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.

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.

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, the waiter greeted us enthusiastically.
  2. I've just bought a new computer. I'm not that happy with the keyboard actually, so I'm thinking of returning it.

Non-structural factors

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.

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.
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.

### Table 1

<table>
<thead>
<tr>
<th>Type</th>
<th>Possessor animacy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>human</td>
</tr>
<tr>
<td>POSS-OFF</td>
<td>7907</td>
</tr>
<tr>
<td>POSS-S</td>
<td>6832</td>
</tr>
<tr>
<td>odds of POS-S</td>
<td>0.864</td>
</tr>
</tbody>
</table>

Table 1. Frequencies of possessives for different levels of possessor animacy

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.

### Table 2

<table>
<thead>
<tr>
<th>Type</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\zeta$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(intercept)</td>
<td>-0.575</td>
<td>0.118</td>
<td>-4.870</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>possessor animacy = body part</td>
<td>-2.321</td>
<td>0.258</td>
<td>-8.990</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>possessor animacy = human</td>
<td>0.429</td>
<td>0.119</td>
<td>3.590</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>possessor animacy = inanimate</td>
<td>-4.110</td>
<td>0.143</td>
<td>-28.760</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>possessor animacy = place</td>
<td>-1.006</td>
<td>0.124</td>
<td>-8.120</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>possessor animacy = time</td>
<td>-0.429</td>
<td>0.126</td>
<td>-3.410</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

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

Interaction

- Body part as possession associated with human as possessor: factors interact.
- Put all of topicality, animacy and number into a single model.
• Now effect of possessum as body part only marginally significant.
• Overwhelming effect of possessor as human is confirmed.

**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.

**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.

**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
**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).

**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.

<table>
<thead>
<tr>
<th>B</th>
<th>SE</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(intercept)</td>
<td>-3.842</td>
<td>0.091</td>
<td>-41.97</td>
</tr>
<tr>
<td>possessum length</td>
<td>2.497</td>
<td>0.090</td>
<td>27.65</td>
</tr>
<tr>
<td>possessor length</td>
<td>-1.578</td>
<td>0.127</td>
<td>-12.41</td>
</tr>
<tr>
<td>possessum length x possessor length</td>
<td>0.352</td>
<td>0.117</td>
<td>3.01</td>
</tr>
</tbody>
</table>

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

**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.

**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.

<table>
<thead>
<tr>
<th>B</th>
<th>SE</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(intercept)</td>
<td>-1.647</td>
<td>0.031</td>
<td>-52.59</td>
</tr>
<tr>
<td>premodification present</td>
<td>-0.848</td>
<td>0.041</td>
<td>-20.87</td>
</tr>
<tr>
<td>postmodification present</td>
<td>-2.477</td>
<td>0.320</td>
<td>-7.73</td>
</tr>
<tr>
<td>premodification present x postmodification present</td>
<td>1.301</td>
<td>0.359</td>
<td>3.62</td>
</tr>
</tbody>
</table>

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

Split possessives

you must put something in a person's mouth that has epilepsy (FBC 109).
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 extrapolation 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?

Conclusions

- 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).  

Table 14 Frequency of possessives for different lengths of possessor premodification

<table>
<thead>
<tr>
<th>Type of possessive</th>
<th>Length of premodification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>POSS-OF</td>
<td>1007</td>
</tr>
<tr>
<td>POSS-S</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 15 Frequency of possessives for different lengths of possessor postmodification

<table>
<thead>
<tr>
<th>Type of possessive</th>
<th>Length of postmodification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>POSS-OF</td>
<td>66</td>
</tr>
<tr>
<td>POSS-S</td>
<td>0</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Split</th>
<th>Length of postmodification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>absent</td>
<td>78</td>
</tr>
<tr>
<td>present</td>
<td>0</td>
</tr>
</tbody>
</table>
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 car's ears ['avoided']

Kreyer’s ‘Proximity principle’

- Processing based explanation, that ‘related constituents should be in the proximity of one another’.
- Any modification should be a 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.

Proximity with postmodification

- Principle should disfavour both postmodification of possessor in POSS-S and of possessum in POSS-OF:
  - POSS-S: POSSESSOR + POSTMOD of 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.

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.

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.
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.

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 0% if discount head + else).
- Clear that there is a special interaction between postmodification and possessors in the POSS-S construction.

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.

Affix and clitic

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

References

(only those cited in presentation)


