

Clarifying Noun Phrase Semantics

Matthew Purver, Jonathan Ginzburg

Department of Computer Science, King's College, London

Abstract

Reprise questions are a common dialogue device allowing a conversational participant to request clarification of the meaning intended by a speaker when uttering a word or phrase. As such they can act as semantic probes, providing us with information about what meaning can be associated with word and phrase types and thus helping to sharpen the principle of compositionality. This paper discusses the evidence provided by reprise questions concerning the meaning of nouns, noun phrases and determiners. Our central claim is that reprise questions strongly suggest that quantified noun phrases denote (situation-dependent) individuals – or sets of individuals – rather than sets of sets, or properties of properties. We outline a resulting analysis within the HPSG framework, and discuss its extension to such phenomena as quantifier scope, anaphora and monotone decreasing quantifiers.

1 Introduction

1.1 Overview

Reprise questions are a common dialogue device allowing a conversational participant to request clarification of some property of an utterance (or part thereof). In this paper we are concerned specifically with those reprise questions which concern the meaning intended by a speaker when uttering a word or phrase (see below). By virtue of this, they can provide us with information about what meaning can be associated with word and phrase types, and therefore provide useful extra evidence for the field of semantics – a domain overfull with theories underdetermined by evidence. We will suggest that this method provides a means of significantly sharpening the principle of compositionality; as well as ensuring that individual constituents combine to give suitable sentence meanings, it allows us to examine the meanings of those constituents directly.

This paper discusses the evidence provided by reprise questions concerning the semantics of common nouns (CNs), determiners and quantified noun phrases (QNPs), and outlines some general implications for NP semantics, together with some implications for semantic representation and inheritance in HPSG and other related underspecified representations. Our central claim is that reprise questions strongly suggest that QNPs denote (situation-dependent) individuals – or sets of individuals – rather than sets of sets, or properties of properties. We develop a witness-set-based analysis which treats all QNPs in a coherent manner, and allows an analysis of reprise questions via grounding. We also show how anaphora and quantifier scope can be accounted for within this analysis, via a view of NPs as functional, and show how non-monotone-increasing NPs can be represented.

1.2 Content of Reprise Questions

Ginzburg and Cooper (2001, 2004) (hereafter G&C) provide an analysis of proper name (PN) reprise questions which treats them as questions concerning the semantic content of the PN (which is taken to be a referential index, the intended referent of the name). In this way, a reprise such as example (1) can be taken to be paraphrasable as shown below. There are two distinct readings, but both concern the content of the PN *Bo*:

- (1) $\left\{ \begin{array}{l} \text{A: Did Bo leave?} \\ \text{B: } \mathbf{BO?} \\ \rightsquigarrow \text{ “Is it } \mathbf{BO}_i \text{ that you are asking whether } i \text{ left?”} \\ \rightsquigarrow \text{ “Who do you mean by ‘Bo’?”} \end{array} \right.$

G&C's analysis (given in section 2.1 below) applies only to PNs. However, it is clear that other nominal fragments¹ can be reprised, and our intention in this paper is to examine such reprises and, where possible,

¹And, indeed, fragments of other categories, but we leave these for future work.

propose a suitable extended analysis. It is also perhaps clear that not all nominal fragment reprises will involve querying a simple referential index as in (1): exactly what a reprise question can query is likely to vary depending on the nature of the fragment itself. However, if reprises² ask about the semantic content of the source fragment, then examining them can give some evidence about what goes to make up that semantic content.

1.2.1 Do Reprises Query Content?

One can imagine an argument that reprises can query any aspect associated with meaning, including perhaps pragmatic inferences, and that it might therefore be difficult to tease apart semantic from pragmatic readings. However, there is good reason to believe that while some reprises may be able to query some material of a pragmatic nature, queries about inferences in general (including implicatures and the like) are very difficult if not impossible to construct.

Pragmatic Readings Some clarification requests³ certainly seem to be able to query the whole intended speaker’s meaning, or even the overall relevance of the utterance to the discourse. In example (2), taken from a corpus study of clarification (Purver et al., 2003a), the question asked seems more about this relevance or intended meaning, than about the utterance’s semantic content or predicate-argument structure:

- (2)⁴ Sheila: ... when Michael’s in she knits him a jumper, the jumper <unclear> <pause>
 Wendy: Best that way then you don’t get sick
 Sheila: **Eh?**
 Wendy: It’ll be better that way if you, like you’re knitting with two different colours
 Sheila: Aye
- ~> “What do you mean by ‘Best that way then you don’t get sick?’”

This example is not a reprise, and it seems much more difficult for reprises to ask this sort of question (the corpus study found no such examples). However, it can be imagined for reprises of whole sentences – a reprise “*Best that way then you don’t get sick?*” in example (2) might serve the same purpose. It seems much harder for sub-utterance fragment reprises such as (1) to ask these sort of questions, though, and the corpus study did not reveal any even though reprise fragments in general are common.

Inferences However, this is a far cry from being able to query inferred pragmatic meaning in general. Examples involving implicatures suggest that it is very difficult for reprise questions to query pragmatically inferred content. It is certainly the case that A’s statement in the invented example (3), taken to be uttered outside a West End theatre currently showing a best-selling musical, could be inferred to be implicating other messages as shown:

- (3) A: I have a ticket for tonight’s performance.
 ~> “I am offering to sell a ticket for tonight’s performance.”
 ~> “Would you like to buy a ticket for tonight’s performance?”

But a reprise of the sentence does not seem to be able to be understood as querying these implicatures, but only the *directly conveyed semantic content*, as shown in (4).

- (4) A: I have a ticket for tonight’s performance.
 B: **You have a ticket for tonight’s performance?**
 ~> “Are you telling me you have *a ticket*?”
 ~> “What do you mean by ‘you have a ticket?’”
 ~> #“Are you offering to sell me a ticket?”
 ~> #“Are you asking if I want to buy a ticket?”

This may be even clearer when considering an answer to such a reprise question (5), which again can

²Or rather, *content readings* of reprises, rather than questions about phonological form – see section 2.1.3.

³A class of utterances (including, but not limited to, reprise questions) which ask about properties of a preceding utterance.

⁴BNC file KR0, sentences 362–366

only be construed as answering a question about this directly conveyed content (see Ginzburg et al., 2003, for a more detailed exposition):

- (5) A: I have a ticket for tonight's performance.
 B: **You have a ticket for tonight's performance?**
 A: Yes.
 ~> "Yes, I am indeed telling you I have a ticket."
 ~> # "Yes, I am indeed offering to sell you a ticket."
 ~> # "Yes, I am indeed asking if you want to buy a ticket."

Note that "Yes, but I'm not offering to sell it" would be perfectly acceptable. Similarly "No" must mean "No, I do not have a ticket", rather than "No, I'm not offering to sell a ticket (although I might have one)". Any inference that B is really asking about buying or selling activities therefore seems to be exactly that – an inference on top of the content of the reprise (a question about content of the original utterance), rather than because the reprise is itself a question about inferred material.

Empirical Evidence The corpus study mentioned above also showed that function words are very unlikely to be reprised. To confirm this, an experimental study (Healey et al., 2003; Purver et al., 2003b) has been carried out involving the introduction of simulated reprise questions into text-based dialogue. Subjects conducted natural conversations in pairs, and at certain points automatically-constructed artificial turns were injected into the conversation. These turns were only seen by one subject, and appeared to them to come from the other subject. The turns took the form of reprise questions: single words repeated from the previous turn. The response of the subject to these turns was then used to judge how they interpreted them: whether they interpreted them as reprises at all, and if so, what question they took them to ask.

Results showed that function word questions were indeed very difficult to interpret as reprises (only 1 of 42 examples was answered as if it was a reprise of the original function word), and apparently impossible to interpret as asking about meaning (the single example seemed to be answered as if it was question about word identity or orthographical form – see section 2.1.3). Instead, most were ignored or explicitly queried themselves. This was in stark contrast to content words (nouns and verbs) which were readily interpreted and answered as reprise questions about meaning.

If reprises ask about semantic content, and in particular some sort of contextually dependent reference, this makes sense. On the other hand, if they could be based on unrestricted contextual inferences, one might expect that such inferences (and resulting reprise readings) would be easily available even for function words, since these do give rise to generalized and particularized conversational implicatures. This expectation is not met.

So while some reprises can be seen as querying pragmatic material (such as overall relevance), they do not appear able to ask about unrestricted pragmatic inferences, and in most cases really do seem to query semantic content (particularly when querying fragments rather than whole utterances). We therefore take it that fragment reprises which appear to query semantic content (rather than, say, phonology – see section 2.1.3) really are doing so.

1.2.2 Strengthening Compositionality

Given this, it seems clear that if a question which reprises a particular phrase asks about a particular semantic object, then that object must be part of the semantic representation of that phrase. In other words, reprise questions must query at least *some part of the semantic content* of the fragment being reprised, and we take this as our basic hypothesis:

Reprise Content Hypothesis (weak version):

- (6) *A nominal fragment reprise question queries a part of the standard semantic content of the fragment being reprised.*

A stronger proposal might be that if a reprise question asks about a particular semantic object, then that object *is* the semantic content of the phrase being reprised:

Reprise Content Hypothesis (strong version):

- (7) *A nominal fragment reprise question queries exactly the standard semantic content of the fragment being reprised.*

While there is (and can be) no independent evidence that this stronger version holds, it is intuitively very attractive, as it provides us with a version of Occam's Razor: it requires that we do not postulate any part of a semantic representation which cannot be observed via a reprise question – in other words, that the semantic representations we do postulate are the simplest possible that can explain the readings of reprise questions. Throughout the paper, then, we will examine the consequences of both versions of this hypothesis for NP semantics, proposing representations which always hold to the weak version, and hold to the strong version wherever possible.

This hypothesis, in either version, provides us with an empirical criterion for assigning denotations that supports, but is stronger than, the usual criterion of compositionality. The standard requirement that the full content of an utterance (or sentence) emerges from the contents of its components often leaves underdetermined the question of which part contributes what. Instead, a semantics that can provide an adequate analysis of reprise questions by holding to the reprise content hypothesis is held responsible for the content it assigns not only to the complete utterance but to each component (or at least each reprisable and semantically potent component). A suitable semantics for NPs must not only allow full sentence content to be built, but be able to explain what it is about NPs that gives NP reprises the meanings that they appear to have. This fits with Montague's overall strategy in assigning a well-defined denotation to QNPs; but as we will see below, it seems to argue against his specific tactics of using higher-order properties-of-properties.

1.3 Corpus Evidence

We have used the British National Corpus (BNC – see Burnard, 2000) to investigate actual occurrences of reprise questions in dialogue. Questions were found using SCoRE (Purver, 2001), by searching for common reprise patterns (e.g. words repeated from the immediately preceding speaker turn). This method means that some examples will have been missed, but provides us with a lower bound: at least those examples that were found must be accounted for by a semantic theory.

The resulting examples were then classified according to possible and impossible paraphrases – we have of course had to construct these ourselves, but have made every effort to infer them not only from the questions themselves but from the dialogue context, particularly the responses of other conversational participants. Possible paraphrases are therefore those which we believe to be consistent with both the question and the context, and impossible ones those which would be inconsistent with either. This method may seem subjective, but is based upon the method used in (Purver et al., 2003a) to classify clarification questions (including reprises): this was shown to have good statistical reliability when the judgements of two independent markers were compared.

Our primary purpose in using a corpus is to provide as many examples as possible, in different situations, with different words and phrases (tokens as well as types) and with different speakers, in order to give us confidence that our claims about possible question readings are not influenced by our own choice of imagined examples. While reprise questions, or more accurately those that fit the patterns which we are able to search for, can be rare for some word/phrase types, the BNC is large enough (the dialogue portion comprises 740,000 speaker turns) to provide a few dozen occurrences for each of the phrase types we are most interested in here – that is, CNs and definite & indefinite NPs (exact numbers are given in the relevant sections below). While this quantity of data is small compared to the samples usually used for statistical studies, it fulfils our main requirement by providing a significant number of examples that must be covered by our analysis. It also provides enough data to ensure that the observed differences in reading distribution for these phrase types are statistically significant according to χ^2 tests, as detailed below.

However, even a corpus of this size yielded very few (< 10) examples of reprises of other classes: NPs with other quantifiers, and determiners. In the corresponding sections we therefore have to augment the sample using our intuition and invented examples, but we have indicated below where this is the case, and have not attempted to draw any conclusions based on statistical distributions or apparent negative evidence,

but only ensured that any observed examples are accounted for.

1.4 HPSG Notation

Our analysis, like G&C's, is based in HPSG (Pollard and Sag, 1994; Ginzburg and Sag, 2000). Although like G&C, we believe that the analysis is applicable to other frameworks, HPSG provides certain features that are advantageous for the analysis of reprise questions (in particular, direct access to phonological, syntactic, semantic and contextual information and the availability of constraints between these levels, and the ability to treat utterances as objects within the grammar).⁵ In an attempt both to save space and to make our examples more readable for those not familiar with HPSG attribute-value matrix (AVM) notation, we will avoid using AVMs wherever possible, and where we they are unavoidable will use some abbreviations throughout. These are shown in table 1.⁶

AVM	Abbreviation
$\left[\begin{array}{l} \textit{parameter} \\ \text{INDEX } x \\ \text{RESTR } \left\{ \left[\begin{array}{l} \text{INSTANCE } x \\ \text{PROPERTY } P \end{array} \right] \right\} \end{array} \right]$	$x : \textit{property}(x, P)$
$\left[\begin{array}{l} \textit{proposition} \\ \text{SOA} \mid \text{NUCLEUS} \left[\begin{array}{l} \textit{verb_rel} \\ \text{ROLE_1 } x \\ \text{ROLE_2 } y \end{array} \right] \end{array} \right]$	$\textit{verb}(x, y)$
$\left[\begin{array}{l} \textit{question} \\ \text{PARAMS } \{ \} \\ \text{PROP } \textit{verb}(x, y) \end{array} \right]$	$?.\textit{verb}(x, y)$
$\left[\begin{array}{l} \textit{question} \\ \text{PARAMS } \{ x : \textit{property}(x, P) \} \\ \text{PROP } \textit{verb}(x, y) \end{array} \right]$	$\begin{array}{c} ?x.\textit{verb}(x, y) \\ \text{or} \\ ?x : \textit{property}(x, P).\textit{verb}(x, y) \end{array}$

Table 1: HPSG AVM Abbreviations

In the next section we give some background on firstly, G&C's analysis of reprise questions, and secondly, traditional views of QNP semantics. The subsequent sections 3 and 4 discuss the content of reprise questions for CNs and QNPs together with a corresponding semantic analysis, and some further issues arising from this are discussed in section 5.

⁵For an alternative formulation of some of G&C's account within Martin-Löf Type Theory, see (Cooper and Ginzburg, 2002). (Poesio and Traum, 1997) also provide a DRT-based framework which includes utterance reference.

⁶As shown in table 1, questions are viewed as abstracts, with a set of queried parameters PARAMS simultaneously abstracted from a propositional body PROP. Polar yes/no questions have empty PARAMS sets. For more detail see (Ginzburg and Sag, 2000).

2 Background

2.1 Reprise Questions

G&C, following on from Ginzburg and Sag (2000), provide an analysis of reprise questions together with a method of resolution of associated elliptical forms. This analysis is couched within a HPSG grammar and a Question-Under-Discussion (QUD) approach to dialogue context (e.g. Ginzburg, 1996; Larsson, 2002).

2.1.1 Contextual Parameters

Standard versions of HPSG directly encode idealized semantic content (that which a speaker would be expected to associate with a sign) within the value for the `CONTENT` feature. Instead, G&C propose a representation which expresses contextual dependence, one which encodes *meaning* rather than *content*: a function from context to fully specified content. In their HPSG terms, contextually dependent parameters such as speaker, hearer, utterance time and (crucially to the reprise analysis) the reference of PNs are abstracted to a set which is the value of a new `C-PARAMS` feature, as shown in AVM (8) for A's original utterance in example (1):⁷

$$(8) \left[\begin{array}{l} \text{CONTENT} \\ \text{C-PARAMS} \end{array} \left[\begin{array}{l} \text{ask}(\boxed{1}, \boxed{2}, ?.\text{leave}(\boxed{3})) \\ \left\{ \boxed{1}[a : \text{speaker}(a)], \boxed{2}[b : \text{addressee}(b)], \boxed{3}[x : \text{name}(x, Bo)] \right\} \end{array} \right] \right]$$

Such representations of meaning can be viewed as λ -abstracts, with the members of `C-PARAMS` simultaneously abstracted over the standard value of `CONTENT`. More specifically, they are interpreted as *simultaneous abstracts with restriction* as shown in (9): $\{ABS\}$ is the set of abstracted indices, $[RESTR]$ a set of restrictions which must be satisfied during application, and $BODY$ the body of the abstract (in this case, the semantic content). For further formal details, see (Ginzburg and Sag, 2000).

$$(9) \lambda\{ABS\}[RESTR].BODY$$

AVM (8) can therefore be rewritten as in (10), or, simplifying even further by omitting the parameters associated with speaker and addressee, as in (11). Wherever possible, we will use these equivalent λ -abstract expressions for readability's sake.

$$(10) \lambda\{a, b, x\}[\text{speaker}(a), \text{addressee}(b), \text{name}(x, Bo)].\text{ask}(a, b, ?.\text{leave}(x))$$

$$(11) \lambda\{x\}[\text{name}(x, Bo)].\text{ask}(a, b, ?.\text{leave}(x))$$

These utterance-level representations are built up compositionally⁸ by the grammar. Lexical items such as PNs are defined to introduce abstracted parameters in `C-PARAMS` – the word *Bo* is given the representation below:

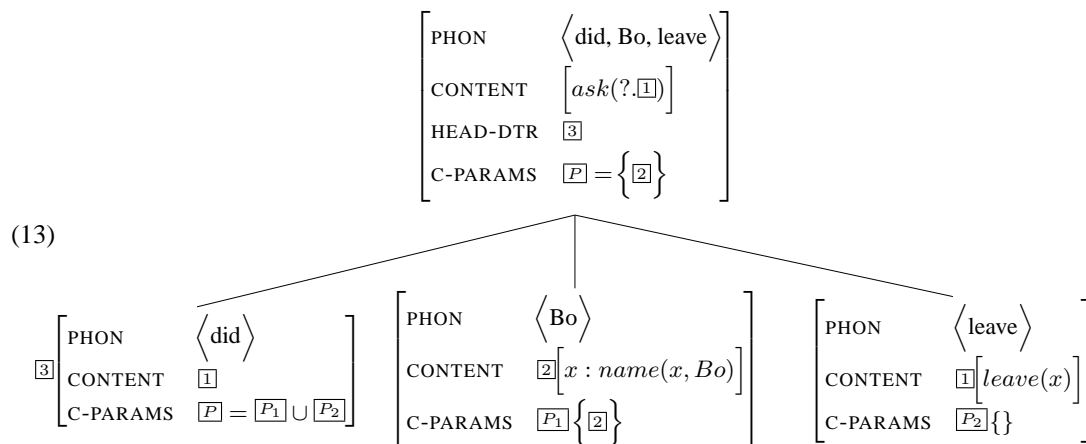
$$(12) \lambda\{x\}[\text{name}(x, Bo)].x$$

These parameters are then inherited via a `C-PARAMS amalgamation principle`: the value of `C-PARAMS` for lexical heads is defined to be the set union of the values of its syntactic sisters, and this is inherited up via heads to the sentence level. This gives the correctly contextually dependent meaning for the whole

⁷Note also that the semantic representation includes the conversational move type *ask*, following Ginzburg et al. (2003) – this is important in order to give the correct interpretation for *clausal* questions (see below).

⁸The grammar uses various *constructions* which define how meaning is built up from constituent parts: this may not be consistent with some strict definitions of the principle of compositionality, but is compositional according to definitions such as that of (Pelletier, 2003) – the grammar gives a principled procedure for establishing utterance meanings given lexical items and their syntactic mode of combination.

utterance, as shown in (13) for the sentence of example (1).



2.1.2 Grounding and Reprises

The grounding process for an addressee can now be modelled as an application of this meaning abstract to the context, establishing the referents of the abstracted parameters such that their given restrictions are satisfied, and resulting in the full fixed semantic content. It is failure to do this for a particular parameter that results in the formation of a clarification question, with the purpose of querying the sub-utterance which contributed that parameter. Failure may be due to, say, the lack of an available referent in context (e.g. no known person named *Bo*), the lack of a unique most salient referent (e.g. two equally salient people named *Bo*), or an available referent which is problematic in some way (e.g. leading to inconsistency in the resulting content).

On this view, then, clarification questions are triggered by parameters which have been abstracted from content. Our basic reprise content hypothesis must therefore require that fragments which are reprised must have contributed part of their content to the abstracted set, as in (14); the strong version will require that they contribute their entire content, as in (15):

$$(14) \lambda\{x\}[\dots].\phi(\dots, x, \dots)$$

$$(15) \lambda\{x\}[\dots].x$$

The resulting clarification question can take many forms: direct non-reprise questions (“*Who do you mean by ‘Bo’?*”) and reprise sluices (“*WHO?*”) are possible, as well as the elliptical reprise fragments that we concentrate on in this paper. G&C give a QUD-based analysis of how their content is derived in context: a conversational participant’s basic dialogue competence includes certain specific contextual update tools or *coercion operations*, which take the utterance being clarified as their input and produce a partially updated context where this utterance is salient and the maximal QUD is a suitable clarification question. Two such operations are possible, as detailed here.

Clausal Readings In the case when a hearer finds a problematic value for a contextual parameter, the question that arises is a *clausal* question, a *polar* (yes/no) question about the parameter’s intended referent,

corresponding to the first of the paraphrases given in example (1) above or to that given in example (16):

- (16) A: Did Bo leave?
 B: **Bo?** / **Bo Smith?**
 A: That's right.
 B: Yes, half an hour ago.
- ↪ “*Is it Bo_x / $Bo\ Smith_x$ that you are asking whether x left?*”

As shown, reprises with clausal readings can repeat the original phrase verbatim (“*Bo?*”) or can use another apparently co-referring phrase (“*Bo Smith?*”). We will call verbatim repeats *direct echoes*.

The coercion operation for these readings produces an updated context where the maximal QUD is the question formed by abstracting the problematic parameter from the original intended content (in this case this resulting QUD is $\{x : name(x, Bo)\}.ask(a, b, ?.leave(x))$, paraphrasable as “*For which Bo_x are you asking whether x left?*”).

This context then allows the bare fragment “*Bo?*” to be resolved as having the content $?.ask(a, b, ?.leave(x))$ (paraphrasable as in example (16) above). Similarly a reprise sluice (a bare *wh*-phrase) “*Who?*” would simply be resolved as having the new QUD as its content.

Constituent Readings In the case where the hearer can find no value for a parameter in context, the question that arises is a *constituent* question, a *wh*-question about the intended content of the problematic utterance, corresponding to the second paraphrase in example (1) or to that given here as example (17).

- (17) A: Did Bo leave?
 B: **BO?**
 A: Bo Smith.
 B: Yes, half an hour ago.
- ↪ “*What is the intended content of your utterance ‘Bo’?*”

For this reading, the coercion operation results in an updated context where the maximal QUD is precisely this question about the content x of a sub-utterance intended by the speaker a , written as $?.spkr_meaning_rel(a, 'Bo', x)$. In this case the elliptical question “*Bo?*” (delivered with suitable intonation) can be resolved as having this question as its content.

While the clausal and constituent readings are distinct, they both involve *querying the meaning* of the relevant sub-utterance, following an inability to find a suitable referent which resolves that meaning in the hearer’s context. It is this property that allows us to use them to investigate what meaning can be attributed to various word and phrase types. Note that the coercion operations described above are mechanisms for updating context so that the elliptical reprise question can have its content resolved, rather than operations on the content of the original sub-utterance being reprised.

2.1.3 Form Identification Readings

As pointed out by G&C, and in more detail by Purver et al. (2003a), reprise questions may have other possible readings apart from the two described above. These possibilities seem to be limited: as already discussed in section 1.2.1, it is not the case that reprise readings can be based on unrestricted contextual inference. But in particular, a *form identification* reading⁹ concerning phonology or orthography of the words used by the speaker is, arguably, available (for example, in situations with high background noise

⁹This reading is variously referred to as a *lexical* reading by Purver et al. (2003a) and a *lexical identification* reading by G&C.

levels).

- (18)¹⁰ Matthew: It wasn't all that bad. At least the pool was clean.
 Lara: **Mr Pool?**
 Matthew: The pool.
 Lara: Oh <laugh>.
- ~> "Did you say the words 'Mr Pool'?"

Whether such readings actually exist or not, we are not concerned with them in this paper as, lacking any reference to word or phrase meaning, they do not shed any light on semantics. When we refer to reprise questions hereafter, this should be taken as referring to content readings only (the clausal and constituent readings described above).

2.2 QNP Semantics

The semantic representation of QNPs has of course been a subject of lively debate for some time, and we cannot hope to do justice to the field here; instead we point out the main differences in currently popular views in the areas on which we hope that the study of reprise questions can shed some light.

2.2.1 The Quantificational View

One view, dating back at least to Russell (1905), holds that QNPs contribute quantificational terms to the semantic representation of a sentence. This is exemplified by Montague (1974)'s PTQ, in which sentences containing QNPs are given representations as follows:

- (19) "every dog snores" $\mapsto \forall x(\text{dog}(x) \rightarrow \text{snore}(x))$

On this view, QNPs therefore denote functions from properties of individuals ($e \rightarrow t$) to truth values (t) (in other words, they are properties of properties ($(e \rightarrow t) \rightarrow t$): The content of a QNP is defined by the properties that hold of some referent contained in it (in the case of "every dog", all those properties which are true or untrue of every dog).

- (20) "every dog" $\mapsto \lambda P.\forall x(\text{dog}(x) \rightarrow P(x))$

Those who adhere strictly to this view take it also to hold for definite descriptions: definites are not considered to be directly referential in the same sense as PNs, but are seen as defined by existential quantification with a uniqueness constraint.

- (21) "the dog" $\mapsto \lambda P.\exists x(\text{dog}(x) \wedge \forall y(\text{dog}(y) \rightarrow y = x) \wedge P(x))$

2.2.2 The Referential View

An alternative view originating with Strawson (1950) and Donnellan (1966) is that some NPs, in particular definites, can be directly referential. Donnellan pointed out that while the Russellian approach covered *attributive* uses well (those described by Russell as "known by description"), it did not appear to cover *referential* uses. Others (e.g. Fodor and Sag, 1982) have also pointed out that indefinites can be used *specifically* (the speaker has a specific individual in mind, although the hearer is not expected to be able to identify it) and *definitely* (expected to be identified by the hearer)¹¹, and that these uses also do not appear to fit with a purely quantificational analysis.

On the quantificational view, this apparently referential nature is argued to follow from pragmatic principles rather than any true semantic reference. This argument originates with Kripke (1977), and a

¹⁰BNC file KPP, sentences 321–325

¹¹A good summary of these terms, with examples, is available in (Ludlow and Neale, 1991).

concise statement is given by Ludlow and Neale (1991) and Ludlow and Segal (2004). Essentially it runs as follows (omitting some steps for brevity's sake here):

1. S has expressed a quantified proposition $\tau x.F(x) \wedge P(x)$.
2. S could not be doing this unless she thought that $P(b)$ where b is some referent.
3. S knows and I know that $b = \tau x.F(x)$
4. Therefore S has implicated that $P(b)$.

Other approaches such as the dynamic theories of Heim (1982), Kamp and Reyle (1993) and possibly Groenendijk and Stokhof (1991) might be said to fall somewhere in between the two camps, with definites having some kind of reference (although this may be to a contextual discourse referent rather than a real-world object). In most views, however, NPs with other quantifiers (*every*, *most* etc.) are seen as quantificational.

2.2.3 Generalized Quantifiers and Witness Sets

The theory of Generalized Quantifiers (GQs) (see Barwise and Cooper, 1981)¹² (hereafter B&C) has been applied to the quantificational view, both to extend the Russellian approach to other natural language quantifiers, and to allow semantics of the QNP constituent to be represented more transparently in the sentence representation:

$$(22) \text{ "every dog" } \mapsto \text{every}(\text{DOG}) \quad \text{where} \quad \llbracket \text{every}(\text{DOG}) \rrbracket = \{X \mid \text{DOG} \subseteq X\}$$

$$(23) \text{ "every dog snores" } \mapsto \text{every}(\text{DOG})(\text{SNORE})$$

$$\text{where} \quad \llbracket \text{every}(\text{DOG})(\text{SNORE}) \rrbracket = \text{SNORE} \in \llbracket \text{every}(\text{DOG}) \rrbracket \\ = \text{SNORE} \in \{X \mid \text{DOG} \subseteq X\} = \text{DOG} \subseteq \text{SNORE}$$

Essentially the quantificational view of QNPs still holds: QNPs are GQs, and as such denote a family of sets (a set of sets, here the set of those sets which contain *DOG*, the set of dogs), rather than being directly referential.

To explain how a hearer can process a GQ without having to determine the identity of this full set of sets, B&C introduce the notion of a *witness set*. For a GQ $D(A)$, this is defined as being any set w which is both a subset of A and a member of $D(A)$. For an indefinite *a dog*, w can be any nonempty set of dogs; for a definite *the dog*, w must be the set containing exactly the contextually unique dog; for the universal *every dog*, w must be equal to the set of all dogs. For monotone increasing (MON \uparrow) quantifiers, the following equivalence holds:

$$(24) \exists w[w \subseteq X] \leftrightarrow X \in D(A)$$

In other words, showing that a predicate X holds of a witness set is equivalent to showing that the corresponding GQ holds of the predicate. We will use this notion heavily below.

In the next section we begin by examining CN reprise questions, and show that G&C's analysis can be extended to account for their apparent meaning in a manner consistent with the traditional view of CN semantics. In section 4 we then discuss QNP reprise questions, and show that their meaning can be more naturally accounted for by the referential view of QNP semantics. Section 5 then discusses some issues raised by the view put forward in section 4.

3 Common Nouns

In this section we examine CN reprise questions, and show that their meaning appears to be entirely consistent with the standard semantic view of CNs as denoting properties of individuals, and with our hypothesis that reprise questions concern the semantic content of the fragment being reprised.

¹²But see also e.g. (Keenan and Stavi, 1986; Keenan and Westerståhl, 1997; van der Does and van Eijck, 1996).

	Pattern	Referent Reading	Predicate Reading
CN Examples	"... <i>DET N</i> ... " / " <i>N</i> ?"	-	58
		-	100%

Table 2: Literal Reprises – CNs

3.1 Nouns as Properties

The semantic content of CNs is traditionally viewed as being a property (of individuals). Montague (1974) expressed this as a λ -abstract, a function from individuals to truth values (e.g. $\lambda x.dog(x)$), and this view is essentially shared by most strands of formal semantics. Variations (especially in representation) certainly exist: in situation semantics (Barwise and Perry, 1983) this might be expressed as a λ -abstracted infon (see Cooper, 1995), in DRT (Kamp and Reyle, 1993) as a predicative DRS (see Asher, 1993), but these approaches share the basic view that CNs are properties of individuals.

Given this, we would expect CN reprise questions to be able to query the property expressed by the noun, and this property only, when the hearer cannot identify this property in context. The clausal and constituent readings may both still be available, but the noun property or predicate should always be the element under question:

Clausal reading: "*Is it the property P about which you are asking/asserting ... P ... ?*"

Constituent reading: "*What is the property P which you intend to be conveyed by the word N?*"

In contrast, it should not be possible for CN-only reprises to be interpreted as questions about e.g. individual referents.

For mass nouns and bare plurals, the picture may not be so simple: these might be expected to refer instead to *kinds* (see e.g. Carlson, 1977; Chierchia, 1998), or in the case of plurals, behave as indefinites (Kamp and Reyle, 1993). We examine both below in sections 3.4 and 3.5.

3.2 Corpus Evidence

Reprises of CNs were identified in the corpus by searching for single-word CN questions where the word is repeated verbatim from the previous speaker turn. To rule out bare mass nouns and plurals, which are discussed separately in sections 3.4 and 3.5, examples were restricted to cases in which the original occurrence of the CN in the previous turn was singular and preceded by a determiner. All examples found confirmed the expectation: as Table 2 shows, a predicate reading seems to be the only interpretation.

Examples are given here together with what appear to be possible and impossible paraphrases:

- (25)¹³
- | | |
|---------|--|
| Monica: | You pikey! Typical! |
| Andy: | Pikey? |
| Nick: | Pikey! |
| Andy: | What's pikey? What does pikey mean? |
| Monica: | I dunno. Crusty. |
| ~> | "Are you saying I am a pikey ?" |
| ~> | "What property do you mean by the word 'pikey'?" |
| ~> | #"Which pikey are you saying I am?" |

¹³BNC file KPR, sentences 218–225

The same appears to be true when the CN reprised forms part of an indefinite NP:

- (26)¹⁴
- | | |
|---------|---|
| Emma: | Got a comb anywhere? |
| Helena: | Comb? |
| Emma: | Even if it's one of those <pause> tremmy [<i>sic</i>] pretend combs you get with a Barbie doll, oh this'll do! <pause> Don't know what it is, but it'll do! |
| ~ | "Is it a comb that you are asking if I've got?" |
| ~ | #"Which comb are you are asking if I've got?" |

And indeed even when the CN is part of a seemingly referential definite NP:

- (27)¹⁵
- | | |
|--------|---|
| Carol: | We'll get the turkey out of the oven. |
| Emma: | Turkey? |
| Carol: | Well it's <pause> it's <pause> er <pause> what's his name?
Bernard Matthews' turkey roast. |
| Emma: | Oh it's looks horrible! |
| ~ | "Are you saying the thing we'll get out is a turkey ?" |
| ~ | "What concept/property do you mean by 'turkey'?" |
| ~ | #"Which turkey are you saying we'll get out?" |
| ~ | #"Is it this/tha turkey you're saying we'll get out?" |

Note that paraphrases which concern an intended referent of the NP containing the CN (e.g. the "Which X..." paraphrases) do not appear to be available, even when the NP might appear to be referential (see example (27)).

3.3 Analysis

As expected, we therefore suppose that the semantic representation of a CN must consist at least partially (and, if we are to hold to our strong hypothesis, solely) of a property of individuals.

An analysis entirely parallel to that of section 2.1 is possible if properties of individuals (which we shall refer to here as *predicates*) are regarded as possible cognitive or contextual referents: that is to say, as entities that must be identified in context.¹⁶ The predicate content of a noun can then be contextually abstracted by being made a member of C-PARAMS; this means it must be grounded by the hearer (by finding the intended predicate referent given its name) or made the subject of a clarification question in case this grounding process fails.¹⁷ Noun content therefore becomes contextually dependent, rather than *a priori* given, as we require for a treatment of clarification.¹⁸

We therefore propose a representation of CNs in which the content (and the sole abstracted parameter) is a parameter whose INDEX value is a named property of individuals:

$$(28) \lambda\{P\}[name(P, dog)].P$$

¹⁴BNC file KCE, sentences 1513–1516

¹⁵BNC file KBJ, sentences 131–135

¹⁶Whether these entities are best taken in a model-theoretic sense to denote atomic concepts (Barwise and Perry, 1983) or sets of individuals (Montague, 1974) is an interesting question in itself, but not one that impacts on the basic analysis here.

¹⁷It may fail for various reasons: with lexically ambiguous words, more than one property with this name will exist; with unknown words, no known property may be found in context; in other cases the hearer may find the apparently intended predicate surprising or impossible.

¹⁸This fact also perhaps offers a way to account for the psycholinguistically observable fact that conversational participants can have different understandings of the predicate being conveyed, and can indeed establish their own agreed meanings (see e.g. Pickering and Garrod, 2004).

3.3.1 Comparison with Standard Approaches

This may seem uncontentious, but note that it does not correspond to the treatment of CNs by standard HPSG approaches to semantics. In the common unification-based approach (Sag and Wasow, 1999; Ginzburg and Sag, 2000), CN content is identified with that of the NP mother, and thus taken to be a parameter whose referent is an individual (the NP referent). Abstracting this parameter to C-PARAMS, as shown in (29), would not give the correct reading for a clarification question, as this individual would become the referent to be grounded and thus the subject of the question (which we have seen is impossible). Avoiding this problem by abstracting only the relevant predicate rather than the entire content, as suggested by (Purver, 2002) and shown in (30), would be possible but no longer holds to the strong hypothesis: as a result, clarification questions would not be able to query the entire semantic content, and we would be left with no explanation as to why not.

$$(29) \lambda\{x\}[dog(x)].x$$

$$(30) \lambda\{P\}[name(P, dog)].x : P(x)$$

Similar problems apply to approaches such as Minimal Recursion Semantics (Copestake et al., 1999) in which the content of a NP mother is constructed by set union (*amalgamation*) over the content of its daughters (sets of *elementary predications*, simple pieces of propositional information). This again results in CN content including the individual referent of the mother NP.

The predicate analysis proposed above seems preferable, as it holds to the strong hypothesis and thus explains why only the observed predicate reading of a reprise question is available. As discussed in section 4.5 below, this has implications for the usual inheritance and amalgamation principles used in HPSG.

3.4 Bare Singulars

As mentioned above, bare singular mass nouns might be expected to refer to kinds or concepts, but again not to individual referents. And again, this did appear to be the case. All reprises of bare singular CNs (i.e. singular CNs where the CN in the original utterance being clarified had no determiner) seemed to fit with this (see table 3).

- | | | |
|--------------------|----------|--|
| (31) ¹⁹ | Richard: | because Donna is high in admir- admiration in fact I |
| | Anon 4: | Admiration? |
| | Richard: | I admire |
| | Anon 4: | I think it's called infatuation |
| | ~ | “Is it the property/concept admiration you're saying Donna is high in?” |
| | ~ | “What property/concept/kind do you mean by ‘admiration’?” |
| (32) ²⁰ | Iris: | Oh you should see <pause> see it! <pause> It has only been <pause> burning coal in it! |
| | Gordon: | Coal? |
| | Iris: | And it's all burnt, it's burnt all the skirting board and er |
| | Gordon: | Good God! |
| | ~ | “Is it the concept/kind/substance coal you're saying was burning?” |
| | ~ | “What concept/kind/substance do you mean by ‘coal’?” |
| | ~ | #“Which individual bits of coal are you saying were burning?” |

Note that we have not attempted to distinguish between concepts, kinds and the properties or predicates discussed above, as this level of distinction does not seem possible from our imputed paraphrases – what is clear is that these sort of paraphrases always seem acceptable.

¹⁹BNC file KSV, sentences 5869–5874

²⁰BNC file KCF, sentences 1573–1577

	Pattern	Referent Reading	Relation Reading	Predicate/Kind Reading
Bare Singular	“... <i>N</i> ...” / “ <i>N</i> ?”	-	-	41
		-	-	100%
Bare Plural	“... <i>Ns</i> ...” / “ <i>Ns</i> ?”	2	1	26
		7%	3%	90%

Table 3: Literal Reprises – Bare CNs

The analysis of mass nouns can therefore take exactly the same form as that for other CNs given above, with the semantic content being a property or kind which must be identified in context:

$$(33) \lambda\{P\}[name(P, admiration)].P$$

3.5 Bare Plurals

With bare plurals, the situation was more complex. Most examples found did seem to follow the same lines, with a property or kind reading being preferred, and often being the only possible reading (see example (34)).

- (34)²¹ John: Now I would like you to tell me about numbers.
 Simon: **Numbers?**
 John: Mhm. What are they?
 Simon: **Numbers** <laugh> erm <pause>
 John: What do we use them for?
- ~> “Is it the property **numbers** you’re saying I should tell you about things with?”
 ~> “Is it the concept/kind **numbers** you’re saying I should tell you about?”
 ~> “Which kind of numbers are you saying I should tell you about?”
 ~> #“Which actual numbers are you saying I should tell you about?”

However, some examples afforded a possible individual referent reading (see example (35)), and one example was best read as querying the plurality relation itself (example (36)).

- (35)²² Dorothy: Anyway, you were telling me about <pause> meals.
 Andrew: **Meals?**
 Dorothy: Mm.
 Andrew: What <unclear>?
 Dorothy: At Pontepool.
- ~> “Which meals are you saying I was telling you about?”
 ~> “Which property/concept do you mean by ‘meals’?”
 ~> ?“Is it the property **meals** you’re saying I was telling you about things with?”

²¹BNC file FMF, sentences 591–596

²²BNC file KBW, sentences 1247–1251

- (36)²³ William: You two
 Unknown: <unclear>
 William: hours ago
 Clare: <laugh> <pause> **Hours?**
 William: Well an hour
 Unknown: <unclear>
 Kim: it wasn't hours
- ~> *"Is it really **more than one** hour ago you're telling me it was?"*

As we will see in section 4.2 below, these are exactly the readings that seem to be available for indefinite NPs (a predicate reading, a logical determiner relation reading, and a (rarer) individual referent reading). This therefore suggests that bare plurals could be represented as indefinites (and we leave the details of this representation to section 4.2). However, as some examples seemed to *only* allow a property/kind reading (e.g. example (34) above), it may be that (as assumed by Kamp and Reyle, 1993) they should be seen as ambiguous between indefinites and kinds.

3.6 Summary

In this section we have presented evidence that shows that CN reprise questions concern a predicate. We have interpreted this as consistent with the common view that CNs denote properties of individuals, and as supporting our hypothesis that reprise questions concern the semantic content of the fragment being reprised.

We have shown how an extension of G&C's contextual abstraction approach allows a corresponding analysis which holds to the strong version of our reprise content hypothesis. We have also noted that standard HPSG analyses are not entirely consistent with the view of CNs as denoting predicates, and therefore would allow only the weaker version of the hypothesis to hold.

Examination of bare singular and plural CNs shows that mass nouns can be represented in a similar way (as denoting properties or kinds), but that some bare plurals must be represented differently, as individual referent reprise questions are possible.

In the next section we examine reprises of QNPs.

4 Noun Phrases

If we hold to the quantificational view of NP semantics, we should find that reprise questions concern a family of properties/sets (those properties which hold of the referent of the QNP). A referential view might instead lead us to expect that reprises of referential definites & specific indefinites should concern the individual referents directly.

4.1 Definite NPs

Taking a referential semantic viewpoint, we might therefore expect reprises of definite NPs to be paraphrasable as follows:

Clausal reading: *"Is it the individual X about which you are asking/asserting ... X ... ?"*

Constituent reading: *"Which individual X do you intend to be referred to by the phrase NP?"*

From a quantificational viewpoint, a paraphrase concerning a set of properties or sets might perhaps be expected:

Clausal reading: *"Is it the set of properties that hold of X about which you are asking/asserting ... X ... ?"*

²³BNC file KBN, sentences 1367–1371

Constituent reading: “Which set of properties do you intend to be conveyed by the phrase NP?”

Our corpus investigation included many types of definite NP: PNs, pronouns and demonstratives as well as definite descriptions. PNs have already been discussed in section 2.1 above – we examine the others here. An overview of results is shown in table 4.

4.1.1 Demonstratives and Pronouns

Perhaps unsurprisingly (many of those who hold to the quantificational view believe demonstratives to be directly referential), our corpus investigation shows that demonstratives license the referential readings, not only via direct echoes as in example (37), but also when reprised with a co-referring PN as in example (38), or with a reprise sluice as in example (39). Both clausal and constituent versions seem available.

- (37)²⁴ | John: Which way’s North, do you know?
 Sara: That way.
 John: **That way?** Okay.
 ~> “Are you telling me **that way there** is North?”
 ~> “By ‘that way’ do you mean that way there?”

- (38)²⁵ | Christopher: What was that lady <pause> <unclear>?
 Dorothy: **Julie?**
 Christopher: Mm.
 Dorothy: She’s been with you, hasn’t she?
 ~> “By ‘that lady’ do you mean Julie?”

- (39)²⁶ | Anon 1: Oh God I hate these lot, they’re so boring.
 Cassie: **What lot?**
 Anon 1: Them!
 Cassie: **Who? What them lot?**
 ~> “What lot are you telling me you hate?”
 ~> “What lot do you mean by ‘these lot’?”

The same also appears to hold for pronouns, although we discuss these in more detail in section 5.4 below:

- (40)²⁷ | Joanne: It’s, how many times did he spew up the stairs?
 Emma: **Julian?** Couple of times.
 ~> “Is it **Julian**_i that you are asking how many times i spewed up the stairs?”
 ~> “By ‘he’ do you mean Julian?”

However, when we look at definite descriptions, the situation appears more complex: while referential readings are common, others are possible which do not appear to be directly referential.

²⁴BNC file JP4, sentences 755–758

²⁵BNC file KBW, sentences 883–886

²⁶BNC file KP4, sentences 1546–1550

²⁷BNC file KCE, sentences 4190–4192

	Pattern	Referent	Functional	CN Predicate
Definite	"... <i>the N</i> ..." / " <i>The N</i> ?"	10 56%	6 33%	2 11%
Indefinite	"... <i>a(n) N</i> ..." / " <i>A(n) N</i> ?"	- -	- -	28 100%

Table 4: Literal Reprises – NPs

4.1.2 Definite Descriptions – Referential Readings

With definite descriptions, over half of the examples of direct echo questions found seemed to query the individual(s) being referred to.²⁸ Examples include constituent readings as in example (41) and clausal readings as in example (42):

- (41)²⁹ George: You want to tell them, bring the tourist around show them the spot
 Sam: **The spot?**
 George: where you spilled your blood
- ~> "Which spot are you referring to by 'the spot'?"

- (42)³⁰ John: they'll be working on the, they'll be working on the kidnapper's instructions though wouldn't they? They would be working on the kidnapper's instructions, the police?
 Sid: **The police?**
 John: Aye
 Sid: On
 Unknowns: <unclear>
 Sid: aye the, the senior detectives
- ~> "Is it **the police** who you are saying would be working ... ?"
 (~> "Who do you mean by 'the police'?"")

Reprises using PNs As with demonstratives, definite descriptions can be reprised with another NP that conveys the same desired referent:

- (43)³¹ Unknown: And er they X-rayed me, and took a urine sample, took a blood sample. Er, the doctor
 Unknown: **Chorlton?**
 Unknown: Chorlton, mhm, he examined me, erm, he, he said now they were on about a slide <unclear> on my heart. Mhm, he couldn't find it.
- ~> "By 'the doctor' do you mean Chorlton?"

This is interesting: not only does it give further weight to the idea that these reprises are genuinely referential (PNs are generally held to be referential even by those who hold to the quantificational view of definite NPs), it also suggests that the referent can be an entity in the world (rather than some kind of discourse object).

²⁸Comparison of the data in tables 2 and 4 shows that the reading distributions for definites and CNs are significantly different: a $\chi^2_{(1)}$ test shows that the probability p that the referent/predicate reading distribution is *independent* of whether the source is a definite NP or a CN is tiny ($p < 0.01\%$). The difference between the distributions for definites and indefinites is similarly significant ($p < 0.01\%$). There is no significant difference between indefinites and CNs, however, as discussed in section 4.2.

²⁹BNC file KDU, sentences 728–730

³⁰BNC file KCS, sentences 660–665

³¹BNC file KPY, sentences 1005–1008

Sluices And again, reprise sluices are available which seem to concern a referent:

- (44)³²
- | | |
|--------|---|
| Terry: | Richard hit the ball on the car. |
| Nick: | What car? |
| Terry: | The car that was going past. |
| Nick: | What ball? |
| Terry: | James [last name]’s football. |
| ~ | “Which car are you saying Richard hit the ball on?” |
| ~ | “Which car do you mean by ‘the car’?” |
| ~ | “Which ball are you saying Richard hit on the car?” |
| ~ | “Which ball do you mean by ‘the ball’?” |

Referential Analysis Two points are perhaps worth reinforcing: firstly, definite descriptions, pronouns, demonstratives and proper names all seem to make the same kind of referential reprise questions available; secondly, it seems very hard to interpret any of these examples as querying a family of sets (a GQ) rather than an individual referent.

It also seems difficult to reconcile these examples with the Kripkean view of reference via pragmatics, as outlined in section 2.2. Firstly, examples like example (43), in which a referential question is asked (and answered) before the sentence containing the original NP has been finished, do not obviously permit an explanation which requires understanding of the proposition expressed as an early step.³³ Secondly, if what is being reprised is the result of pragmatic inference from a GQ, why do readings querying the GQ itself and other associated inferences not seem to be available?

We therefore suppose that the content of definite NPs must at least contain, and perhaps consist entirely of (as sketched out roughly in (45) – we will fill in the details in section 4.4), the intended referent (which in the case of plurals, we assume will be a set). An analysis of these referent reprise questions would then be available along identical lines to that for PNs given in section 2.1 – an identifiable referent for the contextual parameter must be found in context as part of the grounding process.³⁴

$$(45) \lambda\{x\}[the_dog(x)].x$$

4.1.3 Definite Descriptions – Functional Readings

Most of the rest of the examples of direct echoes of definite descriptions did not seem to be querying an individual referent, but rather seemed to be querying a function or its domain. As might be expected, these examples were mostly *attributive* uses, which have long been held up as examples against the referential nature of definite descriptions, but other types that we would expect to behave in this way include *de dicto* uses, *narrow scope* uses, Poesio (1994)’s *weak* definites, and *generic* uses, none of which obviously convey direct reference.

Following Barwise and Perry (1983) we take the function expressed by attributive uses to be one from situations to individuals. Example (46) shows a question which seems to query the identity of the function,

³²BNC file KR2, sentences 862–866

³³Minimally, it would require a radically incremental view of semantic processing.

³⁴Whether this process should be restricted to allow grounding only to unique referents, or to most familiar or most salient referents, is a question we do not address here.

while example (47) seems to have an argument or domain reading available (amongst other possibilities):

- (46)³⁵
- | | |
|---------|--|
| Anon 1: | In those days how many people were actually involved on the estate? |
| Tommy: | Well there was a lot of people involved on the estate because they had to repair paths. They had to keep the river streams all flowing and if there was any deluge of rain and stones they would have to keep all the pools in good order and they would |
| Anon 1: | The pools? |
| Tommy: | Yes the pools. That's the salmon pools |
| Anon 1: | Mm. |
| ~> | "What are you intending 'the pools' to pick out in the situation you are describing?" |
| ~> | #"Which actual entities are you referring to by 'the pools'?" |
-
- (47)³⁶
- | | |
|---------|---|
| Eddie: | I'm used to sa-, I'm used to being told that at school. I want you <pause> to write the names of these notes up here. |
| Anon 1: | The names? |
| Eddie: | The names of them. |
| Anon 1: | Right. |
| ~> | "What situation/notes are you intending me to interpret 'the names' relative to?" |
| ~> | ? "What are you intending 'the names' to refer to in that situation?" |
| ~> | #"Which actual names are you referring to by 'the names'?" |

Again, a reading concerning properties of properties or sets of sets does not seem plausible. We therefore suppose that such uses can be captured by an analysis as sketched in (48), this being the functional equivalent of the version in (45) above, with its constituent function and argument becoming the abstracted parameters:

$$(48) \lambda\{f, s\}[f = the_dog, s \subseteq DOM(f)].f(s)$$

Grounding therefore requires both the function f and the argument s to be found in context. Failure to do so would therefore license clarification questions which can be read as concerning either function or argument/domain, or both. Note that the job of identifying the argument corresponds to Poesio (1993)'s view of definite interpretation as anchoring a parameter corresponding to the resource situation, but that on our view this is not *all* that is required.

We do not insist that the domain of the function is one of situations: indeed, for narrow-scope definites it seems simpler to take the domain as being a set of individuals contributed by a wider-scoping NP (and we set this out in section 5.3). However, the treatment of the semantic content as functional, with the resulting contribution to C-PARAMS, remains.

Strong/Weak Hypothesis This representation does not fit exactly with the strong version of our reprise content hypothesis as it is currently phrased. While both constituent elements of the content (function and argument) are reprisable, a single question might of course query only one of them, thus holding only to the weak version of the hypothesis. However, querying the entire content directly would seem wrong here, as it would necessarily reduce the functional representation to the non-functional version.

Ambiguity Introduction of this alternative analysis means, of course, that we are currently assuming some ambiguity in the representation of definites: but note that this is not an ambiguity of semantic type (the content is still of type e). This ambiguity could be removed by taking *all* definite descriptions to be functional, with referential definites those where the situational argument s is the current utterance situation s_0 (thus resembling von Heusinger (2002)'s analysis of specific indefinites as those functional on

³⁵BNC file K7D, sentences 307–313

³⁶BNC file KPB, sentences 417–421

the speaker).³⁷ In such cases, grounding of the function f in the known current situation s_0 is equivalent to identifying the referent $x = f(s_0)$. As this appears to be a worst-case analysis (over half of our corpus examples appeared to be directly referential), we do not take this step here, but merely note it as an option.

It seems likely that such a step would not be required for PNs and demonstratives in any case, which do not appear to have functional versions (not being able to take narrow scope)³⁸, so these would keep the previous simple referential analysis.

4.1.4 Definite Descriptions – Sub-Constituent Readings

The few remaining examples of definite NP reprises found seemed to be easier to interpret as having a predicate reading, identical to that which would be obtained by reprising the CN alone. No intonational information is available in the BNC, but these readings appear to be those that are made more prominent by stressing the CN (see example (49)).

- (49)³⁹
- | | |
|---------|--------------------------------------|
| Anon 1: | They'd carry the sack on their back? |
| George: | On the back, the bushel, yes |
| Anon 1: | The bushel? |
| George: | <unclear> |
| Anon 1: | <unclear> |
| George: | The corn. |
- ~> *"What are you referring to by 'the bushel'?"*
 ~> *"What property do you mean by 'bushel'?"*
 ~> *"Is it the thing with the property **bushel** that you're saying ..."*

This does not seem to be restricted to definites: in fact, the same readings seemed to be possible for all other NPs we examined (as we will see below). We therefore suppose that this reading is in fact a focussed reprise of the CN rather than the NP as a whole. Examination of sluices reinforces this: where reprise sluices were found with this reading, only the CN was substituted by a wh-word, rather than the whole NP:

- (50)⁴⁰
- | | |
|----------|-------------------------|
| Elaine: | what frightened you? |
| Unknown: | The bird in my bed. |
| Elaine: | The what? |
| Audrey: | The birdie? |
| Unknown: | The bird in the window. |
- ~> *"What property_x is it you're saying the thing with x frightened you?"*

Similarly, although none were found in the BNC, it seems plausible that a reading corresponding to the logical relation expressed by the determiner is possible (again, the reader may find this easier to capture by imagining intonational stress on the determiner).

In other words, the readings available for reprises of sub-constituents of the NP are still available when reprising the NP, especially when the relevant sub-constituent is stressed. This might be expected, given the idea of C-PARAMS inheritance outlined in section 2.1. This leads us to re-formulate our reprise content hypothesis to allow for "inherited" daughter questions:

Reprise Content Hypothesis (revised weak version):

- (51) *A nominal fragment reprise question queries part of the standard semantic content of the fragment being reprised or one of its syntactic daughters.*

³⁷Of course, removing this ambiguity here would lead to more work later. When resolving scope, we will have more arguments which need their reference established – see section 5.3.

³⁸Although possible counterexamples have been proposed for demonstratives – see (Roberts, 2002).

³⁹BNC file H5H, sentences 254–257

⁴⁰BNC file KBC, sentences 1193–1197

Reprise Content Hypothesis (revised strong version):

- (52) *A nominal fragment reprise question queries exactly the standard semantic content of the fragment being reprised or one of its syntactic daughters.*

This has implications for exactly how C-PARAMS inheritance should be reflected in the grammar, and also requires a theory of sub-constituent focussing to explain how the readings arise (see section 4.5).

4.2 Indefinite NPs

So we have seen that the evidence provided by reprises of definite NPs leads us towards a view of them as referential (although possibly functional) rather than quantificational. In this section, we turn to *indefinites*. Again, a referential viewpoint might lead us to expect that reprises of indefinites should involve a referent (perhaps not a specific real-world object but a discourse referent (Kamp and Reyle, 1993), belief object (Zimmerman, 1999) or intentional object (Dekker, 2002)), and that this referent would therefore be queried by a reprise question.

4.2.1 Sub-Constituent Readings

However, if they do exist, such readings seem to be uncommon. All direct echo examples we found were most felicitous when read as the sub-constituent readings described in section 4.1.4 above. For plain singular indefinites (see table 4), all examples seemed identical to the CN predicate reading (whether clausal or constituent). Note that the constituent reading, paraphrased in the examples below as “*What property do you mean by ‘N’?*”, might also be paraphrased “*What is a N?*” – but that this should not be confused with a *referential* constituent reading “*Which N do you mean by ‘a N’?*”.

- | | |
|--------------------|---|
| (53) ⁴¹ | Mum: What it ever since last August. I’ve been treating it as a wart.
Vicky: A wart?
Mum: A corn and I’ve been putting corn plasters on it |
| | ~> “ <i>Is it the property wart_i that you’re saying you’ve been treating it as something with <i>i</i>?</i> ”
~> “ <i>What property do you mean by ‘wart’?</i> ”
~> #“ <i>Which wart are you saying you’ve been treating it as?</i> ” |

- | | |
|--------------------|--|
| (54) ⁴² | Unknown: What are you making?
Anon 1: Erm, it’s a do– it’s a log.
Unknown: A log?
Anon 1: Yeah a book, log book. |
| | ~> “ <i>Is it the property log that you’re saying it’s something with?</i> ”
~> “ <i>What property do you mean by ‘log’?</i> ”
~> #“ <i>Which log are you saying it is?</i> ” |

For plural indefinites the same holds, although a reading querying the determiner rather than the predi-

⁴¹BNC file KE3, sentences 4678–4681

⁴²BNC file KNV, sentences 188–191

cate is also available (as we suggested might be possible for definites in section 4.1.4 above):

- (55)⁴³
- | | |
|---------|----------------------------|
| Anon 2: | Was it nice there? |
| Anon 1: | Oh yes, lovely. |
| Anon 2: | Mm. |
| Anon 1: | It had twenty rooms in it. |
| Anon 2: | Twenty rooms? |
| Anon 1: | Yes. |
-
- | | |
|----|---|
| ~> | <i>“Is it twenty_N that you’re saying it had N rooms?”</i> |
| ~> | <i>“Is it rooms that you’re saying it had twenty of?”</i> |
| ~> | <i>#“Which twenty rooms are you saying are it had?”</i> |

Two approaches therefore present themselves: either the content of an indefinite (be it referential or quantificational) is simply not abstracted to the C-PARAMS set, thus leaving only parameters associated with sub-constituents to be reprised; or the content of an indefinite is in fact identical to that of one of its subconstituents. The second seems problematic: firstly, which sub-constituent would we choose? As seen above (e.g. in example (55)), both determiner and CN content seem to be available. Secondly, it would mean different semantic types for definites and indefinites. There are other problems too, not least for an account of anaphora (see section 5.4 below for more details). In any case, the argument for making this step does not seem strong: after all, the same sub-constituent questions are available for definites.

Sluices This is perhaps reinforced by the fact that reprise sluices which query the CN predicate seem to be equally common for definites and indefinites. As shown in table 5, the same number of “A *what?*” reprises (see example (56) below) were found as “*The what?*” reprises (see example (50) above).⁴⁴ This is hardly strong evidence, but might help us to believe that subconstituent questions are no *more* made available by indefinites than definites, as we might expect them to be if the content of indefinites really was the same as that of one of their subconstituents.

- (56)⁴⁵
- | | |
|---------|---|
| Stuart: | I know it’s good in it? <unclear> but erm, <unclear> bought her, I’ve bought her a Ghost video. |
| Mark: | A what? |
| Stuart: | A Ghost video. |
| Mark: | Oh yeah. |
-
- | | |
|----|--|
| ~> | <i>“What property P is it you’re saying you’ve bought her something with P?”</i> |
| ~> | <i>“What property do you mean by ‘Ghost video’?”</i> |
| ~> | <i>#“Which Ghost video are you saying you’ve bought her?”</i> |

	Pattern	Number in BNC
Definite	<i>“... the N ... ” / “The what?”</i>	10
Indefinite	<i>“... a(n) N ... ” / “A(n) what?”</i>	10

Table 5: Predicate Sluices

It therefore seems more reasonable to take the first approach: that indefinite content is not easily available for reprise, and so sub-constituent readings predominate. But in that case, can we shed any light on whether a referential or quantificational analysis better explains the facts?

⁴³BNC file K6U, sentences 1493–1498

⁴⁴Although definites are more common than indefinites in the BNC (nearly twice as many), there is no statistically significant difference between the relative numbers of predicate sluices shown in table 5 and the relative numbers of overall occurrences.

⁴⁵BNC file KDA, sentences 672–675

4.2.2 Possible Referential Readings

While no clear examples were found in our corpus study, we feel that there *is* a possibility of referential questions with specific indefinites where the hearer realises that the speaker has a particular referent in mind, and intends the hearer to be able to identify it (what Ludlow and Segal (2004) call *definite* indefinites). Some BNC examples, while most felicitous when read as CN predicate queries, do seem to offer a possible referential paraphrase:

- (57)⁴⁶
- | | |
|------------|---|
| Stefan: | Everything work which is contemporary it is decided |
| Katherine: | Is one man? |
| Stefan: | No it is a woman |
| Katherine: | A woman? |
| Stefan: | A director who'll decide. |
| Katherine: | She's good? |
| Stefan: | Hm hm very good. |
| ~> | "Is it a woman you are saying it is?" |
| ~> | ? "Which woman are you saying it is?" |

Sluices If this is the case, we should expect referential reprise sluices "*What/Which N?*" (as opposed to the CN predicate sluice "*A what?*" described in section 4.2.1 above) to be available, if rare. And "*which N?*" examples certainly exist for indefinites, and are indeed rare (about 6 times less common after *a N* than after *the N* – see table 6).⁴⁷

	Pattern	Number in BNC
Definite	"... <i>the N</i> ... " / " <i>What/Which N?</i> "	25
Indefinite	"... <i>a(n) N</i> ... " / " <i>What/Which N?</i> "	4

Table 6: Referential Sluices

However, we must be careful when examining these examples, as it is important to distinguish between *reprise* sluices – questions concerning the directly conveyed content of the utterance, asked by the hearer during the comprehension (grounding) process, and typically delivered with a rising reprise intonation – and the more familiar *direct* sluices – questions asking for more specific information than that directly conveyed, which are not asked during the comprehension process but can be asked even after complete acceptance of an assertion, and which do not appear with the same rising reprise intonation.

Of course, especially given the lack of intonational information in the BNC, it is very difficult to determine the reprise/direct nature of a sluice beyond any doubt – we can merely attempt to fit plausible paraphrases to the dialogue context. In most cases (see example (58)), both interpretations seemed plausible, although the direct version arguably more likely. But one example in particular (example (59)) seemed to support a reprise reading more readily: the speaker appears to be using an indefinite in order to identify a person without mentioning him by name, while the interviewer wants to be sure he has understood the

⁴⁶BNC file KCV, sentences 3012–3018

⁴⁷The referential sluice distribution between definites and indefinites (table 6) is significantly different from the predicate sluice distribution (table 5): a $\chi^2_{(1)}$ test shows probability of independence $p < 1\%$. It is also not merely an effect of the fact that definites are more common ($p < 2\%$).

intended reference correctly.

- (58)⁴⁸
- | | |
|---------|--|
| Nicola: | We're just going to Beckenham because we have to go to a shop there. |
| Oliver: | What shop? |
| Nicola: | A clothes shop. <pause> and we need to go to the bank too. |
| ~> | reprise: "What shop are you telling me we have to go to?" |
| ~> | reprise: "What shop do you mean by 'a shop'?" |
| ~> | direct: "What shop do we have to go to?" |
-
- (59)⁴⁹
- | | |
|-----------------|---|
| Ray: | And of course, when this all happened, and I'm listening to what people are saying tonight, it's it's sort of making me feel a bit sick what they're saying. |
| Nicky Campbell: | Why is that? |
| Ray: | One supports that I lay in the street looking and waiting for a a man they mention tonight and that man is a well known killer of British soldiers. And I'm now asked |
| Nicky Campbell: | Which man? |
| Ray: | I'm now asked to respect him. And I'm sorry, I cannot respect a man |
| Nicky Campbell: | The man who's name has been mentioned tonight? |
| Ray: | Tonight. I cannot say that anybody can respect a man in this country and to run for their country as a well known I R A supporter. And he's up there on one of your pictures. |
| Nicky Campbell: | Mhm. |
| ~> | reprise: "Which man do you mean by 'a man they mention tonight'?" |
| ~> | reprise: "Which man are you telling me you lay waiting for?" |
| ~> | direct: "Which man did you lie waiting for?" |

Again, no examples seemed to support a property-of-properties or set-of-sets paraphrase at all. We take this as at least tentative support for a view that indefinites (a) can be seen as referential, and (b) that this referential term can in certain cases be contextually abstracted, thus being available for reprise questions. We therefore propose that an analysis of indefinites should allow for such readings to be constructed: that as for definites, their content should consist (at least in part, and if holding to the strong hypothesis, entirely) of an individual or set of individuals. The distinction from definites is that in ordinary uses this content is not contextually abstracted, and therefore does not have to be identified during grounding, but instead must be existentially quantified within the sentence (see (60) for a sketch; more details are given in section 5.3). Definite uses are distinguished simply by making the content a member of C-PARAMS as in (61), so that it does have to be grounded in context, and can be reprised.

$$(60) \exists\{x\}\{dog(x)\}.x$$

$$(61) \lambda\{x\}\{dog(x)\}.x$$

This view of indefinites as individuals which are existentially quantified (rather than as generalized quantifiers) is not dissimilar to the *choice function* approach of Reinhart (1997); Szabolcsi (1997), or the *epsilon term* approach of van Rooij (2000); von Stechow (2000); Kempson et al. (2001) – where indefinites denote individuals chosen by some existentially quantified choice function. While these approaches seem perfectly consistent with our observations, we prefer for simplicity's sake to quantify over the individuals directly, although we will use functional versions to express relative scope in section 5.3 below.

This account also allows us to give an analysis of sluicing which expresses the distinction between direct and reprise sluices: direct sluices are those which concern an existentially quantified referent contributed

⁴⁸BNC file KDE, sentences 2214–2217

⁴⁹BNC file HV2, sentences 225–236

by a previous grounded utterance (essentially the analysis of (Ginzburg and Sag, 2000)); while reprise sluices are those which concern the identity of a member of C-PARAMS during grounding, following G&C.

4.3 Other Quantified NPs

We have so far only considered definite and indefinite NPs. What of QNPs which contain other quantifiers?

There are really very few examples of reprises of such QNPs in the BNC⁵⁰, so it is premature to claim strong results; but what indications we could get, together with our intuition, point towards an identical analysis to that proposed above for indefinites. Most examples seem most felicitous when interpreted as concerning sub-constituents (either the CN predicate or the logical relation expressed by the quantifier), but seem to have a possible referential interpretation too:

- (62)⁵¹ Richard: No I'll commute every day
 Anon 6: **Every day?**
 Richard: as if, er Saturday and Sunday
 Anon 6: And all holidays?
 Richard: Yeah <pause>
- ~> "Is it **days**_N that you are saying you'll commute every N?"
 ~> "Is it **every day** that you are saying you'll commute?"
 ~> "Which days do you really mean by 'every day'?"

With universals as in example (62) above, we should perhaps not be surprised by referential readings: it has been suggested that universals should be considered as definites (see e.g. Prince, 1992; Abbott, 2003). They are less clearly available with other quantifiers:

- (63)⁵² Anon 1: Er are you on any sort of medication at all Suzanne? Nothing?
 Suzanne: No. Nothing at all.
 Anon 1: **Nothing?** No er things from the chemists and cough mixtures or anything <unclear>?
- ~> "Is it **no things** that you are saying you're on?"
 ~> ?"Which things do you really mean by 'nothing'?"

As before, imagined examples seem to be possible where referential uses can be made more clear by use of co-referring PNs in the reprise:

- (64) A: I want everyone in here to come with me.
 B: **Everyone? / Me, Carl and Donna?**
- ~> "Who do you mean by 'everyone'?"
 ~> "By 'everyone' do you mean B, C and D?"
- (65) A: Most people came to the party.
 B: **Most people?**
 A: Well, me, Brenda and Carmen.
- ~> "Who do you mean by 'most people'?"

Given this possibility, we propose to analyse these QNPs as existentially quantified sets of individuals, which are not contributed to C-PARAMS under normal circumstances. Referential uses are obtained simply

⁵⁰This is not surprising, as these NPs are relatively rare in the BNC to begin with. They are an order of magnitude less common than "the/a N": there are more than 50 times more sentences containing "the N" as there are containing "every N", and "most N", "many N" and "few N" are even rarer. As we found fewer than 100 reprises of "the N", we would only expect a handful of "every N" reprises, and none for the other quantifiers, and this is what we find.

⁵¹BNC file KSV, sentences 257–261

⁵²BNC file H4T, sentences 43–48

by adding the content to C-PARAMS. In the next section, we outline this approach in more detail.

4.4 Semantic Analysis

If we are to hold to our reprise content hypothesis, the availability of referent readings for QNP reprise questions means that the semantics of QNPs must (at least partially) consist of a referent individual or set. It seems clear that this referent is the witness set of the corresponding GQ (where this set may be functionally dependent on a situation or another set).

Two approaches present themselves. Firstly, we can hold to a standard view of QNPs as denoting GQs, and assume that the witness set forms the parameter to be grounded in context. This will, of course, only hold to the weaker version of our hypothesis. Secondly, as we have been sketching out so far, we can hold to the stronger version by considering QNPs to denote witness sets directly.

4.4.1 QNPs as GQs

The first approach is shown in (66) for the definite NP *the dog*. The content is a GQ, and the abstracted parameters which must be grounded are the witness set w (containing the referent dog to be identified in context) and the parameters contributed by the sub-constituents – the predicate P denoted by the CN *dog* and the logical relation Q denoted by the determiner. An equivalent indefinite version would of course not add the witness set to the abstracted set, leaving only the sub-constituent parameters.

$$(66) \lambda\{w, Q, P\}[witness(w, Q(P)), Q = the, name(P, dog)].Q(P)$$

The relation $witness(w, Q(P))$ is of course defined as:

$$(67) witness(w, Q(P)) \leftrightarrow w \subseteq P \wedge w \in Q(P)$$

This would account for the availability of referential reprise questions: failure to find a suitable witness set in context will result in a clarification question concerning its identity. This solution, however, only holds to the weak version of our reprise content hypothesis, as the reprise question would no longer concern the entire content of the NP, but only a part. As such, it does not offer a clear explanation of why reprise questions can only query this part, rather than the whole GQ content.

4.4.2 QNPs as Witness Sets

Accordingly we take the second approach: to treat QNPs as denoting their witness sets directly. This leads us to a simple representation, using B&C's equivalence stated in section 2.2.3 above, that a verbal predicate holds of a QNP iff the witness set belongs to the set expressed by that predicate.⁵³ The content is therefore a set, which for definites is also a member of the set of contextually abstracted parameters, along with those contributed by sub-constituents:

$$(68) \lambda\{w, Q'', P\}[w = Q''(P), Q'' = the'', name(P, dog)].w$$

Here we define the function the'' which picks out our witness set via the following equivalences:

$$(69) w = Q''(P) \leftrightarrow Q'(w, P) \leftrightarrow witness(w, Q(P))$$

Essentially this will give us a semantic representation of a sentence “*the dog snores*” which can be written as follows:

$$(70) the'(w, P) \wedge dog(P) \wedge snore(w)$$

which is broadly similar to the representation of (Hobbs, 1983, 1996).⁵⁴ Following B&C's equivalence, the sentence is true iff $w \subseteq snore$.

⁵³This could alternatively be thought of as implicitly universally quantifying over the members of the witness set.

⁵⁴Although Hobbs uses the notion of a *typical element* of a set and uses this as the argument of a verb (coercing the predicate into a typical/non-typical version as necessary). We do not take this step.

This solution has the same power to account for clarifications as the previous one (the witness set forms the contextual parameter to be grounded), but also holds to the strong version of our reprise content hypothesis, and therefore straightforwardly explains why reprise questions can only concern this set (or a sub-constituent). However, this version holds only for $\text{MON}\uparrow$ quantifiers: some possible solutions for other quantifiers are discussed in section 5.5 below.

4.5 HPSG Analysis

We can now give a HPSG analysis which shows how the NP's semantic representation is built up from those of its daughters. However, it turns out to be slightly at odds with the usual head-driven principles of HPSG: neither CONTENT nor C-PARAMS is now being directly inherited from or amalgamated across syntactic daughters.⁵⁵

CONTENT Specification As pointed out in section 3.3 above, holding to the strong version of our reprise content hypothesis must mean that NPs do not inherit their content from their head daughter CNs (as in standard HPSG unification-based semantics), or simply amalgamate across daughters (as in Minimal Recursion Semantics): the referential reprises available for NPs are simply not available when reprising the daughters. To specify the content correctly, we must therefore posit a type qnp for all QNPs which specifies how the semantic representation is built:

$$(71) \left[\begin{array}{l} qnp \\ \text{CONTENT} \left[\begin{array}{l} \text{INDEX } w \\ \text{RESTR} \left\{ \begin{array}{l} \text{witness_set_rel} \\ \text{INSTANCE } w \\ \text{PROPERTY } P \\ \text{RELN } Q \end{array} \right\} \end{array} \right] \\ \text{DTRS} \left\langle \left[\begin{array}{l} \text{det} \\ \text{CONTENT | INDEX } Q \end{array} \right], \left[\begin{array}{l} \text{nominal} \\ \text{CONTENT | INDEX } P \end{array} \right] \right\rangle \end{array} \right]$$

(or in abbreviated form):

$$(72) \left[\begin{array}{l} \text{CONTENT} \left[w : w = Q(P) \right] \\ \text{DTRS} \left\langle \left[\begin{array}{l} \text{det} \\ \text{CONTENT } Q \end{array} \right], \left[\begin{array}{l} \text{nominal} \\ \text{CONTENT } P \end{array} \right] \right\rangle \end{array} \right]$$

Note that the constraint expressed above is still monotonic (no semantic information is dropped in construction of the mother) and compositional (the semantics of the mother is obtained purely by functional application of daughters).

C-PARAMS Amalgamation As mentioned in section 4.1.4 above, the availability of sub-constituent readings shows that the C-PARAMS value for a phrase must include the values of its daughters. However, the fact that reprises of head daughters (e.g. CNs) cannot be interpreted as querying the content of their *sisters* (e.g. determiners) means that this inheritance process cannot be via lexical heads (as in the general $\text{NON-LOCAL Amalgamation Constraint}$ assumed to govern C-PARAMS by G&C), but instead must be explicitly specified for the mother. We could therefore express this as a default constraint on the type $phrase$ similar

⁵⁵While we give an analysis here only for our preferred witness-set only approach, the general observations also hold for the GQ approach.

to G&C's CONSTITS Amalgamation Constraint, shown in AVM (73) below.

$$(73) \left[\begin{array}{l} \textit{phrase} \\ \text{C-PARAMS } \boxed{1} \cup \dots \cup \boxed{n} \\ \text{DTRS } \left\langle \left[\text{C-PARAMS } \boxed{1} \right], \dots, \left[\text{C-PARAMS } \boxed{n} \right] \right\rangle \end{array} \right]$$

However, definite NPs must override this default, as they introduce a new contextual parameter as well as amalgamating those of their daughters. Indefinites hold to it, but we must ensure that their content is instead existentially quantified. We can combine these facts into a general *definiteness principle*.

Definiteness Principle In our HPSG terms, indefinites must contribute their content to the STORE feature (which specifies the existentially quantified elements – see section 5.3 for more details), while definites contribute it to C-PARAMS (and this is what distinguishes definite from indefinite uses). We can therefore state a general principle: the content of a NP must be a member of either C-PARAMS or STORE. We can replace AVM (73) with a more general Definiteness Principle, which applies to both words and phrases. For words, it is simply expressed:

$$(74) \left[\begin{array}{l} \textit{word} \\ \text{CONTENT } \boxed{1} \\ \text{STORE } \boxed{2} \\ \text{C-PARAMS } \left\{ \boxed{1} \right\} - \boxed{2} \end{array} \right]$$

For phrases, it also specifies STORE and C-PARAMS inheritance from daughters. The C-PARAMS value of the mother is the union of the daughter values, plus the mother content, unless this is contributed to STORE:

$$(75) \left[\begin{array}{l} \textit{phrase} \\ \text{CONTENT } \boxed{1} \\ \text{STORE } \boxed{2} \cup \boxed{3} \\ \text{C-PARAMS } \left\{ \left\{ \boxed{1} \right\} - \boxed{2} \right\} \cup \boxed{4} \cup \dots \cup \boxed{n} \\ \text{HEAD-DTR } \left[\text{STORE } \boxed{3} \right] \\ \text{DTRS } \left\langle \left[\text{C-PARAMS } \boxed{4} \right], \dots, \left[\text{C-PARAMS } \boxed{n} \right] \right\rangle \end{array} \right]$$

Definites (and referential words such as CNs, which on our account are referential to a predicate) can therefore be specified as having empty STORE values, thus forcing their content to be a member of C-PARAMS. Indefinites can be specified as contributing to STORE, and thus can make no contribution to C-PARAMS.

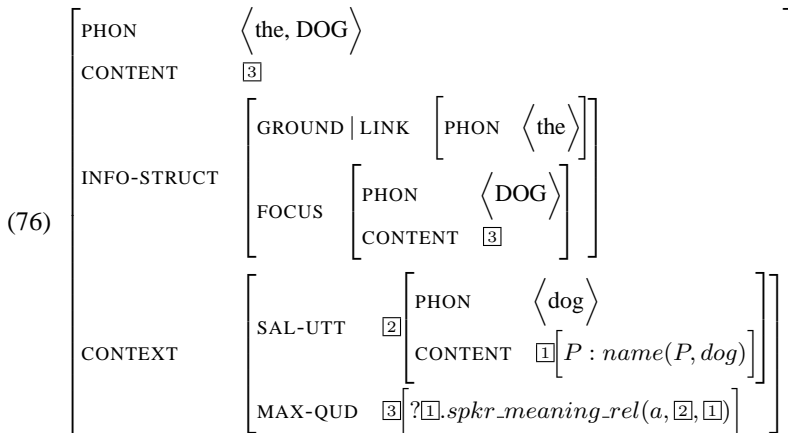
For those NPs for which we have proposed a functional analysis (e.g. attributive definites) a slightly different version of the principle is of course required: the function and argument parameters are treated separately and can be contributed individually to either STORE or C-PARAMS.

Sub-Constituent Focussing This inheritance of C-PARAMS from daughters goes some way towards accounting for the sub-constituent readings that always seem available (especially when a constituent is intonationally stressed), but we also require an explanation of how the sub-constituent becomes focussed in order to assign the relevant content to the reprise question.

We assume Engdahl and Vallduví (1996)'s analysis of information structure in HPSG, with a feature INFO-STRUCT divided into FOCUS and GROUND⁵⁶, with the contents of each linked (in English at least) to intonation. Reprise questions are now taken to be querying the FOCUSSED component (and checking that

⁵⁶There is some redundancy here between GROUND and MAX-QU, as both are expressing contextually given elements. A full account would link the two (see Engdahl et al., 1999).

the GROUND components are indeed given in context by the utterance being clarified).^{57 58}



4.6 Summary

In this section we have shown that no NP reprises appear to query a generalized quantifier or property-of-properties, but that reprises of definite NPs can query an individual (or set of individuals), and have surmised that this may also be true for referential uses of other QNPs.

We have shown how our reprise content hypothesis can be held to in its strong version if a semantic representation of QNPs as denoting witness sets is used. This leads to a relatively simple flat representation, with similarities to that of Hobbs (1983) or the choice function/epsilon term approach. A standard GQ representation can only hold to the weak version of the hypothesis, making it difficult to explain why reprises do not appear to be able to query GQs.

Having outlined our treatment of QNPs, in the next section we take a quick look at some further implications of this treatment (for the semantics of determiners and for a treatment of anaphora), and show how it can be extended to cope with important issues we have so far only mentioned: quantifier scope and non-MON \uparrow quantifiers.

5 Further Issues

5.1 Determiners

Where does this leave us with regard to determiners? A view of NPs as denoting witness sets and of CNs as denoting predicates (properties of individuals) seems to leave us with a view of determiners as denoting functions from the CN predicates to the NP sets (i.e. functions of type $(e \rightarrow t) \rightarrow e$). In a model-theoretic sense, they would therefore denote relations between two sets (the equality relation for *every*, a relation that picks out an epsilon term for *a/some*, a relation that picks out a set of a particular cardinality for *two/three*).

The alternative view of NPs as denoting GQs, on the other hand, would force us to view determiners as denoting functions from CN predicates to GQs (sets of sets) – essentially the Montagovian view of determiners as functions of type $(e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t)$.

Do either of these fit with what determiner reprise questions seem to mean?

⁵⁷We suspect that the GROUND components are present in the reprise either to help disambiguate the exact source constituent being clarified, or just to make the reprise more syntactically palatable.

⁵⁸Note that a complete analysis of this phenomenon may require an account of focus spreading from CN to NP: it seems plausible to us that a reprise even with the CN intonationally focussed may be interpreted as querying the NP referent. This should be possible, again using Engdahl and Vallduví (1996)'s analysis, but we note that the usual assumption that focus spreads from the *most oblique* daughter to the mother would not appear to hold in this case (intuitively at least – as far as we know accounts of focus spreading have never considered phenomena at this low a level, within NPs).

5.1.1 Evidence

Determiner-only reprises seem to be rare: the only suitable examples found through corpus investigation involved numerals (see example (77)).⁵⁹

- (77)⁶⁰ | Marsha: yeah that's it, this, she's got three rottweiler's now and
 Sarah: **three?**
 Marsha: yeah, one died so only got three now <laugh>
 ~> | "Is it **three**_N you are saying she's got *N* rottweilers?"

For these examples, the query appears to concern the cardinality of the set under discussion, which fits quite nicely with the idea of determiners as denoting set relations. For other determiners, we have to rely on our intuition, and on those QNP reprise examples mentioned in section 4 above in which the determiner appears to be stressed, e.g. example (62), repeated here as example (78):

- (78)⁶¹ | Richard: No I'll commute every day
 Anon 6: **Every day?**
 Richard: as if, er Saturday and Sunday
 Anon 6: And all holidays?
 Richard: Yeah <pause>
 ~> | "Is it **every**_N that you are saying you'll commute on *N* days?"

Again, these readings do seem to fit quite nicely with the idea of determiners as denoting set relations, and perhaps less so with that of relations between sets and sets of sets.

Another possible reading seems to be one asking about the situation in which the quantifier relation is being used. This could be accounted for in terms of situated relations (functional on situations), analogous to the functional sets discussed briefly in section 4.1.3 and in more detail in section 5.3 below. However, the sparsity of the evidence and the difficulty of pinning down a definitive paraphrase mean we hesitate to make any strong claims here: but we do claim that determiner reprises provide no counter-evidence to the analysis of section 4.

5.2 WH-Phrases

We have not so far mentioned WH-phrases. How should their semantic content be represented so as to be consistent with what their reprises seem to mean?

Very few examples of reprises of "what/which *N*" phrases were found, so we have also looked at reprises of plain WH-words. Examination of both suggests that the query can concern a property but not a referent. In "what/which *N*" examples (see example (79)) we see the familiar sub-constituent readings (querying the CN predicate or the determiner relation); bare WH examples (example (80)) seem to query a predicate expressed as part of the lexical semantics of the WH-word itself. Referent readings seem impossible in all cases.

- (79)⁶² | Unknown: How many procedures have we actually audited so far Richard?
 Richard: **How many procedures?**
 Unknown: Yeah.
 ~> | "Is it **procedures**_P you are asking about how many *Ps*?"
 ~> | "Is it a **number** of procedures you are asking about?"
 ~> | # "Which procedures are you asking how many of them there are?"

⁵⁹The only non-numerical determiner-only reprises we have found are form identification queries (i.e. regarding surface form rather than semantic content) of a different type, classified by Purver et al. (2003a) as the *gap* reading: the element being clarified is not the original determiner but rather whatever word(s) came immediately after it.

⁶⁰BNC file KP2, sentences 295–297

⁶¹BNC file KSV, sentences 257–261

⁶²BNC file KM4, sentences 920–922

- (80)⁶³
- | | |
|------------|---|
| Charlotte: | Why does the dustman have to take it away? |
| Larna: | No not the dustman, the postman |
| Charlotte: | Why does the postman have to take all the letters away? |
| Larna: | Why? Well he takes them to the post office |
| Charlotte: | Yeah |
| Larna: | then the post office sorts them out |
| ~> | “ <i>Is it a reason you are asking for?</i> ” |
| ~> | #“ <i>Which reason are you asking for?</i> ” |

The simplest and most consistent analysis therefore seems to be that WH-phrases resemble indefinites, in that they represent terms (or sets of terms) which are *not* added to C-PARAMS (hence no referent reprise reading). However, these terms are not existentially quantified within the sentence but queried: on a view of questions as λ -abstracts, they are part of the abstracted set. In our HPSG analysis, this is achieved by adding them to STORE rather than C-PARAMS (like indefinites), but giving them a distinct type which must be discharged into the PARAMS feature, the abstracted set for questions, rather than QUANTS, the existentially quantified set.

5.3 Quantification and Scope

5.3.1 Representation of Scope

Given a representation of NPs as denoting witness sets, we need a way of expressing relative scope between the sets introduced by a sentence, both those sets associated with definites that will be fixed in context, and those associated with non-definites which are existentially quantified over. We cannot therefore use a standard approach of ordering quantifiers; instead we can use the functional representation outlined in section 4.1.3, regarding narrow-scoping NPs as functional on other wider-scoping sets. The alternative readings of “*every dog_d likes a cat_c*” are produced by representing *a cat* either as a simple existentially quantified individual c , or as a functional one $f(d)$, dependent on the set of dogs d via an existentially quantified function f (so giving an analysis similar to the choice function analysis of e.g. Farkas (1997); von Stechow (2002)).

The function f remains a member of C-PARAMS or STORE depending on (in)definiteness, just as described in section 4. The argument d must be bound to the relevant wide-scoping set: where the wide-scope NP is definite and its content is in C-PARAMS, this is achieved by making the narrow-scope argument a member of C-PARAMS and identifying the two during grounding; where the wide-scope NP is indefinite and its content in STORE, it occurs through the anaphoric binding mechanism which we describe in section 5.4.2 below.⁶⁴

5.3.2 Quantifier Storage and Retrieval

As we are representing all non-definites as existentially quantified sets, we need a mechanism for introducing this quantification into the semantic content of the sentence at the appropriate level. For this we use the familiar storage method of (Cooper, 1983), using the feature STORE to which existentially quantified elements are added by lexical/phrasal constituents and from which they are retrieved to form part of the sentence semantics. We keep the lexically-based retrieval mechanism of (Ginzburg and Sag, 2000),

⁶³BNC file KD1, sentences 434–440

⁶⁴We have postulated functional NPs with two types of argument – situations (for attributive definites) and other NP witness sets (for narrow scope). A simpler view with only situations as arguments might be possible: in the case of narrow-scoping elements, the argument would be a situation linked to another NP, directly analogous to Cooper (1995)’s *individual* situation (a situation for each member of the witness set, which supports the proposition expressed by the sentence for that member). The cost of this view would be that sets of individual situations must be provided in C-PARAMS/STORE, either by NPs themselves or by verbal predicates. As we currently have no evidence for this, we leave it aside for now as a possible alternative.

whereby inherited STORE values are discharged into a QUANTS feature by lexical heads.⁶⁵

$$(81) \left[\begin{array}{l} \text{word} \\ \text{CONTENT} \left[\text{QUANTS } \boxed{2} \right] \\ \text{STORE} \left\{ \boxed{1} \cup \dots \cup \boxed{n} \right\} - \boxed{2} \\ \text{ARG-ST} \left\langle \left[\text{STORE } \boxed{1} \right], \dots, \left[\text{STORE } \boxed{n} \right] \right\rangle \end{array} \right]$$

As scope is expressed via functional readings, we can take the members of the QUANTS set to be *simultaneously* quantified over, following Cooper (1993)'s definition of simultaneous quantification for his situation-theoretic reconstruction of DRT (Kamp and Reyle, 1993). A quantified object is viewed as a simultaneous abstract, with the QUANTS set abstracted from the body. Truth conditions are then dependent on the existence of some appropriate assignment for that abstract – one which assigns values to the members of the abstracted set such that the standard truth conditions hold for the body.

5.4 Anaphora

5.4.1 Intersentential Anaphora

An account of anaphora seems to follow simply, whereby anaphoric terms such as pronouns are treated like definites – they have referential C-PARAMS whose reference must be established during the grounding process. The constraints on this identification may be slightly different to those for definites: rather than having to identify a referent in the general context, truly anaphoric uses must have to refer to entities already established in the discourse. Deictic uses can be accounted for by assuming that salient referents are introduced into the discourse (or the general context) by external cognitive means.

Details will depend on the model of context being used, which we do not intend to delve into here.⁶⁶ We just note that the treatment of NPs as denoting witness sets allows these sets to provide potential antecedents for anaphors in future utterances. Where these antecedent sets are associated with definites, it is clear that they are already in the context: for indefinites, a protocol will be required to account for their addition thereto.⁶⁷

One puzzle, however, is the quantifier *every*. In contrast to quantifiers such as *all* and *most* which licence only plural anaphora (which we assume to refer to the witness set denoted by the QNP and therefore introduced to the discourse), *every* also licenses singular anaphora. If we assume that an *every*-QNP denotes a set, it is not clear how a singular individual is provided for reference. If instead we view a singular pronoun as functional on a set, it is not clear why this is not possible for other plural quantifiers.

5.4.2 Intrasentential Anaphora

Accounting for intrasentential anaphora requires a further step. If pronouns (and anaphoric definites) are taken as referring to existentially quantified elements within the same sentence, they can no longer have a contextual parameter associated with them: they do not refer to an element in the context external to the utterance.

We therefore propose that elements of C-PARAMS can be removed if they can be identified with an element of QUANTS – i.e. a binding mechanism similar to Poesio (1994)'s *parameter anchoring* and van der Sandt (1992)'s *presupposition binding* (hence the advantage of our implementation of STORE/QUANTS as parameters rather than quantifiers). This mechanism is implemented via a new feature B(OUND)-PARAMS: referential parameters can be members of either C-PARAMS or B-PARAMS, but membership of B-PARAMS

⁶⁵The details of this mechanism are actually slightly simpler than that of (Ginzburg and Sag, 2000): as only simultaneous existential quantification is used (see below), the order of quantifiers is not important – we can therefore represent QUANTS as a *set* rather than a *list*, thus no longer requiring their *order* operator. We can also treat both STORE and QUANTS as sets of *parameters* rather than *quantifiers*, which simplifies them and turns out to be useful for our treatment of anaphora (see section 5.4 below).

⁶⁶For one thing, a full account will presumably also require some notion of salience or discourse structure.

⁶⁷This cannot be as simple as adding an utterance's existentially quantified sets to a discourse record on acceptance: Ginzburg (2001) gives examples of anaphora to entities from unaccepted assertions and even from ungrounded utterances. One way to take these into account might be to allow for the possibility of pronouns which are functional on (sub-)utterances themselves (or, as Ginzburg suggests, utterance situations).

is limited to those parameters which can be identified with existentially quantified parameters (i.e. members of STORE/QUANTS). This leads us to the final version of our definiteness principle:

$$(82) \left[\begin{array}{l} \text{word} \\ \text{CONTENT} \quad \boxed{1} \\ \text{STORE} \quad \boxed{2} \\ \text{C-PARAMS} \quad \boxed{3} \\ \text{B-PARAMS} \quad \left\{ \boxed{1} \right\} - \boxed{2} - \boxed{3} \end{array} \right]$$

The restriction on B-PARAMS membership is expressed through the final version of our lexical quantifier storage mechanism:

$$(83) \left[\begin{array}{l} \text{word} \\ \text{CONTENT} \quad \left[\text{QUANTS} \quad \boxed{Q} \right] \\ \text{STORE} \quad \boxed{S} = \{ \boxed{1_a} \cup \dots \cup \boxed{n_a} \} - \boxed{Q} \\ \text{B-PARAMS} \quad \{ \boxed{1_b} \cup \dots \cup \boxed{n_b} \} - \text{subset}(\boxed{Q} \cup \boxed{S}) \\ \text{ARG-ST} \quad \left\langle \left[\begin{array}{l} \text{STORE} \quad \boxed{1_a} \\ \text{B-PARAMS} \quad \boxed{1_b} \end{array} \right], \dots, \left[\begin{array}{l} \text{STORE} \quad \boxed{n_a} \\ \text{B-PARAMS} \quad \boxed{n_b} \end{array} \right] \right\rangle \end{array} \right]$$

We ensure that all members of B-PARAMS are thus discharged by specifying top-level sentences (in our grammar, signs of type *root-cl*) as having empty B-PARAMS. Note that this mechanism can also apply to the arguments of narrow-scope functional NPs, thus allowing them to be functional from wider-scoping existentially quantified sets. This includes situational arguments, allowing the argument of an attributive definite to be taken as the situation introduced in the utterance (the described situation).

5.5 Monotone Decreasing Quantifiers

As we mentioned in section 4.4.2 above, B&C point out that it is not sufficient with monotone decreasing ($\text{MON}\downarrow$) cases to show that a predicate holds of a witness set: instead we must show that the witness set contains *all* members of the restriction set of which the predicate holds.

$$(84) \exists w[(X \cap A) \subseteq w] \leftrightarrow X \in D(A)$$

This means that our representation of QNPs as denoting witness sets fails to encapsulate the meaning of $\text{MON}\downarrow$ quantifiers (or non-monotone quantifiers such as *exactly two*). The sentence “*Few dogs snore*” does not only convey the fact that the property of snoring holds of some set w containing few dogs (as our simple representation would – see (85)), but also that the property does not hold of any dogs not in w (e.g. as in (86)):

$$(85) \text{few}'(w, P) \wedge \text{dog}(P) \wedge \text{snore}(w)$$

$$(86) \text{few}'(w, P) \wedge \text{dog}(P) \wedge \text{snore}(w) \wedge \neg \exists w'[(w' \subseteq P) \wedge (w \subset w') \wedge \text{snore}(w')]$$

One solution might be to appeal to pragmatics: Hobbs (1996) solves the problem by use of a pragmatic constraint which strengthens the sentence meaning accordingly: *few dogs snore* is taken just as the assertion that there is a set containing few dogs, all of whom snore, but this is strengthened by an abductive process to the assertion that this set is the maximal set of snoring dogs. Another would of course be to regard the content of QNPs as GQs rather than witness sets, but of course this means only the weak hypothesis can hold (see above). A third possibility is the view of $\text{MON}\downarrow$ quantifiers as the negation of their $\text{MON}\uparrow$ counterparts (*few dogs snore* is truth-conditionally equivalent to *most dogs don't snore*). This has been much explored in the DPL tradition of GQs (see e.g. van den Berg, 1996).

Complement Set Anaphora One of the advantages of this last approach is that it allows for an explanation of the phenomenon of *complement set anaphora* (Moxey and Sanford, 1987, 1993). Kibble (1997a,b) sees sentences with such quantifiers as ambiguous between internal and external negation (*most dogs don't*

snore vs. *it's not true that most dogs snore*), giving rise to the possibility of complement set (the dogs who don't snore) and reference set (the dogs who do) anaphora respectively.

An interesting question is therefore whether reprise questions of $\text{MON}\downarrow$ QNPs can query the reference or complement set. The pragmatic approach would suggest only the reference set is possible, the negation approach the reverse. Sadly, examples of $\text{MON}\downarrow$ QNP reprises are rare. Most of those we have found seem to be best paraphrased as sub-constituent readings, querying either the CN predicate or the logical quantifier relation:

- (87)⁶⁸ | Lorna: Oh shit! I've gotta ring mum. Tell mum no meat.
 | Kathleen: **No meat?**
 | Lorna: I'm not allowed to get meat and stuff.
 | Kathleen: Why?
 | Lorna: Cos we're vegetarians!
- ~> "Is it really **meat**_P you're saying to tell mum no P?"
 ~> "Is it really **no**_N you're saying to tell mum N meat?"

But some do seem to allow for reference set reference, and possibly for complement set reference as well, although this seems less clear:

- (88)⁶⁹ | Anon 1: Did any of them the lads that you the men that you went away with. Did they come back?
 | Richard: Not all.
 | Anon 1: **Not all of them?**
 | Richard: Oh no.
 | Anon 1: Were any of them.
- ~> "Who are you telling me did come back?"
 ~> ?"Who are you telling me didn't come back?"

Kibble gives the following example of complement set anaphora:

- (89) | BBC News: Not all of the journalists agreed, among them the BBC's John Simpson.

where *them* is construed to refer to the group of journalists who did *not* agree. An imagined reprise version seems easier to construe as querying the complement set:

- (90) | A: Not all of the journalists agreed.
 | B: **Not all of them?**
 | A: John Simpson was pretty combative. Marr and Paxman didn't like it much either.
- ~> "Who do you mean **didn't** agree?"

If so, a more consistent approach would be to view $\text{MON}\downarrow$ QNPs as denoting *pairs* of reference and complement sets $\langle R, C \rangle$. The reference set R is, as with $\text{MON}\uparrow$ QNPs, a witness set; the complement set C is $(A - R)$ (for a quantifier living on A). Such a pair might be paraphrased as "*R as opposed to C*", and can be interpreted as follows:

$$(91) \text{snore}(\langle R, C \rangle) \leftrightarrow (R \subseteq \text{snore}) \wedge (C \cap \text{snore} = \emptyset)$$

Most such QNPs will presumably be non-referential and thus will not contribute to C-PARAMS, with the pair of sets instead existentially quantified via STORE (92): what is contributed in any referential cases depends on whether we believe in complement set reprises – if so, the pair $\langle R, C \rangle$ will be made a member

⁶⁸BNC file KCW, sentences 2204–2210

⁶⁹BNC file HEU, sentences 360–365

of C-PARAMS, thus holding to the strong hypothesis (93); if not, just R (94).

$$(92) \lambda\{Q, P\}[Q = few'', name(P, dog)].\exists\{r, c\}[r = Q(P), c = (P - r)].\langle r, c \rangle$$

$$(93) \lambda\{r, c, Q, P\}[Q = few'', name(P, dog), r = Q(P), c = (P - r)].\langle r, c \rangle$$

$$(94) \lambda\{r, Q, P\}[Q = few'', name(P, dog), r = Q(P)].\exists\{c\}[c = (P - r)].\langle r, c \rangle$$

The existence of both members of the pair now helps us explain how they are possible anaphoric referents: and so why (only) $\text{MON}\downarrow$ QNPs license complement-set reference. As it stands, this says nothing about the relative preference for reference set anaphora observed by Nouwen (2003), or the possibility that not all $\text{MON}\downarrow$ quantifiers license complement set anaphora that he also raises, although his approach using inference of non-emptiness seems perfectly applicable to ours. Further investigation of $\text{MON}\downarrow$ reprises, particularly if more data can be obtained, may help us in this direction.

6 Conclusions

In this paper we have introduced the use of reprise questions as probes for investigating the meanings of words and phrases, giving us a strong criterion of assigning denotations which not only combine to make up compositional sentence meanings but explain why individual constituents give their observed reprise readings. We have examined the evidence provided by the apparent interpretation of these questions as regards the denotation of nouns, noun phrases and (very briefly) determiners. This evidence has led us to the following conclusions:

- The commonly held view of CNs as properties (of individuals) seems to correspond well with their reprises.
- The view of NPs as denoting sets of sets, or properties of properties, seems very difficult to reconcile with reprise questions.
- Reprises of NPs all seem to be able to query focussed sub-constituents.
- Reprises of definite NPs suggest that most uses of these NPs are referential to a (possibly functional) individual or set.
- Reprises of indefinite NPs and other QNPs suggest that such referential uses, while rare, are possible.

These conclusions have led us to a representation of NPs as denoting witness sets, and a definite/indefinite distinction expressed by abstraction (or lack thereof) of referential parameters to a contextually dependent set. We have shown how this can take into account relative quantifier scope via a functional view, in-trasentential anaphora via a parameter binding mechanism, and non-monotone-increasing quantifiers via a representation as pairs of sets.

We have outlined a corresponding treatment in HPSG (including a revised quantifier storage mechanism), and noted along the way that this causes us to revise some of the standard assumptions made in HPSG (and its related semantic frameworks such as Minimal Recursion Semantics) about inheritance of content from daughter to mother.

Author's Address

Postal Address Department of Computer Science,
King's College London, Strand,
London WC2R 2LS, UK

Email matthew.purver@kcl.ac.uk

Acknowledgements

The authors are supported by the Engineering and Physical Sciences Research Council (under grant GR/R04942/01) and by the Economic and Social Research Council (under grant RES-000-23-0065) of the United Kingdom. They would also like to thank Robin Cooper, Bill Ladusaw, Gerald Penn, John Beavers and Christian Ebert for useful discussion and comments, and our anonymous reviewer for very helpful and detailed comments from which the paper has benefitted considerably. An early version of this work was presented at the 1999 Amsteloque conference held in Amsterdam. More recently, this work was presented at the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, and at the 5th Tbilisi Symposium on Logic, Language, and Computation.

References

- Abbott, Barbara (2003). 'Definiteness and indefiniteness'. In Horn, L. and Ward, G. (eds.), *Handbook of Pragmatics*. Blackwell.
- Asher, Nicholas (1993). *Reference to Abstract Objects in Discourse*, volume 50 of *Studies in Linguistics and Philosophy*. Kluwer Academic Publishers.
- Barwise, Jon and Cooper, Robin (1981). 'Generalized quantifiers and natural language'. *Linguistics and Philosophy*, 4:159–219.
- Barwise, Jon and Perry, John (1983). *Situations and Attitudes*. MIT Press.
- Burnard, Lou (2000). *Reference Guide for the British National Corpus (World Edition)*. Oxford University Computing Services.
- Carlson, Gregory (1977). *Reference to Kinds in English*. PhD thesis, University of Massachusetts at Amherst.
- Chierchia, Gennaro (1998). 'Reference to kinds across languages'. *Natural Language Semantics*, 6(4):339–405.
- Cooper, Robin (1983). 'Quantification and syntactic theory'. *Synthese Language Library*.
- Cooper, Robin (1993). 'Towards a general semantic framework'. In Cooper, R. (ed.), *Integrating Semantic Theories*. ILLC/Department of Philosophy, University of Amsterdam. Deliverable R2.1.A, Dyana-2.
- Cooper, Robin (1995). 'The role of situations in generalized quantifiers'. In Lappin, S. (ed.), *The Handbook of Contemporary Semantic Theory*. Blackwell.
- Cooper, Robin and Ginzburg, Jonathan (2002). 'Using dependent record types in clarification ellipsis'. In Bos, J., Foster, M., and Matheson, C. (eds.), *Proceedings of the 6th Workshop on the Semantics and Pragmatics of Dialogue (EDILOG)*, Edinburgh. 45–52.
- Copestake, Ann, Flickinger, Dan, Sag, Ivan, and Pollard, Carl (1999). 'Minimal recursion semantics: An introduction'. Draft.
- Dekker, Paul (2002). 'A pragmatic view upon indefinites'. In von Heusinger, K., Kempson, R., and Meyer-Viol, W. (eds.), *Proceedings of the ESSLLI-01 Workshop on Choice Functions and Natural Language Semantics*. Working Papers of the Department of Linguistics in Konstanz.
- Donnellan, Keith (1966). 'Reference and definite descriptions'. *Philosophical Review*, 77:281–304.
- Engdahl, Elisabet, Larsson, Staffan, and Ericsson, Stina (1999). 'Focus-ground articulation and parallelism in a dynamic model of dialogue'. In *Task Oriented Instructional Dialogue (TRINDI): Deliverable 4.1*. University of Gothenburg.
- Engdahl, Elisabet and Vallduví, Enric (1996). 'Information packaging in HPSG'. In Grover, C. and Vallduví, E. (eds.), *Studies in HPSG*, volume 12 of *Edinburgh Working Papers in Cognitive Science*. University of Edinburgh, 1–31.
- Farkas, Donka (1997). 'Dependent indefinites'. In Corblin, F., Godard, D., and Maradin, J.-M. (eds.), *Empirical Issues in Formal Syntax and Semantics*. Peter Lang, 243–267.
- Fodor, Janet and Sag, Ivan (1982). 'Referential and quantificational indefinites'. *Linguistics and Philosophy*, 5:355–398.
- Ginzburg, Jonathan (1996). 'Interrogatives: Questions, facts and dialogue'. In Lappin, S. (ed.), *The Handbook of Contemporary Semantic Theory*. Blackwell, 385–422.
- Ginzburg, Jonathan (2001). 'Fragmenting meaning: Clarification ellipsis and nominal anaphora'. In Bunt, H. (ed.), *Computing Meaning 2: Current Issues in Computational Semantics*, Studies in Linguistics and Philosophy. Kluwer Academic Publishers.
- Ginzburg, Jonathan and Cooper, Robin (2001). 'Resolving ellipsis in clarification'. In *Proceedings of the 39th Meeting of the ACL*. Association for Computational Linguistics, 236–243.
- Ginzburg, Jonathan and Cooper, Robin (2004). 'Clarification, ellipsis, and the nature of contextual updates in dialogue'. *Linguistics and Philosophy*, 27(3):297–365.

- Ginzburg, Jonathan and Sag, Ivan (2000). *Interrogative Investigations: the Form, Meaning and Use of English Interrogatives*. Number 123 in CSLI Lecture Notes. CSLI Publications.
- Ginzburg, Jonathan, Sag, Ivan, and Purver, Matthew (2003). 'Integrating conversational move types in the grammar of conversation'. In Kühnlein, P., Rieser, H., and Zeevat, H. (eds.), *Perspectives on Dialogue in the New Millennium*, volume 114 of *Pragmatics and Beyond New Series*. John Benjamins, 25–42.
- Groenendijk, Jeroen and Stokhof, Martin (1991). 'Dynamic predicate logic'. *Linguistics and Philosophy*, **14**(1):39–100.
- Healey, Patrick, Purver, Matthew, King, James, Ginzburg, Jonathan, and Mills, Greg (2003). 'Experimenting with clarification in dialogue'. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, Boston.
- Heim, Irene (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts at Amherst.
- Hobbs, Jerry (1983). 'An improper treatment of quantification in ordinary English'. In *Proceedings of the 21st Annual Meeting*. Association for Computational Linguistics, 57–63.
- Hobbs, Jerry (1996). 'Monotone decreasing quantifiers'. In van Deemter, K. and Peters, S. (eds.), *Semantic Ambiguity and Underspecification*, number 55 in CSLI Lecture Notes. CSLI Publications, 55–76.
- Kamp, Hans and Reyle, Uwe (1993). *From Discourse To Logic*. Kluwer Academic Publishers.
- Keenan, Edward and Stavi, Jonathan (1986). 'A semantic characterization of natural language determiners'. *Linguistics and Philosophy*, **9**:253–326.
- Keenan, Edward and Westerståhl, Dag (1997). 'Generalized quantifiers in linguistics and logic'. In van Benthem, J. and ter Meulen, A. (eds.), *Handbook of Logic and Language*. Elsevier, 837–893.
- Kempson, Ruth, Meyer-Viol, Wilfried, and Gabbay, Dov (2001). *Dynamic Syntax: The Flow of Language Understanding*. Blackwell.
- Kibble, Rodger (1997a). 'Complement anaphora and dynamic binding'. In Lawson, A. (ed.), *Proceedings of the 7th annual conference on Semantics and Linguistic Theory (SALT VII)*. Cornell University.
- Kibble, Rodger (1997b). 'Complement anaphora and monotonicity'. In Morrill, G., Kruijff, G.-J., and Oehrle, R. (eds.), *Proceedings of the Formal Grammar conference*. 125–136.
- Kripke, Saul (1977). 'Speaker's reference and semantic reference'. In French, P., Uehling, T., and Wettstein, H. (eds.), *Perspectives in the Philosophy of Language*, number 2 in *Midwest Studies in Philosophy*. University of Minnesota Press, 6–27.
- Larsson, Staffan (2002). *Issue-based Dialogue Management*. PhD thesis, Göteborg University. Also published as Gothenburg Monographs in Linguistics 21.
- Ludlow, Peter and Neale, Stephen (1991). 'Indefinite descriptions: In defense of Russell'. *Linguistics and Philosophy*, **14**:171–202.
- Ludlow, Peter and Segal, Gabriel (2004). 'On a unitary analysis of definite and indefinite descriptions'. In Reimer, M. and Bezuidenhout, A. (eds.), *Descriptions and Beyond: An Interdisciplinary Collection of Essays on Definite and Indefinite Descriptions*. Oxford University Press.
- Montague, Richard (1974). 'The proper treatment of quantification in ordinary English'. In Thomason, R. (ed.), *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press, 247–270.
- Moxey, Linda and Sanford, Anthony (1987). 'Quantifiers and focus'. *Journal of Semantics*, **5**:189–206.
- Moxey, Linda and Sanford, Anthony (1993). *Communicating Quantities: a Psychological Perspective*. Lawrence Erlbaum Associates.
- Nouwen, Rick (2003). 'Complement anaphora and interpretation'. *Journal of Semantics*, **20**(1):73–113.
- Pelletier, Francis J. (2003). 'Context dependence and compositionality'. *Mind and Language*, **18**(2):148–161.
- Pickering, Martin and Garrod, Simon (2004). 'Toward a mechanistic psychology of dialogue'. *Behavioral and Brain Sciences*, **forthcoming**.
- Poesio, Massimo (1993). 'A situation-theoretic formalization of definite description interpretation in plan elaboration dialogues'. In Aczel, P., Israel, D., Katagiri, Y., and Peters, S. (eds.), *Situation Theory and its Applications*, volume 3. CSLI Publications, 339–374.
- Poesio, Massimo (1994). 'Weak definites'. In *Proceedings of the 4th Conference on Semantics and Linguistic Theory (SALT-4)*.
- Poesio, Massimo and Traum, David (1997). 'Conversational actions and discourse situations'. *Computational Intelligence*, **13**(3).
- Pollard, Carl and Sag, Ivan (1994). *Head Driven Phrase Structure Grammar*. University of Chicago Press and CSLI, Chicago.
- Prince, Ellen (1992). 'The ZPG letter: Subjects, definiteness and information status'. In Mann, W. and Thompson, S. (eds.), *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text*. John Benjamins, 295–326.
- Purver, Matthew (2001). 'SCoRE: A tool for searching the BNC'. Technical Report TR-01-07, Department of Computer Science, King's College London.

- Purver, Matthew (2002). 'Processing unknown words in a dialogue system'. In *Proceedings of the 3rd SIGdial Workshop on Discourse and Dialogue*, Philadelphia. Association for Computational Linguistics, 174–183.
- Purver, Matthew, Ginzburg, Jonathan, and Healey, Patrick (2003a). 'On the means for clarification in dialogue'. In Smith, R. and van Kuppevelt, J. (eds.), *Current and New Directions in Discourse & Dialogue*. Kluwer Academic Publishers, 235–255.
- Purver, Matthew, Healey, Patrick, King, James, Ginzburg, Jonathan, and Mills, Greg (2003b). 'Answering clarification questions'. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*, Sapporo. Association for Computational Linguistics, 23–33.
- Reinhart, Tanya (1997). 'Quantifier scope: How labour is divided between QR and choice functions'. *Linguistics and Philosophy*, **20**:335–397.
- Roberts, Craige (2002). 'Demonstratives as definites'. In van Deemter, K. and Kibble, R. (eds.), *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*. CSLI Publications, 89–136.
- Russell, Bertrand (1905). 'On denoting'. *Mind*, **14**:479–493.
- Sag, Ivan and Wasow, Thomas (1999). *Syntactic Theory: A Formal Introduction*. CSLI Publications.
- Strawson, Peter (1950). 'On referring'. *Mind*, **59**:320–344.
- Szabolcsi, Anna (1997). 'Strategies for scope taking'. In Szabolcsi, A. (ed.), *Ways of Scope Taking*. Kluwer Academic Publishers, 109–155.
- van den Berg, Martin (1996). 'Dynamic generalised quantifiers'. In van der Does, J. and van Eijck, J. (eds.), *Quantifiers, Logic and Language*, Studies in Linguistics and Philosophy. CSLI Publications, 63–94.
- van der Does, Jaap and van Eijck, Jan (1996). 'Basic quantifier theory'. In van der Does, J. and van Eijck, J. (eds.), *Quantifiers, Logic and Language*. CSLI Publications, 1–45.
- van der Sandt, Rob (1992). 'Presupposition projection as anaphora resolution'. *Journal of Semantics*, **9**:333–377.
- van Rooy, Robert (2000). 'The specificity of indefinites'. In *Proceedings of the 1998 Budapest Workshop on Indefinites*.
- von Heusinger, Klaus (2000). 'The reference of indefinites'. In von Heusinger, K. and Egli, U. (eds.), *Reference and Anaphoric Relations*, number 72 in Studies in Linguistics and Philosophy. Kluwer Academic Publishers, 247–265.
- von Heusinger, Klaus (2002). 'Specificity and definiteness in sentence and discourse structure'. *Journal of Semantics*, **19**(3):245–274. Special Issue on Specificity.
- Zimmerman, Thomas E. (1999). 'Remarks on the epistemic rôle of discourse referents'. In Moss, L., Ginzburg, J., and de Rijke, M. (eds.), *Logic, Language and Computation: Volume 2*, number 96 in CSLI Lecture Notes. CSLI Publications, 346–368.