

# Assessing the relationship between object topicalisation and the grammaticalisation of object agreement

*Stefan Schnell and Geoffrey Haig*

*La Trobe University and Bamberg University*

*s.schnell@latrobe.edu.au*

*geoffrey.haig@uni-bamberg.de*

In this paper we report on a corpus study of narrative texts from different languages, aimed at testing Givón's (1976) topicalisation-origin account of the grammaticalisation of object agreement from free pronouns. In the narrative texts investigated here, we find an exceedingly low proportion of object topicalisation constructions, which suggests that the construction is an unlikely candidate for the origin of object agreement. Moreover, the usage of both bound and unbound object pronouns is independent of object topicalisation, though it is conditioned by animacy in some languages. We conclude that topicalisation and pronominalisation are two distinct operations, with distinct functions. We further suggest that the facts from discourse provide a partial explanation for the findings from language typology, which reveal that canonical object agreement is exceedingly rare in the languages of the world, with some form of conditioned agreement being the norm for objects.

**Keywords:** grammaticalisation, object agreement, topicalisation

## 1. Introduction

Since Givón's (1976) seminal paper, the development of subject and object agreement is often accounted for in terms of 'topicality'. Under this view, subject

*Selected Papers from the 44<sup>th</sup> Conference of the Australian Linguistic Society, 2013*

edited by Lauren Gawne and Jill Vaughan

and object agreement arises in connection with the “depragmatisation” of a topic construction, and is therefore in fact “topic agreement”. Although there are a number of problems with this approach (cf. Lambrecht 1994: 192, Fuß 2005: 6-7), it continues to be exceedingly influential.

Most previous research within the topicality paradigm has focussed on subject agreement, with the recent exception of Dalrymple & Nikolaeva (2011). As for topicalisation-origin accounts of object agreement, despite their continued popularity, there is a notable lack of actual empirical research in support of such a scenario. In our paper, we report findings from a cross-corpus study of object realisation in spontaneous spoken discourse, based on data from five different languages. While there is some support for the notion of gradual degrees of grammaticalisation of object agreement, there is no obvious connection between different kinds of object agreement and prevalence of object topicalisation in our data.

## 2. Topicalisation and object agreement

### 2.1 *Topic shift and pronominalisation: Givón (1976)*

For Givón, object agreement<sup>1</sup> arises via ‘topic-shifting’ constructions under circumstances which we illustrate with the following pseudo-English examples (cf. Givón 1976:157 who proposes this scenario for Swahili and other Bantu languages), involving left-dislocation of a NP and its pronominal resumption *in situ*:

- |     |                               |  |
|-----|-------------------------------|--|
| (1) | a. <i>I saw the man.</i>      | [pragmatically neutral]                        |
|     | b. <i>The man, I saw him.</i> | [pragmatically marked topicalisation strategy] |
|     | c. <i>I saw-him the man.</i>  | [pragmatically neutral, object agreement]      |

The development of object agreement markers thus involves three changes triggered by the dislocation of the object NP: firstly, in order to become a

---

<sup>1</sup> Givón (1976) is primarily concerned with subject agreement, but as Siewierska (2004: 263) notes, his account is “equally valid” for object agreement.

pronominal agreement marker, a personal pronoun must regularly occur in a position adjacent to the verb. Secondly, once this is the case, frequency effects lead to phonological reduction of this person marker and cliticisation with the verb. Finally, in the case of canonical agreement, the marker becomes obligatory and is required on the verb, regardless of the presence or absence of a local coreferent NP (“conominal” in Haspelmath’s 2012 terminology).

Our main concern in this paper is with the conditions that are assumed to precede reduction and cliticisation, namely the more or less regular occurrence of object person markers in actual discourse. It seems clear that for systematic language change to take place, a high frequency in the usage of free object pronouns is a prerequisite for these other processes to catch on (cf. Bybee 2005, 1985; Hopper 1987), as free pronouns would provide the initial input for grammaticalisation. In the spirit of Givón’s (1976) hypothesis though, we should also expect to find topic-shift constructions (left-dislocation), the presumed forerunner constructions of object agreement, at significant levels of frequency in natural discourse, as these constructions motivate the usage of pronouns as resumptive forms of dislocated objects. To date there is very little empirical research focussing on objects in natural discourse; and to our knowledge, neither of the above predictions has been systematically investigated.<sup>2</sup> Our aim, therefore, is to provide some first indications on the different possibilities (expression types) of objects in actual spoken language, and to consider how they might relate to the above scenario.

## *2.2 The typological asymmetry in subject and object agreement*

A number of researchers have pointed to a striking typological asymmetry between transitive subject (A) agreement, and object agreement (P) with regard to the feature of person (e.g. Kibrik 2011, Bickel et al. 2012). Haig (2013a) summarises the broadly observable tendency as follows: With A, agreement in person is more frequently obligatory (i.e. realised independently of the semantics of the argument itself, and of the pragmatic configuration of the clause); P agreement in person (gender and number agreement may differ), on the other

<sup>2</sup> Wald (1979) is one of the few studies to actually test Givón’s hypothesis on spoken language data (Bantu); he finds no support for the dislocated-topic hypothesis.

hand, is typically *conditioned*, that is, dependent on (i) the semantics (e.g. +/-animate) and pragmatic status (definiteness or specificity) of the object itself; (ii) the presence/absence of a local NP object (conominal); (iii) the presence of a higher-ranked recipient/benefactive, which may co-opt the usual marker; (iv) the pragmatic configuration of the clause (topic/focus). Reliable figures are difficult to cite, due to massive differences in terminology. However, the available sources all point in the same direction: Siewierska (2013) considers presence vs. absence of ‘verbal person marking’ with A and/or P in transitive clauses, based on a sample of 378 languages. Eighty-eight of these do not show any agreement at all, and of the remaining 290 languages that do have some kind of agreement, 73 have A-agreement only, while 24 have P-agreement only. The rest (193) have both A and P agreement. Thus, the tendency is that if only one argument shows agreement, this will most commonly be the subject. With regards to the systematicity/obligatoriness, no figures are given, but several of the languages listed as showing object agreement in Siewierska (2013) are best analysed as cases of conditioned agreement: Anejom (cf. Lynch 2000: 84ff.), Roviana (Siewierska (1999: 236), Yimas (cf. Foley 1991: 232f.), Canela-Kraho (cf. Popjes & Popjes 1986: 175ff.), Macushi (Abbott 1991: 24) and Panyjima (Dench 1991: 159). Further cases of conditioned object agreement are discussed in Hopper & Thompson (1980), Dalrymple & Nikolaeva (2011), Haig (2013a), Kibrik (2011: 189) and Witzlack-Makarevich & Iemmolo (2013).

### 3. Object pronominalisation and topicalisation in discourse

#### 3.1 *Corpus data*

Corpus data for this study stem from linguistic fieldwork on 5 different languages.<sup>3</sup> All the texts are spontaneous spoken narratives, typically traditional tales from the respective language communities, recounted by a single speaker, i.e.

---

<sup>3</sup> Data sources: Vera’a: 13 narrative texts collected and annotated during fieldwork in Vera’a (Vanua Lava, Vanuatu) by members of the Vera’a community and Stefan Schnell; Barwar Neo-Aramaic: first part of a narrative in Khan (2008b: 1538-1567), annotated by Roberta Borghero; N. Kurdish: single narrative in Ritter (1968: 17-58), annotated by Geoffrey Haig and Rushen Turgut; C. Kurdish: single narrative from MacKenzie (1962: 5-32), annotated by Geoffrey Haig and Ergin Öpengin; Cypriot Greek: five narratives from Yiangoellis (2009) (Cypriot Folk Tales, spoken by Eleni Satcha (female), born in 1887, and recorded by Andreas Ioannides, between 1960 and 1964), annotated by Harris Hadjidas.

monologues (cf. Haig et al. 2011 on the rationale behind using indigenous narratives). Table 1 gives an overview.

Language	Affiliation	corpus size (clause units)	no. of objects (total)	objects considered (3 <sup>rd</sup> pers. only)
Vera'a	Austronesian	3581	1140	1074
North-East Neo-Aramaic (NENA)	Semitic, Barwar dialect (North Iraq)	864	221	163
North Kurdish	Iranian, IE	1388	414	370
Central Kurdish	Iranian, IE	1233	431	352
Cypriot Greek	Hellenic, IE	1633	866	584

*Table 1. Overview of corpus data from 5 languages.*

The corpora are small, so issues of representativity are obviously relevant. It should be noted, though, that many of the published cross-linguistic studies of spoken discourse (e.g. on Preferred Argument Structure, cf. contributions in Du Bois et al. 2003, on referential density (Bickel 2003), or the figures on object agreement in Ostyak cited in Dalrymple & Nikolaeva 2011: 167 ) are based on corpora well within the range of sizes here. Given the general paucity of studies of this kind, then, the current compilation represents a significant contribution. The sample is also obviously not representative of the languages of the world. However, the five languages concerned represent five different kinds of object indexing systems, as will become apparent. In addition, we considered text data from Savosavo, a language that does show the typologically rare feature of canonical object agreement, but due to restrictions of space we will not discuss these findings in detail. We will take up the Savosavo facts in the discussion in section 4.

With regard to coding procedures, we include all exponents of referential objects present in individual clauses. The following four possibilities are distinguished: (i) lexical NPs, (ii) free person markers within the clause (i.e. not dislocated and not bound in the verbal predicate), (iii) bound person markers (prosodically bound forms, clitics or affixes), and (iv) zero anaphora. The category (iii), bound person marker, requires some comment: it is often a matter of controversy

whether a given instance should be considered “object agreement”, or a “pronoun” (cf. Corbett 2006: 99-112 for discussion). We make no distinction in our counts, and we therefore adopt the more neutral terminology of Haspelmath (2012) and refer to bound person index (BPI) for any bound object marker. It is important to emphasise that our figures include *all* exponents of the object in a clause, thus we also record cases where the object relation is represented by multiple means of expression within the same clause, for instance a bound clitic person marker co-occurring with a free NP, i.e. Haspelmath’s (2012) “conominal”, or Corbett’s (2006) “multi-representation”. How frequently, and under what conditions, such multi-representation occurs is in fact an empirical issue, which needs to be established prior to any discussion on agreement versus anaphora. Below, we present the findings on a language-by-language basis before summing up the main tendencies in Section 4.

### 3.2 *Expression of objects in Vera’a*

Vera’a (vra, Austronesian, Oceanic) is a largely isolating language in which syntactic functions are encoded by rigid SV/AVO word order (Schnell 2010). There are three types of object attested in the corpus: lexical NPs, BPIs and zero. In Vera’a, we consider an object marker to be a BPI when it is part of the verbal complex (VC) and does not bear independent stress. In our data, free pronominal objects are entirely absent, although the language does allow for the possibility of expressing pragmatically marked pronominal objects via a pronominal NP that occurs outside the VC. In Table 2 (and respective tables below), grey-shaded columns indicate expression types that are actually impossible in the language; where a form is possible but unattested in our corpus, zero appears in the appropriate column that is not shaded. In Vera’a then, it is not possible for a BPI to co-occur with a coreferential free pronoun or a NP in the same clause, hence these two columns are grey-shaded. In other languages, on the other hand, this option is available (see below).

	ZERO	BPI	FREE PRO	NP	BPI-	BPI+	TOTAL
+HUM	18	158	0	117	-	-	293
-HUM	206	5	0	570	-	-	781
<b>TOTAL</b>	<b>224</b>	<b>163</b>	<b>0</b>	<b>687</b>	<b>-</b>	<b>-</b>	<b>1074</b>
	<b>(21%)</b>	<b>(15%)</b>	<b>(0%)</b>	<b>(64%)</b>			<b>(100%)</b>

*Table 2. Expression of objects in Vera'a*

Most objects in Vera'a are realised by lexical NPs (64%), and most have non-human referents (781 of 1074, or 73%). The general propensity for non-human reference is even stronger for NPs (580 of 687, or 84%); a typical example is the following<sup>4</sup>:

(2) VERA'A, NP OBJ

*dir'ol ov            ēñ    qe'e wo=k            gin            ēñ    rōrō'*  
 3TL   pull.out   ART **taro** and=TAM2   pinch.off   ART **cabbage**  
 'Then they pulled out **some taros** and picked **some cabbage (leaves).**'  
 AS.1.023

BPIs make up only 15% of all objects. Among the 163 pronominal objects, there is an almost categorical (97%) preference for human reference (cf. Genetti & Crain 2003; Haig et al. 2011 on the *Avoid pronominal P* constraint):

(3) VERA'A, BPI OBJ

*dir=eḱ            qērẽ    ba'a di    sar    lẽ=n            mō-gi*  
 3PL=TAM2    push   into **3SG** in   LOC=ART   POSS.HOUSE-3SG  
  
*=n    nimē̃*  
 =ART   house  
 'They would push **her** into her house.'  
 ISWM.171

<sup>4</sup> Morphological glossing follows the conventions of the *Leipzig Glossing Rules* (LGRs). Other abbreviations used in the Vera'a examples: CC clause-combining particle; DEL delimitative; POSS.EAT possessive classifier for food possession; POSS.GEN possessive classifier general (unspecified) possession; POSS.HOUSE possessive classifier for house possession; TL trial.

Zero objects make up 21% of all objects, and a great majority of them have non-human reference (206 of 224, or 92%):

(4) VERA'A, zero OBJ

*e ruwa mērmēre=k van e=k le*  
 PERS.ART two.people kid=TAM2 go CC=TAM2 take  
*go-r'ol=n gengen dir'ol 'ēqel 'o' kel ma*  
 POSS.EAT-3TL=ART food 3TL descend carry back hither  
 'The kids fetched some food for the three of them, and then they (the three) went back down with (it).'

HHAK.147-148

With regard to Givón's hypothesis concerning the emergence of object agreement, we also examined instances of object topicalisation involving a left-dislocated object. We included as left-dislocated all cases where an object occurs left of the subject. We also noted the kind of resumptive element that occurs *in situ* within the clause. The data are summed up in Table 3.

	LEFT-DISL. NP		
<i>IN SITU</i> :	ZERO	BPI	TOTAL
+HUM	0	5	5
-HUM	29	0	29
<b>TOTAL</b>	<b>29</b>	<b>5</b>	<b>34</b>

**Table 3.** *Left-dislocated objects in Vera'a.*

The first point to note is the overall extreme scarcity of left-dislocation for objects: only 34 instances in a total of 1074 objects, i.e. approx. 3%. Further, we note that the *in situ* form can be either pronominal or zero, with zero overwhelmingly preferred (29 of 34, or 85%). Given that Givón's account of the rise of agreement assumes overt pronominal resumption of a left-dislocated object, the high rate of zero in this construction suggests that left dislocation



would not be a likely source of object agreement, for this language at least. Finally, we note that the choice between an overt BPI and zero *in situ* is determined by animacy, precisely as it is for objects generally, as discussed above. Examples:

(5) VERA'A, left-dislocated OBJ

- a. *ekē*      *e*      *raga*      *anē*      *no*      *me*      *vus*      *dir*  
oh.dear    PERS.ART people    DEM1    1SG    TAM4    kill    3PL

*sivie*  
how

“‘Oh dear, these people how am I ever going to kill them?’” JJQ.401

- b. *riar*      *mu-gi*      [...]      *di=m*      *mom*      *i*  
k.o.bow    POSS.GEN-3SG      3SG(SUBJ)=TAM1    put    DEL

*suwei*  
down

‘His bow, he put (it) down, (he didn’t climb with (it)).’

ISWM.133-134

To sum up, the choice of pronominal versus zero objects is mainly determined by the humanness of the referent in question. This holds for cases with and without topicalisation. The token frequency of topicalisation constructions is extremely low, much lower than pronominal objects in general. Note that Vera’a already seems to have partially “grammaticalised object agreement”, in the sense that its object pronouns are exclusively prosodically bound and incorporated into the verbal complex. But there seems no obvious connection between this fact and the very rare instances of object topicalisation, nor any obvious sense in which object agreement is becoming more strongly grammaticalised.

### 3.3 Expression of objects in Barwar (North East Neo-Aramaic)

The Barwar dialect of North East Neo-Aramaic (Semitic, Aramaic, north Iraq; NENA henceforth) has SV/AVO word order. Nouns inflect for gender and number, but not for case. Direct objects may be flagged in several different ways, depending primarily on discourse pragmatic factors (cf. Coghill, to appear): NP

objects are flagged either through a preposition, or unmarked. Pronominal objects are generally phonologically dependent on the verb, and fuse with a preposition *l-* (Khan 2008a: 805f.), as in: *ʔu triðal-le* ‘and he.drove-**him**’. We refer to these formatives as BPIs. Objects can have a variety of representations: zero, a BPI by itself (without co-nominal NP or pronoun), BPI + pronoun, or BPI + NP. According to the grammar of Barwar Neo-Aramaic (Khan 2008a: 879), “occasionally” an independent pronoun or demonstrative may express a direct object, hence the Free Pro-column is not grey-shaded. However, in our corpus, no examples of free object pronouns were attested.

	ZERO	BPI	FREE PRO	NP	BPI+FREE PRO	BPI+NP	TOTAL
+HUM	2	38	0	10	2	33	85
-HUM	2	14	0	42	1	19	78
<b>TOTAL</b>	<b>4</b>	<b>52</b>	<b>0</b>	<b>52</b>	<b>3</b>	<b>52</b>	<b>163</b>
	<b>(2%)</b>	<b>(32%)</b>	<b>(0%)</b>	<b>(32%)</b>	<b>(2%)</b>	<b>(32%)</b>	<b>(100%)</b>

*Table 4. Expression of objects in North East Neo-Aramaic.*

The NENA corpus is by far the smallest in our data, and any conclusions are correspondingly tentative at this stage. Furthermore, it differs from all other corpora we are aware of in that the proportion of human objects is actually higher than that of non-human ones.<sup>5</sup> Nevertheless, in line with generally observable cross-linguistic tendencies, objects in NENA are mostly lexical, represented either by a NP (only 32%), as in (6a), or by a NP plus a BPI (32%), as in (6b)<sup>6</sup>:

<sup>5</sup> As far as we can ascertain, the statistical tendency for non-human objects to outnumber human objects in connected narrative discourse appears to be universal, though of course this remains to be validated. One reason for the exceptional nature of the NENA-text investigated here is that much of the narrative is concerned with the fate of two children, who, at the order of the King, are taken at birth from their parents and subsequently set adrift in a river. The children are later rescued by the King’s son, who entrusts them to another couple to be raised. A large number of transitive verbs in this short text thus has the children as objects.

<sup>6</sup> The NENA transcription is simplified, and basic glosses are added; BPIs are indicated by :P.

## (6) NENA, lexical objects without (a) and with (b) BPI

- a. *šqila sanduqa' b=iðe diye'*  
 took.3SG **chest** into=hands POSS.3SG

'(He) took **the chest** into his hands.'

- b. *málka yăðéwa=le 'o-nàša.*  
 the.king knew.3SG=3SG:P **that-man**

'The King knew **that man**.'

As for non-lexical objects, zero form is exceedingly rare, attested only in four cases, with two human and two non-human referents. By far the most common form is the BPI-only variant: 88% of all non-lexical objects take this form. A large majority of these (73%) has human reference:

## (7) NENA, BPI only

- yá'ni rába kúlla náše yăðiwa=le.*  
 that.is very.much all people knew.3PL=3SG:P

'That is, all the people knew **him** very well.'

Overall, there is a strong tendency for BPIs to be used when the object is human: of a total of 85 human objects, 73, or 86%, were (co-)expressed through a BPI. Nevertheless, BPIs do get used with non-human objects, but here the corresponding figure drops to 44% (34 of 78). Most Neo-Aramaicists do in fact refer to them as "object agreement". However, agreement is far from obligatory: about 50% of NP objects in our corpus are not accompanied by a BPI, cf. (6a) above; NENA thus exhibits a typical case of conditioned object agreement (or indexing, cf. Haig 2013a). Animacy of the object is certainly a relevant factor in triggering object agreement, but it interacts with considerations of information structure, the details of which go beyond the scope of this paper.

The NENA corpus shows the highest proportion of object topicalisation which is attested in ten cases among the 163 objects (~6%). Three of these topicalised objects have human and seven non-human reference, but in all cases we find a

BPI *in situ*, regardless of humanness. This contrasts with the trend we found in Vera'a, according to which non-human objects are less likely to have BPIs. Examples of both human and non-human dislocated objects are provided below:

- (8) NENA, left-dislocated human object (demonstrative pronoun) + BPI

'áyya      gáni      fadən=na  
**this:P**      SELF.1SG:POSS      ransome=**3SG:P**  
**Her (lit. 'this.FEM'), (I) myself will ransom her**

- (9) NENA, left-dislocated non-human NP objects + BPI

'ánna      kúlla      Óárwa      díya      bšaqlì=la      =w'  
**these(a)**      **all**      **wealth**      of.it      will.take=**3s:P**      =and  
šáwpi      bšaqlì=le=u'  
**my.place**      will.take=**3S:P**=and  
**'All this wealth (of the kingdom) they will take it,**  
**my place they will take it ...'**

The Barwar pattern of left-dislocated objects is in fact reminiscent of Givón's scenario: the left-dislocated object NP is invariably resumed by a BPI on the verb, and this appears to be independent of animacy considerations. Thus, Barwar object dislocation appears to work differently to that of Vera'a, discussed above. We return to the implications of these findings in the discussion.

### 3.4 Expression of objects in North Kurdish

Northern Kurdish (or Kurmanji, Indo-European, Iranian; southeast Turkey) is an OV language; objects are overtly case-marked in present tenses, and unmarked in past tenses (cf. Haig 2008: Ch. 5). In the non-past tenses, verbs consistently agree with S / A, while in the past tenses there is considerable dialectal variation (cf. Öpengin & Haig, to appear): the "standard dialect" arguably shows object agreement, but in many dialects it is inconsistent, with past-tense verbs often showing agreement with a subject rather than the object. Assessing levels of inconsistency is rendered more difficult by the fact that and the relevant form for

the most frequent object type, namely 3SG, is zero; we have coded all non-overt object forms as zero, regardless of tense, though we concede that this is an oversimplification. It is worth noting that historically, the object agreement in past tenses in Northern Kurdish did not arise from the grammaticalisation of bound object pronouns (cf. Haig 2013b, in prep). In Northern Kurdish, object pronouns do not fuse with the verb, so we consider the BPI-option to be unavailable, and the corresponding columns in the table are grey-shaded.

	ZERO	BPI	FREE PRO	NP	BPI+FREE PRO	BPI+NP	TOTAL
+HUM	10	-	21	27	-	-	58
-HUM	53	-	2	257	-	-	312
<b>TOTAL</b>	<b>63</b>	<b>-</b>	<b>23</b>	<b>284</b>	<b>-</b>	<b>-</b>	<b>370</b>
	<b>(17%)</b>		<b>(6%)</b>	<b>(77%)</b>			<b>(100%)</b>

*Table 6. Expression of objects in Northern Kurdish.*

The distribution of object expressions in the North Kurdish corpus is strikingly similar to that attested in the Vera'a corpus (cf. 3.2 above). Note also the very high instance (77%) of NPs as objects, most of which have non-human referents (257 of 284, or 90%). (10) shows a NP object with human reference:

(10) NORTH KURDISH, NP only

*Ūsufšá ji čālê derēxist-in.*  
**Usufsha** from well.OBL out.pulled-3PL  
 '(They) pulled **Usufsha** out of the well.'

Very few objects are expressed through free pronouns (6%), and most of these (91%) have human referents:

## (11) NORTH KURDISH, Free Pronoun

*az vē náqlē wî jî bikúj-im!*  
 I this:OBL means:OBL 3SG:P too may.kill-1SG  
 ‘... (that) I may kill **him** too by these means!’

Objects may have zero expression, and zero objects have predominantly non-human reference (53 of 63, or 84%):

## (12) NORTH KURDISH, zero object

*bika bin čengē xwá da*  
 put.IMP.2SG under arm.of self DRCT  
 ‘Put (**them**) under your arm!’

Left-dislocation with objects is extremely rare in North Kurdish: only two cases of a total of 370 objects, i.e. well under 1%, are attested in the corpus. The following is one example:

## (13) NORTH KURDISH, left-dislocated phrase corresponding to OBJ

*au qat jî mi xwâri-bû*  
**that piece** too I eat.PST-COP.PST  
 ‘**That piece** too I had eaten.’

However, given that the most readily operationalised diagnostic for left-dislocation is object-before-subject, and given that in many transitive clauses, no overt subject constituent is present, it is often difficult to detect left-dislocation reliably. The figures given are therefore probably conservative, and if prosodic or other factors were taken into consideration, the number of left-dislocated objects could be higher. However, given that Northern Kurdish simply has no BPIs, its relevance in assessing possible origins of object agreement is probably restricted anyway.

### 3.5 Expression of objects in Central Kurdish

Central Kurdish (or Sorani, Indo-European, Iranian, North Iraq) is closely related to Northern Kurdish, but differs from it strikingly in the expression of objects. Central Kurdish is OV; non-past verbs, and all intransitives, consistently agree with S / A. Agreement with past transitive verbs remains controversial (Öpengin 2013), and we have followed the same procedure as with Northern Kurdish (see discussion above) in counting all examples of non-overt expression of referential objects as zero. There is no case-marking on objects. However, pronominal objects are most commonly expressed through a special paradigm of clitic pronouns, distinct from the free pronouns, which generally attach to part of the VP; we treat these clitic object pronouns as BPIs (cf. Haig 2008: Ch. 6 and Öpengin 2013). In Northern Kurdish, these BPIs cannot co-occur with a free object constituent in the same clause, hence the grey shading in this part of Table 7.

	ZERO	BPI	FREE PRO	NP	BPI+FREE PRO	BPI+NP	TOTAL
+HUM	3	47	3	48	-	-	101
-HUM	8	38	0	205	-	-	251
<b>TOTAL</b>	<b>11</b> <b>(3%)</b>	<b>85</b> <b>(24%)</b>	<b>3</b> <b>(1%)</b>	<b>253</b> <b>(72%)</b>	-	-	<b>352</b> <b>(100%)</b>

*Table 7. Expression of objects in Central Kurdish.*

Most objects (72%) are realised as NPs only, and the vast majority of these (205 of 253, or 81%) have non-human referents:

(14) CENTRAL KURDISH, NP

*min    ū        tō        šart-ēk                    a-ka-yn.*  
 I        and    you        agreement-INDF        IND-do.PRS-1PL  
 ‘You and I, (we) make **an agreement.**’

Free pronouns are extremely rare, only 3 instances (<1%) are attested, and all have human referents:

(15) CENTRAL KURDISH, free pronoun

*aw=ĩš=ĩ*                      *lagat*    *xōy*                      *hat girt.*  
 3SG:P=and=3SG:A with      self.3SG      took.PST  
 ‘He took **her** with him too.’

After full NPs, by far the most frequent realisation of objects is via a clitic pronoun on the verb (BPI), accounting for 24% of all objects. Clitics in Central Kurdish show no obvious animacy preferences (55% have human referents):

(16) CENTRAL KURDISH, BPI only

*ĩstā*      *a=y=hēn-im,*  
 now      INDIC=3SG:P=fetch:PRS-1SG  
 ‘Now I am fetching **him**.’

Zero objects are rare (3% of all objects). They seem to correlate with non-humanness (8 of the total 11 zero objects have non-human reference), but the absolute figures are too low to draw firm conclusions:

(17) CENTRAL KURDISH, Zero

*a-t-dam-ē*  
 IND-2P:RECP-give.PRS.1SG-to  
 ‘I will give (**it**) to you.’

As far as topicalisation of objects is concerned, there were only two candidate cases in the data, but they were both open to varying interpretations. In general, left-dislocation of objects appears to be very rare in these narrative texts, but similar caveats apply here as were discussed for Northern Kurdish above.



In Central Kurdish, then, clitic BPIs are clearly the default for non-lexical objects, regardless of differences in animacy. But unlike the BPIs of NENA, the Northern Kurdish ones are categorically excluded from co-occurrence with a free object (no “clitic doubling”, no conominals). Obviously, the occurrence of BPIs is entirely independent of topicalisation constructions.

### 3.6 Expression of objects in Cypriot Greek

Cypriot Greek (Indo-European, Hellenic) is the dialect of Modern Greek spoken on Cyprus. It has VO word order, and objects are case-marked. Objects can be realised through all six expression types that we distinguish; object pronouns generally cliticise to the verb, though they are orthographically written as separate words.

	ZERO	BPI	FREE	NP	BPI+FREE PRO	BPI+NP	TOTAL
+HUM	5	106	5	73	9	11	209
-HUM	10	70	1	275	6	13	375
<b>TOTAL</b>	<b>15</b>	<b>176</b>	<b>6</b>	<b>348</b>	<b>15</b>	<b>24</b>	<b>584</b>
	<b>(3%)</b>	<b>(30%)</b>	<b>(1%)</b>	<b>(60%)</b>	<b>(3%)</b>	<b>(4%)</b>	<b>(100%)</b>

*Table 8. Expression of objects in Cypriot Greek.*

As in all languages we have investigated to date, the majority of objects in Cypriot Greek are NPs (348 of 584 objects, or 60%), and most of these have non-human referents (76%):

#### (18) CYPRIOT GREEK, NP only

*efaniskasin      meran-nixtan      pannin*  
weave.PST.3PL   day-night      cloth

‘(They) weaved **cloth** day and night.’

The second most common forms of objects are pronouns cliticised to the verb (30% of all objects). Like the BPIs in Central Kurdish, no strong effects of animacy can be detected (60% human):

(19) CYPRIOT GREEK, BPI only

- a. *epk'aen=ton, epien*  
 took.PST.3SG=3SG:P go.PST.3SG  
 '(She) took **it** (and) went.'
- b. *E, pote ton=ies?*  
 well when 3SG:P=see.2SG  
 'Well, when did you see **him**?'

Free pronouns are extremely rare and almost entirely restricted to human referents. Zero objects are likewise rare, and have predominantly non-human referents:

(20) CYPRIOT GREEK, zero

- E, j enna xoriso je na pao enan meros*  
 So and AUX wear.1SG and to go.1SG INDF place  
 'So, and I will wear (**them**) and go to a place.'

Clitic BPIs can have conominals, either pronouns (21), or NPs (22):

(21) CYPRIOT GREEK, Free Pronoun + BPI

- na to=shonosumen tuton*  
 that 3SG:P=pour.away.1PL 3SG:P  
 'That we pour **it** away.'

## (22) CYPRIOT GREEK, NP + BPI

*E, kalo, pu ta=ivres ta rial'a*  
 well so where 3PL:P=find.2SG ART.PL money.PL.ACC  
 ‘Well, so where did you find **the money(s)**?’

The preceding examples could be interpreted as nascent object agreement, as argued by Charitonidis (2008) for Modern Greek. However, we note that in Cypriot Greek, BPI+conominal is a marginal option: of all the objects that involve an NP expression (322 in total), only 24 are accompanied by a BPI (approx. 7%). With regard to dislocated objects, an object NP is considered left-dislocated when it precedes the verb, yielding the order O(S)V. However, the few cases of left-dislocated objects in the Cypriot Greek corpus (seven, approx. 1% of all objects) make quantitative analysis of this construction of limited value.

#### 4. Discussion

In this section we discuss the issue of object agreement in the light of the findings from the individual languages. For contrastive purposes we also include data on the Papuan language Savosavo (Wegener 2008), based on the Savosavo corpus analysed in Haig et al. (2011). Although the Savosavo corpus is comparatively small, it is nevertheless worth discussing in the context of object agreement, since Savosavo does show the typologically exceedingly rare feature of canonical (i.e. unconditioned) object agreement: every transitive verb carries an affix indicating the person of the object.

The relevant facts are summarised in Table 9 below. The first three columns indicate: 1. the availability and type of BPI; 2. the proportion of human referents among pronominal objects; and 3. the proportion of BPI which co-occur with a lexical NP.

	BPI	+HUM , pro	Co-occurrence with conominal NP	Frequency of object topicalisation
N. Kurdish	none	21/23 (91%)	-	< 1%
Vera'a	inc. pro	158/163 (97%)	-	3%
C. Kurdish	clitic	50/88 (57%)	-	< 1%
Cyp. Greek	clitic	131/231 (57%)	24/200 (12%)	1%
NENA	clitic	73/107 (68%)	52/104 (50%)	7%
Savosavo	affix	54/167 (32%)	100%	1%

**Table 9.** *Object pronominalisation and topicalisation in 6 languages: type of BPI (if any); proportion of human reference among pronominal objects; proportion of cases involving conominal NPs; proportion of topicalisation constructions among all objects.*

The systems of object marking in Northern Kurdish and Vera'a display striking similarities: the pronouns are unreduced forms (though in Vera'a they lack prosodic stress and syntactic mobility), and they are mutually exclusive with conominal NPs. Furthermore, Northern Kurdish and Vera'a permit zero expression of object, and the choice of pronominal versus zero correlates almost perfectly with the feature +/-human. Such systems thus represent a kind of differential object marking (DOM), with pronominal forms highlighting the occurrence of a human referent in the object role (cf. the indexing account of DOM by Næss 2006).

Savosavo at the other extreme has canonical agreement markers: obligatory affixes, with no obvious correlation with human reference (simply because they accompany all object NPs regardless of animacy, and in all corpora investigated to date, with the exception of the NENA corpus as discussed in fn. 4 above, non-human objects are more frequent).

The three remaining languages with clitic object pronouns are situated between the pronoun-like and agreement-like poles: the propensity for human reference is clearly weaker here than with the pronoun-like systems, but clearly stronger than with the agreement-like system of Savosavo. The Central Kurdish clitic BPIs, however, cannot co-occur with object NPs, like the pronominal forms in

Northern Kurdish and Vera'a. In Cypriot Greek the clitics can co-occur with free NPs, but the proportion of NPs accompanied by a clitic is relatively low.

NENA is somewhat special among the languages of our sample: object clitics can and do regularly (but not obligatorily) co-occur with NPs. But they also show a preference for human referents, characteristic of pronoun-like systems. We suspect though that the small size of the NENA corpus and the unusually high overall proportion of human referents in this small corpus may be responsible for this finding.

The rightmost column gives the proportion of object topicalisation constructions. Only in Vera'a and NENA is the proportion higher than 1%. Interestingly, in Vera'a only dislocated human objects are resumed by a BPI, showing the same animacy-conditioning generally attested for BPIs. In NENA though, all topicalised objects – with human or non-human referents – are resumed by a clitic. Thus, NENA does in fact seem to provide support for a connection between object topicalisation and object agreement along the lines of Givón's proposals, but it is the only language in our corpus that does so, and as mentioned, the text is unusual due to the high proportion of human objects.

## **5. Conclusions**

In the light of the data discussed above, we conclude that Givón's (1976) topicalisation-origin of object agreement cannot be maintained, at least if understood as a universal motivation for such a development: topicalisation of objects is so rare in our data that it can hardly provide the motivation for the development of object agreement systems generally.

These findings are particularly interesting because four of the languages investigated here (Vera'a, Central Kurdish, NENA and Cypriot Greek) exhibit what could be interpreted as nascent object agreement: most object pronouns are prosodically bound to the predicate, hence apparently partially "grammaticalised", and free object pronouns are a rarity in all languages investigated. Thus, these languages already exhibit some variant of conditioned object agreement (and as stated, NENA is generally analysed by Neo-Aramaicists as having object agreement). Yet, there is no obvious correlation between these tendencies and the

proportion of topicalisation constructions, and hence there is no obvious way that topicalisation should play a role in the (assumed) development of such systems of conditioned object agreement into the systems of canonical object agreement of the Savosavo type. We thus conclude – contrary to Dalrymple & Nikolaeva (2011), and in line with Næss’ (2012) account of some differential case marking systems – that topicalisation and pronominalisation represent distinct subsystems with distinct functions. What ultimately triggers the emergence of object agreement in fact remains something of a mystery. Recall the two closely-related languages (some would even say: dialects) Northern and Central Kurdish: Northern Kurdish simply lacks bound object pronouns entirely, while Central Kurdish has a special paradigm of bound object pronouns. It seems unlikely that the presence of bound object pronouns of Central Kurdish be linked in some way to higher proportion of object topicalisation in this language when compared to the geographically and genetically closely related Northern Kurdish.

We should, however, point to some obvious drawbacks of the current study. First, it is quite possible that higher proportions of topicalisation obtain in text types other than narratives. Second, for the OV languages, the diagnostic for identifying topicalisation may be too conservative, and closer attention to prosody may reveal much higher levels of topicalisation. Third, our corpora are relatively small, and the scope of languages covered is obviously highly restricted. However, we would nevertheless maintain that our data provide very little support for the object-topicalisation account of the emergence of object agreement. Furthermore, we note that the independence of pronominalisation from topicalisation, and the conditioning role of animacy, is unexpected under the account of Givón (1976). More generally, we consider that a hypothesis that has been so influential for so long should be subjected to more rigorous empirical testing than has previously been the case, and we trust that this contribution may serve to generate further research in this direction.

## References

- Abbott M 1991 ‘Macushi’ in D Derbyshire & G Pullum (eds) *Handbook of Amazonian languages 3* Berlin: Mouton de Gruyter. pp. 23-160.
- Bickel B 2003 ‘Referential density in discourse and syntactic typology’ *Language* 79(4): 708-736.

- Bickel B, G Iemmolo, T Zakharko & A Witzlack-Makarevich 2012 'Patterns of alignment in verb agreement' in D Bakker & M Haspelmath (eds) *Languages across boundaries. Studies in memory of Anna Siewierska* Berlin, Boston: Mouton de Gruyter. pp. 15-36.
- Bickel B, A Witzlack-Makarevich, T Zakharko & G Iemmolo 2013 'The distribution of zeros and alignment patterns in agreement paradigms: exploring diachronic universals' Ms. Available at: [http://www.academia.edu/3814728/The\\_distribution\\_of\\_zeros\\_and\\_alignment\\_patterns\\_in\\_agreement\\_paradigms\\_exploring\\_diachronic\\_universals](http://www.academia.edu/3814728/The_distribution_of_zeros_and_alignment_patterns_in_agreement_paradigms_exploring_diachronic_universals)
- Bybee J 1985 *Morphology. A study of the relation between meaning and form*. Amsterdam-Philadelphia: John Benjamins.
- Bybee J 2005 'From usage to grammar: The mind's response to repetition' *Language* 82(4): 711-733.
- Charitonidis C 2008 'Polysynthetic tendencies in Modern Greek' *Linguistik online* 34(02/08): 17-40. Available at: [http://www.linguistik-online.de/34\\_08/charitonidis.pdf](http://www.linguistik-online.de/34_08/charitonidis.pdf)
- Coghill E. To Appear. 'Differential object marking in Neo-Aramaic' *Linguistics*.
- Corbett G 2006 *Agreement* Cambridge: Cambridge University Press.
- Dalrymple M & I Nikolaeva 2011 *Objects and information structure* Cambridge: Cambridge University Press.
- Dench A 1991 'Panyjima' in R Dixon & B Blake (eds) *Handbook of Australian languages 4* Melbourne: Oxford University Press. pp. 124-243.
- Du Bois J, L E Kumpf & W J Ashby (eds) 2003 *Preferred argument structure: Grammar as architecture for function* Amsterdam, Philadelphia: John Benjamins.
- Foley W 1991 *The Yimas language of New Guinea* Stanford: Stanford University Press.
- Fuß E 2005 *The rise of agreement. A formal approach to the syntax and grammaticalization of verbal inflection*. Amsterdam-Philadelphia: John Benjamins.
- Genetti C & L Crain 2003 'Beyond preferred argument structure: Sentences, pronouns and given referents in Nepali' in J Du Bois, L Kumpf & W Ashby (eds) *Preferred argument structure: Grammar as architecture for function* Amsterdam, Philadelphia: John Benjamins. pp. 197-230.
- Givón T 1976 'Topic, pronoun and grammatical agreement' in C Li (ed.) *Subject and topic* New York: Academic. pp. 151-188.
- Haig G 2008 *Alignment change in Iranian languages. A Construction Grammar approach* Berlin: Mouton de Gruyter.
- Haig G, S Schnell & C Wegener 2011 'Comparing corpora from endangered languages. Explorations in language typology based on original texts' in G Haig, N Nau, S Schnell & C Wegener (eds) *Documenting endangered languages: Achievements and perspectives* Berlin, Boston: Mouton de Gruyter. pp. 55-86.
- Haig G 2013a 'The subject/object asymmetry in bound person indexing: Diachronic and discourse considerations' Presentation at the Workshop on *Agreement in discourse*, University of Bamberg, February 1-2, 2013. Available at: [https://www.academia.edu/2759840/The\\_subject\\_object\\_asymmetry\\_in\\_bound\\_person\\_indexing\\_diachronic\\_and\\_discourse\\_considerations](https://www.academia.edu/2759840/The_subject_object_asymmetry_in_bound_person_indexing_diachronic_and_discourse_considerations)
- Haig G 2013b 'Alignment change in Iranian: What happened to person agreement?' Paper held at the Workshop on *Patterns of alignment in the Indo-Iranian languages: Towards a typology*. LSI, Ann Arbor, July 13-14, 2013.
- Haig G in prep. 'The degrammaticalization of bound object indexing in West Iranian'

- Haspelmath M 2012 'Argument indexing: A conceptual framework for the syntactic status of bound person forms' *Languages across boundaries. Studies in memory of Anna Siewierska*. 197-226. Berlin, Boston: Mouton de Gruyter.
- Hopper P 1987 'Emergent grammar' *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 139-157.
- Hopper P & S Thompson 1980 'Transitivity in Grammar and Discourse' *Language* 56. 251-299.
- Khan G 2008a *The Neo-Aramaic dialect of Barwar, Vol. one: grammar*. Leiden-Boston: Brill
- Khan G 2008b *The Neo-Aramaic dialect of Barwar, Vol. three: texts*. Leiden-Boston: Brill
- Kibrik A 2011 *Reference in discourse* Oxford: Oxford University Press.
- Lambrecht K 1994 *Information structure and sentence form*. Cambridge: Cambridge University Press.
- Lynch J 2000 *A grammar of Anejom*. Canberra: Pacific Linguistics.
- MacKenzie D 1962 *Kurdish dialect studies*, Vol. 2. Oxford: Oxford University Press .
- Næss Å 2006 *Prototypical transitivity* Amsterdam, New York: John Benjamins.
- Næss Å 2012 'Case on the margins. Pragmatics and argument marking in Vaeakau-Taumako and beyond' in S Kittilae, K Vaesti & J Ylikoski (eds) *Case, animacy and semantic roles* Amsterdam: John Benjamins. pp. 305-328.
- Öpengin E 2013 'Clitic/affix interactions. A corpus-based study of person marking in the Mukri variety of Central Kurdish' Unpublished PhD-thesis, Université Sorbonne Nouvelle – Paris 3.
- Öpengin E & G Haig to appear Regional variation in Kurmanji: a provisional classification of the dialects. *Journal of Kurdish Studies*.
- Popjes J & J Popjes 1986 'Canela-Krahô' *Handbook of Amazonian languages 1* edited by D Derbyshire & G Pullum, 128-199. Berlin: Mouton de Gruyter.
- Ritter H 1968 'Kurmāncî-Texte aus dem Tūr'abdin. 1. Kārboran' *Oriens* 21/22: 1-135.
- Schnell S 2010 'Animacy and referentiality in Vera'a' Ph.D. Dissertation, Kiel University.
- Siewierska A 1999 'From anaphoric pronoun to grammatical agreement marker: why objects don't make it' *Folia Linguistica* 33(1-2): 225-251.
- Siewierska A 2004 *Person*. Cambridge: Cambridge University Press
- Siewierska A 2013 'Verbal person marking' in M S Dryer & M Haspelmath (eds) *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at: <http://wals.info/chapter/102>
- Wald B 1979 'The development of the Swahili object marker: a study of the interaction of syntax and discourse' in T Givón (ed.) *Discourse and syntax* Amsterdam: John Benjamins. pp. 505-524.
- Wegener C 2008 'A grammar of Savosavo, a Papuan language of the Solomon Islands' Ph.D. Dissertation, MPI for Psycholinguistics, Radboud Universiteit Nijmegen.
- Witzlack-Makarevich A & Iemmolo G 2013 'When is there agreement? Typologizing suspension restrictions on agreement' Presentation at the *10th Biennial Conference of the Association for Linguistic Typology*, Leipzig, August 15-18, 2013. Available at: [https://www.academia.edu/4392850/When\\_is\\_there\\_agreement\\_Typologizing\\_suspension\\_restrictions\\_on\\_agreement](https://www.academia.edu/4392850/When_is_there_agreement_Typologizing_suspension_restrictions_on_agreement)
- Yiangoullis K 2009 *Kypriaka Paradosiaka Paramythia*. Nikosia: Bibliothiki Kyprion Daikon Poeton.



Minerva Access is the Institutional Repository of The University of Melbourne

**Author/s:**

Schnell, Stefan; Haig, Geoffrey

**Title:**

Assessing the relationship between object topicalisation and the grammaticalisation of object agreement

**Date:**

2014

**Publication Status:**

Published

**Persistent Link:**

<http://hdl.handle.net/11343/40959>

**File Description:**

Assessing the relationship between object topicalisation and the grammaticalisation of object agreement