How UG-Provided Conceptual Structure Restricts the Possibilities for Quantification in Natural Language

David Goss-Grubbs

davidgg@u.washington.edu

1 Introduction

A common way of understanding quantification in natural language is as relations between sets. But it seems that only a highly constrained subset of such possible relations is actually grammaticized in the world's languages. In this paper I propose that at the level of Conceptual Structure, there are very few primitive quantifying relations, and that the narrow range of quantification attested in natural languages follows from the properties of those relations.

In the first section, I review the relational approach to quantification, and how the quantification attested in natural languages is much more constrained than what the theory would potentially allow for. Then, I introduce the Conceptual Structure relations AbsQuant and RelQuant, and show how they account for (at least a large fragment of) the quantifying determiners of English.

In the next section, I use these Conceptual Structure relations to account for various phenomena found cross-linguistically. I offer an explanation for the markedness of distributive-key universals, noted in Gil (1995); the violation of the principle of Quantity by the Dutch determiner *sommige*, noted in de Hoop (1995); and the ambiguity of noun phrases in Warlpiri noted in Bittner and Hale (1995).

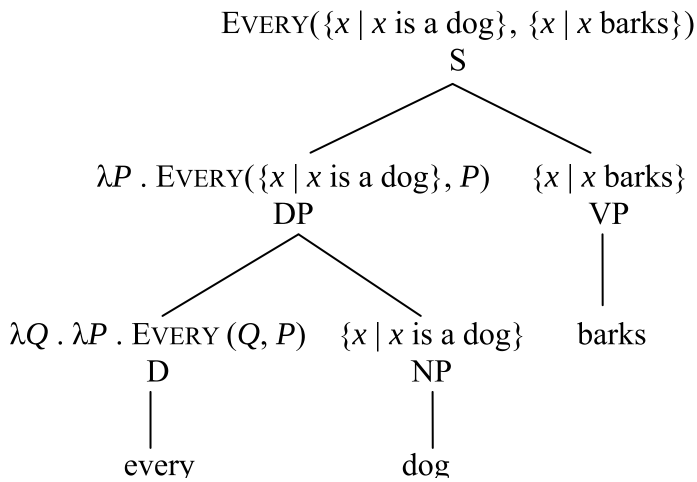
2 The Relational Approach to Quantification

The denotations of determiners (in a language like English) can be viewed as relations between sets. The denotation of a sentence like (1a) is built up compositionally as shown in (1b).

 University of Washington Working Papers in Linguistics, vol. 24 (2005), eds. Daniel J. Jinguji and Steven Moran, pp 15-29, Seattle, WA

(1) a. Every dog barks.

b.



A pair of sets is in the relation EVERY just in case the first set is a subset of the second set. Thus, (1a) is true just in case the set of dogs is a subset of the set of barkers, which is indeed what the sentence seems to mean.

What does a relation between sets look like? In a universe that contains exactly two individuals, *a* and *b*, the extension of EVERY is shown in (2):

(2) $\{\langle\{\,\}, \{\,\}\rangle, \langle\{\,\}, \{a\}\rangle, \langle\{\,\}, \{b\}\rangle, \langle\{\,\}, \{a, b\}\rangle,$
$\langle\{a\}, \{a\}\rangle, \langle\{a\}, \{a, b\}\rangle,$
$\langle\{b\}, \{b\}\rangle, \langle\{b\}, \{a, b\}\rangle,$
$\langle\{a, b\}, \{a, b\}\rangle\}$

It is a set of pairs of sets. As long as the first set in the pair is a subset of the second set, that pair will be in the extension of EVERY.

As another example, the extension of SOME is shown in (3):

(3) $\{\langle\{a\}, \{a\}\rangle, \langle\{a\}, \{a, b\}\rangle,$
$\langle\{b\}, \{b\}\rangle, \langle\{b\}, \{a, b\}\rangle,$
$\langle\{a, b\}, \{a\}\rangle, \langle\{a, b\}, \{b\}\rangle, \langle\{a, b\}, \{a, b\}\rangle\}$

As long as the two sets have a non-empty intersection, the pair is in the denotation of SOME. Thus, a sentence like (4) will be true just in case the set of dogs and the set of barkers have a non-empty intersection:

(4) Some dog barks.

There are quite a few logically possible determiner meanings of this type. In a universe containing *N* individuals, there are 2^N possible sets of individuals. Thus there are $(2^N)^2$ possible pairs of such sets. Finally, that means there are $2^{(2^N)^2}$ possible sets of such pairs. In other words, in a universe containing just 2 individuals, there are 2^{16} or 65,536 logically possible determiner denotations. Add one more individual, and it increases to 2^{64} possible sets of pairs of sets.

Of course, in reality we only find quite a small number of quantifiers. It has often been observed (e.g. de Swart 1998) that the determiners that actually show up universally (or near-universally) exhibit these properties:

Conservativity: A determiner meaning *Det* is conservative if, for any two sets *A* and *B*, *Det*(*A*, *B*) is true whenever *Det*(*A*, $A \cap B$) is true. That is, for a conservative determiner, we don't care about things in *B* that aren't also in *A*.

Extension: A determiner meaning *Det* exhibits extension if, for any two sets *A* and *B*, if *Det*(*A*, *B*) is true in one model, then it will be true in any model with identical sets *A* and *B*. That is, for determiners that exhibit extension, we don't care about things that are in neither *A* nor *B*.

Quantity: A determiner meaning *Det* exhibits quantity if, for any two sets *A* and *B*, *Det*(*A*, *B*) is true in one permutation of the universe whenever it is true in any other permutation. That is, for determiners that exhibit quantity, we don't care *which* things are in the sets, we only care how many.

So for any conservative determiner *Det* that exhibits extension and quantity, the truth value of a sentence like (5) will depend only on the number of dogs and the number of barking dogs.

(5) *Det* dog(s) bark(s).

Barwise and Cooper (1981) propose universals to the effect that all languages have essentially quantificational NPs; and that they all have determiners whose denotations are as I have just described them.

As far as I can tell, these universals are just descriptions of what we observe in human language, but are not explained by anything in UG.

3 Accounting for the Apparent Universals

My hypothesis is that relations like EVERY and SOME are not primitives at the level of Conceptual Structure. Rather, UG provides two very general relations, from which can be built up all and only the sorts of quantifying relations that we actually find in the world's languages. These relations are listed in (6).

(6) a. ABSQUANT(A, B, N)
b. RELQUANT(A, B, X)

The ABSQUANT relation (for "absolute" quantification) is a relation between three things: a set *A*, a set *B*, and an integer *N*. The relation holds whenever the intersection of *A* and *B* contains *N* members. This relation handles the so-called weak[1] determiners, like *some*, *many*, and the numerals. So at the level of Conceptual Structure, the representation for (4) would be (7).

(7) ABSQUANT(A, B, N) $\wedge$ GREATER(N, 0) $\wedge$ DOG(A) $\wedge$ BARK(B)

The sentence is true if the first set (the set of dogs) and the second set (the set of barkers) have at least one element in common.

The RELQUANT relation (for "relative" quantification) is also a relation between three things: a set *A*, a set *B*, and a real number *X* between 0 and 1. This relation holds whenever the cardinality of $A \cap B$ divided by the cardinality of *A* equals *X*. This relation handles the so-called strong determiners, like *every* and *most*. At the level of Conceptual Structure, the (a) sentences are true whenever the conceptual structures in their (b) counterparts are true.

(8) a. Every dog barks
b. RELQUANT(A, B, 1) $\wedge$ DOG(A) $\wedge$ BARK(B)

[1] The interaction between ABSQUANT, RELQUANT and the judgments that give rise to weak vs. strong determiners is an interesting question, but is beyond the scope of this paper.

(9) a. Most dogs bark

b. RELQUANT(*A*, *B*, *N*) ∧ GREATER(*N*, 0.5) ∧ DOG(*A*) ∧ BARK(*B*)

Sentence (8) is true if the number of barking dogs divided by the number of dogs is 1. That is, it is true whenever every dog is in the set of barking dogs. Sentence (9) is true if the number of barking dogs is more than half the number dogs generally.

Other determiners require both these relations in order to get their denotations. I propose the (a) sentences in (10) - (12) are associated with the (b) conceptual structures.

(10) a. The dog barks.

b *content*: RELQUANT(*A*, *B*, 1) ∧ DOG(*A*) ∧ BARK(*B*)

presupposition: ABSQUANT(*A*, *E*, 1)

(11) a. The dogs bark.

b *content*: RELQUANT(*A*, *B*, 1) ∧ DOG(*A*) ∧ BARK(*B*)

presupposition: ABSQUANT(*A*, *E*, N) ∧ GREATER(*N*, 1)

(12) a. Both dogs bark.

b. *content*: RELQUANT(*A*, *B*, 1) ∧ DOG(*A*) ∧ BARK(*B*)

presupposition: ABSQUANT(*A*, *E*, 2)

That is, definites are the same, at the level of Conceptual Structure, as universals with the extra presupposition of an absolute quantification, where the second set is *E*, the set of all individuals. In effect, the presupposition is just specifying the cardinality of the first set. Singular *the* says there is one thing in the set; plural *the* says there is more than one thing in the set, and *both* says that there are exactly two things in the set.

Note that both ABSQUANT and RELQUANT are conservative. That is, they don't care about things in the second set that are not in the first set. In fact, ABSQUANT doesn't even care about things in the first set that are not in the second set. They both exhibit extension in that properties of things outside of either set cannot have any effect on whether those relations hold. They both exhibit quantity in that they are comparing only the cardinalities of sets.

In the rest of the paper, I will use this analysis of absolute and relative quantification to account for various phenomena in the quantification of natural langauges.

4 Distributive Determiners

Gil (1995) proposes a two-way distinction among types of universal quantification. They are *simple* universals and *distributive-key* universals.[2] An example from English is simple *all* versus distributive-key *every*, as seen in (13) - (14).

(13) a. All men gathered at dawn.
b. * Every man gathered at dawn.

(14) a. All men carried three suitcases.
b. Every man carried three suitcases.

Simple determiners can take a variety of scope relations; distributive-key determiners require distributive readings. So (13a) is fine with the collective predicate *gather*, where (13b) is bad. (14a) allows for the possibility that the men carried three suitcases each or that they carried a total of three suitcases between them. (14b) requires that they carried three each.

Many languages make the same distinction. Gil lists examples from Georgian, Tagalog, Russian, Turkish, Lezgian, and Mandarin.

Gil argues persuasively for the position that simple quantifiers are primitive, and that distributive-key quantifiers are portmanteaux which combine a simple quantifier plus some kind of additional information. Thus, distributive-key quantifiers are marked, and will be found only in languages that also have simple quantifiers.

Gil offers a number of kinds of evidence to show that simple quantifiers are basic, and distributive-key quantifiers are marked. Non-distributive readings are preferred even when distributive readings are available. In (15), the most natural reading is that the two men carried three suitcases between them (perhaps one man carried two suitcases, and the other man carried one), rather than two men each carrying three suitcases, or there being three suitcases that each of the two men carried.

---

[2] He also talks about *distributive-share* universals, but space limitations prevent me from giving them a treatment here.

(15) Two men carried three suitcases.

This shows that there is something less natural (more marked) about the distributive reading than the non-distributive reading.

Languages often have constructions that elaborate on simple quantifiers to create a distributive reading. For instance, English (16) uses the word *apiece* to enforce a distributive reading. Maricopa (17) marks the verb with the suffix *-xper*. Turkish (18) marks the numeral in the narrow-scope NP with a suffix *-er*. Tagalog (19) uses an additional distributive-key quantifier within the narrow-scope NP. All examples are Gil's.

(16) Two men carried three suitcases apiece

(17)

ʔipač	xvikk	ʔii	xmokm	paayperšík
man	Two	stick	three	carry-DIST.SHARE-DUAL-REAL

'Two men carried three sticks apiece'

(18)

Iki	adam	üçer	bavul	tasıdı
two	man	three-DIST.SHARE	suitcase	carry

'Two men carried three suitcases apiece'

(19)

Nagdala	ng	bawat	tatlong	maleta	ang	dalawang	lalaki
carry	DIR	all-DIST.KEY	three	suitcase	TOP	two	man

'Two men carried three suitcases apiece'

Again, this shows that distributive quantification is marked.

The distributive-key universal quantifiers typically only appear with count nouns, whereas the simple universals can appear with count nouns or mass nouns. For instance, English *all* can appear with any word that *every* can appear with, as well as with mass nouns, which *every* cannot appear with. The more restricted environment that distributive-key universals can appear in supports the idea that they are marked.

Gil proposes another universal: that distributive-key quantifiers are all universal quantifiers. He gives these English examples.

(20) Two/some/many men { gathered at dawn. / carried three suitcases. }

These determiners work fine with the collective predicate *gather*, and they have quite natural non-distributive readings. Furthermore, English contains no distributive-key counterparts for these determiners. He supports this putative universal with further data from Russian, Turkish, Georgian, Punjabi, Tagalog, and Mandarin.

Gil doesn't try to account for why these patterns might hold. I propose that scope relations are, by default, not explicitly represented at the level of Conceptual Structure. Instead, the normal situation is that the algorithm that determines truth-conditional interpretations from conceptual structures must make choices with respect to scope that are underdetermined by the conceptual structure itself.

For instance, sentences with simple universals, such as English (14a) would get a conceptual structure something like this[3][4]:

(21) $\textsc{Carry}(A, B) \wedge$
$\textsc{RelQuant}(C, A, 1) \wedge$
$\textsc{Man}(C) \wedge$
$\textsc{AbsQuant}(D, B, 3) \wedge$
$\textsc{Suitcase}(D)$

The algorithm for determining truth-conditional interpretations is free to take either quantificational relation as primary or to take them as equal, with the default being to take them as equal. The three options correspond to these three readings:

3 To be complete, I need to spell out in detail the algorithm for how to get from structures like (21) to truth values. This is beyond the scope of this paper.

4 The representation of 'bare plural' noun phrases is an interesting question, but is not addressed in this paper.

(22) a. $\text{R}_{\text{EL}}\text{Q}_{\text{UANT}}(\{x \mid x \text{ is a man}\}, \{y \mid \text{A}_{\text{BS}}\text{Q}_{\text{UANT}}(\{x \mid x \text{ is a suitcase}\}, \{z \mid y \text{ carried } z\}, 3)\}, 1)$

b. $\text{A}_{\text{BS}}\text{Q}_{\text{UANT}}(\{x \mid x \text{ is a suitcase}\}, \{y \mid \text{R}_{\text{EL}}\text{Q}_{\text{UANT}}(\{x \mid x \text{ is a man}\}, \{z \mid z \text{ carried } y\}, 1)\}, 3)$

c. $\text{R}_{\text{EL}}\text{Q}_{\text{UANT}}(\{x \mid x \text{ is a man}\}, \{x \mid x \text{ carried suitcases}\}, 1) \wedge$
$\text{A}_{\text{BS}}\text{Q}_{\text{UANT}}(\{x \mid x \text{ is a suitcase}\}, \{x \mid \text{men carried } x\}, 3)$

On the other hand, sentences with distributive-key universals, like English (14b), would get the same conceptual structure but with the additional information that specifies a particular scoping.[5] The reading that this additional information requires is (22a). So a language is free to have simple universal quantifiers that do not specify this additional information, but if they have structures with this additional information, they are sure to have structures without it.

I still have to answer why distributive-key universals cause the algorithm to make the choices that it does. For instance, it is only when the universally quantified NP is in the subject position that it demands wide scope.

(23) a. Every man carried three suitcases

b. Three suitcases were carried by every man

In (23a), *every man* demands wide scope. In (23b), either scope is available. In any case, one must take scope over the other. We cannot get the three-suitcases-between-them reading.

It also appears as if I may have a problem accounting for why distributive-key determiners are always universals. But really it seems like English *most* is also distributive-key. In (24), *most men* strongly prefers wide scope. At the least, the equal-scope reading is unavailable.

(24) Most men carried three suitcases.

---

[5] For the purposes of this paper, it's not important how this information is expressed.

So it looks like the scope-forcing information can live on any RELQUANT determiner, not just the universals.

5 Dutch *Sommige*

De Hoop (1995) examines the Dutch determiner *sommige*, which is often glossed as English 'some (of the)' or 'certain'. Sentence (25) means that there is a set of unicorns, characterized by some quality, known to the speaker but not necessarily to the hearer, and that the unicorns in that set are white.

(25) Sommige eenhoorns zijn wit.
some unicorns are white.
'Certain unicorns are white.'

This differs from Dutch *enkele*, which is the plain existential quantifier. Sentence (26) merely means that the number of unicorns that are white is greater than zero.

(26) Enkele eenhoorns zijn wit.
some unicorns are white.
'Some unicorns are white.'

Evidence that *sommige* is truly different from *enkele* includes the following pairs of sentences.

(27) a. Er bestaan enkele witte eenhoorns
there exist some white unicorns
'There exist some white unicorns'

b. * Er bestaan sommige witte eenhoorns.
there exist some white unicorns

(28)	a.		Ik	heb	gisteren	enkele	kilometers	gereisd
			I	have	yesterday	some	kilometers	traveled
			'I traveled some kilometers yesterday'					
	b.	*	Ik	heb	gisteren	sommige	kilometers	gereisd
			I	have	yesterday	some	kilometers	traveled

(27) shows that while *enkele* can appear in *there*-sentences, which normally allow only NPs with weak determiners, *sommige* cannot. In (28), the idea is that traveling takes a contiguous sequence of kilometers. It takes a phrase indicating some distance, here measured in kilometers. It does not take (every member of) some subset of the kilometers.

So *sommige* appears not to exhibit the property of Quantity. The truth of (25) depends not just on the number of unicorns and the number of white unicorns. Rather, its truth depends on *which* unicorns are white. How does my proposed system of absolute and relative quantification deal with *sommige*?

I propose the conceptual structure in (29a) for sentence (25), with the semantic interpretation in (29b).

(29) a. $\text{RelQuant}(A, B, 1) \wedge$
$\text{Unicorn}(A) \wedge$
$\text{Salient-Rel}(A) \wedge$
$\text{White}(B)$

b. $\{x \mid x \text{ has the salient quality}\} \cap \{x \mid x \text{ is a unicorn}\} \subseteq \{x \mid x \text{ is white}\}$

That is, *sommige* not only introduces a RelQuant, (like English *all* does), but it also introduces some contextually salient predicating relation. So the actual quantification is still accomplished through means that exhibit the property of quantity, it is just that there is an extra predication thrown into the mix as well.

So my proposal doesn't rule out determiners that don't exhibit quantity, but it does predict that they should be less common than those that do. That is, by default, a determiner will only introduce a quantifying relation (AbsQuant or RelQuant) or other functional relation (e.g. GREATERTHAN) into conceptual structure. But some determiners may, in addition, introduce a predicational relation as well.

6 NP Ambiguity in Warlpiri

Bittner and Hale (1995) argue that Warlpiri has just two major syntactic categories: Noun and Verb. These two classes of words are easily distinguished from one another on the basis of morphology. The main predicate in a sentence may be expressed either by a noun or a verb. Verbs are primarily active and nouns are primarily stative. Nouns can also serve as arguments of predicates, in which case they exhibit pronominal agreement, or as secondary predicates, in which case they exhibit adjective-like agreement.

Bittner and Hale list the following uses for Warlpiri nominals, in order from most argument-like to most predicate-like:

(30) a. Pronouns, demonstratives and other indexicals
b. Names
c. Common nouns
d. Expressions of quality or cardinality
e. Expressions of psychological states
f. Locatives and directionals

An overt NP can be a single noun, or can be constructed by putting together a head noun and one or more modifiers, as in (31). The elements of the NP need not be contiguous.

(31) Maliki wiri-ngki
dog big-ERG
'a/the big dog'

The syntax is the same, no matter what sort of nominals are used. So the single expression of cardinality *jirrima* can be a noun phrase meaning 'two (of them)' or 'the two (of them)'. Determiners do not exist as a separate syntactic category. The noun phrase (31) can either get the weak reading 'a big dog' or the strong reading 'the big dog'.

Bittner and Hale argue that this ambiguity carries over to nouns which are expressions of cardinality. For instance, the word *panu* can be either the weak ‘many’ or the strong ‘all (of them)’. The same sort of syntactic devices are used to narrow down the choice between ‘a dog’ and ‘the dog’ as are used to narrow down the choice between *panu* ‘many’ and ‘all’ or the choice between *jirrima* ‘two’ and ‘both’.

The weak reading can be forced with the suffix -*kari*. Example (32) shows this for common nouns. Example (33) shows this for expressions of cardinality. All examples here are Bittner and Hale’s.

(32)	Jarntu-**kari**	∅-∅	parnka-ja	yatijarra,	jarntu-**kari**	kurlirra.
	dog-**KARI**	PRF-3s	run-PST	north,	dog-**KARI**	south

‘A dog ran north, a dog (ran) south.’

(33)	Panu-**kari**	ka-rna-jana	nya-nyi	panu-**kari**	∅-li	wurulyya-nu.
	many-**KARI**	PRS-1s-3p	see-NPST	many-**KARI**	PRF-3p	hide-PST

‘I see a large group, (but) a large group went into hiding.’

When a noun appears with an obligatorily definite nominal, such as a demonstrative (in bold below), the strong reading is forced. Example (34) shows this for common nouns. Example (35) shows this for expressions of cardinality.

(34)	**Yalumpu**-rra	ka-rna-jana	pura-mi	jarntu
	that-PL	PRS-1s-3p	follow-NPST	dog

‘I am following those dogs.’

(35)	**Yalumpu**-rra	ka-rna-jana	pura-mi	panu
	that-PL	PRS-1s-3p	follow-NPST	many

‘I am following that large group.’

How does this fit with the system I am proposing here? I assume that in the absence of any explicit expression of cardinality, Warlpiri noun phrases just are supplied an ABSQUANT relation, which gives the weak reading. Under the right circumstances, universal-force RELQUANT relations can be added as well, which gives the strong reading. So (36a) gets a conceptual structure like (36b).

(36) a. Maliki wiri-ngki ka-Ø-ju wajilipi-nyi
dog big-ERG PRS-3s_1-1s_2 chase-NPST
'A/the big dog is chasing me'

b. ABSQUANT(*A*, *B*, 1) ∧
BIG(*A*) ∧
DOG(*A*) ∧
CHASE(*B*, *me*)
[∧ RELQUANT(*A*, *B*, 1)]

The RELQUANT predication is optional. With it, you get the definite reading. Without it, you get the indefinite reading.

If a noun phrase contains an explicit expression of cardinality, it is used in an ABSQUANT relation. As before, an optional RELQUANT may be added. An example is sentence (37).

(37) a. panu ka-rna-jana nya-nyi
many PRS-1s-3p see-NPST
'I see a/the large group (of them)'

b. ABSQUANT(*A*, *B*, *N*) ∧
LARGE(*N*) ∧
SEE(*me*, *B*) ∧
[∧ RELQUANT(*A*, *B*, 1)]

Again, the RELQUANT part is optional. If it is added, it changes 'a large group' to 'the large group', which is often just glossed as 'all (of them)'.

7 Conclusion

I have proposed that at the level of Conceptual Structure, UG provides just a very few primitive operators that can be used for quantification. The range of quantification that is possible using them is highly constrained, yet seems to account for the range of quantification that is actually found. I used this system to address various phenomena in quantification cross-linguistically, including the markedness of distributive-key universal quantifiers; determiners that do not exhibit quantity, such as Dutch *sommige*; and the ambiguity of noun phrases in Warlpiri, which do not use determiners for quantification at all.

References

Barwise, Jon and Robin Cooper. 1981. Generalized Quantifiers and Natural Language. *Linguistics and Philosophy*, 4: 159–219.

Bittner, Maria and Kenneth Hale. 1995. Remarks on definiteness in Warlpiri. *Quantification in Natural Languages*. Ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee. Dordrecht: Kluwer Academic Publishers, 81–105.

Gil, David. 1995. Universal quantifiers and distributivity. *Quantification in Natural Languages*. Ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee. Dordrecht: Kluwer Academic Publishers, 321–362.

de Hoop, Helen. 1995. On the characterization of the weak-strong distinction. *Quantification in Natural Languages*. Ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee. Dordrecht: Kluwer Academic Publishers, 421–450.

de Swart, Henriëtte. 1998. *Introduction to Natural Language Semantics*. CSLI Publications, CSLI lecture notes ; no. 80.