

John Benjamins Publishing Company



This is a contribution from *Language Typology and Historical Contingency. In honor of Johanna Nichols.*

Edited by Balthasar Bickel, Lenore A. Grenoble, David A. Peterson and Alan Timberlake.

© 2013. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Marking versus indexing

Revisiting the Nichols marking-locus typology

Nicholas Evans¹ & Eva Fenwick²

¹Australian National University, and ²University of Melbourne

In this article, we build on Johanna Nichols's influential typological distinction between head and dependent marking by supplementing it with a further dimension, that of indexing. We focus on possessive constructions. Whereas marking indicates where morphology about a relationship goes, indexing indicates which syntactic entity the morphology provides information about. We show that, in principle, the two parameters of marking and indexing are logically independent and that for each marking type any indexing option is possible. Passing to an onto-probe of 138 languages, we find that most logically possible combinations are attested, apart from some cases in which the complexity of the combinations required may make instantiations so rare that their absences are likely to be accidental gaps.

1. Introduction: The Nichols marking-locus typology

Great typologists have the ability to see, and to operationalize, the crucial dimensions on which languages can vary. Johanna Nichols's (1986) typological parametrization of languages as head versus dependent marking is on a par with Greenberg's Word Order Typology as one of the most predictive dimensions of language variation – as shown by the frequency with which it crops up in brief typological characterizations of one language after another. It has also been comprehensively surveyed across the world's languages – see, for example, Nichols & Bickel (2005a, b, c) for assemblages of data.

In this paper, we pay tribute to the influence of her 1986 publication by developing a concept she introduced in that same paper but which has not, to our knowledge, been systematically typologized until now: the distribution of “indexing.” Briefly, if “marking” typologizes where inflectional information goes, “indexing” typologizes what that information is about.¹

1. Evans would like, more personally, to pay tribute to Johanna Nichols as a bold and visionary scholar, both for her early encouragement of his descriptive work on Kayardild and Bininj Gun-wok and for the influence that so many of her papers have had on his thinking

1.2 Phrase-level application

Since each element – head or dependent – can receive coding independently of the other, Nichols distinguished four logically possible patterns:

- (1) Types of marking strategy, after Nichols (1986)³
- | | | | |
|-----------------------|-----------|-----------|---------------|
| i. head-marking | N_{H-M} | N_D | HM |
| ii. dependent-marking | N_H | N_{D-M} | DM |
| iii. zero-marking | N_H | N_D | $\emptyset M$ |
| iv. double-marking | N_{H-M} | N_{D-M} | $2M$ |

This categorization abstracts away from constituent ordering, which is not shown here: thus both $N_{H-M} N_D$ and $N_D N_{H-M}$ (with element order meaningful this time) are equivalent for this purpose. It also abstracts away from the exact syntactic levels involved: for example, whether the dependent is a word or a phrase and, if the latter, the most appropriate level for representing it (NP, DP, etc.).

Note also that, within the NP, the semantic relation of head to dependent may be rather varied, even as the syntactic relation of adnominal modification is held constant. The semantic relationships include attributive adjective <> head (Japanese *kirei-na uchi* [beautiful-ATTR house] ‘beautiful house’), proprietive adnominal <> head (Kayardild *wangal-kuru dangka-a* [boomerang-PROP man-NOM] ‘man with a boomerang’), locative adnominals (*a bird in the hand is worth two in the bush*), and adnominal possessives (*woman’s skirt; portrait of a hero*).

It is possessives that form the lion’s share of relevant examples, in both grammatical descriptions and in the typologies that draw upon them.⁴ We continue this practice here, while noting that in principle the typology needs to be surveyed independently for these various types. Focusing on relationships of adnominal possessive modification, it will always be the case that, in possessive phrases, the head noun will be the possessed constituent, and the dependent noun will be the possessor constituent.

3. N=Noun; M=Marking, referring to all methods of marking, morphological or prosodic; H =head, D =dependent.

4. And it can be difficult to call the boundary between possessive and attributive interpretations: compare with the interpretations of Japanese ‘genitive’ clitic =*no* in *gakusei=no itoko* ‘student’s cousin’, *nihongo=no sensei* ‘Japanese teacher, teacher of Japanese’, *eki=no denwa* ‘station phone, phone at the station’, *shikan=no wakamono* ‘modern youth, young of recent times’, *takusan=no purezento* ‘many presents’ (first two examples from Martin 1987:742, others from Kaiser et al. 2001:322–324).

We can now exemplify each of Nichols's four types above, with relevant marking loci shown in bold.

- (2) a. *Sely* ***m-me***
 HM Sely 3SG.F.POSSR-mother
 'Sely's mother' (Maybrat: Dol 1999:93)
- b. *Amah* ***ro-Petrus***
 DM House GEN-Petrus
 'Petrus's house' (Maybrat: Dol 1999:9)
- c. *bíl* *ʔafu*
 ØM house door
 'door of the house' (Dhaasanac: Tosco 2001:255)
- d. *rolu-no* *biy-kûn*
 2M dog-3SG.POSSR man-GEN
 'the man's dog' (Dalabon: Evans field notes)

In many languages, more than one type is available, as already apparent from the alternative Maybrat structures in (2a)–(2b), and further exemplified by the alternate Kayardild structures in (3a)–(3b) which effectively encode a contrast between alienable and inalienable possession, respectively.

- (3) a. *dangka-karra* *wangalk* DM
 man-GEN boomerang.NOM
 'man's boomerang'
- b. *dangka-a* *jar-a* ØM
 man-NOM footprint-NOM
 'man's footprint'

As is well known, the exact semantic content of the alienability versus inalienability contrast varies cross-linguistically. For instance, in Maybrat some kin terms pattern grammatically as inalienables, while in Kayardild all kin terms follow the alienable pattern in requiring the genitive. Thus *dangka-a ngamathu* [man-NOM mother.NOM] cannot mean 'man's mother', which must be expressed with the genitive as *dangka-karra ngamathu* [man-GEN mother.NOM].

In this paper, we abstract away from the various semantic differences expressed by the choice of coding locus, grouping together all adnominal constructions coding possessive relationships in some sense. This means that our survey is counting constructions in languages, rather than languages per se, since the means for expressing possession and other forms of adnominal modification in a given language may split into several constructions.

In the examples so far, the dependents were nouns. But we can apply exactly the same typology to possessive relations in which the possessor is a pronominal element. Note, critically, that in Abkhaz (4a) and Marrku (4c) the pronoun is a generalized first-person pronoun, as these languages do not distinguish possessive pronoun uses from normal

argumental uses as subject or object – for example, Marrku *ɲaʔa ʔawuʔi nkawaʔ* [1SG speak:1SG:PRES language] ‘I speak Marrku’, *ɲirlaku ɲaʔa* [hear:2SG:PRES 1SG] ‘you hear me’. In these languages, the pronoun – that is, the dependent – receives no special marker of possession; in other words, there is no marking on the dependent pronoun in the possessive construction. In other languages, like South Efate and Dalabon (at least in the construction in (4d)), possessive pronouns have a distinctive form.

- (4) a. Abkhaz: HM
sarà sə-y°nə
 1SG 1SG.POSSR-house
 ‘my house’ (Hewitt 1979: 116)
- b. South Efate: DM
nawesien neu
 work 1SG.POSSR (≠ *kineu* ‘I, me’)
 ‘my work’ (Thieberger 2006: 128)
- c. Marrku: ØM
wurat ɲaʔa
 country 1SG
 ‘my country’
- d. Dalabon: 2M
ngey-kûn ngurl-ngan
 1SG-GEN heart-1SG.POSSR
 ‘my heart’

2. Marking versus indexing

We now clarify the difference between marking and indexing.

2.1 Indexing as an orthogonal dimension in phrasal relations

The Nichols typology focuses on where the marking is, not what it is about:

I will speak of morphological forms as MARKING the presence and type of dependency, but as INDEXING various grammatical and lexical properties of the head or dependent on the other. (Nichols 1986:58)

This raises the question of whether we need a separate typological dimension of indexing. The main contention of this article is that we do, and that – perhaps surprisingly – the distribution of indexing is in principle independent of the distribution of marking – that is (5):

- (5) WHERE indexing information is marked
is logically independent of
 WHAT it is ABOUT

Note the qualifier “logically independent.” The claim in (5) pertains to the ontology of the design space, not to the statistical distribution of types through it. Although the languages utilized in this paper do not constitute a well-designed cross-linguistic sample, we shall see in Section 4 that there appears to be a strong principle of complementarity – it is more likely that indexing will contain information about the other element than about its morphological host.

We have a clear heuristic reason for beginning by looking at what is logically independent, before looking at what statistical correlations hold. In typology, the first step is always to get our ontology clear before we can start examining statistical distributions – otherwise we can’t be confident of our base descriptive data, let alone our classificatory decisions in coding cross-linguistic material. We believe, therefore, that the best way to explore the design space is through an “onto-probe” – an opportunistic search through what is known, utilizing whatever leads will help us fill cells of the logical grid. It is only in the second stage of investigation – the checking for statistical distributions (something we don’t tackle here) that sampling issues become important. Despite this, we should reassure the reader that the material here is based on a fairly wide sample of 138 languages from over 35 different language families across six continents. Our sample incorporated any language for which we discovered relevant data, and our exploration included published and unpublished grammars, journal articles, posts on linguistic forums, and personal communication. See Appendix 1 of Fenwick (2007) for a listing of the original sample of 131 languages, since slightly expanded.⁵

2.2 Initial orienting example

It will be helpful to begin with a simple illustration using third-person singular possessive pronouns from some familiar West European languages: English, French, Spanish, and German. For simplicity, we will consider only gender at this stage because, once we include person and number as well, the patterning of these languages is not so clear.⁶ In each of the four languages, the possessive pronouns are distinct from other pronoun forms (*his, her* ≠ *he, she; son, sa* ≠ *il, elle* etc.), and it is in this sense of having a distinct possessive pronoun series that they are all dependent marking. However, the pattern of indexing is different – that is, the possessive pronoun can index information about the dependent (DI), the head (HI), neither the head or dependent (∅I), or both (2I):

5. Through the addition of Nen (PNG: Morehead-Maró), Ekagi (Trans-New Guinea; West Papua, Indonesia; Doble 1987) Marrku (Australian, non-Pama-Nyungan Isolate), and Dusó (Skou, PNG), plus Uruava and Torau (Oceanic, Bougainville Island; Evans & Palmer 2011; Rausch 1912) and Hausa (Chadic; Schuh 1983).

6. For example, once we take number into account, French indexes the number of the possessor as well (*leur* ‘their’) and Spanish the number of the possessed (*sus* ‘his’).

- (6) a. English (DI): indexes gender of possessor (dependent),
not possessed (head)
his son his daughter
her son her daughter
- b. French (HI): indexes gender of possessed (head), not possessor
(dependent)
son fils sa fille
'his/her son' 'his/her daughter'
- c. Spanish (ØI): indexes gender of neither
su hijo su hija
'his/her son' 'his/her daughter'
- d. German (2I): indexes gender of both (possessor through root;
possessed through suffix)
sein Sohn seine Tochter ihr Sohn ihre Tochter
'his son' 'his daughter' 'her son' 'her daughter'

All four languages are dependent marking. Yet each behaves differently in what it indexes – all four possibilities about what is indexed are exemplified within the “dependent-marking cell” of the original Nichols typology.

Strictly speaking, “indexes X” should be replaced with “indexes X for feature α .” In other words, for every inflectional feature in the language, we can ask how its indexing patterns. (For example, our statement that Spanish is zero indexing applies only to gender – were we to pluralize the head to *hijos/hijas* ‘sons/daughters’, we could show that number is being indexed on the dependent – *su hijo/hija* ‘his son/daughter’ but *sus hijos* ‘his sons/daughters’). This is why the above examples, although included because of their pedagogical accessibility, are not all perfect examples of the various possibilities, since a statement that something is not indexed should really mean that it does not index any feature of the relevant constituent. Conversely, in describing relevant phenomena, we should strictly relativize our statements to the feature involved: for example, “head-indexing w.r.t. gender” and so on. In what follows, though, we will draw on whichever inflectional dimensions are helpful in illustrating the pattern under discussion.

2.3 Information in modified NP constructions

In a NP-modifying construction, the marking can potentially contain information about

- i. the relation itself (e.g. possession [or subtypes thereof]; appropriation (‘having’); attribution etc.
- ii. the dependent NP (gender, number, animacy, definiteness, etc.)
- iii. the head NP (gender, number, animacy, definiteness, etc.)

It is information of types (ii) and (iii) that we shall refer to as “indexing,” whereas (i) by itself will be termed “(pure) relational marking.”

2.4 Zero indexing: A Kayardild example

In the simplest case of adnominal modification, the information is confined to (i) – that is, just the nature of the relation – without any further information about either element.⁷ Kayardild (Evans 1995a) is a good example: the genitive suffix is invariant regardless of any aspects of the possessor or the possessed (7a)–(7b).

- (7) a. *maku-karra mijil-d*
 woman-GEN net-NOM
 ‘the woman’s net’
- b. *dangka-karra mijil-d*
 person-GEN net-NOM
 ‘person’s net’

We will say such languages are dependent marking but zero indexing: that is, the marker of the possessive relationship indexes no information about either constituent.

2.5 Double indexing: A Beja example

At the other extreme, the relevant morphology can include information of all three types. Beja (Cushitic; Sudan and Eritrea) is an exuberant example (8).

- (8) *te-mʔa-t-e:-t-a:-kn*
 DEF.F-WOMEN(F).PL-PL-F.POSSR-(GEN)PL.POSSR.F.POSSD-
 F.POSSD-YOUR.PL.NOM-YOUR-YOUR.PL
ʔar dauri:=ta
 daughter(F).PL pretty=PRED.3PL.F
 ‘Your (PL) women’s daughters are pretty’ (Roper 1927: 32)

This is an extreme example of a double-indexing language: the suffix substring *t-e:-tʔ*⁸ includes information about (1) the existence of a possessive relationship (essentially encoded by the presence of suffixal material); (2) specific information about the

7. Of course in languages in which possession is shown just by juxtaposition, as in Indonesian *buku saya* [book 1SG ‘my book’], it is the constructional template itself rather than any morphological sign that encodes (i).

8. There are of course many possible morphological analyses of Beja morphology, and paradigm-based or realizational accounts would avoid many of the morpheme breaks we posit here. The somewhat laboured “item-and-arrangement” divisions we make – based on the principle of postulating a segmented morpheme whenever that slot contrasts with another form of different semantic value – are useful in making explicit how the complex cumulation of indexing is built up.

possessor being (a) feminine (the first *t*), and (b) plural (part of the meaning of *e*); (3) specific information about the possessed element being feminine (part of the meaning of *e*; and the second *t*) – all of this is then suffixed by adding *a:kn*⁹ ‘your (PL)’ as a recursive possessor of the whole.

The logic behind the sequence *te:t* in (8) can be seen in Table 1, which shows that these genitive strings contain three parts:

- a. gender of possessor (M \emptyset versus F *t*)
- b. genitive relationship, plus number of possessor and gender of possessed (PL.POSSR and F.POSSD *e*., versus *i* elsewhere)
- c. gender of possessed (M \emptyset versus F *t*)

Table 1. Number and gender combinations for dependent (possessor) noun suffixes in Beja

Possessor noun	Possessed noun	
	Masculine	Feminine
M.SG	<i>-i</i>	<i>-i:t</i> (9a)
M.PL		<i>-e:t</i> (9b)
F.SG	<i>-ti</i>	<i>-ti:t</i> (9c)
F.PL		<i>-te:t</i> (8)

The feminine plural *-te:t* was already exemplified in (8); (9a)–(9c) below illustrate the other combinations with feminine possessed nouns.

- (9) a. *i-ka:m- \emptyset -i:-t*
 DEF.M-camel. SG-M.POSSR-GEN.SG.POSSR-F.POSSD
sikwena nʔaukw=ti
 foot(F) soft=PRED.3SG.F
 ‘The foot of a camel is soft’ (lit. ‘the camel’s foot is soft’) (Roper 1927: 19)
- b. *yi-ʔar- \emptyset -e:-t* *mak*
 DEF.M.PL-boy(M).PL-M.POSSR-
 (GEN)PL.POSSR.F.POSSD-F.POSSD donkey(F).PL
 ‘the boys’ she-donkeys’ (Roper 1927: 14)
- c. *ʔo(r)-t-i:-t* *de*
 girl(F).SG-F.POSSR-(GEN)SG.POSSR-F.POSSD mother(F)
 ‘a girl’s mother’ (Roper 1927: 14)

9. Itself segmentable into *a-* ‘plural, nominative possessor’, *k-* ‘2nd-person possessor’ and *-n* ‘plural possessor’, though this is irrelevant here.

Thus Beja, like Kayardild, is dependent marking, but, unlike Kayardild, it is double indexing: the marker of possession indexes gender and number information about the possessor and gender information about the possessed. It is crucial to our argument that this is information that only appears in genitive suffixes. In other words, it's not the case that nominative or accusative nouns have affixes giving information about gender – see *sikwena* 'foot', *mak* 'donkeys', and *de* 'mother' in the above examples.¹⁰ Throughout this article, we only count the information as relevant to our indexing typology if it just appears in the possessive construction.

3. Indexing and marking: A more systematic typology

If marking and indexing really are orthogonal, we would expect to find 25 logical possibilities in the design space, made up as follows.

- i. $\emptyset M$, no indexing.¹¹ This has already been illustrated above by Dhaasanac (3) and Kayardild (7b).
- ii. DM , with four indexing possibilities ($\emptyset I$, DI , HI , $2I$)
- iii. HM , with four indexing possibilities ($\emptyset I$, DI , HI , $2I$)
- iv. $2M$, with 16 indexing possibilities – i.e. 4 indexing possibilities on each constituent, multiplied together

This gives a total of 25 logical possibilities: $1 + 4 + 4 + 16$. Below we illustrate these various possibilities systematically, where possible: we discuss the indexing types found with dependent-marking constructions in Section 3.1, those with head-marking constructions in Section 3.2 and those with double-marking constructions in Section 3.3. Once we get to the double-marking types, there are a number of gaps, at least on the basis of data currently known to us.

3.1 Indexing in dependent-marking structures

Initial exemplification was given above using possessive pronouns from Western European languages, but this was a less than perfect example given that it depended on holding some inflectional values constant. Clearer cases are:

10. Gender and number are indexed by the prefixed "article" in all cases, but this is optional, depending on definiteness. A few other cases exhibit a rather disorganized subset of some of the suffixal morphology exhibited here – see Fenwick (2007:29) for discussion – but the resemblances are too weak to invalidate the condition formulated here.

11. Since there is no marking of the possessive relationship, there is no place for any indexing to be encoded, which is why there is only this one possibility for the zero-marking cell.

DM, ØI: Kayardild in (3a) (i.e. pure genitive suffix without indexing of either constituent). Within our sample, 24 languages exemplify this construction type.

DM, 2I: This has already been illustrated with Beja ((8), (9a)–(9c)): a genitive suffix complex indexes the possessive relationship itself but also other information about the gender/number of both constituents. Four languages in our sample have structures of this type: Beja, Kashmiri, Bagwalal, and Upper Sorbian.

This leaves two gaps – DM, DI and DM, HI – which are exemplified in the next two subsections.

3.1.1 DM, DI

This can be illustrated by Somali, in which genitive suffixes index the gender and number of the possessor (Saeed 1999: 64), even though the noun in isolation does not show gender overtly.

- (10) a. *Áf* *shimbir-éed*
 mouth(M).SG bird(F).SG-F.SG.POSSR
 ‘bird’s mouth (beak)’
- b. *hádal* *naagó-od*
 talk(M)SG women:PL-M.PL.POSSR¹²
 ‘women’s talk’

Seven languages in our sample exemplify this pattern.

3.1.2 DM, HI

This can be illustrated by Awngi, a Cushitic language from Ethiopia (Hetzron 1995), in which the suffix on the possessor noun indexes the gender and number of the head in addition to the possessive relationship itself:

- (11) a. *Áf* *shimbir-éed*
 mouth(M).SG bird(F).SG-F.SG.POSSR
 ‘bird’s mouth (beak)’
- b. *muri-t* *ɣuna*
 village-F.POSSD woman
 ‘woman of the village’
- c. *muri-k^w* *aq(ká) / ɣunayúná*
 village- PL.POSSD man:PL woman:PL
 ‘men/women of the village’

12. Somali exhibits “polarity”: plurals of feminine nouns are marked as masculine and vice versa; this is a regular rule independent of the construction being examined here.

Clear examples of this pattern are found in five languages in our sample. There are interesting borderline cases in some additional languages like Torau and Uruava (Rausch 1912:976; Evans & Palmer 2011) in which possessum indexing is found in possessive classifiers preceding the possessed noun.¹³

Concluding this section, we see that all four indexing possibilities with dependent-marking constructions have been exemplified.

3.2 Indexing in head-marking structures

For head-marking structures, we have so far exemplified just one of the four possibilities (see (2a), (4a)), namely the HM, DI pattern where the head is marked with a morpheme indexing person, number, or gender features of the dependent/possessor. Our sample contains 23 examples of this pattern.

We now turn to the other three possibilities.

3.2.1 HM, ØI

Here the head is marked for the relationship alone but not indexed for any features from either constituent,¹⁴ as in Yoruba (12), in which the word for the possessed noun ‘money’ has its vowel lengthened. Such vowel lengthening simply indicates that the noun is possessed by some other entity. Eleven languages in our sample exhibit this pattern.

- (12) *owó-: Dàda*
 money-POSSD Dada
 ‘Dada’s money’ (Awobuluyi 1978:40, cited in Nichols & Bickel 2005a: 102)

Koyukon (Thompson 1996:655) is an interesting variant. For most person/number values of the possessor, the head noun is affixed with a prefix showing possessor

13. Cf. Uruava *egu maru* ‘my wife’, *gogi buri maru* ‘my wives’ (*buri* is a pluralizer with *maru* ‘wife’). The difficulty here is that possessive classifiers are only found in possessive constructions, not in other syntactic positions, so it is difficult to argue that they are the “possessive forms” of any independently existing element.

14. This is close to the situation with so-called “construct state” nouns in many Semitic languages, such as Hebrew *d̄bar hammelek* ‘the king’s word’, in which *d̄bar* is the “construct state” of the noun *d̄bār* ‘word’. Construct state nouns, though, also have additional complicating features or domains of use (e.g. in other types of modification) and it is not the case in all languages with a construct state that there is no marking on the possessor. Perhaps more crucially, “construct state” refers specifically to just the head noun part of the construction, without its being definitionally required that the dependent noun behave in a certain way. Nonetheless, our considerations above suggest that a typologically general definition of what “construct state” would mean is a marked form of a noun that codes it as the possessed element of a possessive construction, without indexing any features of the possessed or possessor nouns.

person/number, plus a suffix *-e'* showing simply that the noun is possessed (13a). However, when the possessor is third-person singular, there is no overt prefix showing person/number categories, but the suffix marking the noun as possessed remains (13b)–(13c).¹⁵ A further complication is that this suffix is only found with alienably possessed nouns – inalienably possessed nouns just take a prefix showing the person and number of the possessor, with no suffix (14). So the best analysis of the *-e'* suffix is that it shows that the noun it suffixes to is alienably possessed without indexing any possessor or possessed features (so: head-marked but zero-indexing), while the prefix indexes the person/number features of the possessor.

- (13) a. *se-tel-e'*
1SG.POSSR-SOCKS-ALIENABLY.POSSD
'my socks'
- b. *Dick ghudl-e'*
Dick sled-ALIENABLY.POSSD
'Dick's sled'
- c. *Dick leeg-e'*
Dick dog-ALIENABLY.POSSD
'Dick's dog'
- (14) *se-tlee'*
1SGPOSSR-head
'my head'

3.2.2 HM, 2I

[N=4]. Here, the head is marked for its participation in a possessive relationship and at the same time indexes information about both entities. This can be exemplified by one Hungarian method for coding possession: in (15a) the suffix *-a* on the head noun *ablak* 'window' encodes (a) the fact that the possessor is third person, and (b) the fact that the possessed noun is singular; if the suffix *-ai* is used instead as in (15b), this now encodes the fact that the possessed noun is plural. Note that this is a distinct system from the normal system of expressing plurality on nouns – the regular (nominative) plural of 'window' is *ablakok*.

- (15) a. *A szoba ablak-a*
DEF room window-3POSSR.SG.POSSD
'the window of the room'

15. When there is no lexical noun possessor, a pronominal prefix for possessor is used, e.g. *be-ghudl-e'* 'his sled' (cf. Thompson 1996:655).

- b. *A szoba ablak-ai*
 DEF room window-3POSSR.PL.POSSD
 ‘the windows of the room’
 (data from Daniel Abondolo, p.c.)

3.2.3 *HM, HI*

[*N=4*]. This pattern, in which head marking indexes both the possessive relationship and information about the head noun, is found in a number of Chadic languages (Schuh 1983), including Hausa (16a)–(16b). Here, the suffixes *-an* and *-aĩ* encode that the possessed noun is masculine and feminine, respectively – compare with the citation forms *sanda* ‘stick’ (masculine) and *goora* ‘cane’ (feminine):¹⁶

- (16) a. *sanda-an* *makaafoo*
 stick(M)-M.POSSD blind.man
 ‘a/the blind man’s stick’
 b. *goora-aĩ* *makaafoo*
 cane(F)-F.POSSD blind.man
 ‘blind man’s cane’

3.2.4 *Summary*

For this set, again, all four cells in the possibility space are attested:

HM, ØI: Yoruba, Koyukon, (construct state in some Semitic) (*N=11*)

HM, DI: Maybrat, Abkhaz, Dalabon, etc. (*N=23*)

HM, 2I: Hungarian (one construction) (*N=4*)

HM, HI: Hausa and various other Chadic (*N=4*)¹⁷

16. In the same paper, Schuh outlines an interesting hypothesis about how these constructions arose diachronically. In the original situation, the possessum was followed by a gender-agreeing determiner (with a showing agreement features on the determiner which are governed by the head noun N_1):

[$N(\alpha)_1$ Det(α)] N_2 for ‘ N_2 ’s N_1 ’, i.e. a/the N_1 of that N_2

This was then reanalyzed from a construction in which the possessive relation was simply marked by juxtaposition (although with the possessum frequently being accompanied by a following determiner) to one in which the erstwhile determiner itself becomes the sign of possession. The erstwhile determiner would then have got suffixed to the head noun and lost its determiner status while still indexing the gender of the head.

17. See Schuh (1983) for further references to other Chadic languages possessing this construction.

3.3 Indexing in double-marking structures

We now turn to the most complicated case, that in which each element is marked. Since both elements are marked, and each can exhibit all four of the indexing possibilities available to marked dependents and marked heads, respectively, there are sixteen logical possibilities (see Table 2). Not surprisingly, given the greater complexity and rarity of double-marking constructions, most cells ($N=10$) are unattested so far (shown by \square in the table). Only cells for which we are aware of exemplars will be illustrated in the sections that follow.

Table 2. The sixteen indexing possibilities for double-marking structures

Information encoded on head	Information encoded on dependent			
	$\emptyset I$	DI	HI	2I
$\emptyset I$	Warlpiri kin possession construction ($N:1$)	\square	\square	\square
DI	Dalabon, Turkish ($N:11$) ¹⁸	Erzya ($N:1$)	\square	\square
HI	\square	\square	\square	\square
2I	Hungarian ($N:3$) ¹⁹	Tundra Nenets ($N:1$)	Komi ($N:1$)	\square

3.3.1 Double marking: $\emptyset I$ on both elements

In this type, each element is marked for participation in the construction – the possessor by a genitive-type marker, and the possessed by a marker indicating that a possessor is present – but neither element indexes any features (e.g. person/number/gender) of either. The sole example we have located so far is the following construction, which can be used to represent kinship possession in Warlpiri:

- (17) a. *Nakamarra-ku ngati-nyanu*
 Nakamarra-DAT mother-POSSD
 ‘Nakamarra’s mother’
- b. *ngaju-ku ngati-nyanu*
 1SG-DAT mother-POSSD
 ‘my mother’

18. Other languages include Huallaga Quechua, Tundra Nenets, Aramaic, Southern Sierra Miwok, Mangarrayi, Jivaro, Komi (Head=Def and Subj), Cheremis (Head=Subj), and Chagatay.

19. Other languages include Cheremis (Head=DirObj) and Tundra Nenets.

3.3.2 *Double marking: DI on head, ØI on dependent*

This type is common. We already presented a Dalabon example in (2d) above, and (18) from Turkish is another. The dependent bears marking just for the relation, without indexing, whereas the head indexes person/number information about the dependent.

- (18) *ev-in kapı-sı*
 house-GEN door-3SG.POSSR
 ‘the door of the house’ (Nichols 1986:65)

3.3.3 *Double-marking: DI on both head and dependent*

This type appears rare: our only example so far is Erzya (Nikolaeva 2002), and even then we have to stretch the indexable feature set to include definiteness.²⁰ The genitive suffix on the dependent noun marks this noun as the possessor and also marks it as definite singular;²¹ the possessive suffix on the head indexes the person and number of the possessor, as in (19).

- (19) *ćora-ńt’ńejavšt’ ašo peje-nze*
 boy-DEF.SG.POSSR saw. REFL.3PL white tooth-3SG.POSSR
 ‘One could see the boy’s white teeth’ (Nikolaeva 2002:6)

3.3.4 *Double-marking: 2I on head, ØI on dependent*

This type is less rare: it combines a genitive-type marker on the possessor, which just indexes the possessive relationship, with head marking that indexes information about both the possessor/dependent and the head, as in (20). Hungarian, again, is an example:²² the head marking indexes person information about the possessor, and number information about the possessed (Rounds 2001:151). (This construction is different, compared with the other Hungarian possessive construction considered in Section 3.2.2, in which the possessor is unmarked, staying in the nominative.) Note here that

20. No other examples of this structure or other genitive forms in Erzya were available from the source text, and therefore this language entry must be taken as tentative at this stage. As mentioned, this example is already unusual in that the relevant indexed feature is definiteness, and since there is a strong correlation between possessed status and definiteness (indicated, for example, by the ongoing reanalysis of *-nya* from possessed marker to definite marker in Indonesian), it is not entirely clear that the definiteness should be imputed to the possessor here, rather than falling out from the possessive nature of the whole construction.

21. The indefinite/definite plural genitive form is *-ń* (Rueter 2010:77).

22. Other marginal examples are Cheremis and Tundra Nenets, although in each instance the indexing on the head is for case and can be better treated as a feature assigned to the whole phrasal constituent rather than to the head in particular. See Fenwick (2007:199–200) for the analysis and Kangasmaa-Minn (1966, 1969), Nikolaeva (2002:5), and Nikolaeva (2005) for the relevant data.

- i. as with the example in Section 3.2.2, the number marking for the possessed/head is separate from the regular method of marking number, and
- ii. the possessive relation is marked with the dative case.

- (20) a. *a fiú-nak a könyv-e*
 DEF boy-DAT DEF book-3 POSSR.SG.POSSD
 ‘the boy’s book’
- b. *a fiú-nak a könyv-ei*
 DEF boy-DAT DEF book-3 POSSR.PL.POSSD
 ‘the boy’s books’
- c. *a fiú-k-nak a könyv-e*
 DEF boy-PL-DAT DEF book-3POSSR.SG.POSSD
 ‘the boys’ book’

3.3.5 Double marking: 2I on head, DI on dependent

Again, this structure is rare. Our only Example (21) is from Tundra Nenets, one among several structures for expressing possession in this language.

- (21) *narey° xob°-q*
 spring skin-PL.POSSR
- meyowa-ko-ryi-doh* *teraø-d°m*
 hardness-DIM-LIM-3PL.POSSR-PL.ACC.POSSD chose-1SG
- ‘I’ve only chosen the firmest of the spring skins’ (Nikolaeva 2005: 543)

Here, the genitive marking *-q* on the dependent noun ‘skin’ indexes the plural number of the dependent noun. Had the dependent noun been singular, the genitive form would have been *-h*. The possessive suffix *-doh* indexes both the person and number of the possessor, and the number and case of the possessed noun.

3.3.6 Double marking, 2I on head, HI on dependent

So far this structure is attested only in Permic languages such as Komi, Permiak, and Udmurt in the particular situation in which a possessed noun functions as the direct object of the clause. (22a)–(22b). In such cases, instead of the usual genitive marker *-lën*, which indexes information only about the possessive relation itself, the dependent noun takes the genitive/ablative marker *-lyś*, which indexes the fact that the possessed noun is the direct object of the clause. In addition, use of the possessive suffix *-së* on the head noun (instead of the usual possessive suffix *-ys*) indexes the person and number of the dependent noun, the fact that the dependent noun is in the accusative case, and also that the possessed noun is a definite direct object (data from Daniel Abondolo, p.c.).

- (22) a. *vok-ys-lys*
 brother-NACC.3SG.POSSR-ACC.POSSD
jaj-së
 flesh/body-ACC.3SG.POSSR.DEF.ACC.POSSD
 ‘(Took a needle and pierced) her brother’s flesh/body’ (Komi)
- (22) b. *orët-i-s* *žmej-liš*
 sever-PST-3SG.DEF dragon-ACC.POSSD
kyl-le-së
 tongue-PL-ACC.3SG.POSSR.DEF.ACC.POSSD
 ‘He cut off the dragon’s tongues’ (Northern Permiak)

4. Conclusions

The data considered in this paper show the need to treat indexing and marking as separate dimensions in typologies of possessive structures – and, ultimately, in NP modification structures more generally.

Elaborating the original four-way Nichols typology by adding in different types of indexing, as argued for in this paper, produces a typological space with 25 logical possibilities, as indicated in Section 3. At this preliminary stage of investigation, based on a sample of more than 135 languages, not all have yet been filled by attested languages. However, for each of the three simplest marking types (that is, types in which no more than one element is marked), all indexing possibilities are attested.

It is only in the most complex marking type (double marking) that we begin to encounter empty cells. Moreover, with 60 percent of cells attested so far, it appears likely that many, perhaps even all, of the gaps would be filled in a much larger language sample. We do not believe there is any in-principle reason for why any cell here would be unlearnable, though some of those for which gaps are found would involve great processing and learning complexity. In other words, we think that these gaps are likely to reflect a ‘diachronic filter’ (Evans 1995b), making some evolutionary pathways highly unlikely, rather than any absolute prohibition on them being possible language structures.

In this paper we have simplified our exposition by speaking of “indexing” in a generalized way. Greater precision could be gained by specifying the relevant inflectional category specified: for example, indexed for gender/number, and so on. Obviously every cell of the design space can then be elaborated according to the inflectional categories specified, an interesting exercise for future research.

Even with these simplifying assumptions, and the relatively small set of languages used in our initial sample, we see very uneven distributions of types across the design space and several instances of areal trends (e.g. Uralic languages and 2I, Chadic languages and HM, HI). Moreover, since our study was an onto-probe aimed at turning up particular cells through queries, it is likely to have inflated the incidence of more

uncommon types. Now that the basic ontology is established, it will be possible to investigate statistical distributions more thoroughly. In the meantime, it is worth commenting on what appear to be some initial general trends:

- i. dependent indexing [39] > zero-indexing [36]²³ > head-indexing [25]
- ii. double indexing is rare overall [13]
- iii. there is a strong trend towards complementarity: in other words, for the indexed information to be about the other constituent to the one it is marked on. That is, if indexing is on the dependent, and about just one element, it is likely to be about the head, and vice versa. The figures are as follows: 23 HM:DI + 5 DM:HI = 28, as opposed to 4 HM:HI + 7 DM:DI = 11.

In terms of historical linguistics, each cell considered here calls for its own investigation in terms of possible historical pathways, as do the pathways between them (i.e. how does a language get from one cell to another). At our present state of knowledge, we have little understanding of most of these, particularly for the very complex case represented by Beja. It would also be interesting to survey whether there are correlations between indexing type and particular semantic distinctions (e.g. alienable versus inalienable possession) in languages permitting more than one construction.

Nichols's 1986 article, by crystallizing the concept of marking-locus typology, immediately made the characterization of a whole range of cross-linguistic phenomena more straightforward. We hope that the elaboration of her typology sketched here will give a useful framework for clarifying an additional and often tricky dimension of descriptive practice and stimulate the search for examples of the cells for which we have so far been unable to find exemplars. Finally, the additional dimension which we have introduced to her typology should open up a number of lines of future research. These include a more detailed analysis of the features indexed, the statistical distribution of the various types, their functional interactions, and the historical origins of each construction type.

Abbreviations

POSSD	Possessed
POSSR	Possessor
PL.POSSR	Plural Possessor
SG.POSSR	Singular Possessor

Other features of the possessor and/or possessed are written immediately before: for example, F.POSSR 'feminine possessor'.

23. Double-marking constructions are only counted as zero indexing if there is indexing on neither head nor dependent.

References

(Examples for Kayardild, Dalabon, Marrku, and Nen are from Evans's own field notes; examples can be found in Evans (1995a) for Kayardild; Evans et al. (2004) for Dalabon; and Evans et al. (2006) for Marrku. Examples in most Australian languages cited employ current practical orthographies, in which digraphs of the form *rX* are used for retroflexion (e.g. *rd* for [d], *rr* for an alveolar tap/trill, and *ng* and *ny* or *nj* for η and η respectively. Dalabon *û* represents a high, unrounded central vowel, close to [i], though realized by some speakers as [u].)

- Awobuluyi, A. Oladele. 1978. *Essentials of Yoruba grammar*. Ibadan: Oxford University Press Nigeria.
- Doble, Marion. 1987. A description of some features of Ekari language structure. *Oceanic Linguistics* 26. 55–113.
- Dol, Philomena. 1999. *A grammar of Maybrat: A language of the Bird's Head, Irian Jaya, Indonesia*. University of Leiden.
- Evans, Bethwyn & Bill Palmer. 2011. Contact-induced change in Southern Bougainville. *Oceanic Linguistics* 50(2). 489–529.
- Evans, Nicholas. 1995a. *A grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Evans, Nicholas. 1995b. Multiple case in Kayardild: Anti- iconicity and the diachronic filter. In Frans Plank (ed.), *Double case: Agreement by Suffixaufnahme*, 396–428. Oxford: Oxford University Press.
- Evans, Nicholas, Francesca Merlan & Maggie Tukumba. 2004. *A first dictionary of Dalabon (Ngalkbon)*. Maningrida: Bawinanga Aboriginal Corporation.
- Evans, Nicholas, Joy Williams Malwagag & Khaki Marrala. 2006. *Majila Inkawart*. Jabiru: Iwaidja Inyman.
- Fenwick, Eva. 2007. Head and dependent indexing within the phrase: Towards a typology, with special reference to Beja. Melbourne: Department of Linguistics & Applied Linguistics, University of Melbourne unpublished MA thesis.
- Hetzron, Robert. 1995. Genitival agreement in Awngi: Variation on an Afroasiatic theme. In Frans Plank (ed.), *Double case: Agreement by Suffixaufnahme*, 325–335. New York, NY: Oxford University Press.
- Hewitt, George. 1979. *Abkhaz*. Amsterdam: North-Holland.
- Kaiser, Stefan, Yasuko Ichikawa, Noriko Kobayashi & Hilofumi Yamamoto. 2001. *Japanese: A comprehensive reference grammar*. London: Routledge.
- Kangasmaa-Minn, Eeva. 1966, 1969. *The syntactical distribution of the Cheremis genitive*, vols. 1, 2. Helsinki: Suomalais-ugrilainen Suera.
- Martin, Samuel E. 1987. *A reference grammar of Japanese*. Tokyo: Charles E. Tuttle.
- Nichols, Johanna. 1986. Head-marking and dependent-marking Grammar. *Language* (62)1. 59–119.
- Nichols, Johanna & Balthasar Bickel. 2005a. Locus of marking in possessive noun phrases. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 102–105. Oxford: Oxford University Press.
- Nichols, Johanna & Balthasar Bickel. 2005b. Locus of marking: Whole-language typology. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 106–109. Oxford: Oxford University Press.

- Nichols, Johanna & Balthasar Bickel. 2005c. Possessive classification. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 242–245. Oxford: Oxford University Press.
- Nikolaeva, Irina. 2002. The Hungarian external possessor in a European perspective. In Cornelius Hasselblatt & Roger Blokland (eds.), *Finno-Ugrians and Indo-Europeans: Linguistic and literary contacts – Proceedings of the Symposium at the University of Groningen, November 22–24, 2001*, 272–285. Maastricht: Shaker.
- Nikolaeva, Irina. 2005. Agreement and situation construal. In M. M. Jocelyn Fernandez-Vest (ed.), *The Uralic languages today. A linguistics and cognitive approach*, 533–546. Paris: Librairie Honoré Champion, Editeur.
- Rausch, P.J. 1912. Die Sprache von Südost-Bougainville, Deutsche Salomoninseln. *Anthropos* 7: 106–134, 585–616, 964–994, 1056–1057.
- Roper, E.M. 1927. *Tu Beḍawi: An elementary handbook for the use of Sudan government officials*. Hertford: Stephen Austin & Sons.
- Rausch, P.J. 1912. Die Sprache von Südost-Bougainville, Deutsche Salomoninseln. *Anthropos* 7: 106–134, 585–616, 964–994, 1056–1057.
- Roper, E.M. 1927. *Tu Beḍawi: An elementary handbook for the use of Sudan government officials*. Hertford: Stephen Austin & Sons.
- Saeed, John. 1999. *Somali*. Amsterdam: John Benjamins.
- Schuh, Russell G. 1983. The evolution of determiners in Chadic. In Ekkehard Wolff & Hilke Meyer-Bahlburg (eds.), *Studies in Chadic and Afroasiatic linguistics*, 157–210. Hamburg: H. Buske Verlag.
- Thieberger, Nicholas. 2006. *A grammar of South Efate: An Oceanic language of Vanuatu*. Honolulu, HI: University of Hawai'i Press.
- Thompson, Chad. 1996. On the grammar of body parts in Koyukon Athabaskan. In Hilary Chappell & William McGregor (eds.), *The grammar of inalienability: A typological perspective on body part terms and the part-whole relation*, 651–676. Berlin: Mouton de Gruyter.
- Tosco, Mauro. 2001. *The Dhaasanac language*. Cologne: Rüdiger Köppe Verlag.