RESEARCH

Universals in possessive morphology

Yvonne van Baal¹ and Jan Don²

¹ University of Oslo, Postboks 1102, Blindern, 0317 Oslo, NO

² University of Amsterdam, Spuistraat 134, 1012 VB, Amsterdam, NL

Corresponding author: Jan Don (j.don@uva.nl)

In this paper, we test the hypothesis that possessive pronouns have the same basic structure containing the genitive pronoun, plus, in some languages, some extra structure, as suggested by Caha (2009). In order to unravel the structure of these pronouns, we use the same logic applied by Caha (2009) and Bobaljik (2012) that excludes so-called ABA-patterns. If possessive pronouns are built on top of the genitive, we derive several predictions.

First, we predict that there are languages in which the possessive pronoun comprises the genitive pronoun plus an extra affix (complex morphology). Furthermore, we predict that there are no possessive pronouns that have the same form as the accusative, or the nominative pronoun, to the exclusion of the genitive (*ABA). And thirdly, we expect that any syncretisms between possessives and other pronominal forms respect the proposed hierarchy in the sense that only structurally adjacent forms may be syncretic.

Our data provide ample evidence for the claim that possessive pronouns are "bigger" structures than the accusative or ergative pronouns, suggesting that the possessives are indeed constructed from these structures. However, the data in our sample do not give crucial evidence for the claim that the possessives are more complex than the genitive. The data leave open the possibility that the genitive is in fact "bigger" than the possessive.

Only in a few languages do we find ABA-patterns. We argue that these ABA-patterns are only apparent counterexamples to the proposed structure. Therefore, we conclude that there is broad typological evidence for the hypothesis that possessives are built from pronouns expressing a dependent (accusative/ergative) case.

Keywords: *ABA; possessive pronouns; case-hierarchy; suppletion; syncretism; typology

1 Introduction

Consider the pronominal paradigm in Dutch (1), where shaded cells indicate syncretisms within the columns. We will use this style of tables throughout the paper.

	1sg	2sg	1pl	2pl
NOM	ik [ɪk]	jij [jɛi]	wij [vɛi]	jullie [jœli]
ACC	mij [mɛi]	jou [jɔu]	ons [ɔns]	jullie [jœli]
POSS	mijn [mɛin]	jouw [jɔu]	ons [ɔns]	jullie [jœli]
DAT	mij [mɛi]	jou [jɔu]	ons [ɔns]	jullie [jœli]

In descriptive grammars of Dutch (cf. Haeseryn et al. 1997; Broekhuis & Den Dikken 2012), the possessive pronoun is considered a separate form that does not belong to the

same paradigm as the other forms of the personal pronoun. However, given the formal correspondence between the possessive pronoun and the accusative form of the pronoun in first and second person (singular and plural), there is reason to believe that the possessive is somehow related to the accusative and that it belongs to the same paradigm as the other forms of the pronoun.

In this paper, we investigate the question whether indeed the possessive pronoun can be considered part of the pronominal paradigm, and if so, how it is related to the other forms of the personal pronoun. The Dutch paradigm in (1) is merely an illustration of what might be a more general pattern in natural languages. In order to see whether it is indeed a general pattern, we have investigated the structure of possessives in a sample of fifty languages. The starting point of our investigation is a suggestion made by Caha (2009) who proposes that possessive pronouns are built from the genitive – in some languages with some extra structure – but not from a more complex case (such as the dative or any other "bigger" cases) or simpler case (such as the accusative or the nominative). We will come back to the missing genitive pronouns in (1), but taking Caha's suggestion at face value, the proposed structure of the possessives in (1) is depicted in (2):

(2) [[[[pronominal features] nominative] accusative] genitive] "possessive"]

This structure is largely motivated by Caha's case-hierarchy.¹ The quotes in (2) around possessive indicate that we have not looked into the specific type of morphology that could express the possessive. As we will see, in some languages there is no evidence for any additional structure in the possessive on top of the genitive. In such languages, the possessive is identical to the genitive. In others, this "possessive" morphology is clearly adjectival or involves adjectival inflection but at this point in our research we cannot be sure that this is a general property, let alone something that is universal. For that reason, "possessive" in (2) should be read as a kind of placeholder for any morphology that makes up the final layer of the morphological structure of the possessive.

If the proposed structure in (2) is indeed a structural universal, it predicts that there are languages that show possessive morphology "on top of" the genitive pronoun. Furthermore, it predicts that syncretisms between the possessive pronouns and the accusatives to the exclusion of the genitive (*ABA) are non-existent. Furthermore, if the genitive happens to be suppletive in a language, then the possessive should also be built from this suppletive form, or the possessive could have its own suppletive form. In this paper, we show that only very few languages show extra morphology on top of the genitive, but that the other predictions are borne out.

The paper is organized as follows. We will first give some further theoretical background detailing our predictions in Section 2. In Section 3 we will present the sample of languages as well as the method we used to construct the sample. In Section 4 we will present the most important results, confirming the expectations that can be based on Caha's proposal and the structure in (2). In Section 5 we will go into some of the apparent counterexamples, and investigate whether they also support the universal case-hierarchy as well as the proposal that possessives are genitives, in some cases with some extra structure. Finally, in Section 6 we conclude.

¹ Recently, Harðarson (2016) has pointed out that there are some empirical problems with the case-hierarchy proposed by Caha (2009). We will come back to his proposal in Section 2.

2 *ABA and the structure of possessives

Throughout the paper, we refer to "case" as the formal marking of the grammatical and/or semantic role of nominal constituents in a sentence (cf. Moravcsik 2009). We will analyse languages without case-marking as languages with one syncretic pronoun that is used for all relevant roles.

We assume a theoretical framework in which the morphology is a post-syntactic interpretative component that realizes the syntactic structure. Rather than proposing a detailed technical implementation in a particular framework, we focus on the patterns in the languages that we have investigated, testing a specific hypothesis that predicts some, and forbids other linguistic patterns to occur.

Our investigation first requires a definition of the notion "genitive" in relation to "possessive". How do we establish that a particular language has a genitive pronoun, and how is that different from establishing that a language has a possessive pronoun? In many languages there is no syntactic difference between a possessive and a genitive pronoun. In these languages, there are no syntactic contexts that require a genitive pronoun that cannot be taken by the possessive pronoun. However, in some languages the two can be syntactically separated since there is a genitive pronoun used in other than possessive contexts. In these languages, the genitive pronoun may be used in combination with certain verbs that require a genitive "object", or a genitive is selected by an adposition or adjective. For example, in Czech the adjective *plný* selects a genitive: *plný peněz* full money.GEN 'full of money' (Caha 2009: 110). We conclude from this that possessives should be separated from genitives syntactically, although they are syncretic in many languages.

Some initial credibility for Caha's suggestion (2009: 287) that possessive pronouns are genitive pronouns plus, in some languages, some extra morphological structure, comes from Czech. Consider the data in (3) (cf. Caha 2009: 284):

(3) *Czech* (Caha 2009: 276)

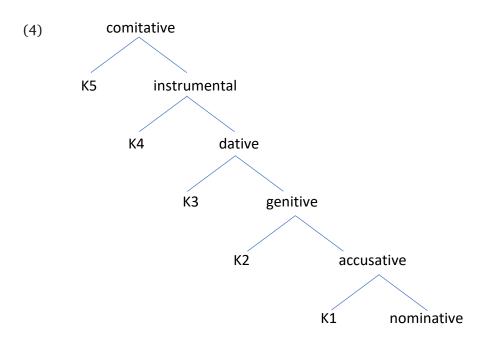
a. jí je- jí she.GEN POSS- she.GEN

b. nás naš- e (=nas-je) we.GEN we.GEN-POSS

Caha observes (ib.: 284): "The possessive pronouns [...] are distinct from, yet clearly based on the genitive form." The element *je* is prefixed to the third person feminine genitive pronoun in (3a) and suffixed to the first person plural genitive pronoun in (3b).

Another example is Lao, a language form the Tai-Kadai family, spoken in Laos and parts of Thailand. In this language (Enfield 2007: 77, 94) the possessive is overtly built from the pronoun that is syncretic for nominative, accusative, dative, and, by hypothesis, the genitive. The possessive consists of this syncretic pronoun preceded by the marker khoong3 (3 indicating the tone), that in its bare form means 'things, stuff'. In this case, the possessive is thus overtly marked by some extra nominal morphology on top of the accusative/genitive. As we will see, there are other languages in which the possessive morphology is adjectival in nature, such as in Old English, which we will discuss in due course.

The structure in (2) is motivated largely by the case-hierarchy of Caha (2009). In this hierarchy cases stand in a containment-relation: the more complex "higher" cases contain the lower, more simplex ones. Caha (2009: 49) proposes the functional hierarchy in (4):



Since our sample contains languages that show an ergative/absolutive case-system, we need to extend this hierarchy in such a way that it may also apply to such languages. We build on Smith et al. (2016) who consider a simplified hierarchy with three positions ranging from bottom to top: unmarked, dependent and oblique. The unmarked cases are the nominative (in nominative-accusative languages) and the absolutive (in absolutive-ergative languages).² The dependent cases are the accusative and the ergative cases, and the other cases are lumped together under the node "oblique".

If we map this proposal onto the hierarchy in (4), we may replace the label "nominative" with "unmarked", so that it applies to nominative and absolutive case. And we replace "accusative" with "dependent" so that it applies to accusative and ergative case. Combining this adapted case-hierarchy with the hypothesis in (2), we arrive at the hypothesized structure for the possessive in (5):

(5) [[[[pronominal features] unmarked] dependent] genitive] "possessive"]

Harðarson (2016) argues that the hierarchy in (4) presents several problems for some Nordic languages (Modern Icelandic, Modern Faroese and Old Norse) that show case-syncretisms between accusative and dative to the exclusion of the genitive. The simplified case-hierarchy used by Smith et al. (2016) solves this problem by leaving the oblique cases unordered. For now, we choose to stick to the hierarchy in (5), and see to what extent the predictions made by this hierarchy are in accordance with the data. We may consider the hierarchy of Smith et al. as an alternative hypothesis. We will come back to this issue in section 6.

Several predictions now follow from this hierarchy given an approach to spell-out that is regulated by the Elsewhere Condition (cf. Halle & Marantz 1993; Caha 2009; Bobaljik 2012), or Subset-principle.³ Put briefly, this principle determines that the lexical item with the most features matching the morpho-syntactic representation wins. In order to clarify these predictions, consider the spell-out of a dative. There are several ways in which this can be done. First, a language could have a dedicated rule that is (only) triggered by the

² See also Zompì (2017) for a proposal for case-decomposition along these lines.

³ The same result can be achieved with the Superset-principle (Starke 2009). We refrain from a discussion of more technical issues involved in the choice between the two.

presence of the feature [K3]. This may result in a so-called complex case in which the special "dative" affix is attached outside the genitive case-marker. Alternatively, there could be a suffix spelling out the "dative" node in (4). In case a language does not have a special rule for the dative (i.e. there is no rule triggered by [K3]) the form of the dative will necessarily be identical to the genitive (which in turn may or may not be syncretic with the accusative). In order to see how this is derived, consider the pronominal system of hypothetical D-Dutch which is identical to the pronominal system of Dutch in (1), but with one crucial difference: the dative forms are identical to the nominative (6).

	1sg	2sg	1pl	2pl		
NOM	ik	jij	wij	jullie		
ACC	mij	jou	ons	jullie		
DAT	ik	jij	wij	jullie		

(6) *D-Dutch (hypothetical)*

The system of D-Dutch, showing a *ABA-pattern, is impossible to derive under these assumptions. The Elsewhere Condition tells us that given a set of rules, the most specific rule that meets its structural description is applied. Given the hierarchy in (4), the structural description for the spell-out of the dative will need to involve the features {K1, K2, K3}. Therefore, any rule that is applicable in the accusative (involving the feature [K1]) will be more specific than the nominative rule, and will thus, as determined by the Elsewhere Principle, have precedence over the nominative rule in spelling out the dative. Syncretism with a "lower" case is therefore only possible, if any intervening cases are also syncretic. The only way the pattern of D-Dutch could be derived is by assuming that the dative forms are coincidentally homonymous with the nominative forms. In the (hypothetical) case at hand this should be rejected since the correspondence between nominal and dative in the pronominal system of D-Dutch is shown to be systematic.

In several cases, the evidence that the possessive is built from the genitive is not so straightforward. Merely inspecting the surface forms does not always suffice to uncover the underlying structure. To illustrate, consider the paradigm in (7) from Old English:

	Ι	Two of us	We
NOM	ic	wit	we
ACC	me (mec)	unc (uncit)	us (usic)
GEN	min	uncer	ure (user)
DAT	me	unc	us

(7) Old English (Flom 1930: 69, cited in Caha 2009: 274)

This paradigm is problematic in view of the hypothesis that syncretisms should always involve adjacent cases in the hierarchy. As can be seen, the paradigm in (7) shows a syncretism between accusative and dative, unexpectedly "skipping" the genitive (mirroring the pattern observed by Harðarson 2016 for the West-Nordic languages). However, if it can be shown that the genitives in (7) are in fact more complex, and include some extra (adjectival) morphology, the prediction is no longer contradicted by these facts. In that case, we may assume a "hidden" genitive that is syncretic to the accusative/dative

pronoun. This is exactly what Caha proposes for Old English. Caha supports his analysis with the following three arguments.

First, pronominal genitives, and not nominal genitives, show adjectival agreement in Old English. Pronominal genitives thus pattern with "true" adjectives in this respect, which suggests that they are something "more" than just a genitive. Second, some genitive pronouns can be shown to be morphologically complex. The form *uncer* is built up from the part *unc*, which is identical to the accusative/dative, and an element *-er* that spells out the adjectival component. Third, independently from the pattern in Old English, in Czech, possessives are also built from the genitive pronoun (which can be observed in adverbial constructions) and an overt suffix *-je* that has the same "adjectivizing" function.

The same kind of analysis seems valid for Dutch. We hypothesize therefore that the possessives in (1) are built from a non-surfacing genitive (identical to the accusative and dative), and that the extra morphological structure is specific to the possessive. We leave open the issue as to the exact nature of this structure, noting that in Dutch the first person plural possessive *ons* 'our' (but not the others) shows adjectival inflection.⁴

Therefore, we may conclude that there is initial evidence (from Dutch, Old English, Czech, and Lao) for the structure of possessive pronouns in (5). Some extra evidence for the "hidden" genitive comes from Albanian (Camaj 1984: 93), one of the languages in our sample.

	1sg	2sg	3sg.masc	3sg.fem
NOM	unë	ti	ai	ajó
ACC	mue/mua	ty	(a)të	atế
GEN	_	_	i (a)tíj	i (a)sáj
DAT	mue/mua	ty	(a)tíj	(a)sáj

(8) Albanian (Can	naj 1984: 93)
-------------------	---------------

In Albanian, only third person pronouns have a separate genitive form. These are also the only parts of the paradigm where the accusative and dative are not syncretic. In other words: Albanian displays a mixed system with in the first and second person a single pronoun for the accusative, "hidden" genitive, and dative, and in the third person, distinct pronouns for all cases. This is exactly what we would predict on the basis of Caha's case hierarchy. These predictions are thus not only valid across languages but also within the system of an individual language, as Albanian illustrates.

Harðarson (2016) provides a different solution for the accusative-dative syncretisms that (seem to) exclude the genitive. Harðarson (2016) argues for flexibility in Caha's (2009) hierarchy, such that the relative order of the genitive and dative is flexible. This means that both hierarchies in (9) are possible in this view.

- (9) a. [[[[pronominal features] unmarked] dependent] genitive] dative]
 - b. [[[[pronominal features] unmarked] dependent] dative] genitive]

In languages with the underlying hierarchy of (9b), the genitive is built from the dative and syncretisms between accusative and dative no longer yield a ABA-pattern. With respect

⁴ Norwegian is another example of a language in which some possessives show adjectival inflection, viz. the 1st and 2nd singular and 1st plural possessive. The adjectival endings -Ø (M, F), -t (N) and -e (PL) are also combined with these possessives, as can be seen in the following examples: *et fin-t hus* ('a nice house'), *mi-tt hus* ('my house'), *hvit-e hest-er* ('white horses') and *din-e hester* ('your horses').

to our hypothesis in (5), this would mean that the possessive can either be built from a genitive that includes the accusative (and nominative) or from a genitive that includes the dative which in turn includes the accusative. As before with respect to Smith et al., we choose to test the stronger hierarchy in (5), considering Harðarson's flexible solution as an alternative. We will come back to this issue in section 6.

Let us now have a look at the third person pronominal paradigm in Dutch (10):

	3sg.masc	3sg.fem	3pl
NOM	hij	zij	zij
ACC	hem	haar	hen
POSS	zijn	haar	hun
DAT	hem	haar	hun

The singular feminine forms are unproblematic: they show the expected syncretism between accusative, possessive and dative. Similarly, the third person plural does not pose any particular problems. Although the accusative differs, the possessive and dative are syncretic.⁵ This is expected if we assume that the "hidden" genitive is syncretic with the dative. The singular masculine forms show the accusative-dative syncretism excluding the possessive. Furthermore, the form *zijn* 'his' (unlike *mijn* 'my' above in (1)) is not phonologically related to the accusative-dative form. In order to account for this, we assume that in this particular example, the extra "possessive" morphology (see the structure in (5)) triggers spell out by a suppletive form that realizes the complex structure including the genitive, a situation similar to what we see in the Old English first person (7). The dative personal pronoun does not include this extra 'possessive' morphological structure, and can therefore be syncretic with the accusative.

Given the structure in (5), we may expect to find the following types of possessives in natural languages, depending on the precise spell-out operations that map the structure in (5) onto the vocabulary items of the language. First, a language may show transparent complex possessives in the sense that the final layer in the possessive is realized by a separate affix added to the genitive pronoun (11a). Second, we may find syncretisms between the genitive or – in the absence of a separate genitive – between the accusative, and the possessive (11b). Third, we may expect to find cases where the possessive is suppletive and shows no resemblance to the genitive (11c).⁶

- (11) Types of possessive pronouns (predicted)
 - (a) genitive pronoun "possessive" (Old English *unc-er*, Dutch *mij-n* 'my')
 - (b) genitive (accusative) = possessive (Dutch haar 'her')
 - (c) possessive is suppletive

It will be clear that cases of type (11c) are not very informative as to the structure in (5). Suppletive forms lower in the hierarchy may tell us something about the structure of the higher forms since it is unexpected that these higher forms "revert" to the form

(Dutch *zijn* 'his')

⁵ The forms given for the 3rd person plural are the normative forms. For many speakers of Dutch there is only a single form for third person pronoun (*hun*) that is used in all four cases.

⁶ Suppletion can be seen at different levels. Many languages show suppletion between singular and plural pronouns (such as Dutch *ik* 1SG versus *wij* 1PL), as described in Corbett (2005). Here we are interested in suppletion between different cases of a single pronoun (as in Dutch *ik* 1SG.NOM versus *mij* 1SG.ACC).

below the suppletive (cf. Bobaljik 2012). However, suppletive forms in themselves are uninformative as to their morphological structure. Since, as far as we know, there are no further more complex forms based on the possessives, we are looking at the highest form in a hierarchy. Suppletive forms are therefore not helpful for our aims here. The other types, however, would fit the structure in (5), and they would count as confirmation of our hypothesis. More interesting would be to find cases that do not fit any of the above types and that therefore falsify our hypothesis. If our hypothesis is wrong it should be easy to find languages in which e.g. possessives are syncretic with the nominal case-form to the exclusion of the accusative, or in which possessives are syncretic with the accusative while there is a different separate genitive. If that would be the case, we would have to revise the structure in (5), or give up the idea of a universal underlying structure for possessives altogether.

However, as we will see in section 4, there is reason to believe that there are no possessives outside the types in (11). That would provide real evidence for the structure in (5) and a universal structure for possessive pronouns. Before we get to this, let us first go into the details of our study.

3 Methodology: a genetically balanced sample

In order to test the hypothesis in (5), we compiled a genetically balanced variety sample containing fifty languages following the sampling procedure of Rijkhoff & Bakker (1998). The rationale in this procedure is that the more genetic variety a language family or subfamily exhibits, the more of its languages should be in the sample, thereby creating a sample free of genetic bias. In other words: the internal complexity of a language family determines the proportion of this family in the sample.

To determine how much variety a language (sub)family exhibits, several factors are taken into account. The first factor is the "depth" of a language family: the number of levels between the top-node (i.e. family name) and the terminal nodes (i.e. the individual languages). Neither the top-node nor the terminal nodes are counted as separate levels, to restrict the influence of the actual number of languages in the family (Rijkhoff & Bakker 1998: 269). The second factor is the "width" of the language family: the number of nodes on a single level. The width and depth of a (sub)family are used to calculate the "Diversity Value", which determines by how many languages each family should be represented in the sample. In calculating the Diversity Value, higher nodes in a language family are assigned more significance, since they represent diachronically older splits and are therefore assumed to have a greater impact on linguistic diversity (Rijkhoff & Bakker 1998: 270, for further details of the procedure).

Calculating the Diversity Value of all language families and sub-families manually is a very challenging task. We therefore used the automatic application developed by Bakker (p.c.) to create a basic sample, which describes the number of languages per family and subfamily based on the total number of languages that will be used. However, in many cases fewer languages are needed than the family has subfamilies, and then the sampling algorithm leaves the final decisions to the researcher. In these cases, the researcher has to select by hand which subfamilies will be included. The actual languages included will also be selected by the researcher rather than by the sampling algorithm.

The sampling procedure and application can be used regardless of the classification system one uses. In this paper, the classification system of Ruhlen (1987/1991) is used, which distinguishes a relatively small number of language families. The sampling procedure requires a minimum of one language per family (Rijkhoff & Bakker 1998: 268, 272), and for a sample containing fifty languages Ruhlen's (1991) classification is the

most convenient to work with. The classification is currently controversial because of its small number of families, but we do not think that this has influenced our results in any meaningful way.

Ruhlen's (1991) classification system contains nineteen families, plus a group of language isolates and a group of pidgins and creoles. These are not actual language families, and Bakker's (p.c.) application treats these as groups with a Diversity Value of 1.5, in order to prevent the sample from being filled with individual language isolates (Rijkhoff & Bakker 1998: 290–292).

From the basic sample calculated by Bakker's (p.c.) application, we created a complete sample, taking several things into account. Firstly, only non-extinct languages were selected. Secondly, the decision of including or excluding a language depended on the availability of good quality descriptions. Only those languages that are described in enough detail with respect to their pronouns, possessives and case-system were included. Apart from this practical restriction, we also took into account the geographical location and size of the language. These factors, however, were not balanced for in a systematic way, since we consider genetic diversity as the most important factor (following Rijkhoff & Bakker 1998). Our method thus differs somewhat from Bobaljik (2012) and Smith et al. (2016). In these studies, the genetic relations between languages were controlled for by considering only one example of a number of cognates.

The complete sample can be found in the appendix, together with a map showing where the languages are spoken.

4 Analysis and results

In the total of fifty languages, only six showed a syntactically separate genitive pronoun. In line with our definition in section 2, we decided whether a language had a separate genitive on the basis of the occurrence of the pronoun in a context different from possessive (such as the genitive object of a verb or in combination with an adposition). Lezgian, for example, has a syntactically separate genitive pronoun since it occurs in phrases such as (12). Note that in Lezgian, the genitive pronoun is overtly built from a smaller case (ergative), which is predicted by (5).

(12) *Lezgian* (Haspelmath 1993: 215) inal abur-u-n wilik here 3PL-ERG-GEN in.front 'Here in front of them'

Evenki is one of the six languages with a syntactically separate genitive pronoun. To illustrate the pronominal paradigm of these languages, consider the following paradigm from Evenki:

(13)	Evenki (Nedjalk	ov 1997:	200-201,	207-208)

	1SG
NOM	bi
ACC	min-e(-ve)
GEN	min-ngi
DAT	min-du

In Evenki, the form *min* underlies the accusative, the genitive and the dative. In all six languages with separate genitive pronouns in our sample, this pronoun is built from the dependent case, as expected. Also, in these six languages, the possessive is syncretic with the genitive, thus following the prediction in (11b). Note that these languages do not provide evidence for an extra morphological "possessive" layer. We stressed above that this extra morphology is present in some languages (such as Czech) but not in all.

This leaves 44 languages that do not have separate genitive pronouns. That is, in these languages the possessive pronoun cannot be syntactically separated from the genitive pronoun. Among these, seven languages show possessive pronouns that are suppletive. For example, in Koasati the nominative and the accusative pronoun share the same base (*isno*-) but the alienable possessive and inalienable possessive are suppletive sharing a base ci.⁷ Lavukaleve is another example of a language with possessive pronouns that are suppletive with respect to the syncretic nominative-accusative-dative pronoun.

(14) Koasati (Kimball 1991: 288–289, 417, 432–433)

SG
sno-k
sno-n
im
i

(15) Lavukaleve (Terrill 2003: 93–96, 170)⁸

	2sg	2dual	2pl
NOM	inu	imil	imi
ACC	inu	imil	imi
DAT	inu	imil	imi
POSS	ngo-	mele-	me-

In Koasati, Lavukaleve, and the other languages with suppletive possessives, the personal pronouns conform to the case hierarchy, but the possessive pronouns, that are the focus of this paper, neither confirm nor disconfirm the structure in (5). These possessives belong to our category (11c).

The remaining 37 languages show non-suppletive possessive pronouns. Of these languages 33 show a pattern that is in conformity with our hypothesis. These languages

⁷ In our data collection, we included both alienable and inalienable possessives to test the hypothesis. In fact, only 5 languages show a distinction between these two types of possessives. This number is too small to draw any conclusions, so we leave the questions about the exact nature of these possessives for further research. We do note, however, that in these languages, both types of possessives follow the patterns predicted by the hypothesis.

⁸ The possessives in Lavukaleve are prefixes, which is indicated by the dash (-) behind the prefixal forms. These prefixes are attached to the noun that refers to the possessed item (Terrill 2003: 93–96).

either display syncretisms between the possessive and the dependent case, or they have some overt morphological structure on top of the dependent case, showing that the possessive is indeed built from this form. There are four languages with patterns that are at odds with our hypothesis. We come back to these patterns in section 5.

In 26 languages, we find possessives with overt morphology on top of the dependent case (and presumed genitive). For example, in Afar (16), we see that there is a syncretism between accusative and dative third person masculine pronouns.

(16) *Afar* (Bliese 1981: 189)⁹

	3SG
NOM	ʻusuk
ACC	'kaa
POSS	'kay ¹⁰
DAT	'kaa

Interestingly, the possessive (being built from a "hidden" genitive) is more complex than the accusative-dative, confirming Caha's suggestion that there is extra morphology involved in the construction of possessives on top of the dependent case.

In Lao (Enfield 2007: 77, 94) the possessive is overtly built from the pronoun that is syncretic for all cases. The possessive consists of this syncretic pronoun preceded by the marker *khoong3* (3 indicating the tone), that in its bare form means 'things, stuff'. In this case, the possessive is thus overtly marked by some extra nominal morphology on top of the dependent case/genitive. Chukchi (17) is an example of an ergative language where the possessive is overtly built from the same basis as the ergative case.

(17) Chukchi (Dunn 1999: 102, 150)

	2sg
ABS	γəto
ERG	yən-an
POSS	yən-in

In seven languages, we find possessives syncretic with the accusative and/or dative pronominals. For example, in Teribe, we find the same form *bor* in accusative, dative and possessive. Again, such cases are in conformity with the structure in (5), but in these cases, unlike the previous ones, there is no overt spell-out of the possessive morphology.

⁹ Unlike Dutch (see (10)), Afar does not show a gender distinction it the pronominal system. This is true for most of the languages in our sample: 70% (35 languages) do not make a gender distinction, and among the remaining languages, a gender distinction is most often found in the third person (as in Dutch). This distribution is comparable to that in the *WALS*-sample, and the genetic and geographic distributions are also comparable between the samples (Siewierska 2013).

¹⁰ We assume that the [y] is added to the form 'kaa with subsequent shortening of the vowel.

(18) *Teribe* (Quesada 2000: 46–47)

	1sg
NOM	ta
ACC	bor
POSS	bor
DAT	bor

Gooniyandi displays a very similar pattern. The possessive is syncretic with the accusative-dative pronominal, as in Teribe, but unlike Teribe, this syncretic form is overtly related to the nominative.

	1pl.incl	1pl.excl	2pl	3pl
NOM	yaadi	ngidi	gidi	bidi
ACC	yad-dangi	ngid-dangi	gid-dangi	bid-dangi
DAT	yad-dangi	ngid-dangi	gid-dangi	bid-dangi
POSS	yad-dangi	ngid-dangi	gid-dangi	bid-dangi

(19) *Gooniyandi* (McGregor 1990: 168–173)

Some languages have a single pronominal form that is used in the nominative, accusative, dative, (and thus also "hidden" genitive) and possessive.¹¹ This pattern can for example be found in Hmong Njua (Harriehausen 1990: 127) and Bambara (Kastenholz 1998: 35–36).

Summarizing this section, of the fifty languages studied 39 conform to the proposed hypothesis. Seven languages (the ones showing suppletive forms in their possessives) are neutral with respect to the hypothesis. Four languages are at odds with the proposed hypothesis. In a way, these are the most interesting cases. We will have a closer look at these languages in section 5.

5 Apparent counterexamples

Four languages in our sample pose problems for our hypothesis. Our hypothesis predicts that there are no languages in which the possessive is related to the nominative, but not to the dependent/genitive. However, this is precisely what we seem to find in Kobon (Davies 1981). The same pattern is also found in Chamorro (Topping 1973) and Atayal (Rau 1992) but only in parts of the paradigm. In Dyirbal (Dixon 1972), the accusative seems to be built from the genitive, while our hypothesis predicts that the genitive is built from the accusative. These languages may thus at first sight be considered counterexamples. However, before we give up our hypothesis, let us look at these languages in some more detail.

Such a closer look into these languages may lead to a more nuanced view. Let us first have a look at the pronominal paradigm of Kobon:

¹¹ In our sample, 20 languages have a syncretic pronoun for nominative, accusative and dative. The sample in the *World Atlas of Language Structures* slows a slightly higher percentage of languages with neutral case (Comrie 2013), but we believe that this difference is not meaningful.

	SG		DUAL			PL			
	1	2	3	1	2	3	1	2	3
NOM	(y)ad	ne/ ni	nipe/ ne	hol	köl⁄ kale	köl/ kale/ kalipe	hon	köl⁄ kale	köl/ kale/ kalipe
ACC = DAT	ip	nöp	nip	halip	kalip	kalip	hanip	kalip	kalip
POSS	(y)ad	ne/ ni	nipe/ ne	hol	köl/ kale	köl/ kale/ kalipe	hon	köl/ kale	köl/ kale/ kalipe

(20) *Kobon* (Davies 1981: 154)

As is clear from the data in (20), the pronominal paradigm of Kobon constitutes a problem for the claim that *ABA holds for nominative-accusative-genitive-dative. The possessive (predicted to be based on the genitive) is structurally syncretic to the nominative (as in hypothetical D-Dutch in Section 2). A closer look into the examples and description in the reference grammar, can help understand the Kobon-system better. It turns out that there are in fact two series of pronouns in Kobon. The first are the so-called "neutral" pronouns (such as 1SG (*y*)*ad*), that are not inflected for case and are used as subject, direct object and indirect object. They are thus syncretic for nominative, accusative and dative, and are also used as possessive. The second class (such as 1SG *ip*) are specialized pronouns that are only used for direct and indirect object, and that are thus a syncretic accusative-dative pronoun (Davies 1981: 147–148).

The important observation here is that the language has two options to express an accusative or dative pronoun: either with the neutral pronoun that is syncretic with the nominative and possessive, or with the specialized pronoun. So, not only is the possessive syncretic with the nominative, it is also identical to the accusative-dative pronoun. Kobon, despite first appearances, thus exactly fits the pattern predicted by our hypothesis.

Let us now have a look at Dyirbal, a language that also seems to contradict our hypothesis. The structure in (5) stipulates that the genitive is built from the accusative, and the possessive is built from the genitive. In Dyirbal the possessive is syncretic to the "simple genitive" (Dixon 1972: 42–43). A more complex form that (possibly) functions as a nominalized possessive can be derived from this simple genitive. However, looking at the pronominal paradigm, it seems that the accusative is more than, and includes the genitive, which would fly in the face of our hypothesis.

	SG		E	DUAL	PL	
	1 2 1 2		1	2		
NOM	naḍa	ŋinda	ŋaliḍi	nubaladi	ŋanaḍi	nuradi
ACC	naygu-na	ŋinu-na	ŋaliḍi-na	nubaladi-na	ŋanadi-na	ɲuraḍi-na
DAT	naygu-ngu	ŋinu-ngu	ŋaliḍi-ngu	nubaladi-ngu	ŋanadi-ngu	nuradi-ngu
GEN = POSS	naygu	ŋinu	ŋaliḍi-nu	nubaladi-nu	ŋanaḍi-nu	ງາມradi-nu

(21) *Dyirbal* (Dixon 1972: 50)

These data show that for the dual and plural persons, the nominative serves as a "base" form, and that the accusative (*-na*), genitive/possessive (*-nu*) and dative (*-ngu*) are built on top of this base form. Each case-head is thus spelled out by its own suffix. In itself this is unproblematic for our hypothesis. However, for the singular persons, the situation is a little different. In these persons, there is a suppletive form in non-nominatives (*naygu* and *ŋinu*) and the genitive does not have its regular suffix *-nu*. These forms, which are problematic for our hypotheses, are boldfaced in the table in (21).

As a result, it may look as if the accusative is built from the genitive in the singular: the genitive contains less material than the accusative. However, we believe that this is not necessarily the correct analysis. As far as we can see, there is at least one other analysis possible.¹² We analyze these data as a case of contextual allomorphy. We may assume that the genitive is realized as a zero-affix, only when it is combined with the features of [1sG] and [2sG], but not in the context of other pronominal features. This solves the problem, as in this analysis there is suppletion of the base-form in singular persons, and this base-form is combined with the regular suffixes in accusative and dative, and with a zero-suffix in the genitive. In this analysis, the genitive is, as expected, built from the accusative, and not the other way around.

Both Kobon and Dyirbal, which at first face seem to violate our hypotheses, have now been explained away. Kobon was shown to support the hypothesis in (5) when both pronominal systems are considered, and Dyirbal can be made compatible with the hypothesis on a certain analysis of the data (i.e. a zero-affix in the genitive 1st and 2nd person).

This leaves two apparent counterexamples to be dealt with, Chamorro and Atayal. In these languages, the same problem arises as in Kobon: they show the D-Dutch pattern in which the possessive seems to be related to the nominative to the exclusion of the accusative/genitive. However, in both languages these violations of our hypothesis are only found in two rows of the paradigm, as can be seen in the tables (22) and (23) below, in which the problematic cells appear in boldface.

	SG			PL			
	1	2	3	1 incl	1 excl	2	3
NOM	hu	un	ha	ta	in	en	ma
ACC DAT	yo'	hao	gue'	hit	ham	hamyo	siha
POSS	-hu/-ku	-mu	-ña	-ta	-mami	-miyu	-ñiha

(22) Chamorro (Topping 1973: 106–108)

(23) Atayal (Rau 1992: 126)

	SG			PL			
	1	2	3	1 incl	1 excl	2	3
NOM	saku', ku'	su'	hiya′	ta'	sami	simu	hga', lhga'
ACC DAT	knan	sunan	hiyan	itan	sminan	smunan	hgan
POSS	maku, mu	su'	nya′	ta'	myan	mamu	nha′

¹² A different solution might be to look for a phonological process that deletes or changes the genitive suffix *-nu* in certain contexts. There is no evidence for such a process in the reference grammar, and therefore, given our data, we believe that the correct analysis is morphological.

In Chamorro, we see that the accusative-dative pronoun is a suppletive form, and the possessive is suppletive again. Only in the first person singular and first person plural inclusive, the possessive seems similar to the nominative. Now, it first has to be noted that the possessive in Chamorro is not a free pronominal element, but rather a clitic that attaches to the noun that expresses the possessed item. This is illustrated in (24) below.

(24) Chamorro (Topping 1973: 108) kareta- hu car- POSS.1SG 'my car'

Up until now, we have considered any element expressing possession as a possessive pronoun. However, it seems reasonable that one should not collapse free pronouns and bound (clitic or affixed) pronominal elements into a single paradigm. If we assume that the structure hypothesized in (5) only holds for free pronouns, Chamorro will be analysed as a language lacking free pronouns in possessive function. The items in the bottom row in (22) are then no longer analysed as part of the pronominal paradigm, and as a result the apparent *ABA-pattern is no longer found.¹³

A similar situation is found in Atayal. As shown in the table in (23), Atayal displays an accusative-dative pronoun that is either suppletive or formally more complex than the nominative, and a possessive that is suppletive. Only in the second person singular and the first person plural inclusive, the possessive is identical to the nominative. However, the nominative and possessive pronouns are clitics, whereas the accusative-dative pronoun is a free pronoun comparable to a full DP (Rau 1992: 126). Besides that, the possessive clitic is often accompanied by a special possessive marker (Rau 1992: 143), and therefore often not syncretic with the nominative clitic. These facts lead us to the conclusion that the apparent *ABA-pattern in Atayal is only apparent and that Atayal is no longer a counterexample to the hypothesis in (5).

If we restrict the structure in (5) to free pronouns, the possessives in Chamorro and Atayal should not be taken into account when testing the predictions in (11).¹⁴ The remaining relevant paradigms do not constitute any counterexamples.

This analysis also influences the analysis of several other languages, since Atayal and Chamorro are not the only languages with bound possessives. In our sample, there are eleven other languages that have bound possessives as well, and four languages show both free and bound possessives. Excluding the bound possessives from these languages, the results described in Section 4 would not change, since in these languages the (bound) possessive is either morphologically related to the genitive (or accusative) pronoun or suppletive.

To summarize, in Sections 4 and 5 we have shown how virtually all languages in our sample follow the predictions in (11) and the structure in (5). If we only include free pronouns in the analysis, Chamorro and Atayal are no longer counterexamples to (5). By taking into account the full pronominal system of Kobon, it was shown that this language

¹³ We thank an anonymous reviewer and Pavel Caha for this suggestion. In theory, there is no reason why the hypothesized structure in (5) should only hold for free pronouns and not for bound morphemes. However, we do believe that these two types of elements should not be collapsed into a single paradigm. In this paper we show that the structure in (5) holds for paradigms of free pronouns. By hypothesis, the same structure would hold for paradigms that consist of only bound pronominal elements. We have chosen our sample in such a way that it does not contain languages with a full paradigm of bound pronouns. Therefore, we leave these paradigms for further research.

¹⁴ The question that rises is what the precise nature is of the bound possessives in these languages. We leave this question for further research, taking note of a suggestion from an anonymous reviewer that these elements might be some kind of agreement on the noun.

in fact does follow the structure in (5), so the found *ABA-pattern turned out to be only apparent. Dyirbal can also be shown to fit the hypothesis under a specific (zero affix) analysis. We summarized our results in the table in (25):

(25)

Languages in the sample (50)						
No syntactically separate genitive pronoun (44)Possessive = genitive (6)						
No supj	Suppletive possessive (7)					
Overt Morphology (26)	Syncretism (7)	Counter-examples (4)				

6 Conclusion

Following a suggestion by Caha (2009) we have tested the hypothesis that possessive pronouns in languages of the world have a uniform structure. This structure includes the genitive form of the pronoun which itself is built from the accusative on top of the nominative. In some languages, added to this structure, we may find specific "possessive" morphology. In order to be able to include languages with an absolutive/ergative case-system, we have extended the hypothesis somewhat by replacing "nominative" by "unmarked", and "accusative" by the label "dependent", rendering the structure in (26).

(26) [[[[[pronominal features] unmarked] dependent] genitive] "possessive"]

In our sample, there were no languages that give us any evidence for the extra layer of "possessive" morphology. We are led to conclude that this is a relatively rare property of languages. Thus, so far, only Czech and Old English provide evidence for this extra layer, and the idea that the possessive is on top of the genitive.

Our data show that the possessive is constructed on top of the dependent case (accusative or ergative), but there is no crucial evidence in our sample that shows that the possessive sits on top of the genitive. The data in our sample are equally compatible with a structure in which the genitive is built on top of the possessive. In the six languages with a syntactically separate genitive pronoun, the possessive is syncretic with this pronoun, which does not give us information about the structural relation between the two. In seven languages, we find suppletive possessives which are again uninformative. In 26 languages, we find possessives that are built from the dependent case, but since there is no separate genitive pronoun, there is no information on the structural relation between the two. Finally, in 7 languages we find possessives syncretic with the accusative and dative which again is uninformative about the relation with the genitive. So, the relation genitive << possessive relies on a few languages outside our sample.

Our data give a firm foundation to the claim that the possessive is more complex than the dependent case. This claim only meets a few potentially problematic cases: Chamorro, Atayal, Dyribal and Kobon, despite the relative ease with which this hypothesis could have been refuted by the data. After close inspection of the pronominal system in Kobon, it turns out that the language has two separate series of pronouns. Properly separating the different forms removes any potential problematic issue, and Kobon possessives easily fit the hypothesis. In addition, some problematic facts from Dyirbal can be made compatible with the structure in (26) once we assume a null morpheme that spells out the 1st and 2nd

person singular genitive. Chamorro and Atayal only form counterexamples when we mix clitics or affixal forms with independent pronouns. When only independent pronouns are considered, the problematic facts disappear. In all other languages, the possessive patterns in the way predicted by our hypothesis.

If these cases can be relegated in this way, this finding provides evidence for the idea that Caha's (2009) hierarchy of case can be extended to pronouns, and that possessives should be included in this hierarchy. We followed the idea that even though the possessive is built from the genitive pronoun, this genitive does not have to exist independently in the language. Rather, it might be syncretic with the accusative and dative, and only serve as the basis for the possessive.

Harðarson (2016) argues for a certain amount of flexibility in the case hierarchy, such that the genitive is either on top of the accusative (with the dative higher than the genitive), or on top of the dative which includes the accusative. Our sample contains 33 languages with an accusative-dative syncretism, which might give rise to the idea that Harðarson's (2016) flexibility stretches to other languages. Caha (2009), however, provides strong arguments for his analysis with a "hidden" genitive. Our data are not decisive with respect to these two analyses. The hypothesis that the possessive pronoun is a complex pronoun including the genitive pronoun, which in turn includes smaller cases, does not find counterexamples in our sample.

Abbreviations

1, 2, 3 = 1, 2, 3 person, ABS = absolutive, ACC = accusative, AL = alienable, DAT = dative, DUAL = dual number, ERG = ergative, EXCL = exclusive, FEM = feminine, GEN = genitive, INAL = inalienable, INCL = inclusive, MASC = masculine, NOM = nominative, PL = plural, POSS = possessive, SG = singular.

Additional File

The additional file for this article can be found as follows:

• **Appendix 1.** Languages in the sample. Overview of where languages in the sample are spoken. DOI: https://doi.org/10.5334/gjgl.395.s1

Acknowledgements

The authors wish to thank Pavel Caha and two anonymous reviewers for help in preparing the final version of this paper. The authors also thank Fenna Bergsma and Fred Weerman for discussion and comments on an earlier version of this paper.

Competing Interests

The authors have no competing interests to declare.

References

Bliese, Loren F. 1981. A generative grammar of Afar. Dallas, TX: Summer Institute of Linguistics.

Bobaljik, Jonathan. 2012. Universals in comparative morphology: Suppletion, superlatives and the structures of words. Cambridge, MA: MIT-Press.

Broekhuis, Hans & Marcel den Dikken. 2012. *Syntax of Dutch: Nouns and noun phrases* 2. Amsterdam: Amsterdam University Press.

Caha, Pavel. 2009. The nanosyntax of case. Tromsø: University of Tromsø dissertation.

Camaj, Martin. 1984. *Albanian grammar: With exercises, chrestomathy and glossaries*. Translated by Leonard Fox. Wiesbaden: Otto Harrassowitz.

- Comrie, Bernhard. 2013. Alignment of case marking of pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (http://wals.info/chapter/99).
- Corbett, Greville G. 2005. Suppletion in personal pronouns: Theory versus practice, and the place of reproducibility in typology. *Linguistic typology* 91. 1–23. DOI: https://doi. org/10.1515/lity.2005.9.1.1
- Davies, John. 1981. Kobon (Lingua Descriptive Studies 3). Amsterdam, North Holland.
- Dixon, Robert M. W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139084987
- Dunn, Michael J. 1999. *A Grammar of Chukchi*. Canberra: Australian National University dissertation.
- Enfield, Nick. 2007. *A Grammar of Lao* (Mouton Grammar Library 38). Berlin: Mouton de Gruyter.
- Flom, George T. 1930. *Introductory Old English grammar and reader*. Boston: D.C. Health and Company.
- Haeseryn, Walter, K. Romijn, Guido Geerts, Jaap de Rooij & Maarten van den Toorn. 1997. *Algemene Nederlandse spraakkunst.* Groningen/Deurne: Martinus Nijhoff Uitgevers.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.
- Harðarson, Gisli R. 2016. A case for a weak case contiguity hypothesis: A reply to Caha. *Natural language & linguistic theory* 44. 1329–1343. DOI: https://doi.org/10.1007/s11049-016-9328-x
- Harriehausen, Bettina. 1990. Hmong Njua: Syntaktische Analyse einer gesprochenen Sprache mithilfe daten-verarbeitungstechnischer Mittel und sprachvergleichende Beschreibung des südostasiatischen Sprachraumes. Tübingen: Max Niemeyer Verlag. DOI: https://doi. org/10.1515/9783111356440
- Haspelmath, Martin. 1993. *A grammar of Lezgian* (Mouton Grammar Library 9). Berlin/New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110884210
- Kastenholz, Raimund. 1998. Grundkurs Bambara (Manding) mit Texten. Köln: Rüdiger Köppe. 2nd edition.
- Kimball, Geoffrey D. 1991. Koasati grammar. Lincoln, NE: University of Nebraska Press.
- McGregor, William. 1990. *A functional grammar of* Gooniyandi. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/slcs.22
- Moravcsik, Edith. 2009. The distribution of case. In Andrej L. Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 231–245. Oxford: Oxford University Press. Nedjalkov, Igor. 1997. *Evenki*. London: Routledge.
- Quesada, Juan Diego. 2000. A grammar of Teribe. München: Lincom Europa.
- Rau, Der-Hwa Victoria. 1992. *A grammar of Atayal*. Ann Arbor, MI: Cornell University dissertation, University Microfilms.
- Rijkhoff, Jan & Dik Bakker. 1998. Language sampling. *Linguistic typology* 2. 263–314. DOI: https://doi.org/10.1515/lity.1998.2.3.263
- Ruhlen, Merritt. 1987. *A guide to the world's languages*, vol. 1: classification. Stanford: Stanford University Press.
- Ruhlen, Merritt. 1991. *A guide to the world's languages*, vol. 1: classification (Second Edition). Stanford: Stanford University Press.
- Siewierska, Anna. 2013. Gender distinctions in Independent personal pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/44.

- Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang & Jonathan David Bobaljik. 2016. Case and Number Suppletion in pronouns. https://ling.auf.net/lingbuzz/003110.
- Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36. 1–6.
- Terrill, Angela. 2003. *A grammar of Lavukaleve* (Mouton Grammar Library 30). Berlin/New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110923964
- Topping, Donald M. 1973. *Chamorro reference grammar*. Honolulu, HI: University Press of Hawai'i.
- Zompì, Stanislao. 2017. *Case decomposition meets dependent-case theories*. Pisa: University of Pisa MA Thesis. https://ling.auf.net/lingbuzz/003421.

How to cite this article: van Baal, Yvonne and Jan Don. 2018. Universals in possessive morphology. *Glossa: a journal of general linguistics* 3(1): 11.1–19, DOI: https://doi.org/10.5334/gjgl.395

Submitted: 01 April 2017 Accepted: 15 October 2017 Published: 19 January 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

]u[Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.

