

# ANAPHORIC *ONE* AND ITS IMPLICATIONS\*

JOHN PAYNE

*University of Manchester*

GEOFFREY K. PULLUM

*Brown University and University of Edinburgh*

BARBARA C. SCHOLZ<sup>†</sup>

*University of Edinburgh*

EVA BERLAGE

*Universität Hamburg*

The nominal anaphoric element *one* has figured prominently in discussions of linguistic nativism because of an important argument advanced by C. L. Baker (1978). His argument has been frequently cited within the cognitive and linguistic sciences, and has provided the topic for a chain of experimental and computational psycholinguistics papers. Baker's crucial grammaticality facts, though much repeated in the literature, have not been critically investigated. A corpus investigation shows that his claims are not true: *one* does not take only phrasal antecedents, but can also take nouns on their own, including semantically relational nouns, and can take various *of*-PP dependents of its own. We give a semantic analysis of anaphoric *one* that allows it to exhibit this kind of freedom, and exhibit frequency evidence that goes a long way to explaining why linguists have been inclined to regard phrases like *the one of physics* or *three ones* as ungrammatical when in fact (as corpus evidence shows) they are merely dispreferred relative to available grammatical alternatives. The main implication for the acquisition literature is that one of the most celebrated arguments from poverty of the stimulus is shown to be without force.

**1. INTRODUCTION.** Suppose it were the case that English anaphoric *one* was required to have a phrasal antecedent headed by N, and was not allowed to take just a noun as antecedent. Suppose further that positive evidence for the possibility of phrasal antecedents existed but was too scarce in children's input to affect acquisition, and that negative evidence provided to children for the prohibition on noun antecedents was nonexistent. And suppose children acquired tacit knowledge of these facts rapidly and easily nonetheless. It would be reasonable to see the situation as enhancing the plausibility of LINGUISTIC NATIVISM—the view that at least some linguistic knowledge is innate. Specifically, we would have support for innate knowledge of certain facts about noun phrase structure, anaphoric elements, and their antecedents.

The idea of an argument along these lines was set out more than three decades ago in a scientifically serious and well-regarded textbook on transformational grammar by C. L. Baker (1978:413–425; see also 1979:571–574), before the coining of the term 'argument from poverty of the stimulus' (Chomsky 1980:34). Baker's argument is clearly presented, and worthy of close attention.

Unfortunately it has not received close attention. Instead, the voluminous linguistic literature on linguistic nativism has simply repeated in abbreviated form what Baker said, virtually always giving the same three or four invented example sentences that Baker relied on in his textbook. There has been no effort to verify the crucial facts about antecedence possibilities, there has been hardly any effort to support the claim that the allegedly scarce data is indeed scarce, or that it is crucially needed; and there has been little attention to the crucial matter of the semantics of anaphoric *one*.

Moreover, the substantial psycholinguistic literature over the last ten years that has attempted to confirm Baker's conjectures empirically has also been based on uncritical adoption of his data and analysis, and has neither probed the reliability of his claims nor developed an accurate picture of what it is that gets acquired.

In section 2 of this paper we review Baker's argument in detail. In section 3 we distinguish three distinct items pronounced *one*, noting that Baker confused two of them in his argument, and we go on to exhibit corpus evidence that Baker's key claim about the facts is false: phrases like *the one of physics* (with *student* as antecedent for *one*) are grammatical and copiously attested. This leads to a further observation: that the empirical facts show the distinction between complements and modifiers of nouns to be unfounded. There is no rational way to motivate drawing the distinction between them.

In section 4 we provide a new syntactic and semantic analysis of anaphoric *one*. We assume no structural differentiation of the phrases formerly classified as either complements or adjuncts: all nouns are treated grammatically as non-relational until they combine with a dependent. The semantic relationship holding between head and dependent in any given context of utterance is determined by a mixture of world and contextual knowledge. Certain relations are more probable than others, and these are the ones which have given rise to the notion of some nouns being inherently relational and taking complements. Anaphoric *one* is just a regularly-inflected noun with a special anaphoric role, and can itself have either a non-relational or a relational meaning depending on the meaning of its antecedent.

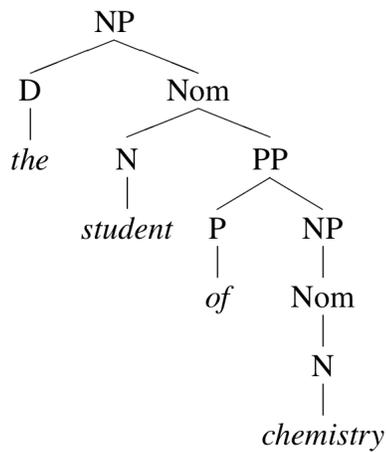
In section 5 we consider why linguists have been so ready to believe that expressions such as *the one of physics* are ungrammatical. The explanation lies largely in frequency effects. In a variety of contexts, anaphoric *one* competes with other anaphoric expressions, and the expressions that are deemed ungrammatical are simply the ones in which anaphoric *one* is a generally less successful, though not impossible, competitor. This explanation applies not only to expressions like *the one of physics*, but also to a variety of other expressions which have at some point been deemed ungrammatical, for instance expressions such as *three ones*, where anaphoric *one* occurs with a numeral determiner.

In section 6, we return to the issue of acquisition, and point out that the falsity of the factual basis is not the only problem: investigators have repeatedly altered their assumptions about what has to be acquired, so they are frequently at cross purposes. None of the works in question have assumed the correct adult system. We conclude (in section 7) with some remarks about the working relationship between the linguistic and psychological sciences that will be needed if we are to develop a proper understanding of the details of first-language acquisition.

**2. THE RECEIVED WISDOM.** Baker (1978) holds that anaphoric *one* can never have a lone N as antecedent. Rather, it must have an antecedent that is a phrasal constituent of a category that he calls Nom (for NOMINAL): the N' of X-bar theory. We follow his notation, which happens to coincide with that of Huddleston & Pullum et al. 2002.<sup>1</sup>

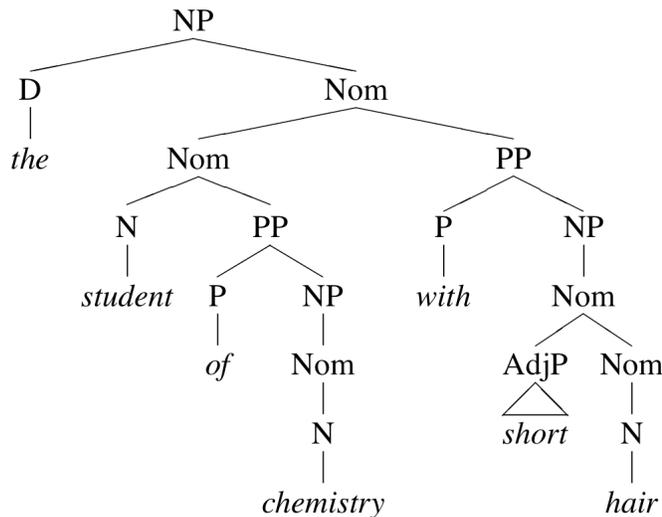
We take phrases like *the student of chemistry* to be labeled noun phrase (NP). An NP has a Nom as head, and *student of chemistry* is a Nom. We posit a category of DETERMINATIVES (D),<sup>2</sup> to which *the* belongs. Thus the structure of *the student of chemistry* would be as in 1.

(1)



In an NP like *the student of chemistry with short hair*, the PP *of chemistry* is standardly taken to be a complement, but *with short hair* is a modifier. Modifiers are not sisters of N, but sisters of Nom, so the structure of this second NP would be as in 2, with an additional Nom constituent.

(2)



Crucially, Baker claimed there was a grammaticality difference between the two sentences in 3.

- (3) a. The student of chemistry was more thoroughly prepared than the one of physics. [example 14b in Baker 1978:415]  
b. The student with short hair is taller than the one with long hair. [example 23 in Baker 1978:419]

The difference is that 3a is supposed to be ungrammatical because it has no Nom to act as antecedent for *one*. The sole Nom constituent in the subject NP (see 1) includes not only the head noun *student* but also the complement, *of chemistry*. By contrast, 3b does have a suitable Nom, because *student with short hair* has the structure [Nom [Nom [N *student* ] ] [PP *with short hair* ] ], and the inner Nom can serve as antecedent so 3b is grammatical.

Baker had actually made his observations almost a decade earlier, and pointed them out to

George Lakoff, who used them as part of the data for section II of his paper ‘Global rules’ (1970). Jackendoff (1977:59), agreeing with Lakoff, proposes a constraint banning a Nom consisting of ‘*one(s) of NP*’. However, he notes, crediting a personal communication from Noam Chomsky, that no similar constraint holds for PP complements headed by prepositions other than *of*. For example, sentences like 4 are grammatical:

- (4) Arguments with Bill are less fruitful than ones with Harry. [example (i) in footnote 4 of Jackendoff 1977:61]

If the *with*-PP after a noun like *argument* is a complement, as the parallel with VPs like *argue with Bill* suggests, then Baker’s general claim that *one* cannot have a complement-taking noun as antecedent had already been shown to be false before he published it. Throughout more than four decades of literature, however, the alleged ungrammaticality of 3a was taken to be secure.

### 3. THE SYNTACTIC FACTS.

**3.1. THE THREE ITEMS SPELLED *ONE*.** English has three distinct lexemes with *one* as their orthographic base form. They differ morphologically, syntactically, and semantically. We summarize their properties in 5.

(5) The three items spelled *one* in English

a. Pronoun

CATEGORY: Regular third-person singular indefinite pronoun  
 INFLECTION: *one* (plain case), *one’s* (genitive case), *oneself* (reflexive)  
 MEANING: ‘An arbitrary person’ (compare French *on*, German *man*)  
 NOTES: As with pronouns generally, no plural form.

b. Determinative

CATEGORY: Indefinite cardinal numeral determinative  
 INFLECTION: Uninflectable  
 MEANING: ‘1’ or ‘some’ or ‘a(n)’ or ‘sole’  
 NOTES: Obligatory when functioning as determiner. Omissible when functioning as modifier with the meaning ‘sole’. Anaphoric to a whole NP when used with no head noun.

c. Noun

CATEGORY: Regular common count noun  
 INFLECTION: *one* (plain sg.), *ones* (plain plur.), *one’s* (gen. sg.), *ones’* (gen. plur.)  
 MEANING: Anaphoric; something like ‘instance thereof’, referring back to some type or class referred to in the discourse or salient in the context.

The item we are concerned with is 5c, the count noun, which is referred to as *one<sub>ct</sub>* in Huddleston & Pullum et al. 2002 (where the determinative is tagged *one<sub>d</sub>*).<sup>3</sup> Like any other regular noun, *one<sub>ct</sub>* has four inflected forms (three of them pronounced identically). Its anaphoric use<sup>4</sup> is illustrated in 6.

- (6)
- a. The art museum in Bilbao is the most impressive one I've seen.
  - b. An honest local government official is harder to find than a corrupt one.
  - c. The long, gently curved Victorian railway station building in York is the finest one in the whole of England.

In 6a the most plausible assumption about the antecedent for *one* would be *art museum* — surely not *art museum in Bilbao* (to call the Guggenheim the most impressive museum in Bilbao would be an understatement). In 6b, the plausible antecedent is *local government official* (not *honest local government official*, which would involve a contradictory interpretation). And in 6c it is *Victorian railway station building*, or perhaps just *railway station building*, but not *long, gently curved Victorian railway station building* (and certainly not *long, gently curved Victorian railway station building in York*, which would render 6c trivially true).

These examples illustrate a point to which we will return when considering the arguments for linguistic nativism: the antecedent of anaphoric *one* can indeed be a multi-word Nom, but it does not have to be the largest Nom available.

**3.2. CONFUSION OF COUNT NOUN WITH DETERMINATIVE.** It has gone unremarked in the linguistic literature, so far as we know, that Baker confused two of the items listed in 5. Although the quantity of data he considered was very small, his original example illustrating a multi-word Nom as antecedent involves the wrong lexical item. The invented sentence he gave was 7:

- (7) John has a blue glass, but Alice doesn't have one.

The occurrence of *one* in 7 is not the noun; it is the determinative. Notice that it does not have a plural form (*\*Alice doesn't have ones*). The constituent whose repetition it avoids is not the Nom *glass* or the Nom *blue glass*, it is the entire NP *a blue glass*. Baker has used the anaphoric noun  $one_{ct}$  to illustrate the claim that  $one_{ct}$  cannot take just a noun without its complement as antecedent, but has used the indefinite determinative  $one_d$  to illustrate the claim that  $one_{ct}$  can have a multi-word antecedent. The sentence in 7 has no bearing on this second claim.<sup>5</sup>

It is not possible to treat  $one_{ct}$  and  $one_d$  as a single lexeme: they are of different syntactic categories, and (as Jackendoff notes) one inflects and the other does not. But in any case, collapsing them would mean changing the claim about anaphoric *one* to a different one: that the antecedent is either a Nom or a full NP. This is not what Baker was proposing.

This descriptive error is not of primary importance: Baker's argument could be rebuilt with different examples (e.g., *John has a blue glass, but we couldn't find another one for Alice*, where the point would be that *another one* can mean "another blue glass"). However, Baker also makes a different descriptive error that is much more serious. It concerns not the permissibility of multi-word Nom antecedents but the alleged impermissibility of noun antecedents.

**3.3. CORPUS INVESTIGATION.** Jackendoff observed that  $one_{ct}$  can indeed replace a lone N before preposition phrases headed by *with* (as in 4). He proposes to preserve Baker's claim about lone N antecedents by narrowing it to PPs headed by *of* (henceforth *of*-PPs). One might think of going further, in fact, and hypothesizing that *of*-PPs are the only true complements of nouns. This

would preserve Baker’s claim that nouns on their own can never be antecedents for *one<sub>ct</sub>*, provided sentences like 3a are genuinely ungrammatical. Unfortunately there is no possibility of maintaining such a thesis: corpus data refute it overwhelmingly.

Payne and Berlage (2009) undertook an extensive study of the relevant data. They had independent reasons — the nativism issue was not on their agenda. They were interested solely in gaining insight into the complement/adjunct distinction. What they found was that nouns serving as antecedents of *one<sub>ct</sub>* in isolation from their *of*-PP complements were abundant. The semantic relation between the head noun and the *of*-PP had clear effects on the frequency of such constructions, as we shall review below, but they concluded that there was no clear place to draw a line between complements and modifiers.

Payne & Berlage’s corpus investigation was based on the British National Corpus (henceforth BNC), a 100-million-word corpus of British English (henceforth BrE) from the later part of the 20th century. Approximately 90% of the texts are written, from a wide-ranging variety of sources, and 10% are spoken. For the purposes of this investigation, the version used of the BNC was BNCweb (CQP Edition).<sup>6</sup>

Two searches were employed. The first extracted all occurrences of the plural form *ones* followed by *of*. After spurious hits were eliminated, this yielded 127 plural tokens of *one<sub>ct</sub>*. Searching for singular tokens of *one<sub>ct</sub>* in the same environment is complicated by the potential confusion with *one<sub>d</sub>*: examples of the type *one<sub>d</sub> of the X* are extremely common. But since *one<sub>ct</sub>* is a count noun it must be preceded by a determiner when singular, and also the form *one* generally cannot represent *one<sub>d</sub>* if preceded by an adjective. Isolating all sentences containing the sequence ‘determinative + adjective + *one* + *of*’ yielded a further 408 genuine singular tokens of *one<sub>ct</sub>*.

From the total of 535 they excluded 6 examples in which *one<sub>ct</sub>* was a non-anaphoric subcomponent of an invented proper name, as in *the Great Ones of the land*. They also excluded 11 examples in which *one<sub>ct</sub>* was followed by an oblique genitive, i.e. an *of*-PP in which the dependent NP stands in the genitive case, since these represent an entirely distinct construction. Each of the remaining 518 tokens of *one<sub>ct</sub>* followed by *of* was then examined to isolate the semantic relation between the antecedent noun and the *of*-PP.

We use the conventions of Huddleston & Pullum et al. (2002: 474–7) as an informal notation for the semantic relations involved:

- (8)            eyes            of            the team manager  
                  *h* (head)                            *d* (dependent)  
                  *d* has body-part *h*

In an example such as 8, *eyes* is the head noun, symbolized by *h*, and *the team manager* is the dependent NP, symbolized by *d*. The term DEPENDENT covers both complements and modifiers, and avoids the necessity for making any prejudgment at this stage as to which semantic relations underlie the syntactic relation of complement. In 8, the semantic relation is then a body-part one: *d* has body-part *h*.

In total, the 518 examples of *one<sub>ct</sub>* followed by a dependent *of*-PP were analysed as representing, at a relatively coarse level, 35 distinct semantic relations between head and dependent. For illustrative purposes, we make the simplification of grouping these into 14 broader

semantic fields, each of which is represented by examples 9–22 below. The antecedent, together with the semantic relation identified between head and dependent, is indicated in braces following each example, followed by the BNC locator in square brackets

(9) Object-like dependent

- a. This interpretation is contrary to an accepted [*one* of wrestling] as a sport.  
 {*interpretation*; *d* is undergoer of *h*} [CGY 1,308]
- b. How the printers had got hold of her photograph she did not know, but they had, and now it was being sold all over London, along with [*ones* of Lillie Langtry and other noted belles].  
 {*photographs*; *d* has depiction *h*} [HGE 1,398]

(10) Function noun

- a. Nephrite contains a high proportion of magnesia and a considerable [*one* of lime].  
 {*proportion*; *h* is amount of *d*} [FBA 470]
- b. Seventy years of Byrd on record must have given us a good 50 versions of *Ave verum corpus* but not a single [*one* of *Deus venerunt gentes*].  
 {*version*; *h* is type of *d*} [J1A 1,344]

(11) Part-whole

- a. ... she gently raised her eyebrows until her eyes met the disconcerted [*ones* of the team manager].  
 {*eyes*; *d* has body-part *h*} [HGM 204]
- b. I hope this little titbit of news about the crews that were formed and especially the [*ones* of Rosie's Riveters]...  
 {*crews*; *d* has associated part *h*} [H5J 59]  
 (Rosie's Riveters was a World War 2 US airplane)<sup>7</sup>

(12) Agentive

- a. Suddenly the river was full of plunging bodies going to the rescue, barking dogs and screaming girls mingling their cries with the masterful [*ones* of the menfolk].  
 {*cries*; *d* is performer of *h*} [ACK 2,535]
- b. The German keyboard tablatures – Elias Ammerbach 's (Leipzig, 1571 and 1575), those of Bernhard Schmid the elder (Strasbourg, 1577) and Jacob Paix (Lauingen, 1583), and the manuscript *ones* of Christoph Loeffelholtz (Tuebingen, Univ. Bibl., Mus. ms. 40034) and August Noermiger (1598, idem, 40098) – consist almost exclusively of vocal transcriptions and dances of various nationalities.  
 {*German keyboard tablatures*; *d* is creator of *h*} [GUH 755]

(13) Control

- ... and to shift the costs from the more visible *budgets* of the services to the less visible [*ones* of the individual] ...  
 {*budgets*; *d* is controller of *h*} [AS6 944]

(14) Content

- ... the decision whether to categorize such questions as [*ones* of law or fact] is a matter on which opinion, both judicial and academic, differs.

- {*questions*; *h* has content *d*} [GU6 948]
- (15) Human properties
- a. ... we invest hospital medicine with technical powers additional to those more home-spun [*ones* of the GP], but we attribute those powers to the institution rather than the person.  
 {*powers*; *d* has human property *h*} [CMS 612]
- b. ... his attitude to women and their problems had always been the conventional [*one* of the young aristocrat he had once been].  
 {*attitude to women and their problems*; *d* has mental response *h*} [HGE 595]
- (16) Context
- a. It is surprising to find that the soft-bodied jellyfish have any fossil record at all, but in fact they have the longest [*one* of the phylum].  
 {*fossil record*; *d* has history *h*} [AMM 139]
- b. ... they point us away from the epistemological frame of reference of this chapter towards the socio-cultural one of the next.  
 {*frame of reference*; *d* has context *h*} [FA3 955]
- (17) Physical content
- a. ... and doors in which the original toughened glass panels have been replaced by more serviceable [*ones* of sturdy plywood].  
 {*panels*; *h* has composition *d*} [GUR 19]
- b. ... drinking from skin water-bottles and smaller stone [*ones* of ale or whisky].  
 {*bottles*; *h* is container of *d*} [A0N 580]
- (18) Time and space
- a. ... constituting a trigger for the crash which separates the period of overheating from the subsequent [*one* of mass unemployment and stagnation].  
 {*period*; *h* is timespan of *d*} [K8U 2,080]
- b. ... that lies between the outer road of St Helen 's and the inner [*one* of Spithead].  
 {*road*; *d* is location of *h*} [BNB 1,115]
- (19) Representative
- Jesus is the Christ, the anointed [*one* of God].  
 {*person* (inferred); *d* has representative *h*} [CEJ 763]
- (20) Causation
- ... the tears, Dexter felt, were as much [*ones* of laughter] as of despair.  
 {*tears*; *h* has source *d*} [G1W 1,995]
- (21) Categorization
- a. The new commercial brewery will be the only [*one* of its kind] in Worcestershire.  
 {*brewery*; *h* has type *d*} [K1R 192]
- b. It might take in all the farms in valley, parish or district. I have been on [*ones* of 100,000 acres].  
 {*farms*; *h* has size *d*} [EER 1,448]

(22) Partitive

- a. She scooped up the bits of spilt polystyrene in her hand and dropped them into the waste-paper basket. I'll get a new [*one of these*] when we move.  
{*waste-paper basket*; *h* is subset of *d*} [ABX 3,324]
- b. The administrator, Tilahu Walle, says they are the lucky *ones* of the 200,000 people in the area who need assistance.  
{*people in the area who need assistance*; *h* is subset of *d*} [B73 1,179]

To summarize, there is an abundance of examples in which *one<sub>ct</sub>* is anteceded by a single noun (or indeed multi-word Nom) followed by an *of*-PP to which it stands in some kind of semantic relation. In order to save a vestige of Baker's claim that *one<sub>ct</sub>* cannot have a complement-taking noun as its antecedent and take complements of its own (already, we remind the reader, delimited by restriction to the preposition *of*), it would be necessary to claim that none of these *of*-PPs is licensed by the antecedent noun or Nom, and that they are all to be treated as modifiers rather than complements.

**3.4. OF-PPS AND SEMANTIC RELATIONS.** A defender of Baker might propose that only those *of*-PPs which stand in an appropriate semantic relation to an inherently relational noun are genuine complements. Thus *student* in Baker's original example 3a would be inherently relational because it is a nominalization of the verb *study*, and *study* is a two-place predicate, one of whose arguments is the entity studied. And indeed, it has occasionally been argued in the post-Baker syntactic literature on *one<sub>ct</sub>*, notably by Oga 2001 and Panagiotidis 2003, that there are two distinct prepositions *of*: the first would be functional/semantically empty and introduce complements of nouns which themselves are claimed to be inherently relational, and the second would be lexical/meaningful and itself denote the appropriate semantic relation with an inherently non-relational noun. Only this second type of *of* would be compatible with *one<sub>ct</sub>*.<sup>8</sup>

However, this defense does not work. If the noun *student* is inherently relational, then nouns such as *interpretation* in our example 9a must also surely be relational. The noun *interpretation* is a nominalization of the verb *interpret*, and what is interpreted presumably has the same argument role for the noun as it does for the verb. Nominalizations such as these have been unequivocally considered as relational in the semantic literature from Grimshaw (1991: 66) onwards.

The same might be said of *photograph* in example 9b: a photograph is an image of something, the thing that has been photographed. So why is *photograph* not also an inherently relational noun? Syntacticians since Jackendoff (1977) have in fact long been aware of the fact (an embarrassing one for Baker's argument) that picture nouns (*picture*, *photograph*, *portrait*, etc.) readily serve as antecedents to *one<sub>ct</sub>* with a dependent *of*-PP denoting the depicted image. The usual response has been not to reconsider Baker's claim, but rather to suggest that the *of*-PP must be a modifier rather than a complement; see e.g. Panagiotidis (2003: 285-6). However formal semanticists, e.g. Vikner & Jensen (2002: 197), who tend not to focus on or even mention the properties of *one<sub>ct</sub>*, have no hesitation in considering picture nouns as relational. Just like other relational nouns, they lexically encode a relationship between two entities

In fact, if we turn to the substantial semantic literature on relational nouns (for a sampling, see DeBruin & Scha 1988, Barker & Dowty 1993, Barker 1995, and Barker 2011), it is not just

nouns in 9 that are treated as standard examples of relational nouns, but also nouns in 10 and 11. These would be nouns which denote functions, e.g. *proportion* in 10a and *version* in 10b,<sup>9</sup> and nouns which are involved in part-whole relationships, including specifically body-part terms, e.g. *eyes* in 11a, and more general part-whole relations, e.g. *crew* in 11b.

Consider next a noun like *cries* in 12a: this takes an agent argument, just like the verb *cry*. Agents, as opposed to patient/theme arguments, are often conceived of as ‘external’ rather than ‘internal’, i.e. standardly realized by subjects in clause structure and ‘s genitives in NP structure rather than by objects in clause structure and *of*-PPs in NP structure. That is, in X-bar theory and its derivatives they would be specifiers rather than complements. But this ignores the fact that agent arguments can be, and often are, expressed by *of*-PPs in NP structure. An example like *the cries of the menfolk* is a case in point. The alternation between the ‘s genitive and *of* constructions is known to be motivated by a variety of disparate factors, of which the semantic relation involved is only one (see for example Rosenbach 2002, Hinrichs & Szmrecsanyi 2007). Other important factors are the length, animacy, and discourse status of the dependent. Payne & Huddleston (2002: 473-8) argue that the set of semantic relations in principle expressible by *of*-PPs is in fact a proper superset of that expressible by the ‘s genitive construction, i.e. there is no ‘s genitive semantic relation which cannot also be expressed by an *of*-PP, given an appropriate combination of other factors.

There is less consensus in the semantic literature as to the status of the other nouns in 12 and 13 where the *of*-PPs stand in a creator or controller relation to the head.<sup>10</sup> At first sight, it might seem that nouns like *tablatures* in 12b or *budget* in 13 must be inherently non-relational, and that the appropriate creator or controller relation is contextually supplied by the *of*-PP. This is indeed the essence of the pioneering analysis of creator and controller relations in ‘s genitive constructions by Barker (1995: 51), and the basis of many formal semantic treatments of creator and control relations since (in particular a series of papers by Partee and Borschev: see their 2003 for discussion).

It is worth pointing out that it is typically assumed that the dependent in such relations must be expressed by an ‘s-genitive. Thus in a simple example like *John’s car*, the noun *car* is taken to be inherently non-relational, and it is the dependent *John’s* which supplies the controller relation. Nouns considered to be non-relational are incorrectly assumed to be unable to take an *of*-PP: thus examples like *the car of John* are asterisked, and contrasted with *John’s car*. This is essentially the same fallacy as the one holding that agents cannot be expressed by *of*-PPs. The reason that *the car of John* seems somewhat unacceptable is simply that one-word dependents generally, but especially in the controller relation, strongly favor the ‘s genitive rather than the *of* construction. Longer and/or indefinite dependents are just fine (e.g. *Gunmen in the Philippines ambushed the car of a university president who police had accused of harboring communist rebels*, from the *Wall Street Journal*, 1987).

We might therefore simply remedy this error and make the claim that *of*-PPs, as in the proposed analysis of ‘s-genitives, contextually supply an appropriate semantic relation to an inherently non-relational head. Note that in 12b, the *tablatures* example, the identical creator relation is expressed first by an ‘s-genitive (*Elias Ammerbach’s*) and then subsequently twice by an *of*-PP, including the one headed by *one<sub>ct</sub>* (*the manuscript ones of Christoph Loeffelholz...*).

However, the non-relational analysis is insecure even with this revision. Vikner & Jensen

(2002: 210) argue that creator and controller relations are too automatic to have to be created anew on each occasion of utterance. For example, *Melissa's dissertation* is automatically interpreted in isolation as involving a creator relation (the dissertation that Melissa wrote), and *Melissa's car* is automatically interpreted in isolation as the controller relation (the car that Melissa controls). More importantly, however, they also argue that these relations cannot simply be supplied by the dependent: they must be allowed to be inherent to the head.

One crucial observation is that a phrase like *Mary's former mansion* has two interpretations. The first (and arguably less likely) interpretation would be straightforwardly obtainable by composing the non-relational meaning of *mansion* with the meaning of *former*, deriving a meaning corresponding to 'entity that used to be a mansion'. That meaning could then be composed with *Mary's* to derive the controller relation: the whole NP would then mean 'the entity under Mary's control which was formerly a mansion'. However, the second (and arguably more likely) interpretation is 'the mansion which Mary formerly controlled'. This cannot be derived compositionally unless the noun *mansion* itself is allowed to have a relational interpretation corresponding to 'mansion controlled by *x*'. The adjective *former* can then apply to the controller relation rather than the building. In other words, there is good reason to think that nouns with control or creation readings must have the potential for a relational interpretation within their semantic representation.

We could discuss further whether the nouns illustrating our other semantic relations are inherently relational or not. In some cases, the relation seems to be quite saliently associated with the noun: questions (example 14) do not exist in the absence of their content, and powers (example 15a) do not exist without the entity in which they are invested. On the other hand, it is perhaps less saliently a property of a brewery (example 21a) that it should belong to a particular type, or of a farm (example 21b) that it should have a particular size. But rather than prolong the exposition by working through all these examples, we propose to move straight to our main conclusion in this section.

We claim that it is simply untenable to argue that all the *of*-PPs in examples 9–22 are modifiers. Wherever anyone might decide to draw a line between complements and modifiers, there will still be examples which are incontrovertibly complements. Baker's claim that *one<sub>ct</sub>* cannot precede complements is simply indefensible.

**3.5. TROUBLE FOR THE COMPLEMENT/MODIFIER DISTINCTION.** We now note a deeper and more radical issue raised by the corpus data. As we have noted, examples in which *one<sub>ct</sub>* is most plausibly treated as having a multi-word antecedent are not at all infrequent; note *German keyboard tablatures* (12b), *attitude to women and their problems* (15b), *fossil record* (16a), *frame of reference* (16b), and in the partitive construction *waste-paper basket* (22a) and *people in the area who need assistance* (22b). However, these are examples which, if viewed syntactically (as Baker viewed them), would require a radically different interpretation of the Nom constituent to the one envisaged by Baker, and indeed all syntacticians who follow the basic tenets of X-bar theory. In Baker's analysis, head nouns combine first with their complements to form a Nom constituent which can then in principle be modified; it is impossible for a complement to be THE COMPLEMENT OF A NOM. But in the data just noted the *of*-PP which follows *one<sub>ct</sub>* must be EXTERNAL to an already formed multi-word unit, whether the dependents involved are themselves

construed as complements or modifiers. None of the proposals we might envisage to handle this data allow Baker's analysis to survive in its intended form.

One possibility is simply to abandon X-bar theory principles concerning the structural distinction between complements and modifiers, and allow dependents of any kind to combine not only with N but also with already formed Noms. The constituent Nom is then simply an NP-internal phrasal category containing a head noun and any number of dependents (except the determiner). This is, for instance, the syntactic structure for English NPs proposed by Payne & Huddleston (2002). In this conception, *one<sub>ct</sub>* can potentially have any single-word N or multi-word Nom as antecedent. The analysis of 12b is 23:

(23) [ [D the ] [Nom [Nom German [Nom keyboard [N tablatures ]]]] of Christoph Loeffelholz...]]

But this bears little resemblance to the conception of syntactic structure at the heart of Baker's claim, namely that Nom is a category which in itself structurally encodes a distinction between complements and modifiers.

X-bar theory principles concerning the order in which complements and modifiers combine might be preserved if *one<sub>ct</sub>* were treated as anaphoric to a unit of a purely semantic nature. In this case 12b would be analysed as having the X-bar-consistent structure in 24, and *one<sub>ct</sub>* would be anaphoric to the logical form of *German keyboard tablatures* (which would not, however, correspond directly to syntactic constituency and would have to be derived by higher-order logical operations).

(24) [[D the ] [Nom German [Nom keyboard [Nom [N tablatures ]] of Christoph Loeffelholz]]]]

By the time Baker published his discussion of *one* (1978), the possibility that anaphoric elements might in general best be analysed as having logical forms rather than syntactic constituents as antecedents had already been proposed by Sag (1976), and this principle forms the basis of many modern theories of anaphor resolution, e.g. Dalrymple et al. (1991). It is a move which we endorse but which Baker did not envisage, and it destroys the basis of any argument based on *one<sub>ct</sub>* concerning the innateness of SYNTACTIC structure.

**4. SYNTACTIC AND SEMANTIC ANALYSIS.** Our syntactic and semantic analysis of *one<sub>ct</sub>* is based on the assumption that it is futile, at least on the basis of the behaviour of *one<sub>ct</sub>*, to draw a binary division at a syntactic level between complements and adjuncts, or correspondingly at a semantic level between inherently relational and non-relational nouns. In conformity with this principle, we will treat all nouns and nominals grammatically as non-relational until combined with a dependent. The semantic relationship which then holds between head and dependent in any given context of utterance is determined by a mixture of world and contextual knowledge. In this conception, then, certain relations are just more probable than others, and these are the ones which have given rise to the notion of nouns as inherently relational and complement-taking.

**4.1. THE OPTIONALITY OF NOMINAL DEPENDENTS.** As an initial observation, we note that our

analysis will neatly account for a fact which is often ignored and which clearly distinguishes nouns as a category from verbs, namely that there are no convincing cases of nouns taking syntactically obligatory dependents.<sup>11</sup> Thus *king* is a classic ‘relational’ noun, but nothing about the grammar of English forces us to specify the king’s realm in a dependent. The BNC examples in 25 illustrate this property:

- (25) a. Along the north Antrim coastal path, you can admire the work of a giant, see the place where a witch turned a king’s daughters into swans and sample the local delicacy, dulce, which resembles burnt tagliatelle but is in fact dried seaweed.  
[A5X 263]
- b. A sculpture representing a king and queen was broken by the builder’s labourer who found it, revealing that the metal of the faces was only about a millimetre thick. [B71 470]
- c. For example: each soloist need not begin with a formal bow to a king or to the audience, nor end with another bow or considered pose; but such behaviour may be included if the choreographer wishes to locate dance in a particular century and probably a palace in which the story unfolds.  
[A12 968]

The fact that kings are conventionally associated with a particular state is part of world knowledge about kings, not something which necessarily forms part of the argument structure of the noun *king*. The particular state involved, in this case Antrim, may be retrievable contextually rather than from a syntactic dependent, as in 25a. But the state itself may not be important, even when there is specific reference. In 25b, a particular king and queen are represented by a sculpture, but what is relevant is simply their royal status, perhaps identifiable by properties such as their regalia. And it is always possible to have non-referential statements about kings, as in 25c, where all that is relevant is the property of being a king, not the properties of a particular king.

Although examples such as 25 show that this is generally not the case, it is sometimes argued that ‘relational’ nouns are odd when presented as first-mention indefinites without an appropriate accompanying dependent. Vikner & Jensen (2002: 209), for example, contrast 26a, to which they prefix a question mark, with 26b:

- (26) a. A brother was standing in the yard.  
b. A car was parked in the yard.

In specific reference, the most likely clue to the identification of a brother is the identification of a relevant sibling. Brothers are not, like kings, identifiable by properties such as their regalia. But this, we argue, is not a GRAMMATICAL fact. Attested examples of the same type as illustrated in 25 are not hard to find, even with a noun like *brother*. Consider the examples in 27, likewise drawn from the BNC:

- (27) a. Another friend, whose husband is a farmer, shares the care of her parents, who live in a neighbouring village, with a married sister, who also lives nearby. Between them

they give their mother the support she needs since their father has had a stroke. But her mother is always distressed when she leaves, and dismayed that she has to go before doing just one more job to help. Fortunately, a brother and his wife take responsibility at nights, when the mother will often ring for reassurance about her husband.

[BLW 761]

- b. One old couple who were village publicans used their house as a shelter for ‘a very composite family’ which included a daughter who did the pub cooking, a brother, and a son who used two rooms as his tailor’s shop. [AP7 852]
- c. A trust can be charged on a brother’s posthumous child: for intention alone is relevant in trusts, and the opinion of Gallus prevailed that the posthumous children of others can also be our own intestate heirs. [B2P 514]

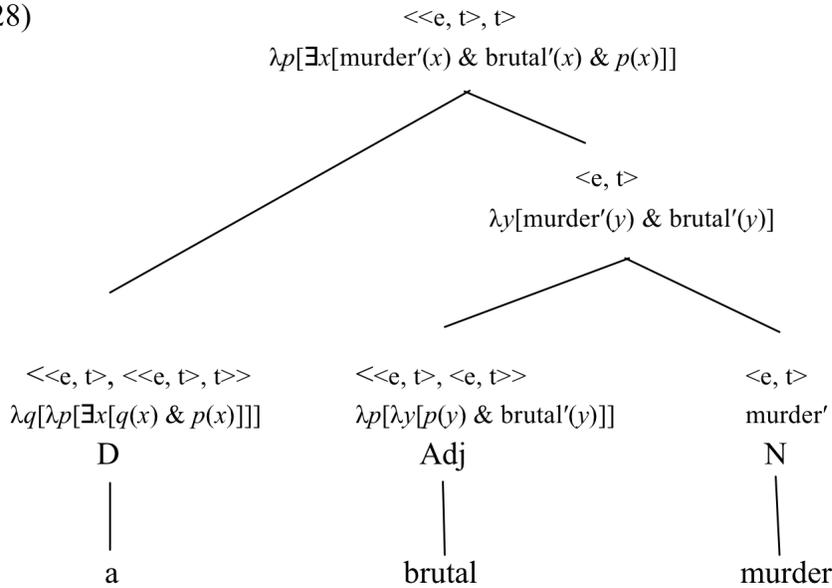
In 27a, the preceding context indeed supplies some information about who the siblings are. However, this contextual information is not even contained within the same sentence as the first-mention indefinite. In 27b, the contextual information that is supplied is not even sufficient to identify precisely whose brother it is: certainly one of the ‘old couple’, but we do not know, or need to know, which one. And in 27c, we have a non-referential statement about a legal property of brothers as such, and no contextual information is necessary.

We therefore regard the optionality of noun dependents as further evidence, on top of the behaviour of *one<sub>ct</sub>*, that a uniform treatment is required for all the semantic relations observed in the corpus. In this treatment, in essence a radical extension of the ideas of Pustejovsky (1991) concerning a generative lexicon, and more specifically the treatment of the control and creator relations in Vikner & Jensen (2002), all nouns can be given either a non-relational or an inherently relational interpretation. The resolution of the semantic relation involved on any particular occasion depends on world knowledge about the relative likelihood of the different types of relation that can hold between the head and dependent, and the precise context.

**4.2. SYNTACTIC AND SEMANTIC ANALYSIS FOR *of*-PPs.** The noun we will use to illustrate our analysis is *murder*, which as a nominalization would standardly be considered to be inherently ‘relational’. In an example such as *a brutal murder of a taxi driver*,<sup>12</sup> the semantic relation between the dependent PP and the head is most likely to be that of undergoer (i.e. *d* is undergoer of *h*), but as shown by the examples in 29, other semantic relations are possible. Also, as argued in the previous section, the dependent PP is not obligatory. We can simply have an NP such as *a brutal murder*.

An analysis tree for the simple NP *a brutal murder* is given in 28:

(28)



For simplicity, we use a standard Montagovian semantics in which NPs are taken to be generalized quantifiers of type  $\langle\langle e, t \rangle, t \rangle$ , and in which the indefinite article has a simple predicate calculus representation. Note two points, however. First, the meaning of the noun *murder* is of type  $\langle e, t \rangle$ : it denotes a function from entities to truth values, not some kind of relation between NP meanings. Second, in order to emphasize the fact that our analysis relies solely on the logical reconstruction of antecedents for *one<sub>ct</sub>*, we do not assign any syntactic category labels to units above the word level. Our analysis tree reflects solely the semantic combinatory potential of the words in the analysed string, and makes no claims as to whether there is any necessity for corresponding syntactic constituents (as might be shown by constituency tests). In particular, our analysis has no need of a syntactic category *Nom* to serve as a structural indicator of the complement/modifier distinction. We do however allow analysis trees to contain a modicum of syntactic information in addition to the specification of word-level categories: they should be construed as indicating word order. We then postulate that units can combine in any order consistent with their typing and the rule of functional application, applied to adjacent units.<sup>13</sup>

In order to construct the logical translation of the NP, this latter principle forces the first step to be the application of the translation of *brutal* to that of *murder*, as in 29a. Then the translation of the indefinite article can apply to the resulting expression, giving 29b. The variables  $p$  and  $q$  are here of type  $\langle e, t \rangle$ , and the variables  $x$  and  $y$  are of type  $e$ .

- (29) a.  $\lambda p[\lambda y[p(y) \ \& \ \text{brutal}'(y)]](\text{murder}')$   
            $= \lambda y[\text{murder}'(y) \ \& \ \text{brutal}'(y)]$   
 b.  $\lambda q[\lambda p[\exists x[q(x) \ \& \ p(x)]]](\lambda y[\text{murder}'(y) \ \& \ \text{brutal}'(y)])$   
            $= \lambda p[\exists x[\text{murder}'(x) \ \& \ \text{brutal}'(x) \ \& \ p(x)]]$

We emphasize that the expression in 29a corresponding to the string *brutal murder* is construed as a logical unit, not a syntactic one.

The analysis of the NP *a brutal murder of a taxi driver* is more complicated. In essence, it



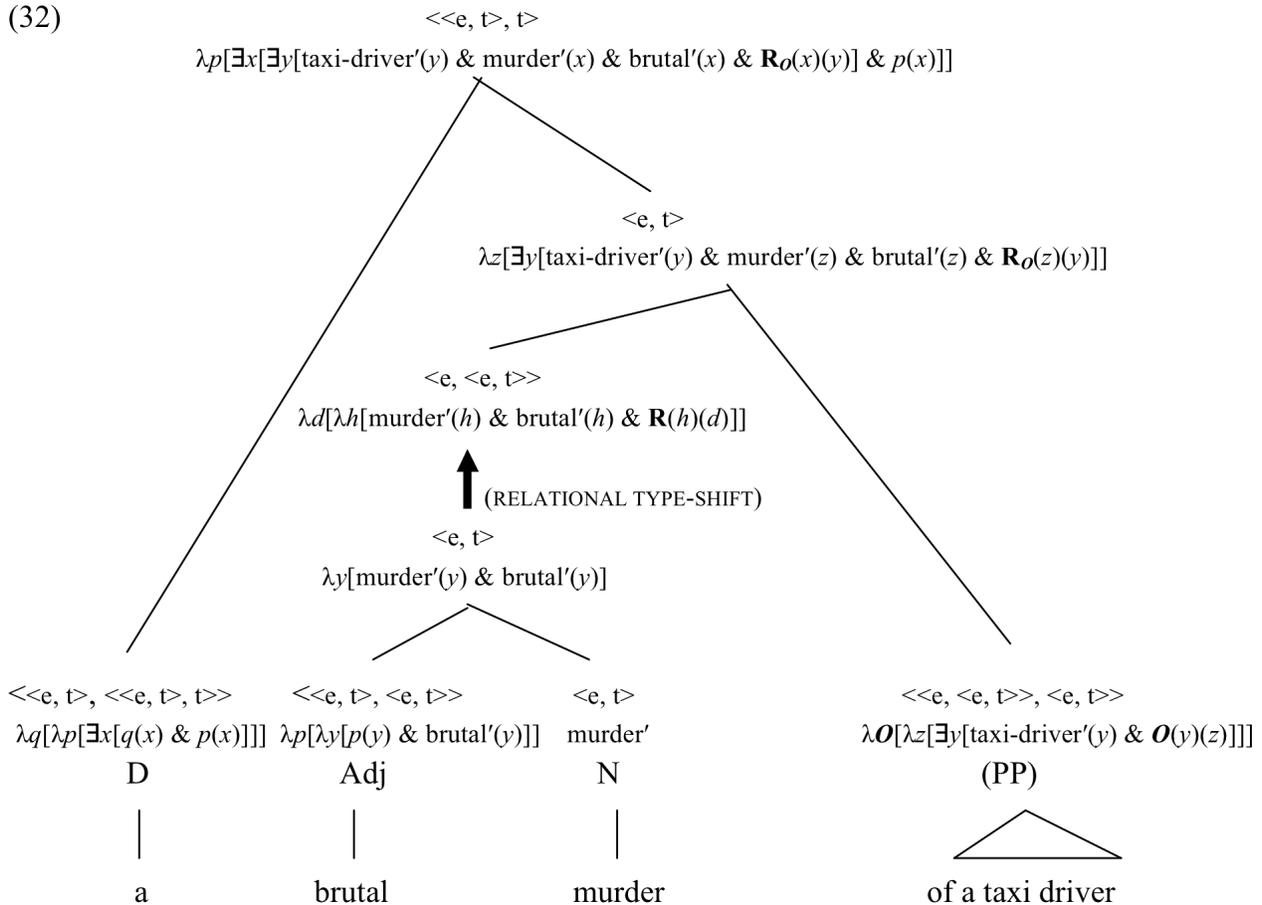
we have an undergoer, and somebody dies. Given that taxi drivers are (unfortunately) more likely to be murdered than to be murderers, this appears to be by far the most likely interpretation of examples like *a brutal murder of a taxi driver*. But many different instantiations of **R** are available for the noun *murder*:<sup>14</sup>

- (31) a. David Peace's Red Riding Quartet, which spins a fictional plot alongside the murders of the Yorkshire Ripper, is all the more potent for its true crime background.
- b. One of two sisters who bombed the Old Bailey in the 1970s is in custody today being questioned about the murders of two soldiers in Northern Ireland in March.
- c. Paul Temple is part of the era between the upper class murders of Agatha Christie and the gritty murders of today.
- d. The driving rhythms of London's fiercely competitive cat-walks may seem a thousand miles away from the cosy cottage murders of Miss Marple, but they provide a perfect environment for the more chilling edge of Agatha Christie's short stories.

In 31a, given the world-knowledge that the Yorkshire Ripper was a notorious murderer, the most probable instantiation of **R** is performer, not undergoer. In 31b, performer and undergoer might in isolation be assigned more equal probabilities, but the wider context suggests that the undergoer relation is intended. In 31c, there are two occurrences. The item of world-knowledge that Agatha Christie is a crime fiction author yields a high probability of the creator relation for the first, but in the second, anything other than a temporal relation is highly improbable. Finally, in 31d, world-knowledge tells us that Miss Marple is a fictional amateur detective, so the performer relation is refined to that of solver rather than committer of the crime.

In order to construct a logical translation for the NP *a brutal murder of a taxi driver*, we propose that the string *brutal murder* of type  $\langle e, t \rangle$  is shifted by the relational operator 30a to the relational type  $\langle e, \langle e, t \rangle \rangle$ . It can then combine with the *of*-PP, which has type  $\langle \langle e, \langle e, t \rangle \rangle, \langle e, t \rangle \rangle$  (a function taking as argument a function from entities to properties and returning as value a new property). The *of*-PP saturates the relational argument and forms a new unit of type  $\langle e, t \rangle$  corresponding to the string *brutal murder of a taxi driver*. This composes straightforwardly with the translation of the matrix determiner. This sequence of operations is reflected in the analysis tree in 32:

(32)



Of note here is that the translation of the preposition *of* contains a variable  $\mathbf{O}$ , mnemonic for *of*, which we take to range over the wide set of semantic relations which this preposition permits, and which we attempt to characterize in section 5.2. The preposition *of* is therefore not simply meaningless. By employing a different variable, the analysis can in principle be extended straightforwardly to any other preposition, for example prepositions such as *for* or *with*. While the range of semantic relations permitted by the preposition *of* is wider than that of any other preposition, it does not include every conceivable relation. For some relations, more specialized prepositions have to be employed, and the preposition *of* cannot in general substitute for these. What the variable  $\mathbf{O}$  does is to place a constraint on the instantiation of the metasymbol  $\mathbf{R}$  introduced by the application of the relational type-shift operator to the string *brutal murder*. This constraint is reflected by the presence of the subscripted metasymbol  $\mathbf{R}_o$  in the final logical translation of the NP. The actual semantic relation instantiated between *brutal murder* and *a taxi driver* must then be one which is permitted by the preposition *of*. Readers who wish to see more detail may consult the appendix.

This (to our knowledge novel) treatment of the way nouns combine with *of*-PPs disposes of many of the difficulties involved in the necessity of pre-assigning nouns categorially to one of two basic syntactic/semantic types: relational and non-relational. As Partee and Borschev (2012: 447) put it: “The distinction is sharp, but the classification of nouns is not”.<sup>15</sup> In particular, we do not need to assume that noun *A* is grammatically non-relational (and requires type-shifting to a

relational type when it combines with an *of*-PP), while noun *B* is grammatically relational (and is type-shifted to a non-relational type when it does not). All nouns can potentially occur in relational constructions in our scenario, and some may more readily do so than others.

**4.3. SYNTACTIC AND SEMANTIC ANALYSIS OF *one<sub>ct</sub>*.** The analysis of *one<sub>ct</sub>* now follows straightforwardly. Like all count nouns, *one<sub>ct</sub>* belongs to the basic type  $\langle e, t \rangle$ . We will write its translation as **Ana** $_{\langle e, t \rangle}$  to suggest its status as a type  $\langle e, t \rangle$  anaphor whose antecedent must be some appropriate logical form of type  $\langle e, t \rangle$ . This antecedent can correspond either to a single noun, or to a multi-word string with a noun as head — what in the earlier sections of this paper was referred to as a nominal. However, what is reconstructed is a logical unit of the requisite type, not a syntactic constituent.

Thus the bracketed NP in an example such as 6b, repeated here as 33a, will have the translation in 33b:

- (33) a. An honest local government official is harder to find than [a corrupt one].  
 b.  $\lambda p[\exists x[\mathbf{Ana}_{\langle e, t \rangle}(x) \ \& \ \text{corrupt}'(x) \ \& \ p(x)]]$   
 c.  $\lambda y[\text{official}'(y) \ \& \ \text{local-government}'(y)]$   
 d.  $\lambda p[\exists x[\lambda y[\text{official}'(y) \ \& \ \text{local-government}'(y)](x) \ \& \ \text{corrupt}'(x) \ \& \ p(x)]]$   
 $= \lambda p[\exists x[\text{official}'(x) \ \& \ \text{local-government}'(x) \ \& \ \text{corrupt}'(x) \ \& \ p(x)]]$

The derivation of the logical form in 33b will follow the same lines as that of *a brutal murder* in 28–29. The resolution of **Ana** $_{\langle e, t \rangle}$  can be the translation of any contextually available string of type  $\langle e, t \rangle$ , in this case most plausibly the translation of *local government official*, given in 33c. This can simply be substituted for **Ana** $_{\langle e, t \rangle}$  as in 33d.

Nothing prevents *one<sub>ct</sub>* from being followed by an *of*-PP. Like any noun, it belongs to the type  $\langle e, t \rangle$  and can be type-shifted to the relational type  $\langle e, \langle e, t \rangle \rangle$  by the relational operator in 30a, as can any larger unit of type  $\langle e, t \rangle$  which contains it. Thus, the bracketed NP in example 34a will have the translation in 34b:

- (34) a. An unprovoked murder in a Bolton back-street last week was followed this week by [a brutal one of a taxi driver].  
 b.  $\lambda p[\exists x[\exists y[\text{taxi-driver}'(y) \ \& \ \mathbf{Ana}_{\langle e, t \rangle}(x) \ \& \ \text{brutal}'(x) \ \& \ \mathbf{R}_o(x)(y)] \ \& \ p(x)]]$   
 c. murder'  
 d.  $\lambda p[\exists x[\exists y[\text{taxi-driver}'(y) \ \& \ \text{murder}'(x) \ \& \ \text{brutal}'(x) \ \& \ \mathbf{undergoer}(x)(y)] \ \& \ p(x)]]$

The derivation of the logical form in 34b will follow the same pattern as that in 32. In 34b, there are now two unknowns which need to be resolved. The resolution of **Ana** $_{\langle e, t \rangle}$  can be the translation of any of the strings *murder*, *unprovoked murder*, *murder in a Bolton back-street* or *unprovoked murder in a Bolton back-street*, all of which will be of the appropriate type  $\langle e, t \rangle$ . In 34d we illustrate this resolution by assuming that this is simply the translation of *murder* as given in 34c. Once this is established, world-knowledge and context will select an appropriate instantiation of **R<sub>o</sub>**, in this case most probably *undergoer* (which we symbolize by the bold relation **undergoer**).

Since  $\text{Ana}_{\langle e, t \rangle}$  can be resolved by any expression of type  $\langle e, t \rangle$ , including single nouns, and  $\mathbf{R}_O$  is any semantic relation permitted by the preposition *of*, nothing remains of Baker's claim that  $one_{ct}$  cannot substitute a lone noun. His second claim, that  $one_{ct}$  can have both single-word and multi-word antecedents, is essentially correct, but not one based on the syntactic postulates of X-bar theory.

**5. FREQUENCY AND GRAMMATICALITY.** What is it which has enabled the received wisdom concerning  $one_{ct}$  to persist for so long? Why do some examples with  $one_{ct}$  continue to be introspectively judged in isolation as ungrammatical by some linguists? One reason for the persistence of the claim that  $one_{ct}$  cannot be followed by an *of*-PP might simply be its frequent repetition, both in syntactic textbooks as a prime exemplar of the supposed rationale for distinguishing complements from adjuncts, and in the psycholinguistic literature as a prime exemplar of the poverty of stimulus argument (see section 1). Linguists with experience of this topic have in effect been trained to believe that 3a is ungrammatical over more than thirty years.

However it is not just to examples like *the one of physics* that ungrammaticality judgments have been applied, but also to other cases such as the supposed incompatibility of  $one_{ct}$  with numeral determiners in examples like *three ones* (Lakoff 1970; Postal 1972).

The answer, we suggest, lies in frequency effects connected with the distribution of  $one_{ct}$ . In a number of environments,  $one_{ct}$  is in competition with at least one alternative anaphoric strategy which has long been established in the language and is arguably simpler. In such environments, we propose that  $one_{ct}$  is not excluded by any grammatical principles concerning its distribution. As argued above, it is just a count noun with the same distribution as any other count noun. But as an anaphor it can lose out to other anaphors and occur with lower frequency than its competitors—sometimes overwhelmingly lower. Nevertheless, in some cases it will occur, and in the presence of other favorable factors it may even become the preferred option. We begin in section 5.1 with some general observations concerning the distribution of  $one_{ct}$  and its anaphoric competitors. In section 5.2 we turn to a detailed account of the frequency effects associated with the occurrence of  $one_{ct}$  preceding *of*-PPs, and in section 5.3 we discuss particular meaning relations involving human head nouns, including *student*.

**5.1. THE ANAPHOR  $ONE_{CT}$  AND ITS COMPETITORS.** The main competitor for  $one_{ct}$  is zero. Almost all determiners (exceptions are *the*, *a(n)* and *every*) can function on their own as anaphors, and where this shorter and arguably therefore simpler strategy is available it is typically the most frequent option.  $One_{ct}$  therefore tends to occur in NPs in which it is not immediately preceded by a determiner. It also does not occur without dependents. In a sentence like *I want one*, *one* must be the determiner, i.e.  $one_d$  rather than  $one_{ct}$ . And *I want ones*, in which *ones* has to be  $one_{ct}$ , genuinely never seems to occur, presumably because it is invariably preempted by *I want some*.

The determiners which do not function on their own as anaphors are of course exceptional. The definite article readily co-occurs with  $one_{ct}$  when there is post-modification, as in *the one over there*, and without post-modification as a predicative NP in examples like *That's the one*, where the pronoun *it* is certainly a competitor but perhaps yields ground to  $one_{ct}$  because of the predicative environment. In non-predicative environments *the one* clearly loses out to simple pronouns, but this is no reason for treating the combination *the one* as ungrammatical. With the

indefinite article, there is a simple and straightforward competitor for *a one*: it is just the determiner *one<sub>d</sub>* as in *I bought one yesterday*. Nevertheless *a one* is not ungrammatical as such: it occurs predicatively in examples like the BrE colloquial *Ooh, you are a one!* – and notice that *every*, which cannot function in isolation as an anaphor, readily co-occurs with *one<sub>ct</sub>*, as in *I counted every one*.

A search of 98 random examples of *one<sub>ct</sub>* in the BNC reveals the prevailing patterns.<sup>16</sup>

Immediately preceded by	Total examples	With post-modification
Adjective	69	10
Participle	5	0
Noun	4	1
None	4	4
<i>the</i>	11	10
<i>these/those</i>	4	3
<i>which</i>	1	0

Table 1: frequency of the dependents of *one<sub>ct</sub>* (sample of 98 examples from the BNC)

In the majority of examples, *one<sub>ct</sub>* is immediately preceded by a modifier belonging to a major category (adjective, noun or participle), as in *the big ones*. The remainder are preceded either by nothing, or by a non-quantificational determiner, and of these almost all contain some form of postmodification (a PP or clause), as in *the ones not in the catalogue*. The examples where we find a determiner preceding a bare *one<sub>ct</sub>* with no post-modification are *the ones* in the predicative function discussed above (*they're the ones*), and *these ones* and *which ones*. In the latter two cases, the determiners could easily function as anaphors without the assistance of *one<sub>ct</sub>*. We can conjecture however that the occurrence of examples like *these ones* may be facilitated by parallelism with the singular *this one*, which focusses on the countability of the identified referent and is thus typically differentiated from *this* on its own. And in the case of *which ones*, *one<sub>ct</sub>* provides an indication of number which would otherwise be lacking. In other words, *one<sub>ct</sub>* has properties which enable it to compete on a reasonable footing with the bare determiner in these cases.

It is difficult to apply this functional account of the distribution of *one<sub>ct</sub>* to the determiners *another*, *each* and *either*, which are not frequent enough to occur in the small sample above, but occur both with and without *one<sub>ct</sub>* with little functional difference. *I want another one* says no more than *I want another*, and is about half as frequent. This behaviour seems idiosyncratic.

The functional account does, however, clearly account for the infrequency with which *one<sub>ct</sub>* co-occurs with possessive determiners like *my*, *you*, *his* etc., or with quantificational determiners such as numerals. In this case, the bare determiner overwhelmingly predominates: we find *mine* and *five* rather than *my ones* and *five ones*. But again this does not entail that the latter strings should be deemed ungrammatical. When a large enough corpus is investigated they do

occur, and there are sometimes obvious motivating factors. Taking the co-occurrence of numerals and *one<sub>ct</sub>* as an example, even the BNC as a whole is not large enough to provide more than the odd example (and they are from the spoken section of the corpus). But a web search readily turns up perfectly natural-sounding examples:<sup>17</sup>

- (35) a. This atoll is on the west of Maldives and has 75 islands—13 of them are inhabited, 57 are uninhabited including the five ones which are currently being developed into resorts.
- b. There were and still are nine bells in a row in the kitchen, about a foot apart, ten feet from the floor, and on enquiry Major Moor learned from the cook that the ones affected were the five ones on the right: these were the ones situated in the dining room, the drawing room over the dining room, an adjacent bedroom, and two attics over the drawing room.

These examples typically involve NPs with post-modification, and the presence of a post-modifier is clearly conducive to the use of *one<sub>ct</sub>* as an overt head to which postmodification can be applied. In examples (35a) and (35b), the analysis could start with either the numeral or the post-modifier being construed as applying to the head first, before the numeral.

In the small sample of 98 examples discussed above, 5 occur with a prepositional phrase, and of these, just a single one is an *of*-PP. If we extrapolate from this sample, we can deduce that *of*-PPs occur in only a very small proportion of occurrences of *one<sub>ct</sub>*. Nevertheless, as we have shown above, they occur quite frequently in a corpus the size of the whole BNC. In the following section, we continue to an investigation of the anaphoric competitors for *one<sub>ct</sub>* in this dataset.

**5.2. FREQUENCY EFFECTS INVOLVING OF-PPS.** It is not the case that all of the possible semantic relations permitted to *of*-PP dependents of *one<sub>ct</sub>* occur with equal frequency in the corpus data. The full set of 35 relations we identified in the corpus is presented in Tables 2 and 3.

Table 2 contains those relations, many involving animate dependents, which we judge to be expressible in principle not just by the construction with *of*, but also by the 's genitive construction, as in *the team manager's eyes*. On the other hand, Table 3 contains those relations which are expressible solely by the *of* construction. The relevance of this division will soon become apparent.

<b>R</b>	<b>Read as:</b>	<b>Type</b>	<b>Example</b>	<b>No</b>
<b>time</b> $(h)(d)$	$d$ is time of $h$	TIME AND SPACE	{acts} of yesterday	<b>20</b>
<b>member</b> $(h)(d)$	$d$ has member $h$	PART-WHOLE	{runner} of Britain's three-strong team	<b>18</b>
<b>depiction</b> $(h)(d)$	$d$ has depiction $h$	OBJECT-LIKE	{photographs} of Lillie Langtree and other noted belles	<b>9</b>
<b>representative</b> $(h)(d)$	$d$ has representative $h$	REPRESENTATIVE	{powerful people} of this city	<b>8</b>
<b>creator</b> $(h)(d)$	$d$ is creator of $h$	AGENTIVE	{tablatures} of Christoph Loeffelholz and August Noermiger	<b>6</b>
<b>performer</b> $(h)(d)$	$d$ is performer of $h$	AGENTIVE	{cries} of the menfolk	<b>5</b>
<b>location</b> $(h)(d)$	$d$ is location of $h$	TIME AND SPACE	{inner road} of Spithead	<b>4</b>
<b>body-part</b> $(h)(d)$	$d$ has body-part $h$	PART-WHOLE	{eyes} of the team manager	<b>3</b>
<b>associated-part</b> $(h)(d)$	$d$ has associated part $h$	PART-WHOLE	{crews} of Rosie's Riveters	<b>2</b>
<b>controller</b> $(h)(d)$	$d$ is controller of $h$	CONTROL	{budgets} of the individual	<b>2</b>
<b>inherent-part</b> $(h)(d)$	$d$ is inherent part of $h$	PART-WHOLE	{action force} of demons and evil spirits	<b>1</b>
<b>undergoer</b> $(h)(d)$	$d$ is undergoer of $h$	OBJECT-LIKE	{interpretation} of wrestling as a sport	<b>1</b>
<b>context</b> $(h)(d)$	$d$ has context $h$	CONTEXT	{frame of reference} of the next chapter	<b>1</b>
<b>history</b> $(h)(d)$	$d$ has history $h$	CONTEXT	{fossil record} of the phylum	<b>1</b>
<b>human-property</b> $(h)(d)$	$d$ has human property $h$	HUMAN PROPERTY	{powers} of the GP	<b>1</b>
<b>mental-response</b> $(h)(d)$	$d$ has mental response $h$	HUMAN PROPERTY	{attitude} of the young aristocrat	<b>1</b>
<b>subperiod</b> $(h)(d)$	$d$ has subperiod $h$	PART-WHOLE	{years} of the 19 <sup>th</sup> century	<b>1</b>
<b>source</b> $(h)(d)$	$d$ is source of $h$	CAUSATION	{voice} of the oral culture	<b>1</b>

Table 2: frequency of *of*-PP dependents of *one*<sub>ct</sub>  
(semantic relations permitted to the 's genitive construction)

In Table 2 we give first a suggested name for the relation, e.g. **undergoer** as in example 51d. This is followed by an informal indication of how this relation is to read, identifying the separate roles played by head and dependent, as in 20 and the illustrative examples in that section. The third column indicates the semantic field to which each relation was assigned for illustrative purposes, and in the final column there is an attested example. In each case, we judge the relation to be expressible in principle not only by the *of*-PP construction, but also by the 's genitive construction), e.g. alongside *powers of the GP*, we also have *the GP's powers*.

Table 3 has an identical structure. However, in this case we judge the semantic relation concerned not to be expressible by the 's-genitive construction.<sup>18</sup> Thus the very frequent partitive construction has no 's-genitive counterpart, and for examples like *questions of law or fact* we do not have a corresponding *\*law or fact's questions*.

<b>Semantic relation R</b>	<b>Read as</b>	<b>Type</b>	<b>Example</b>	<b>no.</b>
<b>subset</b> $(h)(d)$	$h$ is subset of $d$	PARTITIVE	a new one of these {waste-paper basket}	<b>218</b>

<b>content</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has content <i>d</i>	CONTENT	{questions} of law or fact	<b>121</b>
<b>kind</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has kind <i>d</i>	CATEGORIZATION	{brewery} of its kind	<b>36</b>
<b>image</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> is image of <i>d</i>	OBJECT-LIKE	{pictures} of a storm on Saturn	<b>15</b>
<b>size</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has size <i>d</i>	CATEGORIZATION	{farms} of 100,000 acres	<b>9</b>
<b>value</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has value <i>d</i>	CATEGORIZATION	{species} of greater commercial value	<b>7</b>
<b>theme</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has theme <i>d</i>	OBJECT-LIKE	{analysis} of previous authorities	<b>5</b>
<b>cause</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has cause <i>d</i>	CAUSATION	{tears} of laughter	<b>5</b>
<b>composition</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has composition <i>d</i>	PHYSICAL CONTENT	{panels} of sturdy plywood	<b>4</b>
<b>age</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has age <i>d</i>	CATEGORIZATION	{children} of an age to be working	<b>3</b>
<b>timespan</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> is timespan of <i>d</i>	TIME AND SPACE	{period} of mass unemployment	<b>3</b>
<b>container</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> is container of <i>d</i>	PHYSICAL CONTENT	{bottles} of ale or whisky	<b>2</b>
<b>duration</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has duration <i>d</i>	TIME AND SPACE	{pregnancy} of 105 days	<b>1</b>
<b>rank</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> has rank <i>d</i>	CATEGORIZATION	{officers} of much more senior rank	<b>1</b>
<b>amount</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> is amount of <i>d</i>	FUNCTION NOUN	{proportion} of lime	<b>1</b>
<b>collection</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> is collection of <i>d</i>	PART-WHOLE	{group} of two figures	<b>1</b>
<b>type</b> ( <i>h</i> )( <i>d</i> )	<i>h</i> is type of <i>d</i>	FUNCTION NOUN	{version} of Deus venerunt	<b>1</b>

Table 3: frequency of *of*-PP dependents of *one<sub>ct</sub>*  
(semantic relations not permitted to the 's genitive construction)

It is immediately apparent that the total number of examples in Table 2 (85) is much less than that in Table 3 (433). Thus, in a global perspective, *one<sub>ct</sub>* occurs most frequently before *of*-PPs precisely in those semantic relations where there is no alternation with the 's genitive construction. When there is such an alternation, the 's-genitive construction is overwhelmingly preferred when the dependent is short (one word), definite/accessible, and animate. The examples with *one<sub>ct</sub>* following an *of*-PP tend therefore to occur when dependents are either longer, indefinite/inaccessible or inanimate, or embody some combination of these factors.

As an illustration, consider again example 12b, which we repeat here as 36:

- (36) The German keyboard tablatures – Elias Ammerbach's (Leipzig, 1571 and 1575), those of Bernhard Schmid the elder (Strasbourg, 1577) and Jacob Paix (Lauingen, 1583), and the manuscript ones of Christoph Loeffelholtz (Tuebingen, Univ. Bibl., Mus. ms. 40034) and August Noermiger (1598, *idem*, 40098) – consist almost exclusively of vocal transcriptions and dances of various nationalities.

The first underlined phrase *Elias Ammerbach's* illustrates the 's-genitive alternative: it is relatively short (in this case 2 words), definite and animate. As a possessive determiner, it can function on its own as an anaphor, and is not followed by *one<sub>ct</sub>*. The second underlined phrase illustrates the forced use of the *of*-PP alternative when the determiner position is blocked by another determiner (*those*): in this case length of the dependent is not a factor. The third underlined phrase illustrates the combination of *one<sub>ct</sub>*. and an *of*-PP. In this case the dependent is long (5 words not counting the supplementary information in parentheses), and the presence of the

prenominal modifier *manuscript* blocks the use of the zero anaphor strategy.

A combination of factors may thus favour the occurrence of *one<sub>ct</sub>* with a following *of*-PP. Some of these are specific to the selection of *one<sub>ct</sub>* as opposed to zero, in particular the presence of particular determiners or modifiers, as discussed in section 5.1. There are then factors related to the genitive alternation. The selection of an *of*-PP construction may be forced, either because the semantic relation is one of those in Table 3 which does not permit the 's-genitive in principle, or because the use of the 's--genitive is blocked by a pre-existing determiner. Only those *of*-PP examples in which the determiner is the definite article allow substitution by the 's-genitive: *the loud cries of the menfolk* ~ *the menfolk's loud cries*. Where however the genitive alternation applies, factors such as the length, definiteness and animacy of the dependent come into play.

In order to quantify these factors, we analysed the 85 examples of Table 2 using the methodology of O'Connor et al. (2013), a large-scale study of the genitive alternation based on the BROWN corpus of American English.<sup>19</sup> The 85 examples of Table 2 represent the initial dataset in which the semantic relation expressed by the *of*-PP in principle allows the genitive alternation. It is necessary to exclude any examples where the genitive alternation is blocked by the presence of a determiner other than the definite article: there were 10 of these. The remaining 75 examples were then coded for three factors: (a) length of dependent (1 word, 2-3 words, 4 words or more); (b) accessibility of dependent (pronoun, proper noun, common noun);<sup>20</sup> (c) animacy (animate, organization, inanimate).<sup>21</sup> The results are given in Table 4:

ANIMACY					ACCESSIBILITY					LENGTH (words)				
BNC <i>of</i> -PP			BROWN <i>of</i> -PP		BNC <i>of</i> -PP			BROWN <i>of</i> -PP		BNC <i>of</i> -PP			BROWN <i>of</i> -PP	
	freq	%	%	odds		freq	%	%	odds		freq	%	%	odds
animate	24	32	17	1:8	pronoun	5	7	1	1:138	1	15	20	23	1:4
organization	17	23	13	1:1	proper	14	19	18	2:3	2-3	48	64	52	3:1
inanimate	34	45	70	5:1	common	54	74	81	6:1	>4	12	16	25	39:1

Table 4: categorization of *of*-PPs as dependents of *one<sub>ct</sub>*  
(compared to odds ratios for the genitive alternation in the BROWN corpus)

In each case, the frequencies and corresponding percentages are compared to the percentages which O'Connor et al. found for the same factors, considered individually, in the BROWN corpus, based on approximately 2800 examples of *of*-PP. It is striking that the distributions in the larger and smaller datasets are broadly consistent with each other, at least in the ranking of the corresponding factors, and frequently in the closeness of the actual proportions.

The main point of this comparison is that O'Connor et al. also give figures for the occurrence of the alternating 's-genitive construction, based the same factors. We used their figures to calculate the odds ratios for each factor in favor of the occurrence of an *of*-PP as opposed to an 's-genitive. The important dividing line is the bold line, which separates those factors where the odds are strongly against *of*-PP and in favour of the 's-genitive, from those factors where the odds are at least closer to evens, and in some cases strongly in favour of *of*-PPs. It will be noted that the distribution of *of*-PPs in our dataset is consistent with these odds ratios: the majority consist of noun phrases which are either not animate, or are not pronouns, or have a

word length of 2 or greater. The conclusion we draw is that the properties of *of*-PPs as dependents of *one*<sub>ct</sub> are not distinct from those of *of*-PPs in the genitive alternation generally, and that *one*<sub>ct</sub> itself has no special import in this regard.

There are in fact only 3 examples in the dataset where the *of*-PP contains a pronoun of length one word which is categorized as animate (rather than organization or inanimate). This is the kind of example which we would strongly predict to be an 's-genitive rather than an *of*-PP. These 3 examples all represent the DEPICTION relationship: for example *the ones of me*, where *ones* stands for “photos” and *me* represents the person depicted, rather than the controller or creator. This, we believe, is not fortuitous. It is not just the animacy and form of the dependent which is important: the precise semantic relation may also have an effect on the relative frequency of the *one*<sub>ct</sub> + *of*-PP and 's genitive + zero anaphor constructions. Payne & Berlage (2011) investigated the relative weight of a number of semantic relations with respect to the general alternation between the *of*-PP and 's-genitive constructions, with other important factors such as length, animacy and definiteness of the dependent controlled for. They show that, amongst the relations in principle available to both constructions, CONTROLLER most favours the 's genitive while DEPICTION most favours *of*-PP. Thus examples such as *the photos of me*, with pronouns as dependents and depiction as the intended relation, are not improbable.

Payne & Berlage (2011) also found that the BODY-PART relation lies somewhere in between CONTROLLER and DEPICTION. This mirrors the numbers observed in Table 2 for these relations, with DEPICTION showing 9 occurrences for *one*<sub>ct</sub> + *of*-PP and CONTROLLER only 2. The frequencies of these relations in Table 2 are however far too small to support statistical tests of significance, so we leave the correlation between the individual semantic relations and the frequency of *one*<sub>ct</sub> + *of*-PP as a plausible prediction.

If we now turn to the examples shown in Table 3, where the genitive alternation plays no role, the potential for an alternative to the *of*-PP construction is considerably more limited. In the relatively frequent partitive (SUBSET) cases, the presence of *one*<sub>ct</sub> is largely determined by the presence of a premodifier. A typical example is the one in the Table: *a new one of these*. As a consequence of the search methodology (see section 3.3), all the singular examples of *one*<sub>ct</sub> are all of this type. As we would predict, the majority of the plural examples also have premodification, but there are a couple of examples which contain *one*<sub>ct</sub> immediately after a determiner (where it could well have been omitted): *which ones of the eager faces...?*, and *certain ones of these*. The distribution of *one*<sub>ct</sub> in the partitive is thus motivated by factors other than the partitive construction per se, and no different to the distribution of *one*<sub>ct</sub> generally.

In the non-partitive cases there can be an alternation between *of*-PP and a prehead dependent, either nominal or adjectival. Thus, although we do not have *\*law or fact's questions* as an alternative to *questions of law or fact*, we do have *legal or factual questions*. And correspondingly we might have *legal or factual ones* as an alternative to the attested *ones of law or fact*. To our knowledge, there is no previous large-scale investigation of this kind of alternation, but we can make the following observations based on our dataset.

Firstly, the range of semantic relations which is permitted by the prehead dependent construction is very broad and appears in principle to be a superset of the semantic relations permitted by the *of*-PP construction. That is, for each non-partitive semantic relation in Table 3, with the exception of the KIND relation where structural factors (see below) prevent it, we can find

at least one example where the prenominal alternant appears to be grammatical. We illustrate this in Table 5:

Semantic relation	Ellipted nominal	Attested <i>of</i> -PP	Prehead alternant
<b>content</b> (h)(d)	questions	ones of law or fact	legal or factual ones
<b>image</b> (h)(d)	postcard	the other one of New Zealand	the other New Zealand one
<b>size</b> (h)(d)	farms	ones of 100,000 acres	100,000 acre ones
<b>value</b> (h)(d)	condition	the only one of relevance	the only relevant one
<b>theme</b> (h)(d)	impression	a great one of Christopher Watkins	a great Christopher Watkins one
<b>cause</b> (h)(d)	cries	ones of anguish	anguished ones
<b>composition</b> (h)(d)	panels	more serviceable ones of sturdy plywood	more serviceable sturdy plywood ones
<b>age</b> (h)(d)	ball	the only one of similar age	the only similar age one
<b>timespan</b> (h)(d)	period	the early one of railway buliding	the early railway buliding one
<b>container</b> (h)(d)	bottles	smaller stone ones of ale or whisky	smaller stone ale or whisky ones
<b>duration</b> (h)(d)	pregnancy	a short one of 90-105 days	a short 90-105 day one
<b>rank</b> (h)(d)	officers	ones of much more senior rank from the military wing	much more senior rank ones from the military wing
<b>amount</b> (h)(d)	proportion	a considerable one of lime	a considerable lime one
<b>collection</b> (h)(d)	group	a short one of two figures	a short two figure one
<b>type</b> (h)(d)	version	not a single one of Deus venerunt gentes	not a single Deus venerunt gentes one

Table 5: prehead alternants (nominal or adjectival) to the *of*-PP construction

In practice, the prehead alternative is in the majority of the attested cases categorically blocked by structural factors. A prehead alternative does not exist when the *of*-PP is clausal, as in an example such as *the fundamental process of turning customer needs into customer wants*. The prehead constituent also cannot be a determined nominal, nor can it contain any post-head dependent. Thus, as an alternative to *detailed pictures of a storm on Saturn*, we do not have *\*detailed a storm on Saturn pictures*. Correspondingly, there is no prehead alternant to the attested *detailed ones of a storm on Saturn*. It is for this reason that the KIND examples are systematically excluded: the dependent, headed by a noun such as *kind*, invariably contains a determiner. A typical example is *the only brewery of its kind in Worcestershire*.

The length of the dependent is also evidently an important factor. Whilst, we suspect, the vast majority of prehead modifiers are single-word units rather than internally complex ones, the *of*-PPs are typically at least two words long. Of the 179 examples which represent the semantic relations in Table 5 (i.e. the relations in Table 3 excluding SUBSET and KIND), only 35 are one-word long, that is 20%, a figure similar to the proportion of one-word *of*-PPs in the genitive alternation.

In conclusion, we have found in a detailed examination of the *of*-PP dependents of *one<sub>ct</sub>* absolutely no evidence that *one<sub>ct</sub>* itself has any special bearing on the frequency of occurrence of the *of*-PP, let alone its grammaticality. The frequencies which are observed are essentially those

we would expect given the properties of *of*-PPs as dependents of nouns in general.

### 5.3. HUMAN HEAD NOUNS.

We have not yet cited any examples of *one*<sub>ct</sub> + *of*-PP where the antecedent belongs to certain types of human head noun which are usually considered inherently relational. These are nouns denoting interpersonal or kin relations (e.g. *friend*, *brother*), role nouns (e.g. *king*), and indeed agent nominalizations of the *student* type. Such examples do not occur, to our knowledge, in the BNC. But this, we believe, is simply a consequence of the limited size of the BNC rather than grammaticality as such, since natural-sounding examples of the relevant kind are certainly attested in larger corpora. We cite some web examples, identified as almost certainly produced by native speakers, in 37–39.<sup>22</sup> The examples in 37 illustrating interpersonal and kin relations form a new fifteenth semantic field. However, role nouns can be subsumed under the function noun field, and agent nominalizations under the object-like dependent field.

- (37) a. WAGs (wives and girlfriends, usually the badly behaving [ones of English sports stars])  
{*wives and girlfriends*; *d* is interpersonal relation of *h*}
- b. Both the parents of children with difficulties and [the ones of children with a normal evolution] must be contacted to settle educational programs that involve the family.  
{*parents*; *d* is kin relation of *h*}
- (38) Dudley himself was no more eager for the match. Yes, he wanted to marry with a queen, but not [the one of Scotland].  
{*queen*; *d* has role with respect to *h*}
- (39) a. Despite the rivalry between the two sides, supporters, specially [the ones of Real Madrid] are known to show respect to the individual talents in the opposition team.  
{*supporters*; *d* is undergoer / *h* is agent}
- b. A single company, ArkivMusic, has struck deals with all four major publishers (and numerous minor ones) of classical music recordings to make their deleted records available via a burn-on-demand service.  
{*publishers*; *h* is theme, *d* is agent}

The two examples of agent nominalizations in 39 differ in that 39a has an 's-genitive counterpart. It would be possible to say *Real Madrid's* with a zero anaphor, just as well as *the ones of Real Madrid*.

Is it possible to find an example in all relevant respects like Baker's original example 5, with the particular agent nominalization *student*? Users of corpora will know that finding specific strings is virtually impossible even for quite short string lengths (the probability of a possible *k*-word string at any arbitrary point in a text being identical with some specific string is approximately  $1$  in  $10^{2k}$ ). The difficulty of finding an occurrence is further reduced if a specific antecedent (*student*) is called for, given that it may be arbitrarily distant from the occurrence of *one*<sub>ct</sub> + *of*-PP. However, consider example 40:

- (40) In the case of medicine, I think there's no other alternative than the Universidad de la Republica. I would think their classes are equally crowded, but haven't ever heard

any of the medicine students complain as much as the [*ones of computer science*].  
 {*students*; *h* is theme, *d* is agent}

This example — offering advice to a North American about Uruguayan university entrance requirements — may be from a non-native speaker, but the writer’s English betrays no obvious non-native traces and the example sounds entirely natural to us.

As we have shown in section 5.2, the *of*-PP construction has more than just the ’s genitive as a competitor: there is also the possibility in many cases of employing an NP with a simple nominal or adjectival pre-head modifier. As well as *the queen of Scotland* (or *Scotland’s queen*) in 38, we could have *the Scottish queen*. And as well as *the supporters of Real Madrid* (or *Real Madrid’s supporters*) in 39a, we could also have *the Real Madrid supporters*. With agent nominalizations like *student*, it is the only competitor: as well as *the students of medicine*, we could have either *the medicine students* (which the author in 40 actually chooses for the antecedent), or *the medical students*.

A BNC investigation into the relative frequencies of *of*-PPs and pre-head modifiers with the head noun *student* reveals that, at least with single word dependents, the pre-head modifier construction very strongly predominates (see Table 4).

<i>of</i> -PP	EXAMPLES	PRE-HEAD MODIFIER	EXAMPLES
student of physics	1	physics student	7
student of science	2	science student	6
student of chemistry	0	chemistry student	6
student of medicine	0	medical student	64

Table 4: frequencies in the BNC of *of*-PP and pre-head modifier with the head noun *student*

Length of the dependent is a relevant variable: the longer the dependent is, the more likely the *of*-PP construction becomes. A survey of *of*-PP constructions with the head noun *student* reveals that the mean length of the dependent is 2.1 words (number of examples = 423; standard deviation = 1.6). It is in conformity with this length principle, therefore, that in example (40) the author chooses *the ones of computer science* over *the computer science ones*.

In conclusion, there are no grounds for considering examples like *the one of physics* (with *student* as antecedent) to be ungrammatical. No syntactic principle excludes such phrases. They are simply a non-preferred option given a short dependent.

**6. THE IMPLICATIONS FOR ACQUISITION.** The footprint of Baker’s arguments in the literature is huge. Textbook authors often rely on them to motivate X-bar theory (e.g. Radford 1981:92–100, 1988:174ff; Carnie 2002:122; Burton-Roberts 2011:165–170). At least five different works by David Lightfoot and coauthors repeat Baker’s arguments in connection with arguing for universal grammar (Hornstein & Lightfoot 1981:18ff; Lightfoot 1982, chapter 4; Lightfoot 1989:322f; Lightfoot 1991:4–8; Anderson & Lightfoot 2002:196–198). Baker’s thesis is treated as uncontroversially established not only within psycholinguistics (e.g. Hamburger & Crain 1984; Crain 1991:609ff) but also occasionally by philosophers of cognitive science (see e.g. Ramsey & Stich 1991:295). There has been renewed recent theoretical discussion of the facts (Oga 2001; Panagiotidis 2003; Gualmini 2007), and prolonged debate has been stimulated since 2003 by a

series of experimental papers on the early-acquisition claim, and critiques or defenses thereof (Lidz, Waxman and Freedman 2003; Akhtar et al. 2004; Regier & Gahl 2004; Tomasello 2004; Lidz & Waxman 2004; Foraker et al. 2009; Pearl & Lidz 2009).

Unfortunately all of this work has been based on descriptive error. The facts about anaphoric *one* are not as Baker assumed, and once they are properly understood not a trace of Baker's supportive argumentation for innateness survives.

Baker actually supplies two distinct arguments, each associated with a specific fact to be acquired—what Pullum & Scholz (2002) call an ACQUIRENDUM. The two acquirenda are:

- (41) a. A single word of the lexical category N cannot be the antecedent for *one*<sub>ct</sub>.  
b. A multi-word phrase of the category Nom can be the antecedent for *one*<sub>ct</sub>.

Confirming 41a would call for negative information: that *one*<sub>ct</sub> can never be anteceded by a noun that has a complement (as opposed to being anteceded by the whole Nom constituent containing the noun and the complement together). But nobody is ever supplied with this information, so 41a gives rise to what Pullum & Scholz call a STIMULUS-ABSENCE argument for linguistic nativism: nothing in the environment could directly supply the information necessary for learning. But given the evidence we have provided to show that 41a is not true, this collapses. Nothing entailing 41a is acquired by those who become speakers of English — and for anything entailing 41a to be innate would prevent attainment of the adult state of knowledge of language.

To confirm 41b, by contrast, positive information could in principle suffice: if some utterance act could convince you (by occurring in a context where nothing else makes sense) that *one*<sub>ct</sub> must have a multi-word Nom as its antecedent, you would have learned that multi-word Nom antecedents are possible. So 41b gives rise to what Pullum & Scholz call a STIMULUS-POVERTY argument.

Baker gave an example of the kind of rare but in-principle-accessible evidence that would permit 41b to be learned. He pointed out that in a context where Alice has a red glass in her hand, encountering 42 would provide relevant evidence.

- (42) John has a blue glass, but Alice doesn't have one.

*One* cannot mean *glass* here, on pain of contradicting the visible evidence; yet if it is taken to stand for *blue glass*, everything makes sense. Hence multi-word antecedents must be permissible.

Baker made a mistake here. The *one* in 42 is not the noun; it is the determinative. That is why the plural form would be impermissible (*\*John has some toys but Alice doesn't have ones*). And the antecedent in 42 is not a Nom, but the whole indefinite NP *a blue glass*. This can be remedied: we could replace 42 with something like *John has a blue glass, but we couldn't find another one for Alice*, which does have *one*<sub>ct</sub>. It is rather remarkable, though, that through all the repetitions of his point no one ever noticed that Baker's examples did not illustrate his point.

Events like hearing 42 in a context where Alice has a non-blue glass are referred to by Akhtar et al. 2004 as BAKER EVENTS. What Baker says about them is that they 'must certainly be extremely uncommon in a child's early experience.' He offers no support at all for this assertion. And in fact the frequency of Baker events remains unknown. Their frequency might well be

adequate to ensure that random linguistic experience would soon refute the one-word-antecedent hypothesis, but there has been no large-scale study of this; nearly everyone has been content to repeat what Baker said.

Lidz et al. (2003) is an exception. Lidz et al., to their credit, attempted to assess the frequency of Baker events by looking for them in corpora of speech addressed to young children, and they claimed to have found one in the Adam corpus of the CHILDES database and one in the Nina corpus. Unfortunately, both of their examples (which Jeff Lidz kindly showed us) are mistaken diagnoses: they contain *one<sub>d</sub>*. Our own explorations turned up a few apparent Baker events in the Lara corpus (see e.g. lines 441, 770, 912, 1179, and 1218), but we confess it can be very hard to tell from transcripts of interactions with young children, and more work is needed.

It is not difficult to make a preliminary assessment of what the frequency might be in arbitrary text, however. We examined every occurrence of *one* or *ones* in three texts to get a rough sense of how many of them represented Baker events. What we looked for were instances of *one<sub>ct</sub>* for which the only reasonable assumption given the context was to understand them as having multi-word antecedents. The results are in Table 5. What they show is that in each million words of arbitrary text we can expect about 35 utterances that in effect indicate the analog of Baker events. If conversations with children are like other kinds of text in this regard, then since children hear 10 to 30 million words before they are three (Hart and Risley 1995), one might expect 3-year-olds to have encountered between 350 and 1000 Baker events. That is by no means vanishingly small. Understanding some of the utterances involved might be enough to support purely experience-based learning of the fact that multi-word antecedents are possible.

	word count	<i>one(s)</i>	<i>one<sub>ct</sub></i>	Baker events	per million
<i>Wall Street Journal</i> corpus, w7_001	160,000	262	25	6	37.5
<i>Alice's Adventures in Wonderland</i>	26,000	78	4	1	38.5
<i>Anne of Avonlea</i>	90,000	173	22	3	33.3

Table 5: Numbers of apparent Baker events in three texts

We note in addition that there is reason to doubt that a sound stimulus-poverty argument for linguistic nativism can be based on an acquirendum like 41b, and it is important. Whether Baker events are common enough to be relevant or not, if what gets learned is simply that *one<sub>ct</sub>* is an anaphor of type  $\langle e, t \rangle$ , then it is not clear why 41b would ever be doubted by an unprejudiced learner: nouns and Nom constituents can both have that type, so the learner who makes the broadest assumption, namely that anything of type  $\langle e, t \rangle$  will do, will be correct.

A further remark to be made about the developmental-linguistic and psycholinguistic literature is that it is vitiated by frequent shifts in the presumed acquirendum, none of them being accurate. Lidz, Waxman, and Freedman (2003) started out by taking the child's task to be to learn something like Baker's original syntactic claim: that '*one* is anaphoric to the phrasal category [Nom]' — a claim we have shown to be false. Lidz et al. attempted to test whether young children assume that claim. They familiarized 18-month-old infants with a screen display of a yellow bottle accompanied by the utterance 'Look! A yellow bottle!', and then showed both a

yellow bottle and a blue bottle accompanied by either ‘Now look: what do you see now?’ (the control condition) or ‘Now look: do you see another one?’ (the test condition). The idea was that if the children knew that *one<sub>ct</sub>* was an anaphor seeking a Nom as antecedent (and not a noun), *another one* would be interpreted with *one* taking the Nom *yellow bottle* as its antecedent, so the infants would take a longer look at the yellow bottle. (See Akhtar et al. and Tomasello 2004 for detailed criticism.)

The assumption Lidz et al. appear to make is that infants will pick the longest possible antecedent (for notice, *bottle* on its own would also be a Nom). Regier & Gahl (2004) make this explicit in their response, exhibiting a Bayesian strategy that could learn from positive data that *one* must take as antecedent a larger rather than a smaller Nom. But that is not the generalization that competent adult speakers acquire.

Lidz & Waxman (2004) reply to Regier & Gahl, but restate the acquirendum in a slightly different and non-equivalent form: ‘*one* is anaphoric only to syntactic constituents larger than  $N^0$ ’, which neither entails nor is entailed by the former one. Assuming that ‘larger’ means ‘longer’, it entails that *one<sub>ct</sub>* can ONLY have multi-word antecedents, and that is certainly not true: *one<sub>ct</sub>* frequently has one-word antecedents.

Pearl & Lidz (2009) present a fuller response to Regier & Gahl, but change the acquirendum yet again, proposing (p. 239): ‘Anaphoric *one* can take any Nom as an antecedent, but a multi-word antecedent is preferred when it is available.’ They claim that ‘when there is more than one [Nom] to choose from . . . adults prefer the [Nom] corresponding to the longer string,’ and children ‘have the adult pragmatic preference to choose the referent corresponding to the larger [Nom] string when there is more than one [Nom] antecedent.’ This too is false (as well as slightly different from all the earlier work). For instance, it is flatly contradicted by example 23 in Baker (1978:419), *The student with short hair is taller than the one with long hair. If student with short hair were preferred over student as antecedent, the predicted interpretation would be that the one with long hair means ‘the student with short hair who has long hair.’*

Such shifts and inaccuracies wreck the chances of getting a result that bears on the acquisition of *one<sub>ct</sub>* or the issue of linguistic nativism. Unless the participants can agree on what acquirendum they are talking about, they can never succeed in determining whether its acquisition calls for innate linguistic prerequisites. And in this case not only have the parties all picked different acquirenda, but in addition the acquirenda they have picked do not hold in the language to be acquired.

In sum, psycholinguists working on anaphoric *one* have (i) failed to validate the claim that bare noun antecedents are illicit (which they are not); (ii) confused the crucial item with one of its homonyms; (iii) failed to establish that Baker events are rare; and (iv) shifted their assumptions about the acquirendum from study to study. In consequence, the results obtained have agreed neither with each other, nor with linguists’ assumptions about what was to be shown, nor with what (under our analysis) actually has to be acquired.

**7. CONCLUSION.** Nothing remains of the factual basis for an argument from either stimulus absence or stimulus poverty running along the lines Baker suggested. One rests on refuted data and the other is entirely inconclusive.

It is worth reflecting on why anyone could think it likely that a learner would ever assume a

one-word limit on antecedents for an anaphor. What the learner is looking for (if we are anywhere near right) is a meaning to assign. *Glass* and *blue glass* and *pretty blue glass that John is holding* are all expressions of type  $\langle e, t \rangle$ . An unbiased hunt for a type  $\langle e, t \rangle$  antecedent should be content with finding any  $\langle e, t \rangle$  that fits the context. There is no reason to think the word count should matter.

The preferential-looking experiments of LWF and the Bayesian-learning simulations that emerged in the subsequent discussion all involved several shifts in the acquirendum, and all of the investigation was undertaken without any reinvestigation of the relevant English data. In consequence, neither the nativist nor the non-nativist strands of the work arrive at any results that carry conviction. The new puzzle that arises is how  $one_{ct}$  can be promiscuous enough to allow either a complement-taking noun or a whole nominal to be its antecedent and supply its sense. We have provided a formal semantic analysis that answers that question. It leaves us with no specific reason to think that learning Baker's positive acquirendum from the evidence is problematic: given only that  $one_{ct}$  is identified as an anaphor of semantic type  $\langle e, t \rangle$  (and even linguistic nativists have to assume that much can be learned from exposure to speech, since  $one_{ct}$  is not universal), it automatically follows that the meanings of nominals (combinations of nouns with their dependents) will be suitable meaning donors.

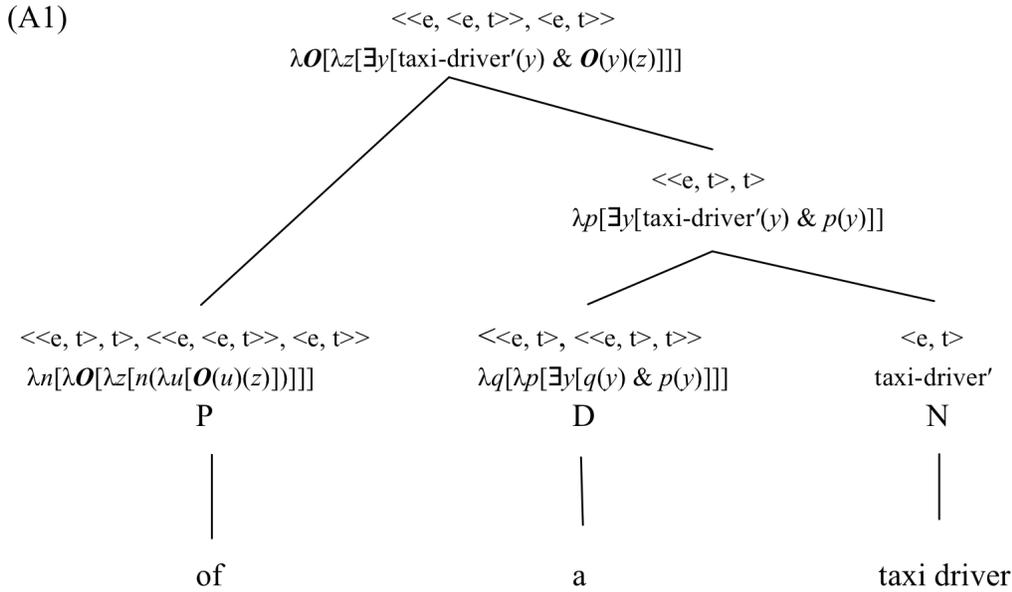
It is unfortunate that the work on anaphoric uses of  $one_{ct}$  began in such a resolutely syntactic mode. No one seems to have thought much about its meaning, or the implications thereof. The fact is that for a child capable of identifying nouns and conjecturing meanings for them, learning  $one_{ct}$  looks rather easy. It is a count noun with hardly any semantic content. A phrase like *a big one* has a meaning something like 'a big thing (of the indicated sort)'. It is scarcely a mystery how a child capable of learning noun meanings could learn a particularly bland and general one such as this.

Learning to use  $one_{ct}$  in particular constructions does not seem to call for anything but positive evidence. No subtle constraint on the category of its antecedent has to be learned: either nouns or multi-word Nom will do. An *of*-PP, or any other PP, can compose with  $one_{ct}$  when a plausible meaning results. On the basis of exposure to the range of alternative constructions like NPs with a genitive determiner or with premodifying nouns and adjectives, the learner will be encouraged to use them in ways that match linguistic experience: use *mine* or *my one* rather than *the one of me* in cases of control or possession; use *the one of me* sometimes with depictive nouns; and so on. The probability matching seen in young language learners' adaptation to the speech of their caregivers is well known. There is no reason to expect this natural process to be switched off when it comes to learning alternations between syntactic constructions. And there is also no reason to consider low-probability examples like *the car of John* or *the ones of physics* as being grammatically 'blocked' by their more probable alternants. No syntactic principle forbids these low-probability constructions; if they sound a bit odd in isolation it is merely because they are less preferred.

Neither theoretical arguments in support of linguistic nativism nor experimental work in developmental psycholinguistics can amount to much if they are based on flawed descriptive linguistics. It is somewhat shocking to reflect on the fact that the syntactic conditions on  $one_{ct}$  have been touted for thirty years as a prime example of a linguistic discovery supporting the plausibility of linguistic nativism when the whole factual basis of the case, presupposed in all the psycholinguistic work, was mistaken.

If language acquisition is ever to be scientifically understood, observation of children's language and child-directed speech will have to proceed in parallel with controlled psychological experiments, and computational modeling, and above all, careful description of the linguistic system that is acquired. But notice, we are not suggesting at all that experimentation and modeling can take over and eliminate the need for theoretical and descriptive linguistics. Having a sound theoretically-based description of the linguistic system to be acquired is surely crucial if progress is to be made on explaining acquisition. Efforts at explaining the acquisition of a linguistic system are doomed to failure if the presupposed description of the acquired system is grossly inaccurate, as has proved to be the situation here.

APPENDIX. Derivation of *a brutal murder of a taxi-driver*. To derive the meaning of the (attested) phrase *a brutal murder of a taxi-driver*, we proceed by first constructing the translation of the *of*-PP:



We straightforwardly apply the translation of the determiner *a* to the translation of the noun *taxi-driver* to derive the translation of the NP *a taxi driver* as in A2a. We then apply the translation of the preposition *of* to the translation of the NP as in A2b. Here, the new variables *u* and *z* are of type *e*, the variable *n* is of type  $\langle\langle e, t \rangle, t \rangle$ , and the variable *O* is of the relational type  $\langle e, \langle e, t \rangle \rangle$ .

- (A2)
- a.  $\lambda q[\lambda p[\exists y[q(y) \ \& \ p(y)]]](\text{taxi-driver}')$   
 $= \lambda p[\exists y[\text{taxi-driver}'(y) \ \& \ p(y)]]$
  - b.  $\lambda n[\lambda \mathbf{O}[\lambda z[n(\lambda u[\mathbf{O}(u)(z)])]]](\lambda p[\exists y[\text{taxi-driver}'(y) \ \& \ p(y)]])$   
 $= \lambda \mathbf{O}[\lambda z[\lambda p[\exists y[\text{taxi-driver}'(y) \ \& \ p(y)]](\lambda u[\mathbf{O}(u)(z)])]]$   
 $= \lambda \mathbf{O}[\lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \lambda u[\mathbf{O}(u)(z)](y)]]]$   
 $= \lambda \mathbf{O}[\lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \mathbf{O}(y)(z)]]]$

The translation of the preposition *of* contains a variable  $\mathbf{O}$  which we take to range over the wide set of semantic relations which this preposition permits. By altering the range of this variable, the analysis can therefore in principle be extended to any other preposition.

We can then represent the structure of the full NP with the analysis tree in 32. The typing here forces first the composition of the translation of *brutal* with that of *murder*, as in A3a.<sup>23</sup> This translation is of type  $\langle e, t \rangle$ , and before it can combine with the translation of the *of*-PP must be shifted to the relational type  $\langle e, \langle e, t \rangle \rangle$  by the relational operator 30a. This type-shifting is shown in A3b.

- (A3)      a.     $\lambda p[\lambda y[p(y) \ \& \ \text{brutal}'(y)]](\text{murder}')$   
                        $= \lambda y[\text{murder}'(y) \ \& \ \text{brutal}'(y)]$
- b.     $\lambda p[\lambda d[\lambda h[p(h) \ \& \ \mathbf{R}(h)(d)]]](\lambda y[\text{murder}'(y) \ \& \ \text{brutal}'(y)])$   
                        $= \lambda d[\lambda h[\lambda y[\text{murder}'(y) \ \& \ \text{brutal}'(y)](h) \ \& \ \mathbf{R}(h)(d)]]$   
                        $= \lambda d[\lambda h[\text{murder}'(h) \ \& \ \text{brutal}'(h) \ \& \ \mathbf{R}(h)(d)]]$

The logical translation in A3b therefore represents a relational interpretation of *brutal murder*.

The translation of the *of*-PP can then apply to the translation of *brutal murder* as in A4:

- (A4)             $\lambda \mathbf{O}[\lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \mathbf{O}(y)(z)]]](\lambda d[\lambda h[\text{murder}'(h) \ \& \ \text{brutal}'(h) \ \& \ \mathbf{R}_{\mathbf{O}}(h)(d)]]])$   
                        $= \lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \lambda d[\lambda h[\text{murder}'(h) \ \& \ \text{brutal}'(h) \ \& \ \mathbf{R}_{\mathbf{O}}(h)(d)]](y)(z)]]$   
                        $= \lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \lambda h[\text{murder}'(h) \ \& \ \text{brutal}'(h) \ \& \ \mathbf{R}_{\mathbf{O}}(h)(y)](z)]]$   
                        $= \lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \text{murder}'(z) \ \& \ \text{brutal}'(z) \ \& \ \mathbf{R}_{\mathbf{O}}(z)(y)]]$

The relational variable  $\mathbf{O}$ , which represents the range of semantic relations permitted by the preposition *of*, imposes a constraint on the instantiation of  $\mathbf{R}$ : whatever semantic relation is chosen to instantiate  $\mathbf{R}$  must lie within this range. In A4 we represent this constraint by subscripting  $\mathbf{R}$  accordingly, i.e.  $\mathbf{R}$  is restricted to  $\mathbf{R}_{\mathbf{O}}$ .

It is then straightforward to derive the translation of the full NP by applying the translation of the indefinite article to the translation of *brutal murder of a taxi driver*, as in A5:

- (A5)             $\lambda q[\lambda p[\exists x[q(x) \ \& \ p(x)]]](\lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \text{murder}'(z) \ \& \ \text{brutal}'(z) \ \& \ \mathbf{R}_{\mathbf{O}}(z)(y)]]])$   
                        $= \lambda p[\exists x[\lambda z[\exists y[\text{taxi-driver}'(y) \ \& \ \text{murder}'(z) \ \& \ \text{brutal}'(z) \ \& \ \mathbf{R}_{\mathbf{O}}(z)(y)]](x) \ \& \ p(x)]]$   
                        $= \lambda p[\exists x[\exists y[\text{taxi-driver}'(y) \ \& \ \text{murder}'(x) \ \& \ \text{brutal}'(x) \ \& \ \mathbf{R}_{\mathbf{O}}(x)(y)] \ \& \ p(x)]]$

The final line of A5 thus corresponds to the translation of the full NP given in 32.<sup>24</sup>

## REFERENCES

- AKHTAR, NAMEERA; MAUREEN CALLANAN; GEOFFREY K. PULLUM; and BARBARA C. SCHOLZ. 2004. Learning antecedents for anaphoric one. *Cognition* 93.141–145.
- ANDERSON, STEPHEN R., and DAVID W. LIGHTFOOT. 2002. *The Language Organ*. Cambridge: Cambridge University Press.
- ARIEL, MIRA. 2001. Accessibility theory: An overview. In Ted Sanders, Joost Schilperoord and Wilbert Spooren (eds) *Text Representation: Linguistic and Psycholinguistic Aspects* [Human Cognitive Processing 8], 29-87. Amsterdam: John Benjamins.
- ASUDEH, ASH. 2005. Relational nouns, pronouns and resumption. *Linguistics and Philosophy* 28, 395-446.
- BAKER, C. L. 1978. *Introduction to Generative Transformational Syntax*. Englewood Cliffs, NJ: Prentice-Hall.
- BAKER, C. L. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry* 10.533–581.
- BARKER, CHRIS. 1995. *Possessive descriptions*. Stanford, CA: CSLI Publications.
- BARKER, CHRIS. 2011. Possessives and relational nouns. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds) *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, 1109–30. Berlin: de Gruyter.
- BÖRJARS, KERSTI; DAVID DENISON; and ALAN SCOTT. 2013. Expression of possession in English. The significance of the right edge. In Kersti Börjars, David Denison and Alan Scott (eds) *Morphosyntactic Categories and the Expression of Possession*, 123–48. Amsterdam: John Benjamins.
- BURTON-ROBERTS, NOEL. 2011. *Analysing Sentences*. Harlow, UK: Pearson, 3rd edition.
- CARNIE, ANDREW. 2002. *Syntax: A Generative Introduction*. Oxford: Basil Blackwell.
- CHOMSKY, NOAM. 1980. *Rules and Representations*. Oxford: Blackwell.
- CRAIN, STEPHEN. 1991. Language acquisition in the absence of experience. *Behavioral and Brain Sciences* 14.597–612.
- DALRYMPLE, MARY; STUART M. SHIEBER; and FERNANDO C. N. PEREIRA. 1991. Ellipsis and Higher Order Unification. *Linguistics and Philosophy* 14, 399–452.
- DALRYMPLE, MARY (ed.). 1999. *Semantics and syntax in Lexical Functional Grammar: The resource logic approach*. Cambridge, MA: MIT Press.
- DAVIES, MARK. 2008–. *The Corpus of Contemporary American English: 450 million words, 1990-present*. Available online at <http://corpus.byu.edu/coca/>
- DEBRUIN, JOS, and REMKO J. H. SCHA. 1988. The interpretation of relational nouns. *26th Annual Meeting of the Association for Computational Linguistics (Buffalo), Proceedings of the Conference*, 25–32. Morristown, NJ.
- FORAKER, STEPHANI; TERRY REGIER; NAVEEN KHETARPAL; AMY PERFORNS; and JOSHUA TENENBAUM. 2009. Indirect evidence and the poverty of the stimulus: The case of anaphoric one. *Cognitive Science* 33.287–300.
- GRIMSHAW, JANE. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
- GUALMINI, ANDREA. 2007. On that One poverty of the stimulus argument. In *Papers from the Language Acquisition Workshop, SCL 2006*, ed. by Merete Anderssen and Marit R. Westergaard, 153–171. Tromsø: CASTL. Published as Nordlyd 34, no. 3.

- HAMBURGER, HENRY, & CRAIN, STEPHEN. (1984). Acquisition of cognitive compiling. *Cognition* 17.85–136.
- HART, BETTY, and TODD R. RISLEY. 1995. *Meaningful Differences in the Everyday Experiences of Young Children*. Baltimore: Paul H. Brookes.
- HINRICH, LARS, and BENEDIKT SZMRECSANYI. 2007. Recent changes in the function and frequency of Standard English genitive constructions: a multivariate analysis of tagged corpora. *English Language and Linguistics* 11.437–474.
- HORNSTEIN, NORBERT, and DAVID LIGHTFOOT. 1981. Introduction. In *Explanation in Linguistics*, ed. by Norbert Hornstein David Lightfoot, 9–31. London: Longman.
- HUDDLESTON, RODNEY, and GEOFFREY K. PULLUM et al. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- JACKENDOFF, RAY S. 1977.  $\bar{X}$  syntax. Cambridge, MA: MIT Press.
- KIMBALL, JOHN. 1973. *The Formal Theory of Grammar*. Englewood Cliffs, NJ: Prentice-Hall.
- LAKOFF, GEORGE. 1970. Global rules. *Language* 46.627–639.
- LDC. 1993. ACL/DCI [CD ROM produced by the Data Collection Initiative of the Association for Computational Linguistics, including the *Wall Street Journal* corpus]. Catalog no. LDC93T1. Linguistic Data Consortium, University of Pennsylvania, Philadelphia PA.
- LIDZ, JEFFREY; SANDRA WAXMAN; and JENNIFER FREEDMAN. 2003. What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition* 89.B65–B73.
- LIDZ, JEFFREY, and SANDRA WAXMAN. 2004. Reaffirming the poverty of the stimulus argument: A reply to the replies. *Cognition* 93.157–165.
- LIGHTFOOT, DAVID. 1982. *The Language Lottery: Toward a Biology of Grammars*. Cambridge, MA: MIT Press.
- LIGHTFOOT, DAVID. 1989. The child's trigger experience: degree-0 learnability. *Behavioral and Brain Sciences* 12.321–334.
- LIGHTFOOT, DAVID. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press.
- LÖBNER, SEBASTIAN. 1985. Definites. *Journal of Semantics* 4, 279-326.
- MITCHELL, JONATHAN. 1986. *The Formal Semantics of Point of View*. University of Massachusetts at Amherst: Ph.D. dissertation.
- O'CONNOR, CATHERINE, JOAN MALING and BARBORA SKARABELA. 2013. Nominal categories and the expression of possession: A cross-linguistic investigation of probabilistic tendencies and categorical constraints. In Kersti Börjars, David Denison and Alan Scott (eds) *Morphosyntactic Categories and the Expression of Possession*, 89–122. Amsterdam: John Benjamins.
- OGA, KYOKO. 2001. Two types of 'of' and theta-role assignment by nouns. In Mamiko Akita and Kyoko Oga (eds) *Newcastle and Durham working papers in linguistics* 6.95–108. Durham: University of Durham, Department of English and Linguistics.
- PANAGIOTIDIS, PHOEVOS. 2003. One, empty nouns and  $\theta$ -assignment. *Linguistic Inquiry* 34.281–292.
- PARTEE, BARBARA. 1989. Binding implicit variables in quantified contexts. *Chicago Linguistic Society* 25, 342-365. Reprinted in Barbara H. Partee 2004. *Compositionality in Formal*

- Semantics: Selected Papers by Barbara H. Partee*, 259-281. Oxford: Blackwell Publishing.
- PARTEE, BARBARA H., and VLADIMIR BORSCHEV. 2003. Genitives, relational nouns, and argument-modifier ambiguity. In Ewald Lang, Claudia Maienborn & Cathrine Fabricius-Hansen (eds) *Modifying Adjuncts (Interface Explorations 4)*, 67–112. Mouton de Gruyter.
- PARTEE, BARBARA H., and VLADIMIR BORSCHEV. 2012. Sortal, relational, and functional interpretations of nouns and Russian container constructions. *Journal of Semantics* 29, 445–86.
- PAYNE, JOHN, and EVA BERLAGE. 2009. Pro-nominal *one* and relational nouns. Paper presented at ICAME 30, University of Lancaster.
- PAYNE, JOHN, and EVA BERLAGE. 2011. The effect of semantic relations on genitive variation. Paper presented at ISLE 2, Boston University.
- PAYNE, JOHN, and RODNEY HUDDLESTON. 2002. Nouns and noun phrases. Ch. 5 of Rodney Huddleston & Geoffrey K Pullum et al., *The Cambridge Grammar of the English language*, 323–523. Cambridge University Press.
- PEARL, LISA, and JEFFREY LIDZ. 2009. When domain-general learning fails and when it succeeds: Identifying the contribution of domain specificity. *Language Learning and Development* 5.235–265.
- POSTAL, PAUL M. 1972. The best theory. P. Stanley Peters (ed.), *Goals of Linguistic Theory*, 131–179. Englewood Cliffs, NJ: Prentice-Hall.
- PULLUM, GEOFFREY K., and BARBARA C. SCHOLZ. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19.9–50.
- PUSTEJOVSKY, JAMES. 1991. The generative lexicon. *Computational Linguistics* 17, 409–441.
- RADFORD, ANDREW. 1981. *Transformational Syntax: A Student's Guide to Chomsky's Extended Standard Theory*. Cambridge: Cambridge University Press.
- RADFORD, ANDREW. 1988. *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.
- RAMSEY, WILLIAM, and STEPHEN P. STICH. 1991. Connectionism and three levels of nativism. In *Philosophy and Connectionist Theory*, ed. by William Ramsey, Stephen P. Stich, David E. Rumelhart, 287–310. Hillsdale, NJ: Lawrence Erlbaum.
- REGIER, TERRY and SUSANNE GAHL. 2004. Learning the unlearnable: The role of missing evidence. *Cognition* 93.147–155.
- ROSENBACH, ANETTE. 2002. *Genitive variation in English: conceptual factors in synchronic and diachronic studies*. Berlin and New York: Mouton de Gruyter.
- SAG, IVAN A. 1976. A logical theory of verb phrase deletion. *Papers from the 12th Regional Meeting, Chicago Linguistic Society*, 533–547.
- STEEDMAN, MARK. 2011. *Taking Scope: The Natural Semantics of Quantifiers*. Cambridge, MA: MIT Press.
- STIRLING, LESLIE, and RODNEY HUDDLESTON. 2002. Deixis and anaphora. Ch.17 of Rodney Huddleston & Geoffrey K Pullum et al., *The Cambridge grammar of the English language*, 1449–1564. Cambridge University Press.
- TOMASELLO, MICHAEL. 2004. Syntax or semantics? A response to Lidz et al. *Cognition* 93.139–140.
- VIKNER, CARL, and PER ANKER JENSEN. 2002. A semantic analysis of the English genitive.

Interaction of lexical and formal semantics. *Studia Linguistica* 56.191–226.

---

\* **Notes**

This paper originates in a confluence of two independent lines of research, one by Pullum and Scholz on stimulus-poverty arguments and the other by Payne and Berlage on *one*-anaphora. We thank the audiences at Brown University, Yale University, **Newcastle University**, the University of Manchester, the University of North Carolina, the Lancaster ICAME conference (Payne & Berlage 2009), and the Boston ISLE conference (Payne & Berlage 2011) for their questions and comments. Pauline Jacobson and Laura Kertz were particularly helpful to us. Barbara Scholz, who was the first to notice the mutual relevance of the two research programs, died in May 2011 before this paper was completed; the other three authors bear all responsibility for remaining errors in the text (and we thank Zoltan Galsi, whose careful reading enabled us to avoid some of them).

- <sup>1</sup> For those who endorse the ‘DP hypothesis’ the relevant constituent is, confusingly, labeled NP; but nothing substantive will hang on the fact that we do not assume that hypothesis here.
- <sup>2</sup> Notice that we distinguish the category D, to which words like *the* and *every* belong, from the function ‘determiner of’, which can be filled by either a D (as in *the house*) or a genitive NP (as in *John’s house*).
- <sup>3</sup> The two items are called *one<sub>N</sub>* and *one<sub>Q</sub>* by Jackendoff (1977:60).
- <sup>4</sup> There are derivative non-anaphoric uses, as in *the great ones of mathematics*, or *military personnel and their loved ones*.
- <sup>5</sup> Jackendoff (1977:60–61) notes the distinction between *one<sub>ct</sub>* (which he calls *one<sub>N</sub>*) and *one<sub>d</sub>* (which he calls *one<sub>Q</sub>*), but makes a double mistake in discussing the facts: he claims that *the quarts of wine and the ones of water* is ungrammatical (we give evidence below that such phrases are well formed) and that it contains *one<sub>d</sub>* (it does not; it contains *one<sub>ct</sub>* — hence the presence of the determinative *the* — so it does not support his claim about why such a phrase would be ungrammatical).
- <sup>6</sup> The CQP edition of the BNC (<http://www.natcorp.ox.ac.uk>) was developed by Sebastian Hoffmann and Stefan Evert.
- <sup>7</sup> *Rosie’s Riveters* is wrongly transcribed in the BNC as ‘[gap:name] Rivetus’. We have replaced this with a corrected transcription.
- <sup>8</sup> Panagiotidis (2003) cites an unpublished 1989 manuscript by Andrew Radford as the source for the idea that the behaviour of *one<sub>ct</sub>* can be explained by assuming that there are two different prepositions *of*. Also, again apparently following Radford, Panagiotidis uses this idea for yet another change to the Baker acquirendum by arguing (correctly) that *one<sub>ct</sub>* belongs to the lexical category N, not the phrasal category Nom. The reason that *one<sub>ct</sub>* supposedly does not occur with a following complement would then be that, as a pronominal, it could not itself inherently be relational.
- <sup>9</sup> **This group is not intended to be coextensive with the set of nouns that Löbner (1985) dubbed “functional nouns”, a type of relational noun whose relational argument represents a unique entity.**
- <sup>10</sup> Following Vikner & Jensen (2002), we use the more general concept ‘controller’ rather than simply ‘owner’. Consider the two examples in the following paragraph of the noun *car* followed by an *of*-PP (both attested examples from the BNC). In the first example, *the car of a passing motorist* [CBC 8327], we might not know whether the driver is actually the legally

---

registered owner of the car, but he/she must certainly be in control of it. And in the second example, *the car of the Spanish consul-general in Rotterdam* [HKX 2612], the consul might not own or even drive the car in question. But he/she controls its use.

- <sup>11</sup> The noun *sake* occurs only in the fixed phrases *for the sake of X* and *for X's sake*, and the noun *dint* is entirely restricted to the fixed phrase *by dint of X*, but these are fossilized idiom parts, not ordinary nouns taking syntactically obligatory complements. Huddleston & Pullum et al. (2002: 440) states that *denizen* is a unique exception, but this is now known not to be true: we have found attested uses of it with no complement.
- <sup>12</sup> This example is based on an example from the BNC (locator [A8F 286]). In order to simplify the translation into predicate logic, we have changed the definite article in the original example (*the brutal murder of a taxi driver*) to an indefinite one.
- <sup>13</sup> The view of semantic composition expressed here has strong affinities to glue semantics, e.g. Dalrymple (1999), or combinatorial categorial grammar (CCG), e.g. Steedman (2011). It could straightforwardly be reformulated in either of these frameworks. However, for the purposes of this paper we do not need to adopt either the Glue semantics assumption that non-adjacent elements can be combined, or CCG assumptions concerning a more extended set of combinatorial operations and a categorial syntax which is tied directly to semantic operations.
- <sup>14</sup> These examples were sourced using Webcorp. URLs, accessed in May 2011, are:  
(31a) <http://www.guardian.co.uk/books/booksblog/2008/jul/16/gruesomecrimesmakegreatboo>  
(31b) <http://www.guardian.co.uk/uk/2009/nov/17/arrests-murders-soldiers-northern-ireland>  
(31c) <http://www.thevervoid.com/columns/inlibtd/paultemple.htm>  
(31d) <http://www.btscene.eu/verified-search/torrent/the-dressmaker/>
- <sup>15</sup> Partee and Borschev (2012: 448) hold as a “working hypothesis” the notion that relational and non-relational nouns are of different syntactic categories and different semantic types. The exemplificatory syntactic diagnostics are however not particularly decisive. For example, both supposed contexts for non-relational nouns (*This is (a(n)) N*; *This/That N is ...* (e.g. *good*)) readily accept supposedly relational ones: *This is a portrait*; *This portrait is good*). Pauline Jacobson (p.c.) has pointed out to us a semantic diagnostic which appears superficially stronger and which she attributes originally to a UMass dissertation (Mitchell 1986, see also Partee 1989, Asudeh 2005), namely that the supposed inherently relational nouns have a hidden argument which is obligatorily bound by quantifiers. Thus in *On Christmas Eve, every boy brought plum pudding to a neighbor*, the interpretation would have to be: “for each boy, *x*, *x* brought plum pudding to *x*'s neighbor”. Even though this interpretation is the most plausible one, we doubt that the bound interpretation it is genuinely obligatory, given the right context. For example, if some particular deserving person is always given plum pudding on Christmas Eve by every boy in the town, then on this particular Christmas Eve because of an address error the pudding was delivered to a neighbor of the deserving person instead. In our account, the bound reading is easily obtained by allowing the type-shifting operator to apply to all nouns in isolation of any dependent.
- <sup>16</sup> For simplicity, the search was restricted to plural examples with *ones*. The original search was for 100 examples, of which 2 were discarded because they represented the plural of the number “1” in multiplication tables.
- <sup>17</sup> The web examples in this section were sourced using Webcorp. URLs, accessed October 2012:  
(35a) <http://famouswonders.com/baa-atoll/>  
(35b) <http://www.spookyisles.com/2012/07/the-bells-of-bealings-house/>  
(35c) <http://www.fmvmagazine.com/?p=7162>

<sup>18</sup> There may of course be idiolectal variation in some of these judgements, but the overall picture is unlikely to be affected by such variation.

<sup>19</sup> We should note that *one<sub>ct</sub>* appears to be distinctly less frequent in American English (AmE) than in BrE. The frequency of the plural *ones* in the BNC (almost all *one<sub>ct</sub>* tokens, sporadic expressions like *two ones are two* being rare) is roughly 117 per million words. The corresponding figure for the *Wall Street Journal* corpus (LDC 1993) is only 37 per million words, and for the Corpus of Contemporary American English (Davies 2008) it seems to be only 7.2 per million. Nevertheless, within these overall lower frequencies, the syntactic behaviour does not deviate markedly from BrE. Sentences with *one<sub>ct</sub>* followed by an *of*-PP represent about 0.6% of COCA, and the range of semantic relations exhibited is much the same. For example, these three sentences from COCA have head nouns frequently considered to be relational (and many more such examples could be cited):

(i) There are those who contend that a trophy property costs at least \$20 million. And that price tends to be the benchmark used when sales, like the recent [*one of a town house on West 10th Street*] for \$20 million, are reported in the media and talked over by those who like to talk about these things (*New York Times*, 30 Jan, 2011).

{*sale*; *d* is undergoer of *h*}

(ii) “Many thanks for sending me the photographs,” he wrote to Sears from Biltmore in Asheville, North Carolina, on August 7, 1895. “The new [*one of Helen*] has a wonderfully fine expression and makes me feel like returning to Boston and putting my umbrella through my portrait.” (*Antiques*, Sep 2001)

{*photograph*; *d* has depiction *h*}

(iii) I have observed individuals of the Negro race in whom the brain was as large as the average [*one of Caucasians*]; (*Natural History* 104, 1995)

{*brain*; *d* has body-part *h*}

<sup>20</sup> Rather than using the binary distinction between definiteness and indefiniteness as a formal proxy for the discourse status of the dependent, O’Connor et al. use a hierarchy of nominal types (pronoun > proper noun > kinship term > common noun definite > common noun indefinite), linking these distinctions to notions of accessibility as in Ariel (2001). In their results, they omit counts for kinship terms, which occur relatively infrequently (we have done likewise, and this is the reason why there are 73, not 75 BNC examples in the accessibility column). They also amalgamate definite and indefinite noun phrases headed by common nouns into a single factor. The basic distinction between definite and indefinite dependents is however shown to be significant in Börjars et al. (2013), another recent large-scale investigation of the genitive alternation, based on the spoken sections of the BNC.

<sup>21</sup> Under “organization”, we have included all animate collective nouns, e.g. nouns like *family*.

<sup>22</sup> The web examples in this final section were again sourced using Webcorp. URLs, accessed variously in May 2009 and May 2011:

(37a) <http://www.tefl.net/alexcase/tefl/vocab/new-words-in-the-shorter-oxford-english-dictionary/>

(37b) <http://www.betterparentingarticles.com>

(38) <http://ladyhedgehog.hedgie.com/mary.html>

(39a) <http://www.wikipedia.org> (old entry for *El Clásico*)

(39b) [http://en.wikipedia.org/wiki/Deletion\\_\(music\\_industry\)](http://en.wikipedia.org/wiki/Deletion_(music_industry))

(40) [http://board.totaluruguay.com/Education/University\\_entrance\\_requirements](http://board.totaluruguay.com/Education/University_entrance_requirements)

<sup>23</sup> Non-intersective adjectives such as *former* will belong to type <<e, <e, t>>, <e, <e, t>>, and

can thus refer to the relation **R**. We ignore this complication here.

<sup>24</sup> The derivation given here in which the translation of the matrix determiner applies last obviously corresponds to interpretations in which this determiner, if scope-bearing, scopes over any determiner in the PP. Thus it is compatible with an interpretation of (say) the NP *every picture of a student* in which a different student is depicted in each picture. Harder to obtain are inverse scope readings, as in the interpretation of the NP *a picture of every student* in which there are different pictures of each student (see the discussion of the analogous scope problem with respect to 's genitive constructions in Vikner & Jensen 2002: 200-204). Our solution, which has affinities with the treatment of inverse scope out of NP modifiers in Steedman (2009: 58-60), has the advantage of generalizing to both the 's genitive and *of*-PP constructions. What is needed is for the PP *of every student* to take the translation of *a picture* as argument, rather than just *picture*. This primarily requires the translation of *of* to be type-shifted from type  $\langle\langle e, t \rangle, t \rangle$ ,  $\langle\langle e, \langle e, t \rangle \rangle, \langle e, t \rangle \rangle$  to type  $\langle\langle e, t \rangle, t \rangle$ ,  $\langle\langle e, \langle\langle e, t \rangle, t \rangle \rangle, \langle\langle e, t \rangle, t \rangle \rangle$ . The translation of *of* will then be  $\lambda n[\lambda \hat{O}[\lambda q[n(\lambda u[\hat{O}(u)(q)])]]]$ , where the variable  $\hat{O}$  has the shifted type  $\langle\langle e, \langle\langle e, t \rangle, t \rangle \rangle$ . When applied to  $\lambda p[\forall y[\text{student}'(y) \rightarrow p(y)]]$ , the translation of *every student*, this yields, after simplification, the translation of the PP *of every student*, viz.  $\lambda \hat{O}[\lambda q[\forall y[\text{student}'(y) \rightarrow \hat{O}(y)(q)]]]$ . The relational translation of *a picture* will be  $\lambda u[\lambda p[\exists z[\text{picture}'(z) \ \& \ \mathbf{R}(z)(u) \ \& \ p(z)]]]$ , yielding an inverse-scope translation for the whole NP, viz.  $\lambda q[\forall y[\text{student}'(y) \rightarrow \exists z[\text{picture}'(z) \ \& \ \mathbf{R}_\delta(z)(y) \ \& \ q(z)]]]$  (where  $\mathbf{R}_\delta = \mathbf{R}_o$ ). In order to derive the relational translation of *a picture*, a type-shift is required for the determiner from type  $\langle\langle e, t \rangle, \langle\langle e, t \rangle, t \rangle \rangle$  to type  $\langle\langle e, \langle e, t \rangle \rangle, \langle e, \langle\langle e, t \rangle, t \rangle \rangle$ . Its translation will be  $\lambda r[\lambda u[\lambda p[\exists z[r(u)(z) \ \& \ p(z)]]]]$ , the variable  $r$  being of type  $\langle e, \langle e, t \rangle \rangle$ . This is applied to the relational translation of *picture*, i.e.  $\lambda d[\lambda h[\text{picture}'(h) \ \& \ \mathbf{R}(h)(d)]]$ .