The impact of semantic relations on grammatical alternation: an experimental study of proper name modifiers and determiner genitives[1]

TINE BREBAN
The University of Manchester

JULIA KOLKMANN
University of York

and

JOHN PAYNE
The University of Manchester
(Received 30 October 2018; revised 20 February 2019)

In this article we investigate the role of semantic relations in grammatical alternations. The specific alternation we look at is that between the proper name modifier construction, e.g. *the Obama government*, and the determiner genitive, e.g. *Obama's government*. Through the use of an experimental study in which participants were asked to rate the naturalness of the two constructions in 20 attested natural language contexts and provide paraphrases of the semantic relations in question, we tested when the two constructions alternate and whether either construction expresses semantic relations that block alternation. Our initial finding is that none of the relations we studied is categorically associated with only one of the constructions, but that certain relations – notably possession and name – are far more preferentially associated with determiner genitives and proper name modifiers respectively. Despite these 'default' associations, participants nevertheless identified a range of possible interpretations for many of the examples, meaning that our study simultaneously supports the opposing theoretical views of default relations and semantic underspecification. Further, our study validates the inclusion of semantic relations in genitive alternation studies as a major factor despite the notorious difficulties in their operationalisation. Animacy distinctions, although more straightforward to codify, appear to be of lesser importance. Methodologically, our study shows the value of an experimental approach as a corrective to researcher intuitions about the identification of semantic relations in context.

Keywords: genitive alternation, proper names, semantic relations, default interpretations, experimental study

1 Introduction

In this article we are concerned with the equivalence of two types of noun phrases (NP) that contain a proper name (PN) and a different head noun (HN): those in which the PN is part of a determiner genitive, e.g. *Obama's administration* (henceforth PNG 'proper name

genitive') and those in which the PN is added as a noun modifier, e.g. *the Obama administration* (henceforth PNM 'proper name modifier'). Rosenbach (2006, 2007, 2010, 2019) observes that, in many cases, the two constructions are interchangeable, e.g. *Obama's administration :: the Obama administration* 'the US administration under president Obama'. The cases in which this type of equivalence is found are subject to certain conditions:

 (i) the PN must have an identifying function;
 (ii) the NP containing the PN must be definite.

The first condition excludes cases in which the PNG or the PNM has a classifying function, denoting a type of HN: e.g. in *the Yorkshire terrier*, *Yorkshire* designates a type of terrier and the combination 'Yorkshire terrier' is an established subclass of 'terriers'. For comparison we can contrast this with *the Yorkshire moors*. Here, *Yorkshire* does not denote a type of moors but rather identifies one instance of moors, those located in Yorkshire. If we look at the equivalence relation, we can see that *the Yorkshire moors* alternates with *Yorkshire's moors*, but that this does not hold for *the Yorkshire terrier*, that is **the Yorkshire terrier :: Yorkshire's terrier*.

The second condition stipulates that equivalence is only possible when the PNM occurs in a definite NP: *the London underground :: London's underground*, but not **a London museum :: London's museum*. The reason for this is straightforward: the identifying determiner genitive implies definiteness and the NP containing it denotes a uniquely identifiable instance.[2]

Yet even when these two structural conditions are fulfilled, equivalence is not guaranteed, e.g. **the Rover car group :: Rover's car group*. The explanation now takes us into the realm of constructional semantics: Rosenbach (2007: 151–3) states that the range of semantic relations expressible by PNGs is smaller than that expressible by PNMs. On the other hand, Rosenbach (2019) identifies a number of corpus examples suggesting that the choice space is larger than previously assumed: these are PNM examples illustrating possession (*the Gerardi residence*) and kinship (*the Obama daughter*), two semantic relations prototypically associated with the determiner genitive. Nevertheless, semantic relations such as the 'named after' relation (e.g. *the Eiffel tower*) are still widely taken to be exclusive to the PNM construction (see Schlücker 2018 for the analogous PNM/PNG alternation in German and references to previous claims that the same holds for English). The aim of this article is to further investigate the role played by semantics in the PNG/PNM alternation and to identify the semantic relations that do and do not alternate between the PNG and PNM constructions. A first research question is then which semantic relations can be expressed by both constructions and which are exclusive to either construction.

The discussion so far has focused on whether and when alternation is *theoretically* possible. In the context of actual examples, a second question that arises is whether alternation preserves the *same* relation. Take for instance:

---

[2] Breban (2018) points out that it is not definiteness of the NP as such that conditions equivalence, but uniqueness. That is, even though *the London museum* is a felicitous NP it does not follow that *London's museum* is.

(1) (a) European leaders gather to discuss **Malta's problem**.
     (b) European leaders gather to discuss **the Malta problem**.

Both the PNG and the PNM construction are successful and, in that sense, possible from an alternation point of view, but they draw out different interpretations: PNG *Malta's problem* 'the problem Malta is experiencing' vs PNM *the Malta problem* 'the problem other European countries are experiencing that involves Malta'. We should therefore refine our first research question to which similar and different semantic relations PNG and PNM constructions can express in particular examples.

The PNG and PNM constructions are just two of a wider family of constructions involving adnominal dependents. These include determiner genitives more generally, noun modifier constructions more generally and the *of*-PP construction. Large-scale corpus and (far more rarely) experimental studies of the genitive alternation have unequivocally demonstrated that, within the choice space between the determiner genitive and the *of*-PP construction, a variety of structural, semantic and pragmatic factors contribute probabilistically to the choices made, both synchronically and diachronically (see Rosenbach 2014 for a general overview). Within this research paradigm, Szmrecsanyi *et al.* (2016) have recently contributed a historical corpus study involving a ternary comparison between determiner genitives, the *of*-PP construction and noun–noun (NN) construction. Interestingly, they find *inter alia* an animacy effect: with an animate dependent the determiner genitive construction is favoured over the NN construction (as it is over the *of*-PP construction). However, this study does not code for the distinction between proper noun and common noun modifiers, nor indeed for semantic relations as such. The probabilities involved in the PNG/PNM alternation, and specifically the probabilities associated with particular semantic relations, remain unexplored.

We aim to fill this gap by means of an experimental study asking participants to complete 20 attested corpus examples containing a blank which stands for the contextually possible PN construction(s). Crucially, we are not interested in alternation possibilities in the abstract but want to probe actual language use. For this reason, we did not use made-up examples but relied on attested corpus data. This experimental method, while not employed in alternation studies with anything like the same frequency as corpus investigations, has a well-established pedigree, beginning with Rosenbach's (2002, 2005) study of the classic genitive alternation and continuing with experimental investigations of the dative alternation (Bresnan *et al.* 2007; Bresnan & Ford 2010; Ford & Bresnan 2013). It has not so far been applied to the PNG/PNM alternation, nor indeed since Rosenbach's study to any of the wider family of genitive alternations. In our study participants were asked to rate the naturalness of both constructions and, since we were specifically interested in semantic relations, also to provide a semantic paraphrase of the relation expressed by the construction(s) they rated as successful. We will discuss this methodological approach and its pros and cons in section 2.1. Our approach is data-driven and exploratory: our goal lies in finding new patterns and associations between constructions and semantic relations. This underlies our operationalisation of semantic relations (see section 2.3) and, in our

discussion of examples (section 4.1), it will become clear that, from that perspective, uniquely interpretable examples are not necessarily the most interesting ones.

The article is structured as follows: in section 2, we present our methodology, including a brief description of the corpus data at the basis of our experimental study (2.1), the design and execution of the experimental study (2.2), the analysis of the data and our semantic annotation scheme (2.3). We present the results of the analysis in section 3, and discuss them in the light of our research questions and interests in section 4. Section 5 is a short conclusion drawing out wider theoretical implications.

2 Methodology

2.1 The examples

We started by searching two corpora of Present-day English (PDE), the *British National Corpus* (100 million words, only British English) and *Collins Wordbanks Online* (550 million words, several varieties of English)[3] for examples of PNMs and PNGs which in the abstract would alternate with the other construction – though not necessarily with the same meaning, e.g. *Edward's affair* 'the affair Edward had' vs *the Edward affair* 'the affair that revolved around Edward'. We set out to collect three sets of examples: two sets in which the context would exclude one of the meanings and therefore make only one of the constructions suitable, and one set in which both meanings were successful or in which the two constructions had the same meaning:

SET I   Only PNG possible in this context; PNM does not express the 'right' relation in context (6 examples)
SET II   Only PNM possible in this context: PNG does not express the 'right' relation in context (8 examples)
SET III Alternation is possible in this context; the relation expressed can be the same one or two different ones that are appropriate in the context (6 examples)

Representative examples of each set in the above order are as follows (the full list of examples used in the study is given in section 3.2):

(2) **Northern Ireland's experience** in Spain and Mexico in the last two World Cups taught them that heat can pose more problems than the opposition for British-based players. (BNC, A5C 385)
(3) **The Malta summit** also discussed an important American proposal to host an international conference on global warming next year. (BNC, A8W 421)
(4) Jesslyn Parkes, **the England goalkeeper**, will be hoping to guide her new team, Middlesex, to a winning start. (BNC, A33 353)

The questionnaire contained a total of 20 examples, to keep the task manageable for the participants. Sets I and III were each represented by 6 examples. We included 8 examples

---

[3] For more information about these two corpora see www.natcorp.ox.ac.uk and https://wordbanks.harpercollins.co.uk

for set II (PNM only) because the semantic relations of the PNM have not been investigated empirically at all in the literature (see section 2.3) and we wanted to establish a benchmark for its meanings.

In our selection of examples, we privileged semantic context, which means that the examples are not controlled for morphological features such as single word vs compound form of the PN and HN, e.g. PNs include *Northern Ireland* and *Malta*, one of the HNs is *sandwich shop*. We did, however, control for final sibilant of the PN as this has been shown to cause avoidance of the *s*-genitive (see, e.g., Szmrecsanyi & Hinrichs 2008; Ehret *et al.* 2014); we should note that, for this reason, the PN in the BNC example which we include as *Sainsbury family* originally was *Guinness family* (see section 3.2.1). By restricting the context to a single sentence, again in order to keep the task manageable, participants did not have access to previous discourse context and therefore did not know whether the referent of the NP was newly introduced or an activated discourse referent. It is unclear to what extent discourse activation weighs on the choice between PNG and PNM; see, e.g., Koptjevskaja-Tamm (2013) and Rosenbach (2007, 2010, 2019).

By using corpus examples, we could provide participants with real contexts, representing the complexities of actual language situations. This advantage of our decision not to use constructed examples is, unsurprisingly, simultaneously a disadvantage: the examples and contexts are more complex and the interpretation is not always straightforward. As we will discuss, in some examples a considerable number of participants interpreted the context differently, sometimes due to a lack of background knowledge. However, in certain cases these misinterpretations actually provided interesting insights into alternation and interpretation (see section 4). Another advantage is that we could use the choice of construction in the corpus example as a benchmark for hypothesis formulation and expectations.

2.2 The experimental study

The experimental study took the shape of an online questionnaire hosted by *LimeSurvey*.[4] The questionnaire was distributed among first-year students at the University of Manchester enrolled on a module called 'Principles of Linguistics'. In total, 110 native English speakers took part in the study for credit.

The examples were presented in the following way. Each example contained a blank representing the original PN construction. Participants were prompted to rate the naturalness of the two constructions as possible fillers on a scale from 1 to 10. If they rated the respective filler construction 4 or higher, they were asked to paraphrase the relation holding between the PN and the HN in the form of a relative clause, i.e. filling in one or both of the blanks in the template *the HN that* (blank) *PN* (blank). If they gave both constructions a rating of 4 or higher, they were asked to paraphrase both separately. If they awarded 3 or lower, representing a point less than one third of the total scale, we

---

[4] www.limesurvey.org

deemed it otiose to request a paraphrase. We note that this cut-off point is somewhat arbitrary and determined by practicality: we did not want participants to waste any of the limited experimental time available on puzzling over how to interpret examples they deemed relatively unacceptable in context. There was of course a concern that by setting such a cut-off point we might have obscured a potential boundary between categorical and low-probability associations. In practice, however, all the associations between semantic relations and constructions turned out in sum to be probabilistic.

We made the following hypotheses related to the alternation between PNG and PNM: (a) when alternation is deemed not possible (sets I + II), participants will rate one of the two constructions very high and the other much lower; (b) when alternation is deemed possible (set III), participants will rate both constructions similarly.

2.3 Semantic analysis: coding for semantic relations

We coded the paraphrases used by our participants as follows: (i) actor, (ii) undergoer, (iii) possessor, (iv) location, (v) name, (vi) involvement, (vii) beneficiary. This set of semantic relations in no way exhausts the potential range of the PNG and PNM constructions.[5] It was, however, sufficient to cover the vast majority of those paraphrases which were explicit enough to be unambiguously coded.

The actor relation was assigned to paraphrases in which the relation between HN and PN was expressed by a transitive verb, either in the active or passive, and where there was an inference that the PN did something. For example, in the case of *Edward's affair*, paraphrases with verbs such as *commit*, *carry out* and *act out* were coded as actor (e.g. *the affair that Edward committed*, *the affair that was carried out by Edward*).

The undergoer relation was treated, as in van Valin (2005), as a relatively broad one. In particular, it was taken to encompass experiencer-like, patient/theme-like and goal-like relations, where the general inference is that something happens to the PN. Instances of undergoer as experiencer were observed for example with the *Northern Ireland experience* and *the Kashmir problem*, where paraphrases with verbs such as *face*, *feel*, *endure* and *go through*, or with the reverse valency *affect* and *happen to*, were used. Undergoer as patient was found in *the Kobe Bryant case* with transitive paraphrases such as *investigate* or *prosecute*, and as goal with paraphrases involving the preposition *against*, e.g. *pursue against*.

The possessor relation was understood in the strict sense of OWNERSHIP, rather than in the more general sense of CONTROL as proposed by Vikner & Jensen (2002). That is, the only verbs which were treated as instantiating the possessor relation were *own* and *possess*, or with reverse valency *belong to*. For example, such paraphrases were used by participants for both constructional variants in *the Subway sandwich shop* and *Subway's sandwich shop*. Paraphrases which might rather be interpreted as control, as for example

---

[5] For attempts to codify the semantics of the PNG construction in the context of the genitive alternation see, for example, Rosenbach (2002), Koptjevskaja-Tamm (2002), Payne & Huddleston (2002), Payne & Berlage (2014) and the survey in Rosenbach (2014).

paraphrases involving *head* or *be in charge of* in the case of *the Thatcher government*, were assigned instead to actor.

The location relation was easily identifiable through paraphrases involving a predicate and a locative preposition, for example *be situated in*, *be found in* or just *be in* for *the Dartmoor scenery*. More rarely location was expressed by transitive verbs such as *surround* or *contain*. Similarly, the name relation was easily identifiable by paraphrases involving forms of the predicates *call*, *name*, *title*, *entitle* and *known as*. An obvious example illustrating this relation is *the Sainsbury family*.[6]

Involvement is not a relation which is found in treatments of the classic genitive alternation, but one which proved very relevant to this study. It was assigned where there was an inference that the HN denotes something 'about' the PNG or PNM. It is expressed by a wide range of predicates such as *involve* itself, e.g. *the case that involves Kobe Bryant*, and naturally *be about*, e.g. *the report that is about McCrone*. Other paraphrases include *concern*, *be associated with*, *deal with*, and richer verbs such as *centre around*, *focus on*.

Beneficiary proved a useful relation to include specifically, rather than as a type of undergoer. The inference in this case is that the HN does something 'for' the PNG or PNM, and paraphrases typically involve the preposition *for*, e.g. in the case of *the England goalkeeper* the predicate *play for* in *the goalkeeper that plays for England*. The beneficiary relation was also assigned to transitive predicates such as *represent*, e.g. *the goalkeeper that represents England*.

This operationalisation of semantic relations was driven primarily by the data, specifically by our judgement about which paraphrases grouped together around a common inference. For each relation, however named, we ended up with a set of predicate-argument structures corresponding to it. This is why we deemed our methodology exploratory: instead of starting with a preconceived set of semantic relations, we constructed each semantic relation based on the paraphrases provided by our participants.

The operationalisation we decided on of course bears some relation to previous classifications of semantic relations in discussions of the PNG/PNM alternation. The most detailed to date, albeit based on an analogous alternation in German, is that of Schlücker (2018). Our possessor relation denotes possession in the narrow sense, rather than the very broad sense of Schlücker's HAVE relation (which also includes subjective/ authorial genitives and objective genitives/genitives of product). Our actor relation is equivalent to the former, and our undergoer relation subsumes the latter, an important difference being that we included experiencers as undergoers rather than as actors (the claim being that experiences 'happen' to people, they don't 'do' them). Our location relation closely corresponds to Schlücker's LOCAL, and our involvement relation is

---

[6] Breban & De Smet (2019) propose that PNMs often occur in onomastic NPs, i.e. NPs that name the referent they refer to and are proper names in their own right. Intuitively, it seems that there is a connection with the name relation, e.g. *the Eiffel tower*, *the McCrone report*, though the connection is not exclusive. *London bridge* is not 'the bridge called London' but the 'bridge *in* London'.

similar to Schlücker's ABOUT. Schlücker lists a variety of specific relations which might be subsumed under 'name', viz. COMMEMORATIVE (e.g. *the Eiffel tower*), LABEL (e.g. *the Gucci bag*) and AFFILIATION (*the FBI agent*). The latter type we have treated rather as beneficiary, i.e. the agent works 'for' the FBI. Unlike Schlücker, we do not treat deverbal heads differently: for example, in our dataset we do not employ a different relation (Schlücker's REL) for cases like *the Italy representative* simply because *representative* is derived from *represent*. In our operationalisation, this is the beneficiary relation (representative 'for' Italy).

Following a preliminary phase in which the operationalisation of semantic relations was decided, each author individually coded a subset of examples, including at least two examples from each of the sets outlined in section 2.1 (expected PNG only, expected PNM only, and expected both PNG and PNM). A list was then independently compiled by a research assistant of the predicates which had been judged to instantiate each relation. The authors then separately examined these lists for (a) inconsistencies in the assignment of relations within and between examples, and (b) the appropriateness of individual assignments. Following the elimination of inconsistencies and further discussion of any disputed cases, the codings were changed accordingly. This methodology ensured agreement by all authors on the individual codings. In the context of our exploratory approach to semantic relations, this opportunity for comparison and discussion was more beneficial than a procedure in which each author coded each example and a score of agreement was calculated.

Paraphrases which did not explicitly point to a single one of the semantic relations discussed above, or were simply irrelevant, were assigned to a 'don't know' category. This included cases in which the paraphrase implicated more than one relation, as well as paraphrases which themselves included a PNG, e.g. *the goalkeeper who was England's*. The verb *have* was also generally assumed to be too vague to implicate a particular semantic relation, except in examples where it functioned as a light verb. For example, in a paraphrase such as *the affair that Edward had* the predicate is not just *have* but the whole light verb construction *have an affair*, for which an actor relation is appropriate. Similarly, in a paraphrase such as *the experience which Northern Ireland had*, *have an experience* is a light verb construction to which, in our classification, it is appropriate to assign an undergoer relation.

3 Results[7]

3.1 *Ratings for PNG and PNM*

Table 1 provides an overview of the average ratings for the different constructions for each of the 20 examples. The examples are grouped in the three sets introduced in section 2.1, and within each set they are ranked from the largest to the smallest differential. Before

---

[7] One participant avoided assigning paraphrases by frequently (not invariably) assigning a low score to both variants. We have nevertheless included this participant in all calculations.

looking at each example individually in section 3.2, we discuss the ratings in light of the hypotheses formulated at the end of section 2.2.

Table 1. *Overview average ratings for PNG and PNM*

	PNG	PNM	Differential
SET I: expected PNG			
Edward's affair	9.4	1.9	7.5
Northern Ireland's experience	9.2	2.3	6.9
Queen Elizabeth's dress designer	9.3	2.5	6.8
Italy's representative	8.9	3.4	5.5
India's game	8.2	4.5	3.7
Yemen's vote	7.5	5.4	2.1
			Mean 5.4
SET II: expected PNM			
The Sainsbury family	3.2	8.8	−5.6
The Everton game	3.8	8.8	−5.0
The McCrone report	4.0	8.7	−4.7
The Malta summit	4.1	8.2	−4.1
The Fergie scandal	5.5	7.7	−2.2
The Kobe Bryant case	5.9	7.8	−2.1
The Glasgow move	5.4	6.9	−1.5
The Kashmir problem	6.2	6.7	−0.5
			Mean −3.2
SET III: expected both PNG and PNM			
The California desert	4.0	8.8	−4.8
The Thatcher government	8.6	4.2	4.4
The Dartmoor scenery	8.5	5.6	2.9
The Subway sandwich shop	5.0	6.9	−1.9
The England goalkeeper	7.7	6.6	1.1
The Interflora website	7.1	6.8	−0.3
			Mean 0.2

These hypotheses were that (a) when alternation was deemed *not* possible by the authors (sets I + II), participants would rate one of the two constructions very high and the other much lower; and (b) when alternation *was* deemed possible (set III), participants would rate both constructions similarly.

A first result of our study is that in several of the examples in which we predicted a clear preference (sets I + II) and in which the semantic relations were also overall interpreted in keeping with context and facts (see section 3.2), the findings are less clear-cut. In particular for the set expected to prefer the PNM variant, average scores for the PNM and PNG showed markedly smaller differentials overall than the PNG-expected set.

Alternatively, in those examples where alternation was judged possible (set III), we very rarely find true alternation (defined for the purposes of this study as a low differential). Low differentials for examples where we predicted that alternation would

be possible were attested only for *Interflora website* (−0.3) and *England goalkeeper* (1.1), while the results for most of the other examples display a preference for one of the constructions, e.g. *California desert* (−4.8) and *Thatcher government* (4.4).

In sum, our hypotheses are less than clearly borne out by the study results. This could be an indication of the fact that there is considerable overlap between the semantic relations expressible by the two constructions.[8] We look at the semantic relations assigned to each example individually to investigate this.

3.2 Semantic relations associated PNM and PNG in individual examples

We look at the examples in the same order as they appeared in table 1 in the three sets. For each example, we list the PN and HN, e.g. 'Edward + affair', and then provide the 'blanked' full context example that participants were presented with in the experiment.

3.2.1 SET I: PNG expected to be preferred

Example I.1 'Edward + affair'

(5) _____ and subsequent marriage to divorced Mrs Simpson had left the family's popularity at an all-time low. (BNC, CBC 9009)

In table 2 (and subsequent tables), the first two columns represent the paraphrases in cases where participants assigned only one construction a score of 4 or higher. Columns 3 and 4 give the paraphrases for those cases in which both constructions were rated 4 or higher.

Table 2. *Semantic relations for* Edward's affair *and* the Edward affair

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
76	actor	1	actor	8	actor	1	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
1	possessor	0	possessor	1	possessor	1	possessor
0	location	0	location	0	location	0	location
0	name	0	name	0	name	2	name
10	involvement	0	involvement	1	involvement	6	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
10	don't know	0	don't know	1	don't know	1	don't know

---

[8] Other potential explanations for the reticence to choose the PNM outright may be that the PNM is a newer construction (e.g. Rosenbach 2007, 2010) and/or that it is associated with particular registers, e.g. news texts and more informal texts (Rosenbach 2006, 2007; Koptjevskaja-Tamm 2013; Breban 2018), as was in fact also occasionally commented upon by some participants. Another suggestion is that the PNM is more likely to be chosen if the referent of the NP and its relation to the PN has been set up in the previous discourse (e.g. Koptjevskaja-Tamm 2013; Breban 2018), in contrast with the PNG which is used more easily to introduce a new referent that is accessible simply because it is anchored by the 'known' PN referent. The limited preceding

In this example, the semantic relation which we anticipated was actor. As table 2 shows, the most common semantic relation attributed to the PNG was indeed the actor relation, reflected most often in the paraphrase *the affair which Edward had*, but also by paraphrases involving explicit actor verbs such as *commit*, *carry out* and *act out*. The actor relation was associated mainly, though not exclusively, with the PNG. A minority of participants instead chose an involvement relation, a typical paraphrase being *the affair that Edward was involved in*. This relation was associated with both the PNG and PNM forms; the latter in particular when participants scored both forms 4 or higher and used distinctive paraphrases for the two forms.

Example I.2 'Northern Ireland + experience'

(6) _____ in Spain and Mexico in the last two World Cups taught them that heat can pose more problems than the opposition for British-based players. (BNC, A5C 385)

Table 3. *Semantic relations for* Northern Ireland's experience *and* the Northern Ireland experience

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
0	actor	0	actor	0	actor	0	actor
78	undergoer	3	undergoer	8	undergoer	6	undergoer
3	possessor	0	possessor	2	possessor	0	possessor
0	location	0	location	0	location	1	location
0	name	0	name	0	name	0	name
0	involvement	0	involvement	0	involvement	1	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
11	don't know	2	don't know	3	don't know	5	don't know

Our anticipation was that the context of this example triggered an undergoer interpretation and table 3 shows that this was the most commonly attested relation in the participant paraphrases, e.g. *the experience Northern Ireland has*, *that Northern Ireland faces*. This relation is strongly associated with the PNG form, though not exclusively. There were a very small number of alternative relations, including some possessor paraphrases (*belongs to Northern Ireland*), all associated with the PNG, and one instance each of location and involvement, both for a PNM paraphrase when participants rated both constructions 4 or higher and may have wanted to distinguish the two semantically.

discourse in our examples may have made very high ratings for the PNM less likely. Our results therefore suggest that further studies looking at genre/register and discourse factors are needed to complete the picture.

Example I.3 'Queen Elizabeth + dress designer'

(7) There was also a regular Mail fashion column by _____ dress designer, Norman Hartnell. (BNC, CDU 705)

Table 4. *Semantic relations for* Queen Elizabeth's dress designer *and* the Queen Elizabeth dress designer

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
9	actor	0	actor	0	actor	0	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
3	possessor	0	possessor	2	possessor	0	possessor
0	location	0	location	0	location	0	location
0	name	0	name	0	name	2	name
1	involvement	0	involvement	0	involvement	2	involvement
30	beneficiary	2	beneficiary	4	beneficiary	5	beneficiary
47	don't know	1	don't know	10	don't know	7	don't know

We predicted that participants would use a beneficiary relation (e.g. *the designer who worked for/designed dresses for QE*). Many participants did indeed choose a beneficiary relation for the PNG both when it was deemed to be the only acceptable variant and one of two acceptable variants; however, the beneficiary relation was also selected for the PNM by a few participants (see table 4). Other relations included actor (e.g. *the designer that QE employed/used*), possessor (e.g. *the dress designer that belongs to QE*), which for this example were only selected for the PNG, and involvement (e.g. *the dress designer that is associated with QE*), which was selected for both PNG and PNM forms. It should be noted that there were a large number of paraphrases of the type *designs/designed QE's dress*, in which the relationship between the HN *dress designer* and the proper noun *QE* conflates two relations: an undergoer relation 'X designs Y' and a (undetermined) relation between *QE* and *dress* expressed by a determiner genitive *QE's*. These examples have been annotated as 'don't know'.[9]

Example I.4 'Italy + representative'

(8) Among other jobs, he has been _____ at the International Monetary Fund. (BNC, CR9 2660)

Many of the paraphrases used the verb *represent*, e.g. *that represents Italy*, or equivalents, *that stands for Italy*. We coded these as indicating a beneficiary relation. This relation was associated with both the PNG and PNM constructions, with a preference for the PNG (see

---

[9] In retrospect we might have predicted this potential confound and avoided this example. Note, however, that it is not simply a matter of avoiding a compound HN: no similar confound arose with *Subway sandwich shop*.

Table 5. *Semantic relations for* Italy's representative *and* the Italy representative

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
12	actor	0	actor	2	actor	1	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
4	possessor	0	possessor	6	possessor	0	possessor
2	location	0	location	1	location	4	location
0	name	0	name	0	name	0	name
0	involvement	0	involvement	0	involvement	1	involvement
31	beneficiary	4	beneficiary	7	beneficiary	12	beneficiary
27	don't know	3	don't know	11	don't know	9	don't know

table 5). Of the remaining paraphrases, actor (e.g. *that is employed by Italy*) and possessor were both near-exclusively cited for the PNG. Location was used in a few instances both for the PNG and PNM variants. There were a relatively large number of paraphrases that did not fit any of the relations, and some participants seemed to struggle with the interpretation of this example.

Example I.6 'India + game'

(9) Two major injustices eventuated in _____ against Australia in Brisbane. (BNC, BN9 546)

Table 6. *Semantic relations for* India's game *and* the India game

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
44	actor	7	actor	23	actor	14	actor
0	undergoer	1	undergoer	0	undergoer	6	undergoer
0	possessor	0	possessor	7	possessor	0	possessor
2	location	0	location	0	location	2	location
0	name	0	name	0	name	0	name
4	involvement	3	involvement	2	involvement	9	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
6	don't know	1	don't know	9	don't know	10	don't know

The intended semantic relation was actor (*the game that India played*). The semantic analysis indicated that the actor relation was the relation most frequently chosen for the PNG without, however, being limited to this variant: notably, 21 participants also assigned this relation to the PNM (see table 6). When both forms were scored 4 or higher, the (unlikely) possession reading (*the game that belongs to India*) clustered exclusively with the PNG, while the location and the undergoer (*the game that was*

against India) readings clustered exclusively with the PNM. The involvement reading was deemed plausible for both variants.

Example I.6 'Yemen + vote'

(10) The move was regarded as a sign of US displeasure at _____ against UN Security Council Resolution 678. (BNC, HL4 2389)

Table 7. *Semantic relations for* Yemen's vote *and* the Yemen vote

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
34	actor	6	actor	33	actor	22	actor
0	undergoer	0	undergoer	1	undergoer	1	undergoer
4	possessor	0	possessor	4	possessor	0	possessor
0	location	4	location	0	location	6	location
0	name	2	name	0	name	0	name
0	involvement	1	involvement	1	involvement	8	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
10	don't know	4	don't know	6	don't know	8	don't know

The context strongly suggested an actor relation based on the Yemen delegation at the United Nations casting its vote (see table 7). A large number of participants, however, appear to have misinterpreted the context and treated the vote instead as one taking place at an election *in* Yemen. For this location interpretation, the PNM variant was selected. The choice of verbs used to indicate the actor relation was relatively extensive, and in a proportion of cases it was clear (or at least possible) that the interpretation assigned was one in which the people of Yemen were carrying out the vote rather than the UN delegation. The relation was cited for both the PNG and the PNM variants, though with a stronger association with the former. The only other relation assigned to the PNG in any quantity was that of possessor, which because of its inappropriateness in the context we interpret as participants resorting to a default paraphrase in a semantically less straightforward example. A number of participants (perhaps for similar reasons) indicated an involvement relation, which was mainly associated with the PNM variant.

3.2.1 SET II: PNM expected to be preferred

Example II.1 'Sainsbury + family'

(11) _____, the last owners of the estate, had also built a fanciful fishing lodge beside the castle. (BNC, ADM 643)[10]

---

[10] As noted earlier, the BNC example had the PN *Guinness* instead of *Sainsbury*; this was changed by us to avoid a final sibilant on the PN.

Table 8. *Semantic relations for* Sainsbury's family *and* the Sainsbury family

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
0	actor	0	actor	1	actor	0	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
0	possessor	1	possessor	8	possessor	1	possessor
0	location	0	location	0	location	0	location
2	name	46	name	0	name	19	name
0	involvement	0	involvement	5	involvement	1	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
4	don't know	28	don't know	12	don't know	4	don't know

By far the most frequent semantic relation was the expected name relation (see table 8). Typical paraphrases were: *the family named Sainsbury*, *the family that were called Sainsbury*, or the *family that has/holds/carries the name Sainsbury*. This relation was strongly associated with the PNM variant. No other specific relations were chosen with any frequency for the PNM, though there were a significant number of 'don't knows'. When participants rated both PNG and PNM 4 of higher, we find possessor and (less appropriately) involvement as relation associated with the PNG.

Example II.2 'Everton + game'

(12) Norwich lost key components of their team in the run-up to _____ – just as they have for tomorrow's tie. (BNC, AHC 286)

Table 9. *Semantic relations for* Everton's game *and* the Everton game

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
3	actor	23	actor	15	actor	12	actor
2	undergoer	32	undergoer	1	undergoer	12	undergoer
0	possessor	1	possessor	4	possessor	0	possessor
0	location	0	location	2	location	3	location
0	name	1	name	0	name	0	name
0	involvement	3	involvement	0	involvement	0	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
2	don't know	10	don't know	11	don't know	7	don't know

The context in this example frames this game as one played by Norwich against Everton. For this reason, we expected the undergoer relation to be the most plausible answer. This was indeed one of the main relations, as reflected by paraphrases in which Everton was either a direct object (*the game that Norwich played Everton in*), or a prepositional

object (*the game that Norwich played against/with Everton*). However, the fact that this example involves a symmetric predicate ('X and Y played' entails 'X played Y' and 'Y played X') evidently led an almost equally large number of participants to suggest a paraphrase indicating the actor relation. Frequently, this was expressed as a coordination (*the game that Norwich and Everton played*), but also simply as *the game that Everton played*. The undergoer relation was strongly associated with the PNM variant. The actor relation was given for both variants and was used to distinguish PNG from PNM when both variants were scored 4 or higher (see table 9).[11] Other minor options are possessor, mainly for the PNG, and location, which featured for both variants.

Example II.3 'McCrone + report'

(13) _____ is a UK Government document which was written and researched in 1974 on behalf of the British Government of the day (Conservative, led by Edward Heath). It was composed by Professor Gavin McCrone, an employee of the Scottish Office, at St. Andrew's House in Edinburgh. (WBO, brregnews 2001)

Table 10. *Semantic relations for* McCrone's report *and* the McCrone report

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
2	actor	32	actor	31	actor	22	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
1	possessor	2	possessor	9	possessor	1	possessor
0	location	0	location	0	location	0	location
0	name	15	name	1	name	11	name
0	involvement	3	involvement	0	involvement	4	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
0	don't know	12	don't know	1	don't know	4	don't know

This example was one in which *McCrone* could be interpreted as the name associated with the report (name relation), or more directly as the author of the report (actor). In the actual corpus example, the PNM *the McCrone report* was used, and we expected this to be the preferred variant given the well-established convention of naming reports after their lead

---

[11] A couple of participants insightfully indicated that the choice of variant depends on perspective ('it depends on which team you support'). In the context, the perspective seems to be that of the Norwich team: Everton are the opponents. Given this perspective, the PNM variant was indeed strongly preferred and the undergoer relation appropriately assigned. However, the attribution to the PNM variant of an actor relation by a significant number of participants suggests that an example like *the Everton game* can indeed be construed simply as one in which Everton is actor and the perspective is neutral. That is, we judge it natural for an Everton supporter to say 'I'm going to the Everton game this afternoon', implying a game in which Everton will play, rather than necessarily a game in which Everton will be played against.

investigators. Table 10 shows that it was in fact the actor relation that was the more frequent one, with paraphrases involving verbs such as *write*, *compose*, *lead* and *publish*. It was given for the PNM variant, and for the PNG – the latter in particular when both constructions were deemed possible. The name relation, with paraphrases involving nouns such as *name* and *title*, or participles such as *entitled* and *known as*, was the second most frequent and was nearly exclusively attributed to the PNM. Some participants assigned the possessor relation to the PNG, which we interpret as a default assignment, and involvement to the PNM.

Example II.4 'Malta + summit'

(14) _____ also discussed an important American proposal to host an international conference on global warming next year. (BNC, A8W 421)

Table 11. *Semantic relations for* Malta's summit *and* the Malta summit

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
0	actor	6	actor	14	actor	4	actor
0	undergoer	1	undergoer	0	undergoer	0	undergoer
2	possessor	2	possessor	10	possessor	1	possessor
2	location	31	location	4	location	18	location
0	name	4	name	0	name	1	name
2	involvement	6	involvement	2	involvement	8	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
4	don't know	14	don't know	6	don't know	4	don't know

The varied semantic relations given show an interesting distribution over the two constructions (see table 11). The expected location interpretation was associated strongly with the PNM. The name relation was only used with the PNM. Involvement, e.g. *that discusses Malta*, which is less appropriate in the context, too was more common for the PNM construction. The relations that featured most frequently for the PNG were possessor (e.g. *belongs to Malta*) and actor (e.g. *that Malta hosted*), in particular, by participants seeking to semantically differentiate between the two constructions. Neither of those relations was limited to the PNG, with the actor relation being expressed for the PNM as only variant rated 4 or higher too.

Example II.5 'Fergie + scandal'

(15) While others at Balmoral were losing sleep over _____, Diana was up at dawn preparing for a surprise visit to terminally-ill cancer patients. (BNC, CH6 8792)

We expected participants to express that this was a scandal involving Fergie and involvement was indeed the most frequent relation provided (see table 12). It was used for both constructions, though more frequently for the PNM. The PNG was often

Table 12. *Semantic relations for* Fergie's scandal *and* the Fergie scandal

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
2	actor	4	actor	23	actor	6	actor
1	undergoer	0	undergoer	0	undergoer	2	undergoer
0	possessor	0	possessor	5	possessor	0	possessor
0	location	0	location	0	location	0	location
0	name	0	name	0	name	5	name
9	involvement	34	involvement	18	involvement	31	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
4	don't know	6	don't know	3	don't know	5	don't know

distinguished by paraphrases instantiating the actor relation, including verbs such as *cause* and *commit*. This relation was also associated with the PNM, though with much less frequency than the involvement relation. Other relations attested are possessor, only for the PNG when distinguishing the variants, and name, only for the PNM in distinguishing. This supports our proposal that these are default relations associated with the respective constructions (section 4.1.3).

Example II.6 'Kobe Bryant + case'

(16) _____ came to a screeching halt when the woman who accused the NBA star dropped the criminal case because of the stress of publicity and the impending trial. She is still pursuing a civil case against Bryant, who said he was sorry for what happened. (WBO, usnews 2004)

Table 13. *Semantic relations for* Kobe Bryant's case *and* the Kobe Bryant case

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
1	actor	0	actor	7	actor	0	actor
2	undergoer	8	undergoer	6	undergoer	17	undergoer
3	possessor	0	possessor	6	possessor	0	possessor
0	location	0	location	0	location	0	location
0	name	0	name	1	name	4	name
5	involvement	20	involvement	18	involvement	19	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
2	don't know	7	don't know	22	don't know	17	don't know

Although the context indicates the case was 'against' Kobe Bryant (triggering an undergoer relation), the results display a variety of relations with involvement (e.g. *the case that involved KB*) being the most frequent choice, followed by undergoer (see

table 13). Both involvement and undergoer are attested for both variants but show a stronger association with the PNM. Undergoer can be contrasted in this respect with a number of actor paraphrases (e.g. *the case that KB fought*) which exclusively feature with the PNG. In this example, actor, undergoer and involvement paraphrases only present slightly different semantic views on the situation. Once more, there are some possessor instances for the PNG and some name instances mainly for the PNM.

Example II.7 'Glasgow + move'

(17) If _____ does go ahead, it will mean the loss of 180 jobs in Chester. (BNC, K3K 1012)

Table 14. *Semantic relations for* Glasgow's move *and* the Glasgow move

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
19	actor	7	actor	20	actor	15	actor
1	undergoer	1	undergoer	2	undergoer	2	undergoer
0	possessor	0	possessor	4	possessor	0	possessor
1	location	21	location	3	location	4	location
0	name	0	name	0	name	3	name
0	involvement	5	involvement	4	involvement	9	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
3	don't know	12	don't know	6	don't know	6	don't know

We anticipated location (*the move to Glasgow*) as main relation. However, as table 14 shows, location was the second most attested relation only after the puzzling actor relation (e.g. *the move that Glasgow made*). While location is strongly associated with the PNM, actor is somewhat more frequent for the PNG. This may explain why the PNG scored an average of 5.4 (see table 1) and why the differential with the PNM, which scored 6.9, was the second lowest of all 8 examples in this category. Other notable semantic relations included involvement, e.g. *the move Glasgow are participating in*, which was more frequently attributed to the PNM, and possessor and name used when both variants were deemed appropriate for respectively PNG only and PNM only.

Example II.8 'Kashmir + problem'

(18) President Bhutto of Pakistan was seeking to adopt a more conciliatory approach to the resolution of _____. (BNC, HKX 1187, adapted, original has just 'Bhutto')

In this example too, participants' unfamiliarity with the topic influenced the results (see table 15). The expected interpretation was involvement, 'the problem that involves Kashmir', and we anticipated participants would avoid undergoer, 'the problem that Kashmir has'. However, results show the two relations are frequently attested. As

Table 15. *Semantic relations for* Kashmir's problem *and* the Kashmir problem

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
1	actor	2	actor	2	actor	4	actor
15	undergoer	7	undergoer	20	undergoer	8	undergoer
1	possessor	0	possessor	7	possessor	0	possessor
5	location	4	location	5	location	6	location
0	name	0	name	0	name	5	name
2	involvement	17	involvement	1	involvement	14	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
4	don't know	5	don't know	10	don't know	8	don't know

expected, the involvement relation was largely associated with the PNM variant, whereas the undergoer relation was associated with the PNG in twice as many cases as with the PNM variant. Perhaps unsurprisingly, location is attested with some frequency too. It distributes evenly over the two constructions. The actor relation, as in *the problem that Kashmir causes*, is given for the PNG and for the PNM. Other less semantically fitting paraphrases are possessor and name, exclusive to the PNG and the PNM respectively.

3.2.3 *SET III: PNG and PNM expected to be possible*

Example III.1 'California + desert'

(19) Last year, I travelled on my bike through _____. (WBO, usnews 2004)

Table 16. *Semantic relations for* California's desert *and* the California desert

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM ≥ 4			
relation		relation		relation PNG		relation PNM	
0	actor	0	actor	0	actor	0	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
0	possessor	1	possessor	12	possessor	0	possessor
5	location	51	location	19	location	28	location
0	name	0	name	0	name	3	name
0	involvement	0	involvement	0	involvement	0	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
1	don't know	11	don't know	10	don't know	10	don't know

The preferred semantic relation is location with paraphrases such as *that is / is situated / is found in California* (see table 16). There is a strong but not exclusive association of the location relation and the PNM construction. Only two of the other relations are found:

possessor, which is given by participants attempting to disambiguate the PNG variant (and by 1 'stray' participant for the PNM), and name (*is called California*), which was attested for the PNM when both constructions received a score of 4 or higher.

Example III.2 'Thatcher + government'

(20) Critics tried to rescue older notions of 'community' or 'citizenship' to set against the competitive values of _____. (BNC, A66 570)

Table 17. *Semantic relations for* Thatcher's government *and* the Thatcher government

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
39	actor	3	actor	20	actor	17	actor
0	undergoer	0	undergoer	1	undergoer	1	undergoer
8	possessor	0	possessor	11	possessor	2	possessor
0	location	0	location	0	location	0	location
0	name	0	name	0	name	2	name
1	involvement	1	involvement	1	involvement	3	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
12	don't know	2	don't know	7	don't know	14	don't know

The actor relation, marked by paraphrases such as *the government that Thatcher led*, *the government that Thatcher headed*, or *the government that Thatcher controlled*, accounted for the vast majority of assignments to both the PNG and the PNM variants (see table 17). We also find possessor, marked by paraphrases such as *the government that belonged to Thatcher* or *the government that Thatcher owned*, mainly for PNG, though not exclusively. With only 2 participants selecting a name relation, judgement seems to be that name is an inappropriate relation here: a government is not named as such after its leader.

Example III.3 'Dartmoor + scenery'

(21) Dartmoor has a long history of use by humans. There are burial chambers and standing stones from the Bronze Age, Iron Age forts, and tramways from mining and quarrying in the last century. Clapper or rough stone bridges are also a familiar part of _____. (BNC, CMD 1379)

The relation most frequently used by participants to paraphrase this example was the predicted location relation, which was associated with both PNG and PNM. In fact, the same number of participants (i.e. 40) selected it for each (see table 18). However, and particularly when both variants were deemed possible, the possessor relation was chosen for the PNG to distinguish its semantics from that of the PNM, which was less likely to be associated with a different relation than location. A minority of participants opted for

Table 18. *Semantic relations for* Dartmoor's scenery *and* the Dartmoor scenery

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
0	actor	0	actor	0	actor	0	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
7	possessor	0	possessor	20	possessor	3	possessor
20	location	4	location	20	location	36	location
0	name	0	name	0	name	2	name
0	involvement	1	involvement	0	involvement	1	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
11	don't know	3	don't know	22	don't know	20	don't know

name (e.g. *the scenery that is named after Dartmoor*) and involvement relations (e.g. *the scenery that Dartmoor was linked to*) for the PNM. We also counted a sizeable number of unidentifiable responses, e.g. *the scenery that Dartmoor has/offers/displays*.

Example III.4 'Subway + sandwich shop'

(22) Lee locked up the shop herself, swung by _____ and picked up four combination sandwiches containing everything but the kitchen sink. (WBO, usbooks 1994)

Table 19. *Semantic relations for* Subway's sandwich shop *and* the Subway sandwich shop

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
1	actor	2	actor	1	actor	1	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
9	possessor	10	possessor	24	possessor	2	possessor
0	location	0	location	0	location	0	location
4	name	22	name	3	name	17	name
0	involvement	0	involvement	0	involvement	2	involvement
0	beneficiary	0	beneficiary	0	beneficiary	0	beneficiary
9	don't know	14	don't know	6	don't know	12	don't know

For this example, we anticipated competition in the assignment of semantic relations being between the name relation, appropriate for the PNM, and (less likely in the context) the possessor relation for the PNG. This generally turned out to be the case: the name relation was more strongly associated with the PNM and the possessor relation with the PNG, though in neither case was the association exclusive (see table 19).

Example III.5 'England + goalkeeper'

(23) Jesslyn Parkes, _____, will be hoping to guide her new team, Middlesex, to a winning start. (BNC, A33 353)

Table 20. *Semantic relations for* England's goalkeeper *and* the England goalkeeper

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation			relation PNG		relation PNM
1	actor	0	actor	2	actor	1	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
3	possessor	0	possessor	12	possessor	1	possessor
0	location	2	location	0	location	1	location
0	name	0	name	0	name	0	name
0	involvement	0	involvement	0	involvement	1	involvement
17	beneficiary	8	beneficiary	37	beneficiary	46	beneficiary
5	don't know	6	don't know	15	don't know	16	don't know

The majority of participants provided beneficiary paraphrases for both the PNG and the PNM variants (see table 20). Our semantic analysis of the other paraphrases further resulted in a number of 'don't know' responses, e.g. *the goalkeeper England has/that is for England*. Other than this, 15 participants provided possessor responses for the PNG, compared with only one for the PNM. We also counted a negligible number of actor relations such as *the goalkeeper England have employed/use* which occurred for both PNG and PNM.[12]

Example III.6 'Interflora + website'

(24) According to _____: "With the exception of funeral orders we are unable to carry out AM or PM deliveries." (WBO, brmags 2004)

We had expected a range of relations including involvement, e.g. 'the website associated with Interflora' and possessor 'that belongs to Interflora'. The lack of a single straightforward interpretation was visible in the results, which show a relatively large number of often vague paraphrases which we could not assign to one of our relations and classed as 'don't know' (see table 21): examples are *the website that Interflora has/uses* or *the website that displays information about Interflora*. However, the paraphrases that could be classified display some clear patterns. Possessor, e.g. *the website that Interflora owns/that belongs to Interflora*, is the most frequently selected (48 times for the PNG and 18 times for the PNM). The second most frequent relation

---

[12] A few participants commented that the use of the PNG suggests that Jesslyn Parkes 'is the only option' while use of the PNM implies that she is 'the regular first choice on the England international team' but that there may be other goalkeepers. This ties in with the observation that the combination PNM and definite article does not imply uniqueness, contrary to the PNG (see footnote 2).

Table 21. *Semantic relations for* Interflora's website *and* the Interflora website

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
2	actor	3	actor	16	actor	13	actor
0	undergoer	0	undergoer	0	undergoer	0	undergoer
12	possessor	5	possessor	36	possessor	13	possessor
0	location	0	location	0	location	0	location
0	name	0	name	0	name	6	name
1	involvement	0	involvement	0	involvement	7	involvement
1	beneficiary	3	beneficiary	1	beneficiary	2	beneficiary
6	don't know	11	don't know	11	don't know	23	don't know

is actor (e.g. *the website Interflora created*), which distributes more or less equally over both variants. Involvement and name are used in disambiguating the PNM when both variants are deemed acceptable.

4 Discussion

4.1 The role of semantic relations in the PN alternation

In this section, we look at the semantic relations participants associated with each of the constructions across all 20 examples. We focus on general conclusions that can be drawn from the aggregate data before moving on to the theoretical implications of our findings.

4.1.1 Aggregate data for semantic relations associated with PNG and PNM

Table 22 summarises the total number of paraphrases by all participants assigned to each semantic relation, as well as the total number of 'don't knows'.

Generally, we find that all seven semantic relations are in principle expressible by both constructions, with none constituting a categorical context for either construction in the sense of Rosenbach (2002: 28). Our data thus concur with Rosenbach's (2019) observations that even the possessor relation is also expressible by the PNM, e.g. for *the Subway sandwich shop* and *the Interflora website*.[13] Conversely, we find that the name relation is also expressible by the PNG, e.g. *Subway's sandwich shop* and *Sainsbury's family.*

---

[13] Kinship was not included in the set of relations we studied. It seems to us that it is possible that this relation is more restricted to the PNG construction than possession. In possible counterexamples such as (i), it seems to us that the relation is not 'son of Beckham' or 'child of Ramsay' but 'son in the Beckham family' and 'child in the Ramsay family'. In both cases, it is not just a single person that is denoted/famous, but the entire family, e.g. Gordon Ramsay's wife published her own cookery book and daughter Matilda (Tilly) and the rest of the family feature in a TV programme *Matilda and the Ramsay Bunch*. Alternatives such as *the David son* or *the David Beckham son* in which the PN unambiguously refers to a single individual strike us as less felicitous.

(i) **The Beckham son** pulled his grey hood up and layered up with a huge coat while **the youngest Ramsay child** stayed warm with her beanie and thick coat. (www.dailymail.co.uk/tvshowbiz/article-5298495/Cruz-Beckham-plays-Tilly-Ramsay-snow.html)

Table 22. *Semantic relations for all examples added up*

only PNG scored ≥ 4		only PNM scored ≥ 4		both PNG and PNM scored ≥ 4			
relation		relation		relation PNG		relation PNM	
246	actor	96	actor	218	actor	133	actor
99	undergoer	53	undergoer	39	undergoer	55	undergoer
61	possessor	22	possessor	190	possessor	25	possessor
37	location	117	location	54	location	109	location
6	name	90	name	5	name	82	name
35	involvement	94	involvement	53	involvement	127	involvement
79	beneficiary	17	beneficiary	49	beneficiary	65	beneficiary
178	don't know	152	don't know	182	don't know	190	don't know

We do note, however, that the proportions assigned for each relation to each construction in some cases differ where only one relation was chosen as predominant (first two columns) and in the case where an attempt was made to assign a paraphrase to both constructions, potentially discriminating between them (second two columns). There is in particular a very high figure (190) associated with the possessor relation and the PNG variant in the case where the need for discrimination might have been felt: this may be put down not just to the particular examples involved, but to the possessor relation being construed by participants as the one most strongly associated with the PNG construction. Similarly, the name relation shows a preference (82) for the PNM in such contexts. The same goes for the location relation, which shows a clear bias (109) towards the PNM. These findings are observed at the level of individual examples as well, as was evidenced in section 3.2.

4.1.2 Hierarchy of semantic relations and the role of animacy

Summing the totals in table 22 for each relation abstracts away from the individual examples as far as possible, given our data, and gives us the proportion of assignments to each of the two constructions. This we express in figure 1 as odds ratios: the likelihood of each relation to be expressed by one construction rather than the other.

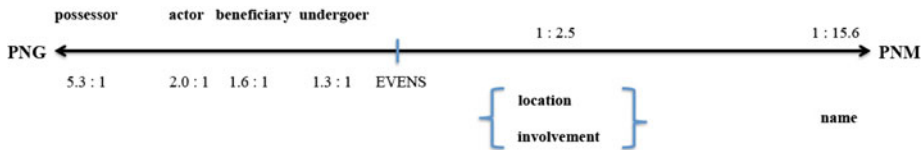Figure 1. Odds of a given semantic relation being expressed by the PNG and PNM constructions

As can be seen from figure 1 (not to scale, and with figures rounded to one decimal place), the odds ratios enable us to situate the different semantic relations in a hierarchy. The two endpoints of the hierarchy are defined on the one hand by the

name relation, exceptionally strongly associated with the PNM construction (odds of 15.6:1), and on the other hand by the possessor relation, most strongly associated with the PNG construction (odds of 5.3:1). A particularly strong association of the possessor relation with the determiner genitive was also observed by Szmrecsanyi (2013) in a corpus-based regression analysis of the genitive alternation. Location and involvement are still strongly associated with the PNM construction (odds of 2.5:1), while actor and beneficiary gravitate towards the PNG (odds of 2.0:1 and 1.6:1 respectively). Undergoer is closest to evens, the actual odds ratio being 1.3:1 in favour of the PNG. This ratio would move in favour of the PNM construction if we were to separate experiencer from other types of undergoer. This appears to be supported by individual examples. *Northern Ireland's experience* shows a strong association between experiencer and PNG in its congruent interpretation, and also for *Kashmir's problem/the Kashmir problem*, the – in this case less suitable – experiencer interpretation is associated more strongly with the PNG variant. By contrast, examples that show an association of undergoer and PNM are *the Kobe Bryant case* and *the Everton game*, where the meaning is the case/game 'against' KB/Everton.

It is possible to compare the odds ratios obtained above with those calculated simply for animacy of the dependent. Many authors have argued that animacy is a highly significant factor favouring the PNG over alternatives, notably the *of*-PP construction but also the PNM construction (see Rosenbach 2006, 2007 for discussion and references). In an expanded form, the proposed animacy hierarchy (as in Rosenbach 2007) contains six subcategories: human > animal > (human) collective > temporal > locative > (inanimate) common, with human most strongly favouring the PNG. The examples in our experiment represent three of these: (i) human (Edward, Queen Elizabeth, Thatcher, McCrone, Fergie, Kobe Bryant); (ii) collective (Northern Ireland, Italy, India, Yemen, Subway, England, Interflora, Sainsbury, Everton); and (iii) locative (California, Dartmoor, Malta, Glasgow, Kashmir). In the case of (iii), we judge that primary reference is to a geographical entity rather than a people, organisation, team or family. The outcome is indeed consistent with the animacy hierarchy as defined above. The odds of a human dependent favouring the PNG construction over the PNM construction are 1.4:1, compared with 1.2:1 for a collective dependent. A locative dependent on the other hand favours the PNM construction (1:1.4).

However, the difference between the odds ratios for humans and collectives is notably small, suggesting that in this case semantic relations rather than fine-grained animacy distinctions play a prominent role. Note that there can be large discrepancies within the human subcategories as to which construction is favoured, for example (i) *Edward's affair* vs *the Fergie scandal*; (ii) *Northern Ireland's experience* vs *the Everton game*. Overall, we obtain a considerably greater range of odds ratios by examining semantic relations as a dominant factor. This result provides an endorsement of the necessity to include semantic relations in the study of alternations such as the one in this article.

4.1.3 Underspecification and default relations

One main conclusion of our study is that the association between semantic relations and particular constructions cannot be considered categorical, just as in the genitive alternation the association between semantic relation and construction is one of a statistical nature. The probabilistic nature and the variety of semantic relations attested for both of the constructions in our 20 examples suggest that the sets of relations expressible by the PNG and the PNM constructions are wide-ranging. Their semantics is underspecified and actual interpretations are derived in actual usage situations based on these potential relations (e.g. Barker 1995; Vikner & Jensen 2002; Peters & Westerståhl 2013). Their interpretation is thus properly considered a matter of pragmatic inference: background and discourse knowledge are essential for the correct interpretation. This provides support for similar conclusions in Kolkmann's (2016) study of the PNG and is consistent with Breban's (2018) analysis of the PNM construction.

At the same time, the aggregate numbers and the fact that participants use the possessor and name relations to disambiguate the PNG and PNM respectively when both constructions were rated sufficiently high are powerful evidence for recognising a strong association of PNG and possession and PNM and name. It points to the existence of default associations between particular constructions and particular relations, e.g. Nikiforidou (1991), Taylor (1996), Stefanowitsch (2003) for the classic genitive alternation. However, contrary to these authors, we have shown these defaults to be a matter of preference rather than categorical association.

Our study thus provides support for the simultaneous validity of two views on the semantics of constructions. This *prima facie* contrast between default and underspecification echoes Bauer's (2017) discussion of the process by which a listener discerns possible semantic relations obtaining between elements making up compounds. Bauer argues that the 'listener has to use whatever strategies are available, and knowledge of known structures will be one of those strategies. However, that does not imply that there is a fixed and known set of relationships which are possible between compound elements. All the listener needs are hints which push in the appropriate direction for interpretation … The listener is playing the odds' (Bauer 2017: 105).

5 Conclusion

Our experimental study asked participants to rate the naturalness of the PNG and PNM constructions in 20 attested contexts and to paraphrase semantic relations holding between the PN and HN. Our study firstly showed that true alternation is rare even for examples deemed possible alternates in context. Most notably, we found that the association between semantic relations and PN constructions is not categorical but rather probabilistic in nature. When we compared the odds ratios for different values on the animacy hierarchy with those for the semantic relations we distinguished, the latter provide a larger range. Moreover, semantic relations explain the findings for individual

examples better than animacy. This indicates the desirability of including semantic relations in the study of grammatical alternations.

In the past, semantic relations have been considered difficult to operationalise. Our study puts forward a method for doing this. Our methodological approach was data-driven, and yielded insights that may not have been brought out in a different, more controlled approach. In several cases, it was precisely participants' misinterpretations of PNGs and PNMs in their natural language context that drew out semantic differences and preferences. Similarly, our choice to ask participants who rated both constructions highly to provide two paraphrases invited consideration of a potential semantic difference between the PNG and the PNM in the context at hand, thereby providing insight into potential semantic contrasts between the two constructions at large. In this way the set-up of our study stimulated explanation and hypothesis generation in addition to hypothesis testing.

Finally, our study simultaneously supported two opposing theoretical views on semantic relations. Our results lend new evidence to the claim that PN constructions are underspecified and receive their referential interpretations in context. This does not, however, preclude the existence of default relations that are probabilistically associated with each construction, in particular possession – PNG and name – PNM.

Authors' addresses:

Department of Linguistics and English Language
The University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom
tine.breban@manchester.ac.uk, john.payne@manchester.ac.uk

Department of Language and Linguistic Science
University of York
York YO10 5DD
United Kingdom
julia.kolkmann@york.ac.uk

References

Barker, Chris. 1995. *Possessive descriptions*. Stanford: CSLI Publications.
Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge: Cambridge University Press.
Börjars, Kersti, David Denison & Alan Scott (eds.). *Morphosyntactic categories and the expression of possession*. Amsterdam and Philadelphia: John Benjamins
Breban, Tine. 2018. Proper names used as modifiers: A comprehensive functional analysis. *English Language and Linguistics* 22(3), 381–401.
Breban, Tine & Hendrik De Smet. 2019. How do grammatical patterns emerge? The origins and development of the English proper noun modifier construction. *English Language and Linguistics* 23(4), 879–900.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Kraemer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.

Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1), 186–213.

Ehret, Katharina, Christoph Wolk & Benedikt Szmrecsanyi. 2014. Quirky quadratures: On rhythm and weight as constraints on genitive variation in an unconventional data set. *English Language and Linguistics* 18(2), 263–303.

Ford, Marilyn & Joan Bresnan. 2013. Using convergent evidence from psycholinguistics and usage. In Manfred Krug & Julia Schlüter (eds.), *Research methods in language variation and change*, 295–312. Cambridge: Cambridge University Press.

Kolkmann, Julia. 2016. The pragmatics of possession: Issues in the interpretation of pre-nominal possessives in English. PhD dissertation, The University of Manchester.

Koptjevskaja-Tamm, Maria. 2002. Adnominal possession in the European languages: Form and function. *Sprachtypologie und Universalienforschung* (*STUF*) 55(2), 141–72.

Koptjevskaja-Tamm, Maria. 2013. A *Mozart sonata* and the *Palme murder*: The structure and uses of proper-name compounds in Swedish. In Börjars, Denison & Scott (eds.), 253–90.

Nikiforidou, Kiki. 1991. The meanings of the genitive: A case study in semantic structure and semantic change. *Cognitive Linguistics* 2(2), 149–206.

Payne, John & Eva Berlage. 2014. Genitive variation: The niche role of the oblique genitive. *English Language and Linguistics* 18(2), 331–60.

Payne, John & Rodney Huddleston 2002. Nouns and noun phrases. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 323–524. Cambridge: Cambridge University Press.

Peters, Stanley & Dag Westerståhl. 2013. The semantics of possessives. *Language* 89(4), 713–59.

Rosenbach, Anette. 2002. *Genitive variation in English: Conceptual factors in synchronic and diachronic studies*. Berlin and New York: Mouton de Gruyter.

Rosenbach, Anette. 2005. Animacy versus weight as determinants of grammatical variation in English. *Language* 81(3), 613–44.

Rosenbach, Anette. 2006. On the track of noun + noun constructions in Modern English. In Christoph Houswitschka, Gabriele Knappe & Anja Müller (eds.), *Anglistentag 2005 Bamberg: Proceedings of the conference of the German Association of University Teachers of English*, 543–57. Trier: Wissenschaftlicher Verlag Trier.

Rosenbach, Anette. 2007. Emerging variation: Determiner genitives and noun modifiers in English. *English Language and Linguistics* 11(1), 143–89.

Rosenbach, Anette. 2010. How synchronic gradience makes sense in the light of language change (and vice versa). In Elizabeth C. Traugott & Graeme Trousdale (eds.), *Gradience, gradualness and grammaticalization*, 149–79. Amsterdam and Philadelphia: John Benjamins.

Rosenbach, Anette. 2014. English genitive variation – The state of the art. *English Language and Linguistics* 18(2), 215–62.

Rosenbach, Anette. 2019. On the (non-)equivalence of constructions with determiner genitives and noun modifiers in English. *English Language and Linguistics* 23(4), 759–96.

Schlücker, Barbara. 2018. Genitives and proper name compounds in German. In Tanja Ackermann, Horst Simon & Christian Zimmer (eds.), *Germanic genitives*, 275–99. Amsterdam and Philadelphia: John Benjamins.

Stefanowitsch, Anatol. 2003. Constructional semantics as a limit to grammatical alternation: The two genitives of English. *Topics in English Linguistics* 43, 413–44.

Szmrecsanyi, Benedikt. 2013. The great regression: Genitive variability in Late Modern English news texts. In Börjars, Denison & Scott (eds.), 59–88.

Szmrecsanyi, Benedikt, Douglas Biber, Jesse Egbert & Karlien Franco. 2016. Toward more accountability: Modeling ternary genitive variation in Late Modern English. *Language Variation and Change* 28(1), 1–29.

Szmrecsanyi, Benedikt & Lars Hinrichs. 2008. Probabilistic determinants of genitive variation in spoken and written English: A multivariate comparison across time, space, and genres. In Terttu Nevalainen, Irma Taavitsainen, Päivi Pahta & Minna Korhonen (eds.), *The dynamics of linguistic variation: Corpus evidence on English past and present*, 291–309. Amsterdam and Philadelphia: John Benjamins.

Taylor, John R. 1996. *Possessives in English: An exploration in cognitive grammar*. Oxford: Oxford University Press.

van Valin, Robert D. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Vikner, Carl & Per A. Jensen. 2002. A semantic analysis of the English genitive. Interaction of lexical and formal semantics. *Studia Linguistica* 56(2), 191–226.