Computing Genitive Superlatives

Johan Bos University of Rome "La Sapienza" bos@di.uniroma1.it

Superlative expressions (such as "the tallest boy") single out entities by comparing them to others with respect to a set of properties. Therefore, a proper semantics for a superlative expression requires a correct computation of the comparison set. In most of the cases, this can be done in a compositional fashion (leaving contextual influences aside). When superlative adjectives occur in prenominal genitive constructions (such as "his most expensive car", this isn't so straightforward anymore. Three solutions are proposed, that all change the underlying syntactic structure of the genitive construction in order to arrive at a correct comparison set. The solution where ordinary nouns are analysed "relationally" is preferred, because it requires the least number of modifications of the grammar. Under certain conditions, superlatives also exhibit so-called comparative readings (as opposed to absolute readings). With the current machinery it is however impossible to account for such readings in a compositional way.

1 Introduction

The superlative construction singles out entities as having some property to the greatest (or smallest) degree. In English, superlative adjectives are normally preceeded by the definite article or a possessive; one-syllable adjectives and some two-syllable adjectives end in -est, some words of two syllables and all words of three or more syllables are preceeded **most** (or **least**). Here are some examples (superlatives underlined):

- (1) ... the <u>tallest</u> boy ...
- (2) ... AS Roma's quickest player ...

(3) ... the Big Board's <u>most respected</u> floor traders ...

An adequate semantics of the superlative construction presupposes a proper comparison set — the entities that are compared. In this paper we discuss the interaction between superlative adjectives and genitive constructions and their effect on the construction of the comparison set. We argue that the mainstream analysis of superlatives in prenominal genitive expressions is troublesome.

We provide the various syntactic analyses in Combinatory Categorial Grammar (CCG, [Ste01]) with Discourse Representation Theory [KR93] functioning as model-theoretical framework. We demonstrate the syntax-semantics interface for a sentence with simple superlative adjective in Section 2 and then show why possessives complicate the formation of a correct comparison set in Section 3. We discuss so-called comparative readings in Section 4 and show what role the genitive plays in blocking certain interpretations.

2 Superlatives and the Syntax-Semantics Interface

It is well known that superlatives can be analysed in terms of comparative constructions [Wal72]. Accordingly, "the R-est X" denotes an X such that there is no Y different from X that is R-er. This is also the approach we will follow in our theory of superlatives, in which we use CCG (Combinatory Categorial Grammar, [Ste01]) to provide us with a syntactic analysis, and DRT (Discourse Representation Theory, [KR93]) to produce semantically interpretable structures.

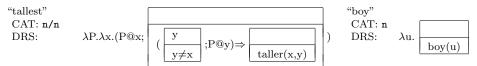
In what follows we will illustrate the formalism with an example. We will give a CCG derivation and introduce some of the relevant terminology. Then we will show how the syntax-semantics interface works by compositionally constructing a semantic representation for a sentence containing a superlative.

The version of CCG that we use in the context of this article comprises four atomic categories (s, np, n, pp), functor categories such as np/n (determiner) and $(s\np)/np$ (transitive verb) and the combinatory rules [fa] (forward application), [fc] (forward composition), and [ba] (backward application). We indicate lexical categories by [lex] in derivations. Note that slashes indicate directionality: X/Y denotes a functor category looking for an argument of category Y on its *right*, yielding category X; X\Y is a functor category looking for an argument Y on its *left*. Now consider a basic example of a sentence with a simple superlative adjective, and its analysis in CCG:

			tallest	boy
			[lex]	[lex]
		the	n/n	n
		[lex]		[fa]
	saw	np/n	n	
	[lex]			[fa]
I	(s\np)/np	np		
[lex]				[fa]
np	s\np			
				[ba]
S				

On the basis of the syntactic structure that CCG gives us we can build a semantic representation in a bottom-up fashion. We do so by employing a version of lambda-DRT [Mus96, vEK97]: all lexical entries are associated with lambda-DRSs. The number of abstracted variables mirrors the syntactic type of the CCG category.¹

The derivation given above tells us that we combine "tallest" and "boy" first, under guidance of the [fa] rule. Consider first the lambda-DRS for each of these words separately (we use the ; operator to indicate a merge (signifying dynamic conjunction) between DRS, and we take the @ symbol to mean function application):



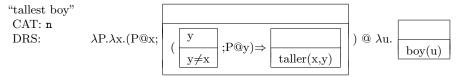
The superlative introduces a complex condition, stating that whatever entity is selected with property P, it is taller than all other entities with property P. We simply assume that superlatives introduce a two-place ordering relation, a relation which can be characterised as asymmetric and transitive.² In our example "tallest" introduces taller(x,y), a relation that orders entities in the domain with respect to their size. (It is a transitive relation, because

¹In terms of Van Eijck and Kamp's semantic type system [vEK97], with the basic types e for entities and T for state transitions, the category **s** corresponds to T, **n** and **pp** correspond to $\langle e, T \rangle$, and **np** to $\langle \langle e, T \rangle, T \rangle$. So there is strict type-transparancy between the syntactic types of CCG and the semantic types of lambda-DRT expressions.

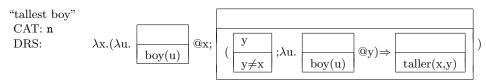
 $^{^{2}}$ A *bad* semantics, in Van Benthem's view [vB83], because it wouldn't explain and only describe the behaviour of comparatives. That might be true from a philosophical point of view — from a computational semantics perspective it is quite practical.

if X is taller than Y, and Y is taller than Z, then X is taller than Z; it is asymmetric, because if X is taller than Y, then it is not the case that Y is taller than X.)

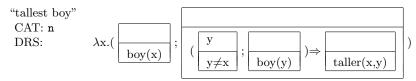
Going back to our example, the interpretation of the [fa] rule tells us that the lambda-DRS of the functor category should be applied to the lambda-DRS of the argument category. So we get this as intermediate result:



If we carry out the substitutions (for the lambda-bound variable P) in the process of β -conversion we obtain the following lambda-DRS:



Once more we can perform β -conversion (for the lambda-bound inner variables u) yielding the following lambda-DRS:



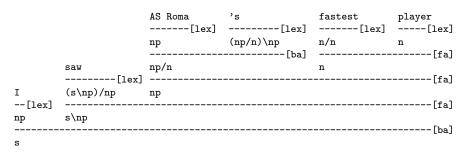
Finally we can carry out merge-reduction (the process of joining two DRSs into one, by taking the unions of the discourse referents and the conditions, respectively) and arrive at a DRS which meaning can be glossed as "the boy which is taller than every other boy". In this example, the comparison set of the superlative expression is just the set of boys. Its semantic scope is established in the antecedent of the DRS conditional introduced by the lexical semantics of the superlative.

3 Superlatives in Genitive Constructions

3.1 The Problem

As shown in the previous section, the analysis given to a superlative adjectives seems to be satisfactory: it can be done in a compositional way, and it yields the desired interpretation. However, as noted in [BN06], this analysis does not easily carry over to superlative expressions that combine with prenominal genitive expressions (constructions marked by the clitic -'s or possessive pronouns), such as "AS Roma's fastest player", "his least expected answer", and "London's most expensive restaurant". Why is this the case?

Let's consider the comparison set of AS Roma's fastest player, which is *not* just a set of players, but the set of players from AS Roma. It seems here that the superlative adjective out-scopes the possessive noun phrase. But this is not the effect that we get if we follow the standard analysis in a categorial grammar [Ste01, Hoc03, Car97]:



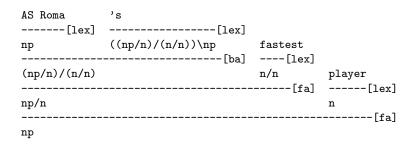
It is not difficult to see why. Once we've analysed the sub-phrase "fastest player", we've basically fixed the scope of the comparison set in the lambda-DRS. The antecedent DRS of the implicational condition, representing the comparison set, is completed, and there is nothing we can do about it anymore — unless you're in for some ad-hoc post-processing activities. We argue that the syntactic analysis for phrases such as the above is insufficient to get a correct semantic analysis, and therefore any compositional solution will involve work at the level of the syntax-semantics interface.

3.2 Possible Solutions

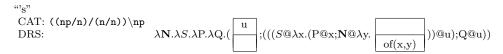
Within the framework of CCG, we present three ways to change the analysis to reach a proper treatment of prenominal genitives that interact with superlatives. The first is Bos & Nissim's solution to the problem [BN06]. The second and third are inspired by work of Partee and Carpenter, respectively, who introduce different interpretations for the genitive to account for the ambiguities that appear with inherently relational nouns — solutions that also seem to work well for our problem with genitive superlatives.

3.2.1 Proposal I: Bos & Nissim 2006

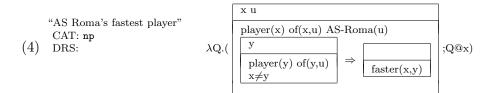
Bos & Nissim note the problem only manifests itself with prenominal genitives [BN06]. To fix it, they add a new syntactic (and semantic) type for the prenominal possessive. A way of doing this is to recognise phrases such as "his most expensive" and "AS Roma's fastest" as determiner-like phrases, with CCG category np/n. Under this view, let's consider our problematic case again.



Now consider the lambda-DRSs for the different lexical items in this example. The only new category is the one for the clitic -'s, which is given the following lexical semantics:



That is, the lambda-DRS for the possessive takes three arguments: the noun phrase, a superlative adjective, and a noun. This results in a representation for "AS Roma's fastest player" shown as the lambda-DRS in (4).



This analysis yields the desired comparison set: we quantify over "players of AS Roma", not just 'players". But arguably, the solution involves overloading the possessive construction with a new CCG category that only is used in superlative constructions, lacking independent motivation. So let's look at an alternative.

3.2.2 Proposal II: Partee 1997

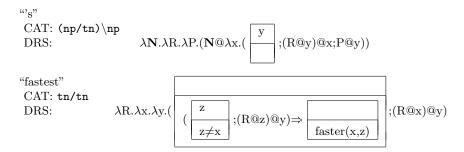
An alternative way of computing the comparison set for superlatives in genitive expressions builds on Partee's analysis for relational nouns [Par97, PB99]. Partee enriches the syntax by including a "transitive common noun phrase", along the usual common noun phrases. As a result she also distinguishes between two types of determiner in the genitive: one that selects transitive nouns, and one that selects ordinary nouns. Semantically, the former receives an inherently relational interpretation (expressed by a two-place relation), and the latter a free interpretation, filled in by the context.

We argue that almost any noun can be interpreted "relationally" when it is part of a genitive. By making the noun relational, we can leave the creation of the comparison set to the clitic -s, by giving it a category that subcategorises for a relational noun. This ensures that all relevant material enters the comparison set, as we will see in the example below.

In order to achieve this, we introduce a new basic CCG category tn (transitive noun, Partee's TCN) and an additional category tn/tn for superlative adjectives. As in Partee's grammar, we introduce an additional category for the clitic -'s as well. Hence, superlative adjectives can be assigned two categories: n/n if they modify non-relational nouns, and tn/tn if they modify relational nouns. This gives us the following Partee-style analysis for AS Roma's fastest player:

AS Roma	's	fastest	player
[lex]	[lex]	[lex]	[lex]
np	(np/tn)\np	tn/tn	tn
	[ba]		[fa]
np/tn		tn	
			[fa]
np			

Since we have several new CCG categories here, we need to propose corresponding lambda-DRSs too. (We assign the syntactic category tn the semantic type $\langle e, \langle e, T \rangle \rangle$.)



"player"		
CAT: tn		
DRS:	$\lambda x. \lambda y.$	
	-	player(x,y)

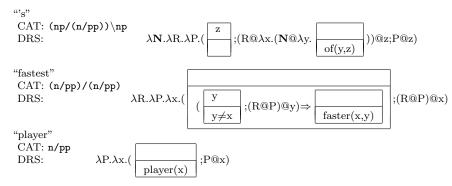
The lambda-DRS that will result from the derivation given above is tantamount to the one presented in (4). So this proposal computes the correct comparison set, but at the cost of introducing a new atomic CCG category. Could we have avoided this?

3.2.3 Proposal III: Carpenter 1997

The third propopal resembles the second, but it doesn't introduce new basic CCG categories. It adopts the idea of Partee to distinguish nouns from relational nouns, but instead of giving it a new basic category, we give it the functor category n/pp, inspired by Carpenter [Car97], who introduced a second lexical category for the possessive that "consumes a nominal complement that itself seeks a complement". In Carpenter's type-logical framework this interpretation of the possessive receives the category n/(n/np:of). Translated into CCG we get the following CCG derivation for our running example:

AS Roma	's	fastest	player
[lex]	[lex]	[lex]	[lex]
np	(np/(n/pp))\np	(n/pp)/(n/pp)	n/pp
	[ba]		[fa]
np/(n/pp)		n/pp	
			[fa]
np			

There are no new basic CCG categories (the category pp is used independently in CCG for prepositional complements [Ste01]), but the structure of the derivation is identical to that of Proposal II. Semantically, however, it is slightly different, and in a way closer to Proposal I, because the possessive relation is introduced by the clitic -'s and not by the relational noun:



3.3 Evaluation

Which proposal is better? The first proposal changes the structure of the derivation for basic superlatives as given in the previous section, but doesn't introduce new basic categories. The second and third proposal have the same syntactic structure as the original CCG analysis; but the former does so at the cost of introducing a new basic CCG category. Arguably, the third proposal has a more intuitive semantics. In sum, Proposal III seems the best choice for analysing superlatives in possessive constructions. As an added bonus, a reformulation of the postnominal genitive construction will yield a semantic representation tantamount to the prenominal one:

			of [lex]	AS Roma [lex]
the	fastest	player	pp/np	np
[lex]	[lex]	[lex]		[fa]
np/n	n/n	n/pp	рр	
	[fc]			[fa]
np/n		n		
				[fa]
np				

Furthermore, under Proposal II and III one could argue that the prenominal genitive *always* subcategorises for a relational noun, independently from the presence of superlatives. This would result in only one syntactic category for the clitic -'s, rather than the two categories we need to distinguish between relational and non-relational nouns.

Finally, we are well aware that analysing the possessive relation for a noun N as "N(x) & of(x,y)", as in Proposal I and III, is problematic for various reasons. First of all, it doesn't explain what kind of relation there is between the two entities. Secondly, it will give sometimes the wrong predictions; we would wrongly predict that "Vincent's boss is Mia's husband" entails that Mia's boss is Vincent's husband!³ It is a problem independent from the objectives of the paper. A solution would be to analyse phrase such as "husband of Mia" as $\lambda x.[person(x) \& \exists y(role(x,y) \& husband(y) \&$ of(y,mia))], paraphrased as the "person who is playing the role of Mia's husband". However in the context of this paper we will analyse the possessive using the simplified (but semantically inadequate) analysis.

³This was kindly pointed out to me by Steve Pulman, Yuliya Lierler, Vladimir Lifschitz and anonymous reviewers.

4 Comparative Readings

An interesting behaviour of superlative expressions is their capacity to invoke so-called comparative readings, at least under certain conditions. As the examples and discussion in this section demonstrate, this is a complex issue and we don't have a computational analysis to offer. Nevertheless we do make some interesting new observations and present some fresh data.

4.1 Comparative Readings and the Genitive

There are other linguistic means to restrict the comparison set of superlative expressions, as is well known from the formal semantics literature on superlatives [Hei99, Sza86, Gaw95, FK00]. Consider the following examples:

- (5) Callum climbed the highest mountain.
- (6) Who climbed the highest mountain?

It has been observed that (5) and (6) have two distinct interpretations: an *absolute* reading and a *comparative* one. The absolute reading of (5) is where Callum climbed a mountain higher than any other mountain (i.e. the highest mountain in the domain). The comparative reading describes a situation where Callum climbed a higher mountain than anyone else climbed (but not necessarily the highest mountain in the context).

The comparative reading is not always available, and even when it is, the absolute reading seems to be the preferred one. Hence, an issue of importance is to know under what linguistic conditions the comparative reading is licensed. As has been pointed out by [Sza86], the comparative reading requires association with focus (5) or a wh-question (6). A further observation is that a superlative expression such as "the highest mountain" has a definite interpretation giving rise to the absolute reading, or can be understood indefinitely, in which (5) can be paraphrased as "Callum climbed a mountain higher than anyone else climbed" [Sza86]. We note that this view correctly predicts that the comparative reading is not available when the superlative is part of a genitive construction (which are usually intrinsically definite), as in (7) and (8):

- (7) Callum climbed Scotland's highest mountain.
- (8) Callum climbed the highest mountain of Scotland.

There are, by and large, two different proposals to account for the comparative reading in superlative noun phrases: the movement theory, and the in-situ theory. Neither of them is easy to implement. Following the *movement* theory of superlative interpretation, the superlative quantifier -est can move out of its noun phrase to a higher position in the sentence, and then associates with focus to build the comparison set [Hei99, Sza86]. The *in-situ* theory of superlatives [Hei99, FK00, Mat05] accounts for the comparative interpretation by treating the superlative as a quantifier that is contextually restricted.

As CCG is a lexicalised theory of grammar, it is natural to adopt an in-situ approach. Let's assume that nouns can either get the non-relational, "free", interpretation (via the category n) or the relational interpretation (n/pp) as in Proposal III. The latter category will always give rise to the absolute reading, explaining the absence of the comparative reading in (7) and (8), because genitive contexts trigger the n/pp interpretation for nouns.

4.2 Comparative Readings and Relational Nouns

The transitive verb "to have" shows peculiar behaviour with inherently relational noun phrases as object [Par97]. It seems to be able to take bare relational noun phrases as argument, unlike other transitive verbs. It also blocks noun phrase objects where relational nouns are part of a genitive. Consider the following contrasting cases (stars indicate non-grammatical sentences):

- (9) John has a sister.
- (10) * John has his sister.
- (11) John has the smartest sister.
- (12) * John has his smartest sister.
- (13) * John has the smartest sister of Bill.
- (14) * John has Bill's smartest sister.

Intriguingly, the only possible reading of (11) is the comparative one: for some people with sisters, John is is the one that has a sister smarter than every other person with a sister. There is no absolute reading in which "the smartest sister" has a definite interpretation. Partee wonders whether "to have" has a second possible interpretation besides the transitive verb reading. In terms of CCG, this verb would receive the category $(s\p)/(np/pp)$ in addition to $(s\p)/np$. Assuming that this would be the case, it would be possible to come up with a semantics for "have" that would yield the desired interpretation for (11). It would be something like:

$$\lambda R. \lambda N. (N@\lambda x. ((R@\lambda y. \boxed{x} of(y,x)))@\lambda y. \boxed{have(x,y)}))$$

However, other examples prove that this would not be sufficient:

- (15) What city in Africa has the greatest population?
- (16) What golf course has the highest slope rating?
- (17) What cities have the worst automobile traffic?

In (15), for example, the members of the comparison set are populations of African cities, not populations in general. Similarly, in (16) we talk about slope ratings of golf courses, not of, say, amplifiers or agricultural vehicles; in (17) finally, the comparison set contains traffic in cities, not in countries or regions.

Given that this information is all provided within the sentence, one would expect the compositional semantics to take care of that. However, it is unclear how the in-situ theory would get a correct interpretation of the superlatives using a compositional semantics. We leave it as an open problem.

4.3 Comparative Readings and Most

The availability of the absolute and relative reading is not a phenomenon that manifests itself only with superlative adjectives. A very similar effect can be seen with the generalised quantifier most. Consider (18) and (19):

- (18) Gloria sold most cars.
- (19) Most babies are born in March.

Example (18) has besides the absolute reading (for instance, in a situation with twenty cars, and Gloria sold more than ten of them), also a comparative reading (for example, in a situation with twenty cars, where Gloria sold five, Kenny three, and James four). A similar story can be told for (19): the

absolute reading, could be glossed as "more babies are born March than in all other months together", which doesn't make much sense. The comparative reading, where the number of babies born in each month are compared, is preferred here.

These are interesting examples, as a standard generalised quantifier semantics for "most", meaning "more than half", would not suffice to get a correct interpretation for (18) and (19). Again, we obviously don't have a solution for dealing with this problem, but there is a clear correspondence of behaviour with that of the superlative adjective.

5 Discussion

It is perhaps surprising that none of the available theories of superlatives ([Hei99, Sza86, Gaw95, FK00]) say anything on how the genitive should be integrated in a theory of superlatives, at least as far as the absolute reading is concerned. Given this, it is perhaps *not* so sursprising, that most of the currently existing computational grammars for English do not have a correct treatment of this phenomenon. This does not only hold for the CCG grammar on which this work was inspired [Hoc03], but also for other well-known wide-coverage parsers and grammars. Analysing nouns as "relational" in genitive constructions would solve as least part of the problem.

A correct semantic analysis of comparative readings requires further work: we still know little under which conditions such readings are present — a thorough corpus study might give inspiration and shine new light on the matter. In addition, various other problems triggered by the superlative are not dealt with in this paper. One is the interpretation of plural entities in superlative expressions and interaction with ordinals. Another issue concerns over-generation of the categorial grammar that we use. The current analysis incorrectly supports extra-grammatical expressions such as "the tall fastest player", which would be of particular importance for generation tasks [vD06].

References

[BN06] Johan Bos and Malvina Nissim. An empirical approach to the interpretation of superlatives. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, pages 9–17, Sydney, Australia, July 2006.

- [Car97] Bob Carpenter. Type-Logical Semantics. The MIT Press, 1997.
- [FK00] Donna F. Farkas and Katalin E. Kiss. On the comparative and absolute readings of superlatives. Natural Language and Linguistic Theory, 18:417–455, 2000.
- [Gaw95] Jean Mark Gawron. Comparatives, superlatives, and resolution. Linguistics and Philosophy, 18:333–380, 1995.
- [Hei99] Irene Heim. Notes on superlatives. MIT, 1999.
- [Hoc03] Julia Hockenmaier. Data and Models for Statistical Parsing with Combinatory Categorial Grammar. PhD thesis, University of Edinburgh, 2003.
- [KR93] Hans Kamp and Uwe Reyle. From Discourse to Logic; An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and DRT. Kluwer, Dordrecht, 1993.
- [Mat05] Ora Matushansky. The DP and the Deepest. Ms. Université Paris 8/CNRS, 2005.
- [Mus96] Reinhard Muskens. Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy*, 19:143–186, 1996.
- [Par97] Barbara Partee. Appendix B. Genitives A case study. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*, pages 464–473. Elsevier, MIT, 1997.
- [PB99] Barbara H. Partee and Vladimir Borschev. Possessives, favorite, and coercion. In Proceedings of ESCOL99, pages 173–190. CLC Publications, Cornell University, 1999.
- [Ste01] Mark Steedman. The Syntactic Process. The MIT Press, 2001.
- [Sza86] Anna Szabolcsi. Comparative superlatives. In N. Fukui et al., editor, *Papers in Theoretical Linguistics, MITWPL*, volume 8. MIT, 1986.
- [vB83] Johan van Benthem. Five Easy Pieces. In Alice G.B. ter Meulen, editor, Studies in Modeltheoretic Semantics, pages 1–17. FLORIS, Dordrecht, 1983.

- [vD06] Kees van Deemter. Generating referring expressions that involve gradable properties. *Computational Linguistics*, 32(2):195–222, 2006.
- [vEK97] Jan van Eijck and Hans Kamp. Representing Discourse in Context. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*, pages 179–240. Elsevier, MIT, 1997.
- [Wal72] John Wallace. Positive, comparative, superlative. Journal of Philosophy, 69(21):773–782, 1972.