POSS-S vs POSS-OF revisited

Kersti Börjars, David Denison and Grzegorz Krajewski
University of Manchester

Workshop on Genitive variation in English
ISLE 2, Boston, 18 June 2011

1. Introduction
2. Non-structural factors
3. Structural factors
4. Conclusions

Acknowledgements

- Maciej Baranowski
- Inbal Arnon
- Stephanie Dipper
- Alan Scott
- Project 'Germanic possessive *-s*: an empirical, historical and theoretical study' funded by the Arts and Humanities Research Council 2006-9. We gratefully acknowledge their support.

Arts & Humanities Research Council

Definition of terms

Types of possessive discussed

- Usual alternation:
- POSS-S
  Obama's government
- POSS-OF
  the government of Obama

- Here *Obama* is **possessor**, *(the/a) government* is **possessum**.

Types of possessive ignored

- Dataset excludes pronominal possessors, e.g.
  his in POSS-S
  of him in POSS-OF

- Other possessive relations:
- compound possessive
  the Obama government
- double possessive
  a government of Obama's
- Nothing to say here about these variants.

Dataset

- All spoken data in BNC (ca. 10m words).
- All possessive NPs containing POSS-S or POSS-OF
- Database contains 43,151 British English possessive NPs, reduced to 41,738 when descriptive genitives (*women's magazines*) are stripped out.
- Other kinds of example also removed during detailed statistical analysis, further reducing dataset to 40,354 tokens.

7

Dataset

- Before analysis, we revisited an earlier decision. Measure possessives included in earlier analysis:

  a delay of about twenty minutes (KRT 5372)

  about twenty minutes' delay
- Payne & Huddleston (2002: 470) class POSS-S measure genitives, like descriptive genitives, as modifiers which do not confer definiteness. They often lack a POSS-OF alternative, and distribution will be very different, so for this analysis we have removed them.

8

Non-structural factors

Animacy, number, topicality

- Definiteness used as proxy for topicality, though clearly an oversimplification, esp. with respect to so-called 'first-mention definites' (e.g. Fraurud 1990, Poesio & Vieira 1998).
- Some may have some level of topicality when there is an associative relation with a previous referent:

1. We had walked for ages when we finally found a restaurant we both liked. As we entered, <u>the waiter</u> greeted us enthusiastically.
2. I've just bought a new computer. I'm not that happy with <u>the keyboard</u> actually, so I'm thinking of returning it.

10

Topicality

- Possessum topicality is excluded:
- If definiteness is a proxy for topicality, this is a knockout context, since in POSS-S construction, possessum is always definite.

11

Values

- Animacy had 9 possible values, collapsed to 6:
  - human
  - animal
  - time
  - place
  - body part
  - inanimate
- Number collapsed to singular vs. plural.
- All unclear or unclassified examples of animacy, number and topicality were removed.

12

Logistic regression

- Simple model, illustrated with possessor animacy as predictor.
- Figures from database plotted in Table 1, with odds of POSS-S occurring simply calculated as a ratio.

13

Type	Possessor animacy					
	human	animal	time	place	body part	inanimate
POSS-OF	7907	199	1989	4265	362	16893
POSS-S	6832	112	729	878	20	156
odds of POSS-S	0.864	0.563	0.367	0.206	0.055	0.009

Table 1 Frequencies of possessives for different levels of possessor animacy

14

Logistic regression

- 'Animal' is (arbitrarily chosen) reference value. Intercept in Table 2 represents the logarithm of odds of POSS-S for an animal possessor.
- A 'coefficient' (*B* in column 2) computed for each of the 5 remaining values of animacy = log odds from Table 1 shifted by amount of intercept.
- A positive coefficient means increased odds compared to the reference level, i.e. a greater likelihood of POSS-S compared with animal possessors, while a negative coefficient means decreased odds, and zero means no difference from the reference level. The bigger the absolute number, the greater the difference from the reference level.

15

	B	*SE*	*z*	*p*
(intercept)	-0.575	0.118	-4.870	<.001
possessor animacy = body part	-2.321	0.258	-8.990	<.001
possessor animacy = human	0.429	0.119	3.590	<.001
possessor animacy = inanimate	-4.110	0.143	-28.760	<.001
possessor animacy = place	-1.006	0.124	-8.120	<.001
possessor animacy = time	-0.429	0.126	-3.410	<.001

Table 2 Coefficients of the model with possessor animacy as a predictor

16

Interaction

- Body part as possessum associated with human as possessor: factors interact.
- Put all of topicality, animacy and number into a single model.

17

	B	*SE*	*z*	*p*
(intercept)	-1.138	0.298	-3.820	<.001
possessor animacy = body part	-2.392	0.263	-9.090	<.001
possessor animacy = human	0.668	0.130	5.150	<.001
possessor animacy = inanimate	-4.032	0.152	-26.500	<.001
possessor animacy = place	-1.010	0.136	-7.440	<.001
possessor animacy = time	-0.396	0.137	-2.890	<.005
possessum animacy = body part	0.569	0.286	1.990	<.047
possessum animacy = human	-0.855	0.274	-3.120	<.003
possessum animacy = inanimate	-0.066	0.273	-0.240	<.808
possessum animacy = place	0.240	0.277	0.870	<.388
possessum animacy = time	-0.727	0.294	-2.470	<.014
possessor topicality = indefinite	-0.194	0.039	-5.000	<.001
possessor number = singular	0.877	0.041	21.470	<.001
possessum number = singular	-0.126	0.034	-3.670	<.001

18

- Now effect of possessum as body part only marginally significant.
- Overwhelming effect of possessor as human is confirmed.

19

Structural factors

Weight

- Many studies look at weight of possessor and possessum.
- Few distinguish as to how weight distributed in phrase.
- Jucker (1993) found that postmodification on possessor decreases chance of POSS-S.
- In Kreyer's (2003) smaller, also written dataset (*n*=698), a postmodified possessor makes POSS-OF compulsory.

21

Weight ~ length

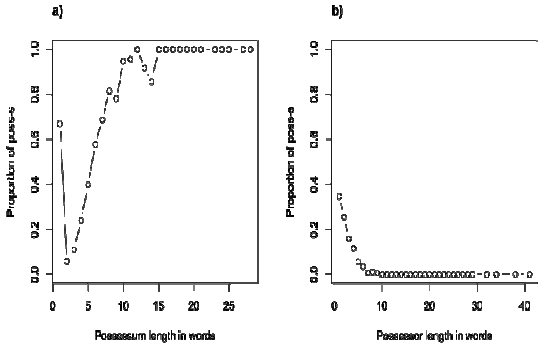
- Structural complexity not straightforward. Is a modified adjective more or less complex than 2 unmodified ones?
  - the incredibly stupid dog
  - the stupid dirty dog
- Similarly, 2 PPs vs. 1 more complex PP:
  - the student of chemistry from Bristol
  - the student from the mayor's estate
- We used length as a proxy. Close correlation between length in words and length in syllables, so use former.

22

Principle of end-weight

- General tendency to prefer long(er) constituents at end of phrase:
  - a football player's performance
    - POSS-S: POSSESSOR < POSSESSUM
  - the performance of a football player
    - POSS-OF: POSSESSUM < POSSESSOR
-

23



24

Possessum length = 1

- In poss-of, the possessum is a standard NP.
- A one-word non-pronominal NP only grammatical if N = proper noun, non-count singular or plural, otherwise a determiner is required.
- But in POSS-S, the possessor phrase acts as determiner, so e.g. singular count nouns are OK as possessum
  - we lived in my aunt's house (D90 109)
  - *we lived in house of my aunt
- For them it is a knockout context.
- So we excluded all one-word possessums (n = 8994).

25

Other extreme lengths

- We also excluded possessum length > 15, since invariably POSS-S, also possessor length > 10, since invariably POSS-OF (n = 384).
- Also datapoints for which possessum or possessor length was 0 (n = 566, n = 1, respectively).
- Modelled possessor length and possessum length and their interaction.

26

	B	*SE*	*z*	*p*
(intercept)	-3.842	0.091	-41.97	<.001
possessum length	2.497	0.090	27.65	<.001
possessor length	-1.578	0.127	-12.41	<.001
possessum length × possessor length	0.352	0.117	3.01	<.003

Table 12 Coefficients of the model with length of possessum and possessor

27

Effect of length

- If length of possessum increases by one word, log-odds of POSS-S increases by 2.497, and if length of possessor increases by one word, log-odds of POSS-S decreases by 1.578. Directionality as predicted.
- But as length of possessor increases, the *effect* of the length of possessum increases as well. Analogically, as length of possessum increases, the (negative) effect of the length of possessor decreases (remaining, however, significant for all datapoints).
- What this means is that of the two length variables, possessum length is the more potent both in itself and in combination with possessor length.

28

Pre- vs postmodification

- Created two binary variables for their presence/absence.
- Presence of premodification has much weaker effect than presence of postmodification and disappears completely when postmodification is present.
- Conversely, effect of postmodification survives (though weaker) when premodification is present.

30

	B	*SE*	*z*	*p*
(intercept)	-1.647	0.031	-52.59	<.001
premodification present	-0.848	0.041	-20.87	<.001
postmodification present	-2.477	0.320	-7.73	<.001
premodification present × postmodification present	1.301	0.359	3.62	<.001

Table 13 Coefficients of the model with presence of premodification and postmodification

31

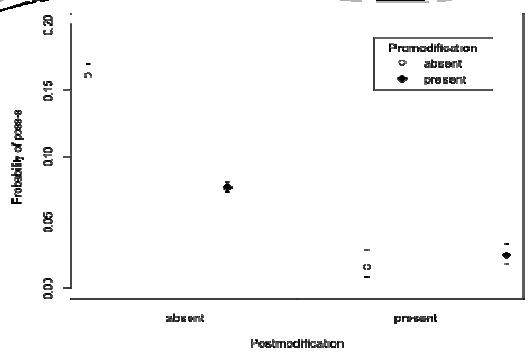


Figure 2 Predicted probability of POSS-S as a function of presence of premodification and postmodification

32

Length of modification

- Does effect of postmodification remain strong when actual length (rather than mere presence) taken into account?
- Data too sparse: variation in type of possessive only found for two values of premodification length (1 and 2) and for three values of postmodification length (2, 3 and 4), and of these five cases, three have only a handful of POSS-S each.
- Even restoring length=0 datapoints (excluded because so disproportionally frequent) would not help much.

33

Type of possessive	Length of premodification						
	1	2	3	4	5	6	7
POSS-OF	1007	330	116	36	12	1	1
POSS-S	37	3	0	0	0	0	0

Table 14 Frequency of possessives for different lengths of possessor premodification

Type of possessive	Length of postmodification						
	1	2	3	4	5	6	7
POSS-OF	66	391	489	265	153	96	43
POSS-S	0	33	6	1	0	0	0

Table 15 Frequency of possessives for different lengths of possessor postmodification

34

Split possessives

you must put something in a person's mouth that has epilepsy (F8C 105)

We don't know the gentleman's name with the tape recorder. (FM7 8)

- Generally said to be ungrammatical (Quirk et al. 1985: 1282, Payne & Huddleston 2002: 479 n. 65).
- We argue that not ungrammatical, part of a more general process of extraposition or production errors (Denison, Scott & Börjars (2010). Rather a strategy to avoid standard POSS-S constructions with postmodified possessor.
- Is split possessive more likely the longer the postmodification?

35

Split	Length of postmodification							
	1	2	3	4	5	8	9	10
absent	78	83	21	1	0	0	0	0
present	0	2	8	1	1	1	1	1

Table 16 Presence of split as a function of postmodification length.

- Clear relationship between the presence of split and length of postmodification.
- Relationship even more pronounced when actual length of split part taken into account: the longer the postmodification, the longer the split part (regression coefficient B = 1.276, SE = 0.2, $p < .001$).

36

Conclusions

Postmodification

- Weight is an independent factor and has more effect than information structural status.
- Why should weight in postmodification have more of an effect than in premodification?
- Quirk et al. (1985) cite ambiguity avoidance if postmodification ends in a noun:

  the man with the car's ears

  the man with the cat's ears ['avoided']

38

Kreyer's 'Proximity principle'

- Processing based explanation, that 'related constituents should be in the proximity of one another'.
- Any modification should be as close to its head as possible. In Kreyer's terminology, the possessor modifies the possessum and hence the two should be in proximity to each other. Similarly, any modification of the possessor or the possessum should stand in proximity to its head.

39

Proximity with postmodification

- Principle should disfavour both postmodification of possessor in POSS-S and of possessum in POSS-OF:

POSS-S: POSSESSOR+*POSTMOD*'s POSSESSUM

POSS-OF: POSSESSUM+*POSTMOD* *of* POSSESSOR

- Both proximity requirements satisfied in

  the *car* [of the *man* that is talking to you]
- So why should 'group genitive' ever occur?

  [the *man* that is talking to you's] *car*
- No explanation for selective application of proximity principle.

40

Proximity with premodification

POSS-S: POSSESSOR's *PREMOD*+POSSESSUM

POSS-OF: POSSESSUM *of* *PREMOD*+POSSESSOR

- So premodified possessors should favour POSS-S as strongly as postmodified possessors prefer POSS-OF.
- Not so in Kreyer's data or ours. Natural, attested examples:

the general *knowledge* [of the sixth form science *teacher*]

the *productivity* [of the independent middle *peasant*]

rather than predicted

[the sixth form science *teacher*'s] general *knowledge*

[the independent middle *peasant*'s] *productivity*

41

Affix and clitic

- The six criteria posited by Zwicky & Pullum (1983: 503–4) are generally used to distinguish between affixes and clitics. Subsequent literature frequently implies that there is a clear dichotomy resulting in two distinct categories. In synchronic descriptions, if elements show inconsistent behaviour with respect to these criteria they have been deemed untypical affixes or untypical clitics.
- But a simple dichotomy between affix and clitic is as inappropriate for the English POSS-S (Börjars & Vincent 2011, Denison, Scott & Börjars 2010) as for the broadly similar Swedish POSS-S (Börjars 2003)
- It does not allow a proper description of either its current properties or its historical development.

42

Multi-dimensional distinction

- We take suggestion that POSS-S is a phrasal affix (Miller & Halpern 1993, Payne 2009, Zwicky 1987) and extend argument to allow for multi-dimensional distinction.
- Now clitic and affix may represent fairly common clusterings of properties, but other constellations are possible.
- A diachronic argument for clitic status of POSS-S is that its precursor in Old English was clearly an affix (or an inflection), and since PDE POSS-S behaves differently, therefore it must be a clitic – if simple dichotomy.

43

OE vs. PDE

- The most common changes are (i) (*e*)*s* was one exponent in a paradigm, *'s* has only one form; (ii) GEN in Old English was an agreement feature, *'s* is marked once only; (iii) (*e*)*s* occurred on the head, *'s* occurs on the right edge of the phrase.
- Neither (i) or (ii) impinges on the issue of whether the description of POSS-S as a clitic is appropriate; this rests on (iii). The evidence we have presented in this paper shows that this is not as clear-cut an issue as has generally been assumed.

44

Postmodified heads in NPs

- Evident that speakers avoid realising POSS-S on the right edge when the right edge is not also the head.
- This goes beyond any general process of extraposition (Denison, Scott & Börjars 2010: 555–6).
- In spoken ICE-GB, 14.8% of all NPs have postmodification, whereas in our corpus, the proportion of possessors in the POSS-S construction with postmodification is about 2.2% (or 1% if discount head + *else*).
- Clear that there is a special interaction between postmodification and possessors in the POSS-S construction.

45

Persistence

- In Hopper's (1991: 28–30) principle of PERSISTENCE, a grammaticalised element retains some evidence of its original lexical meaning. Likewise – we would argue – with structural properties (see also Breban 2009 for another example).
- So head placement of OE genitive (*e*)*s* persists to some extent in POSS-S, even though it has developed into an edge-based once-only marking element.
- Only when possessor NP is head-final can both constraints on placement of POSS-S be satisfied.
- Explains low rate of properly postmodified possessors in the POSS-S construction and resort to POSS-OF even if other factors would militate against it, or to the split construction.

46

Affix and clitic

- In sum, these are idealised, "pure" categories.
- The behaviour of most bound elements will be messier.

47

References

(only those cited in presentation)

Börjars, Kersti. 2003. Morphological status and (de) grammaticalisation: The Swedish possessive. *Nordic Journal of Linguistics* 26, 133–63.

Börjars, Kersti & Nigel Vincent. 2011. Grammaticalization and directionality. In Heiko Narrog & Bernd Heine (eds.), *Handbook of grammaticalization*, 163–76. Oxford: Oxford University Press.

Breban, Tine. 2009. Structural persistence: A case based on the grammaticalization of English adjectives of difference. *English Language and Linguistics* 13, 77–96.

Denison, David, Alan Scott & Kersti Börjars. 2010. The real distribution of the group genitive. *Studies in Language* 34, 532–64.

Fraurud, Kari. 1990. Definiteness and the processing of NPs in natural discourse. *Journal of Semantics* 7, 395-433.

Hopper, Paul. 1991. On some principles of grammaticalization. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization.* Vol 2, *Focus on types of grammatical markers*, 17–35. Amsterdam: John Benjamins.

Jucker, Andreas. 1993. The genitive versus the *of*-contruction in newspaper language. In Andreas Jucker (ed.), *The noun phrase in English. Its structure and variability*, 121–36. Heidelberg: Carl Winter.

49

Kreyer, Rolf. 2003. Genitive and of-construction in modern written English. Processability and human involvement. *International Journal of Corpus Linguistics* 8, 169-207.

Miller, Philip & Aaron Halpern. 1993. English possessives and the syntax of morphological features. In Laurel Smith Stvan & et al (eds.), *FLSM III: Papers from the third Annual Meeting of the Formal Linguistics Society of MidAmerica*, 219–34. Bloomington, In: Indiana University Linguistics Club.

Payne, John. 2009. The English genitive and double case. *Transactions of the Philological Society* 107, 322–57.

Payne, John & Rodney Huddleston. 2002. Nouns and noun phrases. In Rodney Huddleston & Geoffrey K Pullum (eds.), *The Cambridge Grammar of the English Language*, 323–523. Cambridge: Cambridge University Press.

Poesio, Massimo & Renata Vieira. 1998. A corpus-based investigation of definite description use. *Computational Linguistics* 24, 183-216.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Zwicky, Arnold M. 1987. Suppressing the Zs. *Journal of Linguistics* 23, 133-48.

Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59, 502-13.

50