A construction grammar account of possessive constructions in Lancashire dialect: some advantages and challenges

ABSTRACT

This study investigates reduction of 1Sg possessives in possessive-noun constructions in Lancashire dialect. On the basis of a corpus of 26 interviews we show that reduction patterns according to the (in)alienability hierarchy. This dialectal evidence runs counter to the normal assumption about English, i.e. that there is no such effect. Following work by Haspelmath (2006b) that reinterprets iconicity effects in terms of frequency we proceed to show that frequency may indeed underlie alienability/iconicity in our data as well. Relative frequency seems more useful in capturing the correlation with reduction than absolute frequency. For a few [1Sg POSS-N] combinations the reduction facts are problematic for the frequency-based account we offer. These difficulties might seem to disappear in light of the construction grammar notion of schemas, but we point out that this notion itself has serious theoretical problems associated with it. Future theory-driven work on dialect grammar may help resolve these issues.

1 INTRODUCTION

This study deals with 1st person singular possessive nominal constructions in Lancashire dialect, exemplified by (1-4) below:*

(1) … I couldn't play for them because they couldn't afford *my football shoes* (JA)[1]

(2) I was so young then like and er *me brother* took the opportunity and he went. (HF)

(3) when I was four I used to go round this house with *my eyes* closed (RG)

(4) I remember *my father* coming out a small room (CS)

As is obvious from the transcription of example (2), there is variation in the realisation of the possessive pronoun: while it is realized as [maɪ] in (1), in (2) we find the shorter form [mi]. Examples (3-4) display additional reduced variants: in (3) the realisation is [ma], in (4) we get [mə]. (Since unlike in the case of [mi] there is no conventional way to represent them as distinct from [maɪ], the conventional spelling *my* is retained.)

It is not clear whether in the speakers' grammars the form *me* is essentially the same form as the objective personal pronoun. There would be some evidence for this hypothesis if in our corpus we found *us* used as the 1Pl possessive, but that is the case for only one speaker:

(5) But if we had er *us* clothes wet… (ED)

The status of [mə] is also uncertain. Given that schwa is the vowel that requires least articulatory effort it may be a reduced form of any or all of the three other variants. For the purpose of this paper we will remain agnostic with regard to these aspects of the status of these forms — we will analyse [maɪ] as the fullest variant and the other forms as reduced, with [mə] being more reduced that [mi] and [ma].

The purpose of this article is to account for the variation in the realisation of the 1Sg possessive pronoun, both across and within speakers. Taking our cue from the typological literature on (in)alienability effects (see section 2, below), we will show that reduction is more frequent in constructions where the noun is a kinship or body part term (see exx. (2-4) and (3), respectively) than where it is not (see ex. (1)). This is a remarkable finding inasmuch as English has never been characterized as a language where (in)alienability plays a structural role. Contrary to traditional work in typology we will not automatically assume that this effect is based on iconicity (the closeness of the relation between possessor and possessed, see e.g. Haiman, 1985) but instead investigate the possibility that the underlying factor may be token frequency, i.e. of the possessor-possessed strings in question (for the correlation between token frequency and morphophonological reduction see the usage-based work by e.g. Schuchardt, 1885, Zipf, 1935, Fidelholtz, 1975, Hooper, 1976, Bybee & Scheibman, 1999, Berkenfield, 2001). In this connection we will be drawing on Haspelmath (2006b), who has shown that a frequency-based account of possessor encoding makes better predictions than a more traditional view based on the semantics of (in)alienability. Some aspects of the variation observed are not easy to explain in terms of the token frequencies of the patterns in question — at least not if we define token frequency in absolute terms, i.e. as the raw frequency of a given pattern. Despite the fact that many studies in the usage-based approach (including all the ones mentioned above) rely on absolute frequencies, there is a long tradition in linguistics — going back at least to Jespersen (1923) — that argues that frequencies of certain patterns must be evaluated relative to other frequencies. The work on reduction phenomena by Jurafsky and his colleagues is a good more recent example. In a study on lenition in 10 frequent function words (not including *my*) by Jurafsky, Bell, Fosler-

Lusier, Girana & Raymond (1998) predictability is one the factors studied. Predictability is operationalized in their study as the conditional probability that a certain word will occur given the two words immediately preceding it. Predictability is found to be a significant factor (although, surprisingly, for the function words *I* and *you*, Jurafsky et al. find the opposite effect: high predictability leads to less rather than more reduction, see 1998: 3113). Haspelmath's (2006b) study is especially relevant to our study. He assesses the frequency of nouns occurring in the possessive relative to their occurrence in all other constructions. This method appears to explain some of the data better, although some other aspects of the data seem more susceptible to an analysis in absolute not relative terms.[2] An alternative solution offered here is to adopt the construction grammar notion of schemas (see e.g. Langacker. 1987; Goldberg; 1995; Croft, 2001). Since the data that we have at our disposal here cannot but lead us to equivocate between absolute/relative token frequencies and schemas as the most appropriate explanation, we conclude that further research is called for to shed light on these aspects of the usage-based model and construction grammar.

The structure of the article is as follows. Section 2 throws the present study into crosslinguistic relief, showing how contrary to what is traditionally assumed about English in many languages of the world (in)alienability does have structural implications. In section 3 we describe our corpus, and discuss the way in which we coded and retrieved the data. Section 4 presents the results, and discusses them in terms of (in)alienability, token frequency (absolute and relative) and schemas. Section 5 ends the discussion with some general conclusions and pointers for future research.

2 (IN)ALIENABILITY EFFECTS IN THE LANGUAGES OF THE WORLD

A number of cross-linguistic studies of substantival possession (e.g. Ultan, 1978; Seiler, 1983; Nichols, 1988; Chappell & McGregor, 1996a; Heine, 1997) have shown that one of the major factors underlying how this relation is expressed is the distinction between alienable and inalienable possession. Inalienable possession is generally seen as involving a fairly stable relation over which possessors have little or no control, alienable possession as comprising a variety of less permanent, more controlled relationships. Whether the relationship between the possessor and possessed is alienable or inalienable depends to some extent on the possessor (only humans and higher animates are typically seen as capable of exerting control) but primarily on the semantic properties of the possessed. Most commonly the inalienable nouns encompass some set of nouns referring to body parts, kinship terms, spatial terms and part-whole relations. The inalienable/alienable distinction may affect the formal realisation of the possessor and possessed in several ways all of which concern the linguistic proximity between the possessor and the possessed. If there is a difference between inalienable and alienable possession in this respect, it is always the case that the possessor and possessed are located closer to each other in inalienable possession than in alienable possession. This is typically attributed to the workings of iconicity (Haiman, 1985a, b; Croft, 2003; Haspelmath, 2005), in particular the iconicity of distance, i.e. the tendency for the conceptual distance between concepts to be reflected in the linguistic distance between the linguistic expressions of these concepts.

Of the various manifestations of the iconicity of distance in regard to alienable and inalienable possession of relevance to the current discussion are only those involving person forms. Assuming that proximity of forms is a precondition for

fusion, as suggested by the grammaticalisation literature (see e.g. Hopper & Traugott, 2003: 8-13 for some examples of periphrastic constructions coalescing over time), a particularly common reflection of the greater conceptual proximity between the possessor and possessed in inalienable constructions than in alienable ones involves person marking of the possessor on the possessed. If a language employs some type of bound person marking (by means of affixes, clitics or weak forms) of the possessor on the possessed with alienable nouns, it also uses such bound person marking with inalienable nouns, but not necessarily vice versa. Observe, for example, the presence of a person suffix on the possessed in (6a) as compared to the use of an independent person form in (6b) in Hoave, an Austronesian language spoken in the Western Solomon Islands.

(6) Hoava (Davis, 2003: 98, 105)

(a) sa kalu-na

ART hair-3SG

'his hair'

(b) nana siki

3SG:POSS dog

'his dog'

A similar contrast is to be found in Paamese and many other Oceanic languages, in which bound person possessors are used in both alienable and inalienable constructions, but while in the inalienable construction the bound person possessor is attached to the possessed (7a), in the alienable construction it is attached to a special

classifier (7b) rather than the possessed.

(7) Paamese (Crowley, 1996:389, 411)

(a) vati-n

head-3SG

'his head

(b) ani emo-n

coconut CLF-3SG

'his drinking coconut'

Thus the possessor and possessed are closer together linguistically in the inalienable construction than in the alienable one. In contrast to Hoava and Paamese, Udihe, a Tungusic language of the Russian Far East, displays bound person marking of the possessor on the possessed in both inalienable and alienable possession. There is nonetheless a difference in the linguistic proximity of the possessor and possessed in the two constructions in line with the iconicity of distance. In inalienable possession the bound person possessor is directly affixed to the possessed, in alienable possession the two are separated from each other by the additional suffix *–ni.* Compare (8a) and (8b).

(8) Udihe (Nikolaeva & Tolskaya, 2001: 481, 505)

(a) bi anda-i

1SG friend-1SG

'my friend'

(b) nuani ja:-ni-ni

3SG cow-AL-3SG

'his cow'

Thus the linguistic distance between the bound person possessor and possessed is shorter in the inalienable construction than in the alienable one.

Morphological fusion of the possessor and possessed as illustrated above is often preceded or accompanied by phonological reduction of the possessor. It should therefore come as no surprise that the other common manifestation of the alienable/inalienable opposition with respect to person forms concerns the phonological size of the possessor. In languages exhibiting differences in the form of the person markers found in alienable and inalienable possession it is typically the case that the forms used in inalienable possession are shorter or morphologically simpler than those which occur in alienable possession. This is obviously so in languages in which alienable constructions require the presence of a free pronoun while inalienable ones have person markers attached to the possessed, the latter being invariably shorter than free forms. Recall the situation in Hoava illustrated in (6) above. But the same difference may involve bound person forms. Nichols (1988: 564) states that among the languages in her corpus which use bound person forms in both alienable and inalienable possession, the ones which occur in inalienable constructions are shorter than the ones found in alienable constructions. A case in point is Paumari, an Arauan language spoken in the State of Amazonas in Brazil, which uses the prefixes in (9a) for alienable possession and the discontinuous prefix and suffix in (9b) for inalienable possession.

(9) Paumari (Chapman and Derbyshire 1991: 256-7)

(a)	1Sg	kodi-	(b)	1Sg	o- -na
	2Sg	kada-		2Sg	i- -ni
	3Sg	kidi-		3SgF	Ø- -ni
	Gen	ka-		3SgM	Ø- -na
	1Pl	akadi-		1Pl	a- -na
	2Pl	avakadi-		2Pl	ava- -ni
	3Pl	vakadi-		3Pl	va- -na

Needless to say, this cross-linguistic tendency for person forms in inalienable possession to be phonologically reduced relative to those found in alienable possession plays a major role in our analysis of the factors underlying the realisation of first person singular in Lancashire dialect.

In discussing the formal reflexes of the alienable/inalienable distinction we have side stepped the issue of the type of nouns that tend to emerge as inalienable. Cross-linguistic investigations clearly show that there is quite some variation in this regard. Nonetheless, several linguists, most notably Seiler (1983: 13), Haiman (1985: 136), Nichols (1988) and Chappell and McGregor (1996b: 26) have suggested that the type of nouns most likely to be inalienable may be seen to form a typological hierarchy. For example, Nichols (1988: 572; 1992: 160) has suggested the inalienability hierarchy in (10).[3]

(10) The inalienability hierarchy

body parts and/or kinship terms > part-whole > spatial relations > culturally basic possessed items > other

The inalienability hierarchy is slightly unusual as far as typological hierarchies are concerned, in that it is headed jointly by two items, body parts and kin terms, connected by an inclusive disjunction. The 'or' part is a reflection of the fact that there are languages in which only body parts are treated as inalienable (e.g. Dizi, Paumari, Tauya, Worora) and also languages in which only kin terms emerge as such (e.g. Dongolese Nubian, Mumuye and Wappo). The 'and' part, in turn, caters for the languages in which inalienability embraces exclusively just body parts and kin terms (e.g. Haida, Maung, Washo and Yuchi).

The alternative categorisations of inalienability captured in the inalienability hierarchy suggest that reflexes of the alienable/inalienable distinction may well be more diverse and wide spread than has been previously assumed. It is precisely in this spirit that we approach our dialect data.

3 THE CORPUS

The corpus is made up of 26 interviews with different speakers which we obtained from the North West Sound Archive.[4] After we had them transcribed we had the 1Sg possessive pronouns checked and corrected by Claire Dembry, one of our PhD students in Lancaster. The interviews run to a total of almost 230,000 words, although this includes the interviewer's (brief) questions as well. The number of words produced by the interviewees is around 200,000. While this is a fairly small corpus,

the frequency of 1st Sg possessive pronouns, especially in the kind of interviews in the NWSA, means it is large enough to draw valid conclusions.

The possessive tokens were all tagged according to their phonetic realisation ([maɪ], [mi], [ma], [mə]), as well as for the type of possessed noun (kinship term, body part, other). For our searches we used MonoConc Pro.

4 RESULTS AND EXPLANATION

4.1 By-items and by-subjects analysis: the (in)alienability effect in Lancashire dialect

Table 1, below, presents the results for the entire corpus:

TABLE 1 HERE

Chi-square analysis reveals that the distribution of the data is very highly significant, as the alienability hierarchy (see section 2) would lead us to expect ($\chi^2 = 80.71$, df = 6, critical value at $p \leq .001$ = 22.46). More specifically, the significant correlations, in order of importance, are the overreprestation of other nouns with [maɪ], the underrepresentation of kin nouns with [maɪ], the underrepresentation of other nouns with [mɪ], the overrepresentation of kin nouns with [mɪ] and finally the overrepresentation of body part nouns with [mɪ]. (If the three reduced forms are collapsed the effect is still significant at the $p \leq .001$ level: $\chi^2 = 73.76$; df = 2, critical value = 13.82).

Gries (2006) and Hollmann & Siewierska (2006) have argued that in corpus studies the focus is too often on differences between constructions in the corpus as a whole (what Gries calls by-items analysis) and not on differences between and within individual speakers (so-called by-subjects analysis). Indeed, a by-subjects analysis is very revealing in the present study as well, as we now proceed to show.

Not all speakers in our corpus have all four variants — only four of them do. Most speakers have two or three variants, while two speakers consistently use [maɪ] and one always produces [mi]. If we plot the number of speakers who produce one or several of the reduced forms against the noun type(s) they combine them with, an interesting picture emerges; see table 2, below: The + sign indicates that the speakers in question produce a reduced variant for the relevant noun category, at least some of the time. The – sign means that they use the full form. The Ø, finally, represents the absence of the given category in the speech of the speakers' in question.

TABLE 2 HERE

This table shows that all speakers behave in accordance with the alienability hierarchy. Another interesting way to look at the data is to ask whether there are speakers who consistently produce reduced possessive forms, and for which of the three noun types they do so. Table 3 presents the relevant numbers of speakers:

TABLE 3 HERE

Again, the data pattern in the way that the alienability hierarchy would predict. The one speaker who always uses the reduced possessive pronoun for body parts but not

always for kin terms is not an exception, as these two noun types are not in any way ordered with respect to each other on the alienability hierarchy.

4.2 What lies beneath: iconicity or token frequency?

We saw in section 2, above, that (in)alienability effects have often been explained with reference to iconicity: possessed entities that are somehow conceptually closer to their possessor are coded with less linguistic distance between them. Haspelmath (e.g. 2005, 2006a) has recently started looking critically at functional-typological notions such as iconicity and markedness, and has presented strong evidence that the phenomena these notions are supposed to explain are actually more straightforwardly and accurately accounted for in terms of frequency. Of particular relevance for the present article is Haspelmath's (2006b) paper, which demonstrates the superiority of a frequency-based account in relation to (in)alienability effects. It is widely recognized, at least among linguists subscribing to the usage-based model (Langacker, 1987; Barlow and Kemmer, 2000; Croft, 2000), that high token frequency constructions are more likely to undergo reduction, and to a higher degree, than rare ones (see section 1 for some references). Almost every relevant study that we are aware of (Jurafsky et al.'s work mentioned in section 1, above, being a notable exception) defines token frequency in absolute terms. That is to say, in order to explain why construction X has undergone more reduction than construction Y the tokens of each in a corpus are counted, and it is shown that X is more frequent than Y. We may carry out this exercise for our possessives data. In order to get a sense of the frequencies in ordinary speech we decided against using our NWSA corpus: the very specific genre of the

interviews might have led to undesirable biases in terms of the nouns used. Instead, in order to get a handle of token frequencies in ordinary spoken discourse we used the 10-million word spoken subcorpus of the British National Corpus.[5] Table 4, below, gives the raw frequencies of 1Sg possessive constructions with the kinship and body part terms included in our Lancashire data, and with a number of other nouns as well. The other nouns consist of two groups. The higher group of seven nouns are all the nouns that in our Lancashire corpus occur with a reduced form of the possessive at least twice, and never with the full form. The lower group of eighteen nouns are the ones that in our data never feature any reduction, whilst also occurring at least twice.[6, 7]

TABLE 4 HERE

To a large extent the BNC data are as we would expect. The mean frequency of 1Sg possessive with a kinship noun is 106.65, the median frequency being 38. which The mean and median frequencies for body part nouns are 60.00 and 41.5. For the other nouns (combining both groups), finally, the respective frequencies are 32.83 and 15. This makes it unsurprising that the kinship and body part terms should be reduced more often in our Lancashire corpus than the other nouns. There is also a difference between the first group of other nouns (which do display some reduction) and the second group (which do not): the mean frequencies are 33.29 vs. 20.00, the median frequencies 20 vs. 10. Again, the reduction effects appear to pattern in the way that one expects on the basis of the frequencies. Nonetheless, in light of the low numbers of nouns in each of the twin sub-groups we would be hesitant to attach to much importance to this observation.

A few questions emerge as well. The kinship nouns *father in law*, *niece*, *sister in law, stepfather, stepmother* and *stepsister* are actually very infrequent, as is the body part term *milk tooth*. Yet despite this our Lancashire speakers reduce them quite often: *father in law* 1 out of 3 times; *niece*, 3 out of 3 times; *sister in law*, 3 out of 3 times; *stepfather*, 2 out of 2 times; *stepmother*, 4 out of 4 times; *stepsister*, 3 out of 3 times; *milk tooth* 1 out of 1 time.[13] The example containing *milk tooth* can probably be explained straightforwardly:

> (11) …and there was, one of *me*, this milk tooth must have been, you know, troubling me in some way or another" (HF).

The speaker may have had in mind the phrase *my* (or in this case *me*) *teeth*, which *is* frequent (153 tokens in the spoken part of the BNC).

As we discussed in section 1, above, Haspelmath's (2006b) explanation of (in)alienability effects relies on relative not absolute frequencies. Relative frequencies for the kinship, body part and other nouns included in Table 4 are presented in Table 5. The percentages show how often, out of all occurrences of the relevant nouns, they occur in the 1Sg possessive construction. If the total number of noun tokens falls below 30 it becomes dangerous to calculate the relative frequency — in these cases the table simply reports the total number.[14]

TABLE 5 HERE

Interestingly, it looks as though relative frequencies do a somewhat better job at explaining the "problem case" *niece* : the high proportion of this noun occurring in

the 1Sg possessive construction may have led to its frequent reduction. The spoken part of the BNC contains some examples of *stepfather, stepmother* and *stepsister* but unfortunately none of them with a 1Sg possessive. This does not mean that this pattern is not relatively frequent (in the sense of Haspelmath): the total number of examples is so low that we cannot predict what kind of percentage would occur in the possessive construction in question. A larger corpus might be useful in this regard. As regards *father in law* and *sister in law*, here the total number in the BNC of the nouns in question is too low for the calculation to be entirely reliable. The percentages we get, i.e. 37.50% and 34.78% may nevertheless be seen as suggestive. Again, a larger corpus would be desirable.

Relative frequency does seem to run into a problem in relation to *child*: while in absolute terms the possessive construction occurs with this noun quite frequently (90 tokens is almost equal to the mean frequency of the kinship group (i.e. 106.65) and considerably higher than its median frequency (i.e. 38), the relative frequency drops to only 1.85%. One could suggest that the relative frequency of *my child* in the BNC is not be an accurate reflection of the speech of our Lancashire speakers in this respect. After all, about 60% of the spoken part of the BNC is made up of text-types from the education/informative, business, public/institutional and leisure spheres (see e.g. the on-line manual at http://www.natcorp.ox.ac.uk/docs/userManual/design.xml.ID=spodes [accessed 25 February 2007]). In these spheres one might perhaps reasonably expect speakers to refer less to their own children, and more to children in relation to education, legal issues, and so on. However, even if we filter out the texts in question (the so-called "context-governed" part of the corpus) and run a search on the remaining, more informal, conversations of the so-called demographically sampled remainder of the

spoken sub-corpus, the relative frequency does not rise to more than 2.22% (or 2.31% if we further exclude the youngest group of under fifteen-year-olds). This is similar to the median relative frequency for the category "other" (i.e. 1.70%), and lower than the mean relative frequency for that category (4.61%).

We conclude that whichever definition of frequency one chooses, a lot of the data are accounted for. Token frequency, then, whether absolute or relative, seems a viable alternative to an iconicity-based explanation of the Lancashire patterns. In view of Haspelmath's (2006b) findings it would even appear to be preferable.

4.3 Schemas: type frequency and productivity

Given the fairly long tradition, in usage-based linguistics, of relying on absolute not relative token frequencies in accounting for reduction phenomena, one may wonder whether there is another way to explain the problem cases identified above, viz. the reduced possessive pronouns with *father in law*, *niece*, *sister in law, stepfather, stepmother* and *stepsister*. Taking a construction grammar view of the issue, we argue that there is.

Construction grammarians reject the mainstream view of Generative Grammar and its derivatives that our linguistic knowledge consists of a lexicon plus separate components governing the traditional levels of language, phonology, semantics, syntax (and according to some scholars, morphology and information structure, see Croft and Cruse, 2004: Ch.9 for an overview of so-called componential models). Instead, our knowledge is made up of form-meaning pairings — i.e. constructions — of various levels of specificity and complexity, ranging from fully specific and simple

monomorphemic words to abstract, complex constructions such as, say, the transitive clause construction. The "words and rules" model is thus replaced by a model with a " **uniform representation of all grammatical knowledge** in the speaker's mind" (Croft and Cruse, 2004: 255, emphasis original).

Zooming in on possessive constructions, the idea is that on the basis of our use of and exposure to tokens of 1Sg possessive pronoun – possessed noun sequences we build up mental representations such as [*my mother*], [*my head*], [*my money*], and so on. As hinted above, these constructions would not only contain phonological information, but semantic information as well, linking the elements of the linguistic structure (the possessive pronoun and the possessed noun) to our knowledge of what it means to "possess" the relevant entities (viz. mothers, heads, money). All these constructions obviously share some semantic and formal similarity: they all portray possessive relations, and they all comprise a possessive and a noun. Human beings are very good at recognising patterns and similarities, and are therefore likely to build up more schematic representations that generalize over the various types of possessive relation. Given the semantic similarity between possessive relations involving members of kin, they may thus form a schema [*my* KIN] (likewise for [*my* BODY PART]). On an even higher level an additional schema [*my* X] may emerge, which is entirely unspecified as regards the type of entity possessed. Going back down to the intermediate level of [*my* KIN] and [*my* BODY PART], one wonders whether there is a parallel schema for the 'other' category, which we will refer to here as [*my* OTHER]. The case for this schema seems to us less convincing than for kinship terms and body parts, in view of the lack of semantic cohesion of the 'other' category: the kinds of entities included in this category are obviously very diffuse indeed (see e.g. Bybee, 1985: 118 for the importance of semantic similarity in the emergence of

categories). At any rate, whether or not such a schema exists is not that relevant to our argument. We suggested in section 4.2 that the relatively high frequency (whether defined absolutely or relatively) of most combinations of possessive and kinship or body part noun leads to reduction of the possessive pronoun. Now it is possible that on the basis of these frequent strings speakers abstract a schema for kin and body part terms that features a reduced version of the possessive. To the extent that this is plausible, then given that schemas may be productive, that would help explain the formation of strings involving a reduced possessive where the possessed noun is not found in the possessive construction very often, such as *father in law*, *niece*, *sister in law, stepfather, stepmother*, *stepsister* (which are not frequently found in the construction in absolute terms) and *child* (which is infrequently attested in relative terms).[15]

It is important to point out that it is not necessarily the case that every speaker undergoes this sort of schema development, and consequent phonological reduction of the possessive in the relatively infrequent combinations we have just listed. Once a number of speakers have gone through this and have started producing these reduced tokens, other speakers in the community may simply store them as such. It is not the purpose of this article to try and establish when this development may have taken place.

The explanation involving schemas is attractive because it allows one to explain away the few problem cases encountered with the purely frequency-based perspective (in its relative guise and perhaps more strongly in its absolute guise). The notion of schemas has proven useful in other studies on productivity as well (for a well-known example consider Goldberg's (1995) study of novel uses of the caused-motion construction such as *I sneezed the napkin off the table*). Yet it, too, is not

without problems: the usage-based view of schema productivity raises several hard questions (which were also raised, in a different context, in Hollmann, 2003). The extent to which schemas are productive is said to be a function of the size of the category — in Bybee's terms, of the schema's "type frequency" (1985: 132-4). The more members a class already has, the more likely it is to attract new ones. The difficulty lies in the definition of high frequency: thus far it has not been determined what counts as high enough frequency for a class to become productive. The number of constructions categorized by [*my* KIN] and [*my* BODY PART] for most speakers will be a dozen or two for each. That may seem like a reasonably high type frequency, but how can we be sure? The issue is aggravated by the fact that Bybee argues that if certain members of a category have a very high token frequency, then they are unlikely to contribute very much, if at all, to the schema — instead they are more or less "autonomous" (see Bybee, 1985: 129-134 for discussion). This is how Bybee's usage-based model is able to explain the observation that crosslinguistically, irregular patterns are able to resist regularisation as long as they are sufficiently frequent (e.g., in many languages, the paradigm of the verb *be*). The problem is that the combinations of possessive and kinship or body part noun that underlie the schemas [*my* KIN] and [*my* BODY PART] *are* relatively frequent. It is again unclear what would constitute sufficiently high token frequency for a construction to be fully or largely autonomous. Croft and Cruse (2004: Ch.11), incidentally, are not convinced by Bybee's evidence for the correlation between token frequency and autonomy, but unfortunately they do not offer more precision either.

5 CONCLUSION

This article has shown that contrary to received wisdom on the English language, it does actually display what may be labelled (in)alienability effects — at least in its Lancashire variety. The by-subjects analysis in particular made it clear that there is not a single speaker in our corpus whose reduction behaviour in relation to 1Sg possessive pronouns does not follow the well-known (in)alienability hierarchy. Given the privileged position of English in terms of the amount of scholarly attention it has received compared to other languages, our finding is rather surprising, and it suggests that the recent trend towards putting dialect grammar in a typological perspective (see e.g. Kortmann, 2003, Kortmann, Herrmann, Pietsch & Wagner, 2005) is likely to be a fruitful one.

From the point of view of linguistic theory the present study has also made some interesting points. First of all, with Haspelmath (2006b) we have shown that frequency effects provide an important explanatory tool in language description, so important that the status of more traditional notions such as in this case iconicity will merit careful reconsideration. On a more critical note, we observed that the notion of token frequency is still somewhat underdeveloped. We were admittedly able to account for the facts of reduction in Lancashire pretty comprehensively. Slightly better (though not perfect) correlations were obtained by using Haspelmath's (and some others') relative, as opposed to the more widespread absolute, understanding of token frequency. We have argued that construction grammar may provide a framework that allows us to account for the aspects in the data that posed difficulties for the purely frequency-based approach. In this respect the model certainly seems attractive. However, here too we identified several theoretical problems, especially

concerning the notion of schema productivity and its relation to type and token frequency.

In order to resolve these theoretical problems it seems obvious that we need some appropriate test cases. It will not be easy to find these as it will require a lot of data, but the chances of doing so may improve as the theoretical linguistics community is getting increasingly interested in dialect grammar and new data are forthcoming.

REFERENCES

Aston, G. & L. Burnard (1998). *The BNC handbook: exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.

Barlow, M. & S. Kemmer (eds.) (2000). *Usage-based models of language*. Stanford: CSLI.

Bavin, E. (1996). Body parts in Acholi: alienable and inalienable distinctions and extended users. In Chappell, H. & W. McGregor (eds.) *The grammar of inalienability*. Berlin: Mouton de Gruyter. 841-64.

Berkenfield, C. (2001). The role of frequency in the realization of English *that*. In Bybee, J. L. & P. J. Hopper (eds.) *Frequency and the emergence of linguistic structure*. Amsterdam/Philadelphia: John Benjamins. 281-308.

Bybee, J. L. (1985). *Morphology*. Amsterdam / Philadelphia: John Benjamins.

Bybee, J. L. & J. Scheibman. (1999). The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics* 37: 575-96.

Chapman, S. & D. C. Derbyshire (1991). Paumari. In Derbyshire, D. C. & G. Pullum (eds.) *Handbook of Amazonian languages*, *vol. 3*. Berlin: Mouton de Gruyter. 161-351.

Chappell, H. & W. McGregor (eds.) (1996a). *The grammar of inalienability*. Berlin: Mouton de Gruyter.

Chappell, H. & W. McGregor (1996b). Introduction. In Chappell, H. & W. McGregor (eds.) *The grammar of inalienability*. Berlin: Mouton de Gruyter. 3-30.

Croft, W. (2000). *Explaining language change*. Harlow: Longman.

Croft, W. (2001). *Radical construction grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, W. (2003). *Typology and universals*. 2nd edn. Cambridge: Cambridge University Press.

Croft, W. & D. A. Cruse (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.

Crowley, T. (1996). Inalienable possession in Paamese grammar. In Chappell, H. & W. McGregor (eds.), *The grammar of inalienability*. Berlin: Mouton de Gruyter. 383-464.

Davis, K. (2003). *A grammar of the Hoava language, Western Solomons*. Pacific Linguistics 535. Canberra: Australian National University.

Fidelholtz, J. L. (1975). Word frequency and vowel reduction in English. *CLS* 11: 200-14.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Gries, S. Th. (2006). Some proposals towards more rigorous corpus linguistics. *Zeitschrift für Anglistik und Amerikanistik* 54: 191-202.

Gries, S. Th., B. Hampe & D. Schönefeld (2005). Converging evidence: bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16: 635-76.

Gries, S. Th., B. Hampe & D. Schönefeld (to appear). Converging evidence II: more on the association of verbs and constructions. In Newman, J. & S. Rice (eds.) *Empirical and experimental methods in cognitive/functional research*. Amsterdam / Philadelphia: John Benjamins.

Haiman, J. (1985). *Natural syntax*. Cambridge: Cambridge University Press.

Haspelmath, M. (2005). Frequency vs. iconicity in explaining grammatical asymmetries", Paper presented at the University of Jena. [Handout available http://email.eva.mpg.de/~haspelmt/Jena05.pdf, accessed 10 November 2006.]

Haspelmath, M. (2006a). Against markedness (and what to replace it with). *Journal of Linguistics* 42: 1-46.

Haspelmath, M. (2006b). Explaining alienability contrasts in adnominal possession: economy vs. iconicity. Paper presented at the 2nd Conference on the Syntax of the World's Languages, Lancaster University. [Handout available http://email.eva.mpg.de/~haspelmt/2006swl.pdf, accessed 10 November 2006.]

Heine, B. (1997). *Possession. Cognitive sources, forces, and grammaticalization*. Cambridge: Cambridge University Press.

Hoey, M. (2005). *Lexical priming: a new theory of words and language.* London: Routledge.

Hollmann, W. B. (2003). Synchrony and diachrony of English periphrastic causatives: a cognitive perspective. Ph.D. dissertation Manchester: University of Manchester.

Hollmann, W. B. & A. Siewierska (2006). Corpora and (the need for) other methods in a study of Lancashire dialect. *Zeitschrift für Anglistik und Amerikanistik* 54: 203-16.

Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In Christie, W. (ed.) *Current progress in historical linguistics*. Amsterdam: North Holland. 96-105.

Hopper, P. J. & E. C. Traugott (2003). *Grammaticalization*. 2nd edn. Cambridge: Cambridge University Press.

Jurafsky, D., A. Bell, E. Fosler-Lusier, C. Girana & W. Raymond (1998). Reduction of English function words in Switchboard. *Proceedings of ICSLP 98*: 3111-4.

Kortmann, B. (ed.) (2003). *Dialectology meets typology. Dialect grammar from a cross-linguistic perspective*. Berlin: Mouton de Gruyter.

Kortmann, B., T. Herrmann, L. Pietsch & S. Wagner (eds.) (2005). *A comparative grammar of British English dialects. Agreement, gender, relative clauses*. Berlin/New York: Mouton de Gruyter.

Langacker, R. W. (1987). *Foundations of cognitive grammar. Vol. I. Theoretical prerequisites*. Stanford: Stanford University Press.

Nichols, J. (1988). On alienable and inalienable possession. In Shipley, W. (ed.), *In honor of Mary Haas: from the Haas Festival Conference on Native American Linguistics*. Berlin: Mouton de Gruyter. 557-609.

Nichols, J. (1992). *Linguistic diversity in space and time*. Chicago / London: University of Chicago Press.

Nikolaevna, I. & M. Tolskaya (2001). *A grammar of Udihe*. Berlin: Mouton de Gruyter.

Oakes, M. P. (1998). *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.

Renouf, A. & J. Sinclair (1991). Collocational frameworks in English. In Aijmer, K. & B. Altenberg (eds.) *English corpus linguistics*. London / New York: Longman. 128-44.

Schuchardt, H. (1885). *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: R. Oppenheim.

Seiler, H. (1983). *Possession as an operational dimension of language.* Tübingen: Narr.

Siewierska, A. & W. B. Hollmann (2007). Ditransitive clauses in English with special reference to Lancashire dialect. In Hannay, M. & G. J. Steen (eds.) *Structural-functional studies in English grammar*. Amsterdam / Philadelphia: John Benjamins.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Stefanowitsch, A. & S. Th. Gries (2003). Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8: 209-43.

Stubbs, M. (1996). *Text and corpus analysis: Computer assisted studies of language and institutions*. Oxford: Blackwell.

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.

Ultan, R. (1978). Toward a typology of substantival possession. In Greenberg, J. (ed.) *Universals of human language. Vol. 4. Syntax*. Stanford: Stanford University Press. 11-49.

Zipf, G. K. (1935). *The psycho-biology of language*. Boston: Houghton Mifflin.

* We are grateful to two anonymous reviewers for their comments on this paper. All remaining errors are of course our own.

[1] The provenance of the Lancashire examples in this article is the North West Sound Archive (NWSA). The speakers' names are represented as initials. More details about our corpus will be given in section 3 of the main text.

[2] Another approach that sets out to go beyond raw frequencies is the collocational method, developed and used largely within British corpus linguistics (see e.g. Hoey, 2005; Renouf & Sinclair, 1991; Sinclair, 1991; Stubbs, 1996, 2001; for a useful overview of methods for calculating collocational strength see Oakes, 1998: 162-193). A recent addition to the corpus-based approach to associations between words and constructions in particular is the so-called collostructional method, pioneered by Stefanowitsch & Gries (2003). In using this method to assess the frequency of a word W in construction C not only is the raw frequency of this pattern considered, or just this frequency in combination with the frequency of W in all other constructions, but also the frequency of C with other words, and the total frequency of other constructions that do not contain W. These frequencies are entered into a table and some distributional test, preferably the Fisher exact test (see Stefanowitsch & Gries, 2003: 217-8 for the reasoning behind this), is then used to compute the measure of association of W in C. The present authors are impressed with the fact that at least in the case of one construction (the English *as*-predicative, see Gries, Hampe & Schönefeld, 2005, to appear) this method of analysis makes better predictions concerning the degree of association between the words in that construction and the construction itself than is possible on the basis of relative frequencies as defined by e.g. Haspelmath (2006b) and in the main text. However, there is one aspect of collostructional analysis that makes us unsure about its value in relation to our possessives data. Stefanowitsch & Gries (2003) and Gries et al. (2005) focus on a number of verbal constructions, where the association between different verbs and such constructions is calculated. In all cases the total number of constructions that do not include the verb in question is "approximate[d] by using the token frequency of all verbs" (Gries et al., 2005:645). The problem here is that it is not apparent to us what, in our case, constitutes the correct level of generalization in terms of the constructions to use: all possessed NPs? All definite NPs? All NPs with initial determiners? All common NPs? All NPs including pronominals? In fact it is also not fully clear to us exactly what in the analysis of verbal constructions it means to "approximate" the number of relevant constructions. There are often multiple verbs within phrases — should these indeed

all be counted or, depending on the construction, should some attempt be made to differentiate between main and auxiliary verbs? We would very much like to see the sophisticated triangulation method of Gries et al. (2005, to appear) being used to decide empirically between these slightly different ways of calculating collostructional strength.

[3] Unlike some other scholars, Nichols (1988) considers the alienable/inalienable opposition to be lexical rather than semantic.

[4] See http://www.gmcro.co.uk/other/NWSA/nwsa.htm (12 October 2005); cf. also Siewierska & Hollmann (2007).

[5] The BNC is a 100 million word corpus of spoken and written Present-day English; for more information see e.g. Aston & Burnard (1998).

[6] The table only gives the singular forms. If and where plural forms occurred in our data (as in the case of e.g. *auntie/aunties*), these have been included in the relevant BNC searches as well. In cases where the noun only occurred in the plural in our corpus (e.g. *grandparents*), this is the form we supply in the table. In these cases we restricted our search in the BNC accordingly.

[7] We only include nouns that immediately follow the possessive. Thus, for example, *birthday* — which in our corpus occurs as "my twenty-first birthday" (JAsh) and "my fourth birthday" (DG) are ignored. The reason for this decision is that here it is not clear whether the reduction (or full realisation) of the possessive is due to the noun or to the modifier(s). Some nouns (e.g. *big toe*, *right shoulder*) we have treated as compounds.

[8] There are two variant spellings in the BNC.

[9] See fn.8.

[10] This is a more common variant in Lancashire than the BNC figure might seem to suggest.

[11] See fn.8.

[12] See fn.8.

[13] *Grandparents* also has a low text frequency but all 5 examples in our corpus have the full possessive. The sole instances of *great grandfather*, *great aunt, nostrils* and *right shoulder* also co-occur with a full possessive.

[14] Many of the words in question occur not only as nouns but also as verbs, adjectives or adverbs. We have restricted our BNC search to nouns only, but it must be noted that in so doing we have relied entirely on the CLAWS tagging. A manual check would no doubt reveal some misclassifications.

[15] One of the anonymous reviewers suggests that "there is actually no need to invoke "construction grammar"" here", going on to point out that "[w]hat must be at work here is simply good old analogy, which nobody ever managed to explain away". We fully agree that analogy is at work, but we note that this notion has a central place in the versions of construction grammar associated with Langacker, Croft, Goldberg, and others. This is in contrast to many other linguistic theories, where instead of incorporating analogy, there is indeed a concern to "explain it away". Of course we could choose to invoke analogy without adopting construction grammar, but we feel that it is important to present the status of analogy in this theory as a merit.

	Kin	Body	Other
[maɪ]	137 (29%)	11 (28%)	224 (56%)
[mi]	246 (51%)	25 (63%)	118 (30%)
[ma]	83 (17%)	3 (7%)	51 (13%)
[mə]	16 (3%)	1 (2%)	4 (1%)
Total	482 (100%)	40 (100%)	397 (100%)

Table 1. Realisation of 1Sg possessive pronoun for kinship terms, body parts and other nouns.

Number of speakers who produce one of the reduced forms	Kin	Body	Other
12	+	+	+
10	+	Ø	+
1	+	+	-
1	+	Ø	-

Table 2. Number of speakers who produce one of the reduced forms and the nouns the combine them with

Number of speakers who consistently produce one of the reduced forms	Kin	Body	Other
2	+	+	-
4	+	Ø	-
1	-	+	-
1	+	+	+

Table 3. Number of speakers who consistently produce one of the reduced forms for any or all of the three noun types and the nouns the combine them with

Kin	absolute frequency	Body	absolute frequency	Other 1	absolute frequency
aunt(ie)	34	back	130	bag	77
brother	203	big toe	0	corner	1
child	95	bottom	16	football	1
cousin	60	eyes	89	money	112
dad	338	face	67	notice	2
daughter	133	feet	83	pension	20
family	86	finger	40	wage	20
father	378	fist	4		
father(-)in(-)law[8]	6	hand	208	Other 2	
gran(ny.nie)	27	head	171	base home	0
gran(d)dad[9]	23	knee	43	case	34
grandfather	38	leg	118	chief superintendent	0
grandma	11	milk tooth	0	day	47
grandmother	45	neck	46	desk	39
grandparents	7	nose	35	digs	0
great grandfather	8	nostrils	1	fault	86
great aunt	1	right shoulder	1	home	33
husband	210	shoulder	28	job	116
mam[10]	35			mate	50
mother	432			men	10
mum/mom[11]	463			opinion	120
niece	8			pet sayings	0
parents	95			shop	4
sister	176			street	6

sister(-)in(-)law[12]	8	study bedroom	0
son	143	wardrobe	10
stepfather	0		
stepmother	0		
stepsister	0		
uncle	38		
wife	205		

Table 4. Absolute frequencies of 1Sg possessive constructions with kinship terms, body parts and some other nouns in the spoken part of the BNC

Kin	relative frequency	Body	relative frequency	Other 1	relative frequency
aunt(ie)	10.29%	back	4.53%	bag	5.59%
brother	19.03%	big toe	n/a (4 tokens)	corner	0.10%
child	1.85%	bottom	1.26%	football	0.12%
cousin	32.26%	eyes	7.39%	money	1.70%
dad	12.08%	face	5.69%	notice	0.44%
daughter	20.72%	feet	4.74%	pension	2.13%
family	3.91%	finger	6.32%	wage	2.70%
father	24.31%	fist			
father-in-law	n/a (16 tokens)	hand	6.39%	Other 2	
gran(ny/nie)	n/a (27 tokens)	head	8.78%	base home	n/a (0 tokens)
gran(d)dad	n/a (23 tokens)	knee	14.79%	case	0.93%
grandfather	36.89%	leg	11.41%	chief superintendent	n/a (9 tokens)
grandma	3.79%	milk tooth	n/a (1 token)	day	0.43%
grandmother	50.56%	neck	13.37%	desk	12.11%
grandparents	20.59%	nose	8.66%	digs	n/a (29 tokens)
great grandfather	n/a (13 tokens)	nostrils	n/a (7 tokens)	fault	18.98%
great aunt	n/a (2 tokens)	right shoulder	n/a (6 tokens)	home	1.11%
husband	23.68%	shoulder	9.33%	job	2.26%
mam	25.18%			mate	8.85%
mother	21.20%			men	0.15%
mum/mom	10.61%			opinion	24.39%
niece	16.00%			pet sayings	n/a (0 tokens)

parents	6.67%	shop	0.15%
sister	21.86%	street	0.36%
sister(-)in(-)law	n/a (23 tokens)	study bedroom	n/a (0 tokens)
son	16.51%	wardrobe	5.08%
stepfather	n/a (2 tokens)		
stepmother	n/a (3 tokens)		
stepsister	n/a (0 tokens)		
uncle	15.02%		
wife	16.98%		

Table 5. Relative frequencies of 1Sg possessive constructions with kinship terms, body parts and some other nouns in the spoken part of the BNC.