sense of Barwise & Perry (1983). The uniqueness condition is encoded as a presupposition, indicated by the underlined portion in (3).

$$(3) \llbracket the \rrbracket = \lambda s. \lambda P : \underline{\exists! x P(x)(s)} . \iota x P(x)(s)$$

On the other hand, under (strong) *familiarity*-based accounts (Christoffersen 1939; Heim 1982; Kamp 1981; Roberts 2003), the definite article instantiates an anaphoric meaning, picking out a uniquely salient or *familiar* referent from within the narrow discourse context. The canonical way in which referents become *familiar* is by explicit mention of the referent in the prior discourse. Note this is the case in (2a), indicating that this example may be accounted for under *familiarity*-based analyses of *the* just as well as under *uniqueness*-based ones. The infelicity of (2b) and (2c) is also predicted, as in these cases, there is no uniquely salient mouse mentioned in the preceding discourse. Formally, familiarity theories are implemented by means of an anaphoric index, as shown in (4), also adapted from Schwarz (2009):

$$(4) \llbracket the_i \rrbracket = \lambda P : \exists! x [P(x) \wedge x = g(i)] . \iota x [P(x) \wedge x = g(i)]$$

According to (4), it is presupposed that there exists a previously mentioned (and therefore indexed) discourse referent that uniquely satisfies the descriptive content P of the *the*-phrase.

Notice the lack of domain restriction *s* in the above lexical entry. The domain here is understood to be fixed to the narrow discourse context containing all and only previously indexed discourse referents. There is thus no need for any kind of context-specific domain computation.¹

While both the *uniqueness*-based and *familiarity*-based ideas cover a significant amount of overlapping ground in explaining the distribution of the definite article in English, neither is fully comprehensive. Familiarity theories, for instance, are challenged by descriptions targeting referents that are uniquely described but not necessarily familiar (e.g., an out-of-the-blue utterance of *the tallest person in Williamsburg*). Uniqueness theories are challenged by anaphoric epithets, where it is not so clear that the descriptive content when taken on its own accurately describes any known entity (e.g., *Hans called. The poor guy is in trouble again!*). Furthermore, certain data points are problematic under both theories. In so-called *weak definites* (Klein et al. 2013; Carlson & Sussman 2005; Aguilar-Guevara & Zwarts 2010; Schwarz 2014), *the* does seem to not pick out any identifiable referent at all, contrary to what is expected in either *uniqueness* or *familiarity*-based accounts. For example, (5) can mean that Sara and Ibrahim listened to different radios, neither uniquely identifiable to the hearer.

(5) Sara and Ibrahim listened to songs on **the radio**.

These challenges notwithstanding, the two competing analyses continue to hold sway not only in discussions concerning the definite article in English, but also in discussions of definiteness crosslinguistically. Similar ideas have been extended to kind-denoting occurrences of the definite article as well, which we discuss next.

1.2. KIND-DENOTING USES OF *the*. Kind-denoting uses of *the*-phrases occur quite productively, some examples of such occurrences are shown in (6). In all these cases, the intended referent is not an individual entity but a whole class or kind. (6a)-(6b) illustrate reference to a naturally occurring kind, while in (6c), the definite expression stands for a constructed kind.

- (6) a. **The African lion** is a magnificent animal.
b. At the zoo, we saw **the (kind/type/sort of) owl that screeches**.
c. People want to work with **the kind of person that they can be friends with**.

In this article, we assume without argument that the definite article retains a common denotation across its individual-denoting and kind-denoting uses, following the proposal in Dayal (2004) who adopts a uniqueness-based denotation for *the* across all of its occurrences. Instead, it is the noun phrases or NPs that appear as complements to the article which represent not only properties of individual entities, but may also represent – as appropriate – properties of kinds.

A key prediction concerning the distribution of the definite article in individual-denoting *vs.* kind-denoting contexts follows from this view. Since the definite article is semantically identical everywhere, we expect it to be licensed across roughly the same range of uses in both types of contexts. That is, for any context where a previously mentioned individual may be anaphorically identified using a *the*-phrase, we expect anaphoric reference to be possible in maximally similar

¹ In its lack of the *s* argument, the entry in (4) diverges from Schwarz (2009). However, as Schwarz notes (p. 264, footnote #16), such an argument is likely superfluous under an anaphoric story for the definite article.

kind-referring contexts as well, as long as the NP within the *the*-phrase has a valid kind denotation. However, we have already seen an example that violates this expectation in (1). Section 2 introduces further examples of such divergences in the ability to anaphorically refer to individuals vs. kinds using *the*-phrases.

2. Core data. The central data puzzle to be presented in this section and analyzed in forthcoming ones is as follows. In contrast to individual-denoting contexts, the anaphoric potential of the definite article in kind-denoting contexts is significantly limited. This contrast was exemplified in (1) – where we see that the previously mentioned kind is not the preferred interpretation of the anaphoric *the*-phrase in (1a), despite the fact that its lexical content *mouse* is capable of denoting at the kind level, as seen in a sentence like *The mouse is a rodent*, as well as in (1c). A similar contrast is once again demonstrated in (7):

- (7) a. A monarch butterfly_k regularly visits our garden. **The butterfly_k** feeds on the nectar from many plants.
- b. The monarch butterfly_k is quite a remarkable creature. #**The butterfly_k** feeds on the nectar from many plants.
- c. The monarch butterfly_k is quite a remarkable creature. ***The kind of butterfly_k** feeds on the nectar from many plants.

In (7a), which instantiates an individual-denoting, episodic context, the first sentence of the discourse introduces a (specific) monarch butterfly. The subsequent sentence is naturally understood to provide more information on this very same butterfly, that it feeds on the nectar of plants in the speaker's garden. In other words, anaphoric reference to the previously introduced individual butterfly is successfully established. Let us now turn to the kind-referring contexts in (7b)-(7c). Analogous anaphoric reference to the previously introduced monarch butterfly kind is obtained in neither of these examples. The highlighted description in (7c), where *the* appears with the modifier *kind of* but no other descriptive content, is odd to the point of ungrammaticality. The highlighted description in (7b), while not ungrammatical, is nevertheless not readily interpreted as referring to the monarch butterfly. Instead, the most natural kind-level interpretation here is one where *the butterfly* encompasses all kinds of butterflies, including but not limited to monarch butterflies.

It is moreover possible to observe that such resistance to anaphoric reference is not in any way inherent to the NP with lexical content *butterfly*, which is capable of denoting at the kind level. Sure enough, anaphoric reference is readily achieved when it is attempted with referring expressions containing alternative determiners (e.g., *this*):

- (8) The monarch butterfly_k is quite a remarkable creature. **This (kind of) butterfly_k** feeds on the nectar from many plants.

Taken together, (7)-(8) indicate that the resistance to anaphoric reference with *the*-phrases in kind-denoting contexts is crucially connected to the type of determiner within the anaphoric phrase, and does not extend to individual-denoting contexts. This is precisely unexpected under our assumption following Dayal (2004) that *the* has a unified denotation across individual-denoting and kind-denoting contexts, as noted in Section 1.2.

We continue to observe similar contrasts across other occurrences of kind-denoting *the*, establishing its productivity:

- (9) I read a book on a kind of macau_k yesterday.
- #**The (*kind of) macau_k** is only found in the Amazons.
 - This (kind of) macau_k** is only found in the Amazons.
- (10) I met a committee of students_k yesterday.
- The (committee of) students_k** had a lot to say.
 - This committee of students_k** had a lot to say.

Again, anaphoric reference is freely achieved with *the* in the non-kind-denoting context (10). However, structurally parallel uses of *the* within kind-level contexts, like in (9a), are ungrammatical or infelicitous or both².

(11) shows that the anaphoric limitation with kind-referring *the* extends to non-natural kinds as well³.

- (11) Do you know [the kind of person that sings in the shower]_k?
- I live with #**the (*kind of) person_k**.
 - I live with **that (kind of) person_k**.

In (11), interestingly, the oddness of the highlighted *the*-phrase in (12a) remains unmitigated even in the absence of the modifier *kind of*, unlike what we observed with natural kinds. This is presumably due to the NP *person* being independently incapable of denoting at the kind level, as observed by the contrast in (12)⁴:

- (12) a. # **The person** is a social creature.
b. The human being is a social creature.

Reed (2024) provides the example below involving an intended bridging relationship between the antecedent sentence and the subsequent *the*-phrase to illustrate a similar anaphoric limitation, not observed with the demonstrative:

- (13) (passion) is the most important quality in the making of a Major leaguer.
- Kevin Youkilis is **that/ #the type of guy**.
 - Aaron Hill is **that/ #the type of guy**. (ex. 20a in Reed 2024)

² We thank an anonymous reviewer for the examples (9)-(10).

³ We use the term *non-natural* kinds to refer to kinds that do not belong to a natural / biological taxonomy.

⁴ It is an independently interesting question why certain NPs readily allow for kind-level interpretations while others do not. One reason cited in the literature is *well-establishedness*, where the kind referred to cannot be an ad-hoc kind, it must be well-established. It is unclear that this explanation applies to *person* vs. *human being*, however.

Finally, we note that this type of contrast between anaphoric uses of the definite article *vs.* demonstrative has been noted at least once even prior to Reed (2024). Gundel (2010) introduces the data point in (14), where the observation remains identical to what we have noted in the preceding discussion – that it does not quite work to use the definite article to refer back to a previously mentioned kind, a demonstrative determiner is necessary.

- (14) A restudy of pareiasaurs_k reveals that ??**the primitive reptiles_k / these primitive reptiles_k** are the nearest relatives of turtles.

As per the explanation in Gundel (2010), the oddness of the anaphoric interpretation in (14) is due to an ambiguity surrounding the identity of the intended referent. While it could be that the expression *the primitive reptiles* intends to pick out the salient, previously mentioned pareiasaurs, this is not the only possibility. It could alternatively be understood to describe the larger class of all primitive reptiles taken together. The referential ambiguity here is said to arise as a direct consequence of the semantics of the definite article, whereby the definite article only signals that the hearer must associate a unique representation for the intended referent that is accurately described by the descriptive content of the accompanying NP. In (14), this unique representation could be at the level of pareiasaurs or at a higher level involving all primitive reptiles. The ambiguity is bypassed with demonstrative expressions such as *these primitive reptiles*, as demonstratives assume salience of the intended referents in addition to accuracy of description. In this case, the previously mentioned pareiasaurs are more salient in the current discourse context than the broader class of all primitive reptiles. Anaphoric reference with demonstratives is thus more easily achieved.

Despite the success of Gundel’s explanation for (14), it nonetheless falls short for some of the other examples discussed above. Specifically, with non-natural kinds like in (11) or (13), the only viable referent corresponding to the descriptor (*kind of*) *person* or (*type of*) *guy* would seem to be the previously mentioned kind, as these lexical descriptors do not independently denote at the kind level (see 12). However, an anaphoric interpretation with the *the*-phrases in these examples yet fails to obtain. In light of this, a more comprehensive explanation of these facts is yet to be discovered. We attempt such an explanation in Section 3.

3. Proposed resolution and related discussion. In Section 2, we observed that the anaphoric potential of *the*-phrases within kind-denoting contexts is significantly limited when compared to individual-denoting ones. We moreover concluded that such anaphoric limitation must have to do with the semantics of the definite article itself, since anaphoric reference to kinds is not similarly obstructed with demonstrative phrases. The question arises: why is anaphora blocked only in contexts where *the* combines with a kind-level NP, but not elsewhere when these two elements (*the* and kind-level NPs) do not co-occur?

One kind of response to this question, of course, would be to propose that the semantics of the definite article are not in fact identical across individual-denoting and kind-denoting contexts. One could claim, for instance, that the hearer only has access to the preceding discourse context when *the* appears in individual-denoting contexts and not in kind-denoting ones, so that it is unable to establish anaphoric links in the latter case alone. Effectively, this explanation posits a *bona fide* lexical ambiguity, whereby *the* in individual-level contexts is associated with a separate lexical denotation from the one in kind-level occurrences. We do not pursue this route here, instead pursuing an attempt to maintain a unified semantics for *the* throughout the grammar –

following Dayal (2004) – but one that nevertheless anticipates a particular kind of interaction between *the* and kind-level contexts that does not apply when these elements occur individually⁵.

In particular, we argue that the desired unified analysis for *the* that successfully reconciles the facts from Section 2 can be successfully formulated under the umbrella of *uniqueness*-based accounts for *the* that make crucial use of domain restriction, but would be incompatible under a *familiarity*-based view. The anaphoric limitation observed in kind-denoting uses of *the* is categorically unexpected under a *familiarity*-based account for the definite article such as in (4), where the ability to establish a link to a previously mentioned referent is lexically specified regardless of any other considerations. On the other hand, a *uniqueness*-based account in which the domain restriction associated with kind-denoting uses of *the* is importantly different when compared to its individual-denoting uses is exactly capable of introducing an interaction only when *the* occurs with kind-level NPs, and not elsewhere. In what follows, we elaborate on such an account.

3.1. TAXONOMIC DOMAIN RESTRICTION IN KIND-REFERRING USES OF *the*. Under Dayal’s (2004) analysis of kind-referring, singular *the*-phrases, the NP complements to the definite article *the* denote properties of kinds located on a taxonomic hierarchy (as opposed to properties of individuals). An example of a taxonomic hierarchy consisting of higher-level kinds and lower-level sub-kinds is represented in Figure 1, reproduced from Dayal (2004). Sub-kinds are characterized by the presence of a dominating parent node representing a higher-level kind, and often also by the presence of sibling nodes representing other sub-kinds at the same level. We take the sibling nodes at any level within the taxonomic hierarchy to constitute a *contrast set*.

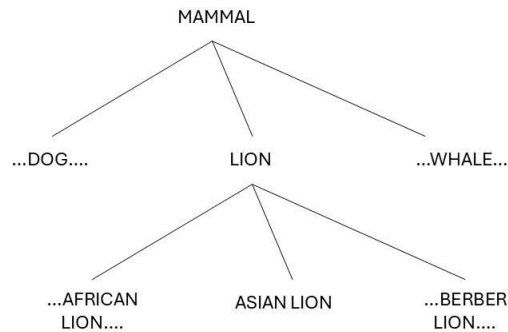


Figure 1. A taxonomic hierarchy consisting of a highest-level kind (mammal), and two hierarchical levels of sub-kinds. Every sub-kind belongs to a contrast set that also contains its siblings.

Dayal proposes the denotation in (15) for the definite article in kind-level constructions. Notice that the semantics assumed are *uniqueness*-based, though no domain restriction is explicitly represented. A “taxonomic domain” is referenced, however, and it is noted that *X* ranges over entities in this domain.

$$(15) \quad \llbracket the \rrbracket = \lambda P. \iota XP(X)$$

⁵ Though see 6 for a brief consideration of an alternative account.

Dayal’s analysis of the definite article in kind-denoting contexts can be straightforwardly reinterpreted to incorporate an explicit domain restriction by simply replacing the denotation for *the* in (15) with the one in (3), reproduced below – along with the additional stipulation that the domain of interpretation in kind-denoting uses of *the* is fixed to a taxonomic domain.

$$(3) \quad \llbracket the \rrbracket = \lambda s. \lambda P : \underline{\exists ! x P(x)(s)} . \iota x P(x)(s)$$

Overall then, the following story holds of the English definite article. It is associated with the meaning represented in (3) within both individual-denoting and kind-denoting contexts. The domains of interpretation differ, however, between the two types of occurrences. In individual-denoting contexts, the referent is identifiable within a contextually salient domain containing individual objects or entities. In kind-denoting contexts, the referent must be identified from within a taxonomic domain induced by the hearer in context. The taxonomic domain represented in Figure 1, for instance, may be assumed in interpreting an utterance such as (16):

(16) **The African Lion** is slowly becoming extinct.

We do not address here the exact mechanism by which a taxonomic domain is induced by the hearer in any given context, but note only that the induced domain is required by definition to comprise of relationships among the kind-level elements within it. This means, crucially, that we cannot have a taxonomic domain with a single kind in it – a contrast set or a dominating node is minimally required.

The enforcement of a taxonomic domain restriction in the interpretation of kind-denoting uses of *the* may now be understood to contribute to their limited anaphoric potential. According to (3), successful interpretation of the intended referent in *the*-phrases relies on the identification of a unique referent satisfying the lexical property *P* within the domain *s*. As stated above, the domain *s* in kind uses of *the* is restricted to the taxonomic domain, consisting at the very least of one or more members in the contrast set of the kind being referenced or its parent kind (or often, both). In this scenario, the descriptive completeness of the lexical property *P* becomes key to the interpretation of the *the*-phrase. Critically, underspecifying *P* within the anaphoric description as in (1) or (7) does not result in successful identification of the referent, since neither the descriptive content in the NP complement of the anaphoric *the*-phrase in (1): *mouse* nor (7): *butterfly* identifies a unique kind within the relevant taxonomic domain, represented in Figures 2 and 3 respectively. This is because more than one element in each taxonomic domain satisfy the property of being a mouse or a butterfly respectively (in fact, all of them do).

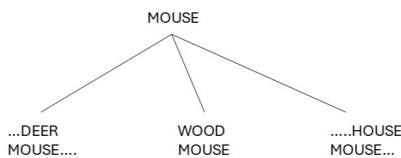


Figure 2. Taxonomic domain for (1a)

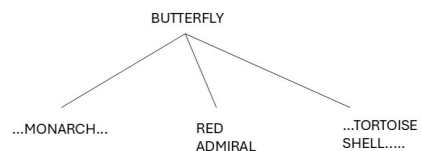


Figure 3. Taxonomic domain for (7b)

We also find a potential explanation for the gradience in oddness of the anaphoric interpretation in (1a), (7b) *vs.* (7c), where the latter sounds much worse. In (1a) and (7b), the lexical content within the highlighted *the*-phrase – *mouse* in (1a) and *butterfly* in (7b) – corresponds directly to the name of one of the members (i.e., the parent) within the taxonomic domain indicated in Figures 2 and 3 respectively. It is not therefore surprising that the preferred kind interpretation of the *the*-phrase in these instances is one that refers to the parent kind. On the other hand, in (7c) where the lexical content within the highlighted *the*-phrase is of the form *kind of butterfly* without any additional descriptive qualifications, no such advantage is forthcoming. In this case, the lexical content does not obviously apply to any one of the nodes within the hierarchy any better than others, which could be why (7c) sounds so much worse than (1a) or (7b).

Let us now consider the case of (11a) and (13). The highlighted *the*-phrases in these examples do not represent natural kinds. Nevertheless, a taxonomic contrast set, and thereby a taxonomic domain of interpretation, can be identified by the hearer. For (11a), the contrast set includes the kind of person who does sing in the shower in addition to the kind of person who does not. For (13), it includes the type of guy that has passion as well as the type of guy who does not. The assumption of these contrast sets allows us to anticipate the oddness of the highlighted *the*-phrases in (11a) and (13). The underspecified lexical content in these phrases accurately describes all elements within the contrast set in each case, thus failing to identify a unique kind.

In sum then, we have proposed here that the limited anaphoric potential of the English article *the* can be accounted for by adopting a denotation for the article as in (3), along with the assumption that the domain restriction s in kind-referring uses is restricted to the taxonomic domain.⁶ The resulting lack of flexibility in the domain with respect to which kind-referring uses of *the*-phrases are interpreted creates a “marked need for lexical content”, as Reed (2024) puts it. An underspecified NP is not guaranteed to uniquely describe an entity within the taxonomic domain. On the other hand, in individual-level uses, the domain restriction is more flexible and context-specific⁷.

Before closing this section, a brief word on how the proposal advocated here differs in its explanation of the data point in (14), reproduced below, from the one in Gundel (2010). Recall

⁶ A word on a possible alternative to the analysis for kind-denoting uses for *the* and what it means for our current proposal. In this article, we have more or less adopted without question the view in Dayal (2004): that *the* denotes identical meanings across its individual-referring and kind-referring uses, and it is the accompanying NP that is ambiguous between the two levels. An alternative possibility for the semantics of *the* in kind-level uses exists, however. In the alternative analysis, *the* is associated with the standard *iota* operator only in individual-referring uses, while kind-referring interpretations are obtained by way of the inherently intensional \cap operator from Chierchia (1998). The intensionality of the \cap operator means that its interpretation is necessarily not restricted to a singular domain or situation of interpretation. Rather, it returns the maximal entity satisfying the lexical content P within *any* situation. In this type of account, there is in some sense no need at all to induce a domain of interpretation, as opposed to the specific need under our proposal to induce a taxonomic domain. Under a domain-restriction based account, as it turns out, such a non-necessity of the domain turns out to be as problematic to anaphoric interpretation as restriction to the taxonomic domain. As the individual returned by \cap is domain-agnostic, hearers may not confine themselves to a particular domain (such as the narrow discourse domain) in interpreting this kind term. Consequently, they must rely on a fully specified lexical restriction to correctly interpret the intended kind referent. In this way, we believe that a domain-restriction based account serves the data in Section 2 better under either view of kind-referring uses of *the*, though we have only seriously considered one of them here.

⁷ While we do not address here the mechanism by which the domain of interpretation is inferred for individual-denoting uses of the definite article, this issue has been discussed extensively by other authors including Wolter (2006), Schwarz (2009), Srinivas (2021). We point the interested reader to these works.

that according to Gundel (2010), the oddness of the highlighted *the*-phrase in (14) is due to more than one potential referent fitting the descriptive content of the phrase. One of these referents (the *pareiasaur* kind) is available by virtue of prior mention, while the other (the whole class of primitive reptiles) is available due to the lexical descriptive content in the NP. Rephrased in terms of domains of interpretation, the story in Gundel (2010) may be taken as one in which the inferred domain in which the referent is identified could just as well be the narrow discourse context as the taxonomic domain.

- (13) A restudy of *pareiasaur*_k reveals that ??**the primitive reptiles**_k / **these primitive reptiles**_k are the nearest relatives of turtles.

Under the proposal advocated here, which incorporates the idea of a restricted taxonomic domain of interpretation in kind-denoting uses of *the*, the oddness of the anaphoric *the*-phrase in (14) receives a somewhat different explanation. The relevant taxonomic domain of interpretation assumed by a naive listener in this case may be taken to have at least the structure represented in Figure 4. Here as well, the oddness results from the presence of more than one possible alternative satisfying the lexical content of the *the*-phrase – only all potential referents are now located within the taxonomic domain

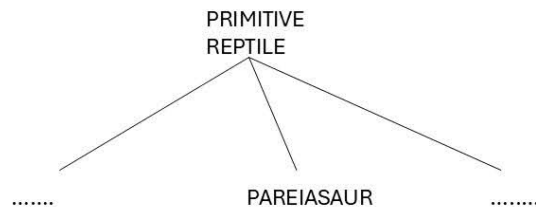


Figure 4. The relevant taxonomic domain for the interpretation of (11)

A starker divergence between the predictions of Gundel (2010) and the current proposal is obtained, as briefly discussed in Section 2, in examples that refer to non-natural kinds. In the case of (11a), for example, the only potential kind-level referent available under Gundel (2010) is the one that has been previously mentioned, as the NP *person* lacks an interpretation at the kind-level. It is therefore expected that anaphoric reference should be successfully achieved with the *the*-phrase. This expectation, however, is not borne out. Our current proposal, by contrast, makes the right prediction. Both elements in the contrast set making up the taxonomic hierarchy match the descriptor *person* or *kind of person* equally well, violating the uniqueness presupposition.

3.2. DEMONSTRATIVES ENCODE ANAPHORIC INDICES. The discussion in the preceding section addresses why kind-denoting uses of *the* exhibit limited anaphoric uses when compared to individual-denoting ones. We claimed this is because of the strict need to interpret kind-referring terms within the taxonomic domain, while individual referents may be identified in any contextually-determined domain of objects/entities. We are however yet to address why other types of determiners – such as the demonstratives *this* and *that* – do not give rise to a similarly limited anaphoric potential with kind-level NPs. In other words, how do the semantics of demonstrative determiners

differ from those of the definite article stated as in (3), so that a wider anaphoric distribution is supported with the former when compared to the latter?

Our answer to this is simple. We propose that the reason we do not observe any hinderances to anaphoric interpretations with demonstrative determiners is because they explicitly encode anaphoricity-based semantics, as in (4), repeated below:

$$(4) \quad \lambda P : \exists !x[P(x) \wedge x = g(i)] . \iota x[P(x) \wedge x = g(i)]$$

The denotation in (4) effectively dictates that the domain of interpretation be restricted to the discourse context built up within the preceding discourse moves, so that the intended referent must be identified from the set of previously indexed discourse referents.

This type of an analysis for demonstrative determiners is further consistent with recent ideas proposed in Ahn (2023). Using a wide range of data pertaining to demonstrative descriptions in English, Ahn concludes that demonstratives are binary operators that encode not only the lexical restriction carried by the complement NP but also an anaphoric “linker”. What is relevant for our purposes is that formal anaphoric indices – underlyingly construed as a deictic reference to a linguistic entity – are one way in which these linkers are realized. On the other hand, it is proposed that *the* is a unary operator, involving only lexical restriction (alongside contextual reasoning). We see the ideas presented in the current article as bolstering and empirically complementing a conclusion along these lines⁸.

4. Conclusion and future steps. This article takes the first step towards unpacking an interesting yet understudied observation made in Reed (2024), that kind-referring uses of the definite article *the* show starkly limited anaphoric potential in comparison to their individual-level occurrences, as well as to demonstrative determiners. Building on existing ideas surrounding the semantics of the definite article and their interaction with this central empirical observation, we have advocated for a *uniqueness*-based semantics for *the* that incorporates a specific view of domain restriction in kind uses – that they are always interpreted within the taxonomic domain. In essence, this relatively strict restriction imposed on the domain restriction in kind-denoting uses places the burden on the lexical content within the NP to uniquely identify the intended kind referent, resulting in what Reed describes as a “marked need for lexical content”. Crucially, we have argued that the contrast between the distribution of the definite article in individual-level *vs.* kind-level uses cannot be explained under a *familiarity*-based semantics that explicitly encodes an anaphoric index.

Much scope for further work remains, building on the ideas presented here. Most urgently, the empirical picture needs to be further fleshed out by way of controlled experimental studies, complementing the initial observations made in Reed (2016) and in the current article. Two issues, in particular, stand out from among the ones we have been unable to address here in depth. The first regards the modifier *kind of*. What exactly is its semantic contribution such that the presence of this modifier in many kind-level anaphoric contexts without any additional descriptive

⁸ The precise denotation that Ahn adopts for demonstratives is not, however, identical to (4). She proposes (17):

$$(17) \quad \lambda P . \lambda R . \iota x . \forall y [P(y) \wedge R(y) \leftrightarrow y \subseteq x]$$

Here P corresponds to the descriptive content within the NP. R is the linker restriction, which may correspond to an anaphoric index – in which case (17) is effectively equivalent to (4). Alternatively, R can encode a deictic pointing gesture, a relative clause restriction (*that which rolls*) or a prepositional restriction (*that in the corner*).

content leads to worse infelicity compared to its absence? A related set of questions pertains to larger differences between natural and non-natural kinds. For instance, are natural kinds more amenable to anaphoric interpretations than non-natural kinds? How do they differ, if at all, in terms of the process by which hearers infer the structure of the taxonomic domain during real-time comprehension?

References

- Aguilar-Guevara, Ana & Joost Zwarts. 2010. Weak definites and reference to kinds. *Semantics and Linguistic Theory (SALT)* 20. 179–196. <https://doi.org/10.3765/salt.v20i0.2583>.
- Ahn, Dorothy. 2023. Definite expressions with and without deixis. *West Coast Conference on Formal Linguistics (WCCFL)* 41.
- Barker, Chris. 2004. Possessive weak definites. *University of Massachusetts Occasional Papers* 29. 89.
- Barwise, Jon & John Perry. 1983. *Situations and attitudes*. Cambridge, MA: MIT Press.
- Carlson, Greg & Rachel Shirley Sussman. 2005. Seemingly indefinite definites. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 71–86. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110197549.71>.
- Chierchia, Gennaro. 1998. Reference to kinds across language. *Natural Language Semantics* 6(4). 339–405. <https://doi.org/10.1023/A:1008324218506>.
- Christophersen, Paul. 1939. *The articles: A study of their theory and use in English*. Copenhagen: Einar Munksgaard.
- Dayal, Veneeta. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27(4). 393–450. <https://doi.org/10.1023/B:LING.0000024420.80324.67>.
- Gundel, Jeanette. 2010. Reference and accessibility from a givenness hierarchy perspective. *International Review of Pragmatics* 2(2). 148–168. <https://doi.org/10.1163/187731010X528322>.
- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst: University of Massachusetts dissertation.
- Kamp, Hans. 1981. A theory of truth and semantic representation. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings*, 189–222. London: Wiley.
- Klein, Natalie M., Whitney M. Gegg-Harrison, Greg Carlson & Michael K. Tanenhaus. 2013. Experimental investigations of weak definite and weak indefinite noun phrases. *Cognition* 128(2). 187–213. <https://doi.org/10.1016/j.cognition.2013.03.007>.
- Löbner, Sebastian. 1985. Definites. *Journal of Semantics* 4(4). 279–326. <https://doi.org/10.1093/jos/4.4.279>.
- Reed, Ann. 2024. Procedural, underspecified, but good-enough, *the*. Manuscript. William and Mary.
- Roberts, Craige. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26(3). 287–350. <https://doi.org/10.1023/A:1024157132393>.
- Russell, Bertrand. 2005. On denoting. *Mind* 114(456). 873–887. <https://www.jstor.org/stable/3840617>.
- Schwarz, Florian. 2009. Two types of definites in natural language. Amherst: University of Massachusetts dissertation.
- Schwarz, Florian. 2014. How weak and how definite are weak definites. In Ana Aguilar-Guevara, Bert Le Bruyn & Joost Zwarts (eds.), *Weak referentiality*, 213–235. Amsterdam: John Benjamins.

- Srinivas, Sadhwi. 2021. *The semantics of (in) definiteness in bare vs. non-bare nominals: A study of Kannada and English*. Baltimore: Johns Hopkins University dissertation.
- Strawson, Peter F. 1950. On referring. *Mind* 59(235). 320–344.
- Wolter, Lynsey. 2006. *That's that: The semantics and pragmatics of demonstrative noun phrases*. Santa Cruz: University of California dissertation.