

**LESS-TRAVELLED PATHS FROM PRONOUN TO
AGREEMENT: THE CASE OF THE URALIC
OBJECTIVE CONJUGATIONS**

Elizabeth Coppock and Stephen Wechsler
Lund University University of Texas at Austin

Proceedings of the LFG10 Conference

Miriam Butt and Tracy Holloway King (Editors)

2010

CSLI Publications

<http://csli-publications.stanford.edu/>

Abstract

Building on Bresnan and Mchombo's (1987) theory that the transition from pronoun to agreement marker constitutes the loss of a PRED 'pro' specification on an affix, we explore the idea that the historical path from pronoun to agreement marker can involve the loss of person and number feature specifications as well. We apply this idea to object agreement in the Uralic languages, with particular attention to Ostyak and Hungarian, and propose that person and number specifications on object agreement affixes, historically derived from bound pronouns, were lost independently at different stages. We then consider the more general hypothesis that the special distribution of person agreement can be explained as a consequence of its historical origin in incorporated pronouns, with loss of the person feature as a complicating factor. Preliminary typological evidence supports this view over Baker's (2008) theory of person agreement.

1 Introduction

Agreement is sometimes in person and sometimes not. One factor is the lexical category of the target: while verbs tend to agree in person, adjectives tend not to. For example, in French, verbs agree in person with the subject but adjectives do not:

- (1) a. Je suis belle.
 I be.PRES.1SG beautiful.FEM.SG
 'I [female] am beautiful.'
- b. Tu es belle.
 You be.PRES.2SG beautiful.FEM.SG
 'You [female, singular] are beautiful.'
- c. Elle est belle/*beau.
 She be.PRES.3SG beautiful.FEM.SG/beautiful.MASC.SG
 'She is beautiful.'

The finite verb registers the person and number values of the subject, while the adjective shows the subject's number and gender, but not its person. As revealed by broad typological studies, this pattern is common among the world's languages: verbs generally take prominence over other categories for the marking of person agreement (Stassen 1997; p. 38).

The distribution of person agreement can be traced to its historical origin: agreement inflections that encode person derive from incorporated personal pronouns, which therefore can include the pronominal features of person, number and gender (Lehmann 1988, Wechsler to appear). In the Index/Concord theory of agreement (Wechsler and Zlatić 2003), these pronoun-derived morphemes are

[†]We thank the audience at the LFG 2010 conference for helpful feedback, especially Farrell Ackerman, Miriam Butt, György Rákosi, Mary Dalrymple, Ida Toivonen and Ash Asudeh.

Index agreement inflections. In contrast, Concord agreement comprises number, gender and case, and is derived from other sources: for example, gender marking is thought to be derived from classifiers. Since verbal agreement tends to derive from pronoun incorporation, it tends to involve person; adjectival agreement derives from other sources and therefore does not.

Nonetheless, sometimes verbal agreement does not include person among its features. One example of this is object agreement in Northern Ostyak. In Northern Ostyak (a.k.a. ‘Khanty’; Uralic), verbs agree with their subjects in person and number, but with their objects only in number, and not person. For example, a plural object triggers the same plural suffix *-l* on the verb regardless of whether the object is in third person (2) or first person (3) (Nikolaeva 1999):

- (2) Ma täm käläng wel-sə-l-am
 I this reindeer kill-PAST-PLOBJ-1SGSUBJ
 ‘I killed these reindeer.’
- (3) Xünsi näng müng-iluw xälśa want-lə-l-an?
 when you we-ACC where see-PRES-PLOBJ-2SGSUBJ
 ‘When did you see us where?’

Baker (to appear) takes aim at the distinction between Index and Concord, arguing that it cannot explain the situation in Northern Ostyak:

Wechsler and Zlatić (2003) divide the phi-features into two groups, the Concord features and the Index features, and stipulate that adjectives access the first set and verbs the second set. This account has no obvious way of saying why verbs can access the person feature when agreeing with their subject, but not when agreeing with their theme object... [T]here is no reason to expect verbs in a language to show Index agreement with their subjects and Concord agreement with their objects.

Our answer to Baker’s objection, which we put forth in this paper, is that the verb-object agreement in Northern Ostyak should not be analyzed as Concord agreement. Instead, both subjects and objects trigger Index agreement, and both subject and object verbal inflections descend from incorporated pronouns. But in the case of objects, only third person pronouns were incorporated, and the third person specifications on those affixes were subsequently lost, giving rise to the Northern Ostyak system. Different features were lost in different Uralic languages, leading to variation in the systems of object agreement in Eastern Ostyak, Samoyedic languages, and Hungarian.

Baker (2008, to appear) proposes a rather different explanation for the distribution of person agreement: he posits that the person feature is subject to a special universal structural condition. In Section 6 we broaden our scope to look at typological predictions of the two theories. We argue that the cross-linguistic facts support our view, namely that the distribution of person agreement can be explained as

Stage 1: Pronoun	Stage 2: Loss of reference	Stage 3: Loss of person
(↓PRED) = ‘pro’	(↓PRED) = ‘pro’	(↓PRED) = ‘pro’
(↓INDEX PERS) = <i>p</i>	(↓INDEX PERS) = <i>p</i>	(↓INDEX PERS) = <i>p</i>
(↓INDEX NUM) = <i>n</i>	(↓INDEX NUM) = <i>n</i>	(↓INDEX NUM) = <i>n</i>
(↓INDEX GEND) = <i>g</i>	(↓INDEX GEND) = <i>g</i>	(↓INDEX GEND) = <i>g</i>

Table 1: Stages of lexical specification in the transition from pronoun to agreement

a consequence of its historical origin in incorporated pronouns: pronoun-derived agreement includes person, except when the person feature is lost over time.

2 Feature loss

Feature loss can be modeled in LFG as loss of f-description equations, as Bresnan and Mchombo (1987) demonstrate (see also Butt 2007). They analyze the object markers in Chicheŵa as morphologically bound pronouns, having a lexical entry as in the first column of Table 1. The PRED ‘pro’ specification of such a morphologically bound pronoun may then become optional; this is how Bresnan and Mchombo analyze Chicheŵa’s subject markers. If the PRED ‘pro’ is eliminated entirely, as in the second column of Table 1, the result is a grammatical agreement marker.

In this paper, we explore the idea that this process may target other features as well, such as person or number features. For example, the person feature could be eliminated, as in the third column of Table 1. Toivonen (2001) uses this idea to explain dialectal variation in Finnish possessive suffixes, showing that both number and person feature specifications can be lost. The suffix *-nsa*, which is third person in Standard Finnish, is unrestricted with respect to person in one dialect of Finnish. The suffix *-ni*, which is restricted to first person singular in the standard dialect, is unrestricted with respect to person *and* number in some South-West dialects. In other South-West dialects, *-ni* has retained its first person specification, while losing its number specification. In the same dialects, the second person singular suffix *-si* has lost its number specification, retaining its person specification. These facts show that feature loss can target either person or number specifications, and that the loss of one feature may take place independently of the loss of another. We will propose a similar account of variation in object agreement across the Uralic languages.

3 Object agreement in Uralic

Among the Uralic languages (comprising the Finno-Ugric and Samoyedic families), object agreement is found in Samoyedic (Nenets, Enets, Nganasan, Selkup), Mordvinian, and Ugric (including Hungarian, and the Ob-Ugric languages Ostyak

	SUBJECTIVE	SG.OBJ	DU.OBJ	PL.OBJ
1SG	tulêm	tulim	tulêylam	tullam
2SG	tuwên	tulin	tulêylan	tullantul
3SG	tuwêl	tultê	tulyêl	tulêl(lê)
1DU	tulmên	tulmên	tulêylamên	tullamên
2DU	tuletên	tulêtên	tulêylin	tullin
3DU	tulyên	tulêten	tulêylin	tullin
1PL	tuloy	tuluy	tulêyloy	tulloy
2PL	tultêy	tulêtên	tulêylin	tullin
3PL	tuwêlt	tulil	tulêylal	tullal

Table 2: Northern Ostyak conjugations for *tu-*, *tuy* ‘bring’ (Honti 1984; 107, taken from Kortvély 2005)

and Vogul). These languages vary in the feature specifications of their object agreement morphology. In Hungarian (É. Kiss 1987) and Samoyedic (Hajdú 1968), there are two conjugations, subjective and objective, and the objective conjugation is used when the object is definite and third person. Ob-Ugric languages have a subjective conjugation and several objective conjugation paradigms, one for each of the possible number values (singular, dual, or plural) of the object (Honti 1984, Kálman 1965). In Northern Ostyak, the use of one of these objective conjugations is further conditioned by whether the object is third person (Gulya 1966). In Mordvinian, the verb genuinely agrees in both person *and* number with the object (Keresztes 1989, see also Béjar 2008). We will address Northern and Eastern Ostyak, Hungarian, and Samoyedic in detail here. For reasons that we do not have space to go into in this paper, we believe that the Mordvinian system arose independently, and we will leave it out of the discussion for now. We begin with Northern Ostyak.

3.1 Northern Ostyak

3.1.1 Object agreement in number but not person

As shown above in (2) and (3), verbs agree in number but not person with the object in Northern Ostyak: a third person plural object and a first person plural object trigger the same plural marking on the verb. The full paradigm is given in Table 2, which shows that there is a subjective conjugation along with three objective conjugations, one for each number value (singular, dual, and plural).

As Nikolaeva (1999) documents, use of an objective conjugation requires that the object be topical. Focussed objects, as in (4), appear with the subjective conjugation:

- (4) tam kalang we:l-əʂ / *we:l-s-əlli
 this reindeer kill-PAST.3SGSUBJ kill-PAST-SG OBJ.3SGSUBJ
 ‘It was this reindeer that he killed.’

Given that the objects that trigger the objective conjugation are topical, one might suspect that they are really floating topics, anaphorically linked to a bound pronominals realized by the objective conjugation affixes. Nikolaeva (1999) shows that this is not the case, however; the objective conjugation suffixes have, in effect, lost their PRED ‘pro’ specification. Her evidence is as follows: (i) external NPs receive case, (ii) when the external NP is a pronoun, it does not have a contrastive discourse function, (iii) the external NP may not appear in the periphery of the clause, and (iv) the verb does not specify all of the features of the pronoun, only number.

In summary, the basic facts about Northern Ostyak object agreement we propose to account for are: that it is sensitive to number, without being sensitive to person; that objects that trigger the objective conjugation must be topical; and that the objective conjugation suffixes are not pronominal.

3.1.2 Analysis

We propose to analyze Northern Ostyak object agreement as Index agreement without person; the object agreement affixes in Northern Ostyak derive from incorporated third person pronouns whose person specifications were subsequently lost.

Our proposed development begins with an incorporated pronoun stage. We suggest that there were three third person pronouns that were incorporated, one for each number value (singular, plural, and dual). This is schematized in (5). We crucially assume that only third person pronouns were incorporated in this ancestor of Northern Ostyak. Evidence for that assumption is that in its neighbor, Eastern Ostyak, the objective conjugation is restricted to third person objects (see §3.2 below).

(5) **Stage 1: Incorporated pronouns (only in third person)**

$$\begin{aligned}
 V_{aff} \quad (\uparrow \text{OBJ}) &= \downarrow \\
 (\downarrow \text{PRED}) &= \text{‘pro’} \\
 (\downarrow_{\sigma} \text{DF}) &= \text{TOPIC} \\
 (\downarrow \text{INDEX PERS}) &= 3 \\
 (\downarrow \text{INDEX NUM}) &= n \quad \text{where } n \in \{\text{SG, DU, PL}\}
 \end{aligned}$$

We assume that the bound pronominal also comes with a topicality condition, encoded by the third equation. Indeed, it has often been observed that it is highly topical pronouns that can be reduced to bound forms. The notation for the topicality condition follows Dalrymple and Nikolaeva (to appear); it specifies that the value of the DF (‘discourse function’) attribute in the σ (semantic) projection of the node corresponding to the affix is TOPIC. This will cause its semantics to appear as the value of the TOPIC attribute at the level of information structure, by a principle relating semantic structure to information structure.

Then, we propose, the PRED ‘pro’ specification was lost, leaving the topicality condition in place.

(6) **Stage 2: Loss of PRED ‘pro’, retention of topicality condition**

$$\begin{aligned}
 V_{aff} \quad (\uparrow\text{OBJ}) &= \downarrow \\
 (\downarrow\text{PRED}) &= \text{‘pro’} \\
 (\downarrow_{\sigma}\text{DF}) &= \text{TOPIC} \\
 (\downarrow\text{INDEX PERS}) &= 3 \\
 (\downarrow\text{INDEX NUM}) &= n \quad \text{where } n \in \{\text{SG, DU, PL}\}
 \end{aligned}$$

This explains the modern-day topicality condition on the use of the objective conjugation, and the fact that the objective conjugation suffixes are not pronominal.

Next, the person specification was lost.¹

(7) **Stage 3: Loss of person specification:**

$$\begin{aligned}
 V_{aff} \quad (\uparrow\text{OBJ}) &= \downarrow \\
 (\downarrow\text{PRED}) &= \text{‘pro’} \\
 (\downarrow_{\sigma}\text{DF}) &= \text{TOPIC} \\
 (\downarrow\text{INDEX PERS}) &= 3 \\
 (\downarrow\text{INDEX NUM}) &= n \quad \text{where } n \in \{\text{SG, DU, PL}\}
 \end{aligned}$$

At this stage, the objective conjugation suffixes start being used with first and second person objects as well.

To summarize, we propose the following events in the development of Northern Ostyak: (i) incorporation of third person pronouns; (ii) loss of PRED ‘pro’, retention of topicality; (iii) loss of the person specification. This accounts for the main facts about object agreement in Northern Ostyak.

3.2 Eastern Ostyak

As noted above, one reason to believe that a person specification was present at an earlier stage of Northern Ostyak is that it is reportedly still present in its neighbor, Eastern Ostyak. Gulya’s (1966) grammar of Eastern Ostyak says (p. 115), “the definite [i.e. objective] conjugation ... expresses not only a definite object of *the third person*, but its number as well” (emphasis added). This implies that the objective conjugation appears with third person objects (as in the Northern Ostyak example (2)), but not first or second person objects, unlike Northern Ostyak (recall example (3)).

We hypothesize that Eastern Ostyak instantiates Stage 2 (given in (6)), and that Stage 1, where third person object pronouns were incorporated into the verb, is a common ancestor of Northern and Eastern Ostyak. When these incorporated pronouns lost their PRED ‘pro’ specification, they retained their third person and number specifications. The result was Eastern Ostyak: object agreement in number, and only in third person.

¹Assuming that there are distinct affixes for each number value, this means that person was lost independently in each affix. Support for this idea could come from the existence of dialects where the person specification is lost in some but not all numbers. However, we know of no such evidence.

	SUBJECTIVE	OBJECTIVE
1SG	-ok/-ek/-ök	-om/-em/-öm
2SG	-(a)sz/-(e)sz or -ol/-el/-öl	-od/-ed/-öd
3SG	∅	-ja/-i
1PL	-unk/-ünk	-juk/-jük
2PL	-(o)tok/-(e)tek/-(ö)tök	-játok/-itek
3PL	-(a)nak/-(e)nek	-ják/-ik

Table 3: Verb conjugations in Hungarian

3.3 Hungarian

We posited above that the Northern Ostyak system arose from the loss of the person feature specification from those (formerly) third person pronouns, with the retention of number. Now we suggest that the Hungarian system is also derived from an Eastern Ostyak type system through feature loss, like Northern Ostyak. However, in Hungarian, the *person* specification was retained, and the *number* specification was lost.

Modern Hungarian has two subject-verb agreement paradigms, subjective and objective, shown in Table 3. The choice of conjugation is conditioned by the definiteness of the object. If the object is definite, as in (8), the objective conjugation is used; if the object is indefinite, as in (9), then the subjective conjugation is used.

(8) Lát-om a madar-at
 see-1.SG.DEF the bird-ACC
 ‘I see the bird’

(9) Lát-ok egy madar-at
 see-1.SG.INDEF a bird-ACC
 ‘I see a bird’

Somewhat mysteriously, the use of the objective conjugation is also sensitive to the person of the object. While a third person pronoun triggers the objective conjugation, as in (10), first and second person pronouns do not, as shown in (11).

(10) Lát-ják őt/őket
 see-3.PL.DEF it/them
 ‘They see it/them’

(11) Lát-nak engem/téged/minket
 see-3PL.INDEF me/you/us
 ‘They see me/you/us’

We will refer to this as the *third person restriction* in Hungarian. Note that the Hungarian objective conjugation is not sensitive to the number value of the object.

We suggest that the third person restriction can be viewed as the vestige of an earlier stage at which third person object pronouns were incorporated. To be precise, we suggest the following development. At Stage 1, third person pronouns are incorporated, as suggested above for Northern and Eastern Ostyak (repeated from (5)).

(12) **Stage 1: Pronoun incorporation (only in third person)**

$$\begin{aligned}
 V_{aff} \quad (\uparrow \text{OBJ}) &= \downarrow \\
 (\downarrow \text{PRED}) &= \text{'pro'} \\
 (\downarrow_{\sigma} \text{DF}) &= \text{TOPIC} \\
 (\downarrow \text{INDEX PERS}) &= 3 \\
 (\downarrow \text{INDEX NUM}) &= n \quad \text{where } n \in \{\text{SG, DU, PL}\}
 \end{aligned}$$

Then the PRED 'pro' specification was lost, leaving a topicality condition in its place.

(13) **Stage 2: Loss of PRED 'pro', retention of topicality condition**

$$\begin{aligned}
 V_{aff} \quad (\uparrow \text{OBJ}) &= \downarrow \\
 (\downarrow \text{PRED}) &= \text{'pro'} \\
 (\downarrow_{\sigma} \text{DF}) &= \text{TOPIC} \\
 (\downarrow \text{INDEX PERS}) &= 3 \\
 (\downarrow \text{INDEX NUM}) &= n \quad \text{where } n \in \{\text{SG, DU, PL}\}
 \end{aligned}$$

The idea that PRED 'pro' was lost in Hungarian is not entirely uncontroversial; den Dikken (2006) argues for an analysis of the objective conjugation as an incorporated clitic pronoun. Coppock and Wechsler (under revision) argue against this view, using evidence from binding, islands, and the fact that while clitic doubling is sensitive to semantic properties like definiteness, specificity and topichood, Hungarian object agreement is predictable solely based on the form of the object.² Coppock and Wechsler conclude that the objective conjugation does not represent an incorporated pronoun; rather that there is agreement (in definiteness) between the verb and the object (or that the choice of verb conjugation is *conditioned* by definiteness, as Corbett (2006) suggests).

What we have in Stage 2 is a system like Eastern Ostyak's. One piece of evidence in favor of the idea that Hungarian has such a stage in its history is that the Samoyedic languages have object agreement of the Eastern Ostyak type as well (agreement in number, restricted to third person), and the object agreement systems of Samoyedic are thought to be related to Hungarian's. Although the Samoyedic languages are genetically quite distant from Hungarian, Helimski (1982) reconstructs the objective conjugation in Hungarian to an areal feature shared by the

²This is not strictly speaking true, because there are cases where the same object can take either the objective or the subjective conjugation, with a concomitant change in meaning, e.g. *Ismerem/ismerek néhány titkodat* 'I know-DEF-INDEF some secret-yours-ACC'; 'I know some of your secrets' (cf. Szabolcsi 1994, Bartos 2001). But the fact remains that it is not a semantic property that determines whether or not the objective conjugation may be used.

Samoyedic languages, carefully showing that the common features of their objective conjugations cannot have been present in proto-Uralic.

After PRED ‘pro’ was lost and replaced by a topicality condition, we propose that the number restriction was eliminated, but the person restriction was retained:

(14) **Stage 3: Loss of number**

$$\begin{aligned}
 V_{aff} \quad (\uparrow\text{OBJ}) &= \downarrow \\
 (\downarrow\text{PRED}) &= \text{‘pro’} \\
 (\downarrow_{\sigma}\text{DF}) &= \text{TOPIC} \\
 (\downarrow\text{INDEX PERS}) &= 3 \\
 (\downarrow\text{INDEX NUM}) &= n \text{---where } n \in \{\text{SG, DU, PL}\}
 \end{aligned}$$

This represents a system quite similar to modern Hungarian, where the objective conjugation is restricted to third person objects. However, one difference is that the use of object agreement in Hungarian is conditioned by definiteness rather than topicality. We suggest that the topicality condition was reanalyzed as a definiteness condition, due to the large overlap in distribution between these two features (Givón 1976):

(15) **Stage 4: Reanalysis of topicality as definiteness**

$$\begin{aligned}
 V_{aff} \quad (\uparrow\text{OBJ}) &= \downarrow \\
 (\downarrow\text{PRED}) &= \text{‘pro’} \\
 (\downarrow_{\sigma}\text{DF}) &= \text{TOPIC} \quad (\uparrow\text{OBJ DEF}) = c + \\
 (\downarrow\text{INDEX PERS}) &= 3 \\
 (\downarrow\text{INDEX NUM}) &= n \in \{\text{SG, DU, PL}\}
 \end{aligned}$$

The [DEF] feature is a formal grammatical feature and not a semantic feature. As noted above, Hungarian object agreement is predictable based on the form of the object, but not its meaning: designated forms, such as certain determiners and possessives, are marked as [DEF]. (See Coppock and Wechsler (under revision) for detailed discussion.)

The Stage 4 scenario, depicted in (15), captures the facts of modern Hungarian very well except for one wrinkle: First and second person *reflexive* pronouns trigger the objective conjugation.

(16) (Én) szeret-em magam-at
 I love-3SG.DEF myself-ACC
 ‘I love myself’

(17) (Te) szeret-ed magad-at
 You love-2SG.DEF yourself-ACC
 ‘You love yourself’

Since they are not third person, it will not do to say that the objective conjugation requires a third person object. We propose that the third person restriction was reanalyzed such that reflexive pronouns of all person values count as [DEF]. Meanwhile, for non-reflexive pronouns, third person pronouns count as [DEF] and first and second person pronouns do not.

$$(18) \quad \check{o} \quad (\uparrow\text{PRED}) = \text{'pro'}$$

$$(\uparrow\text{PERS}) = 3$$

$$(\uparrow\text{NUM}) = \text{SG}$$

$$(\uparrow\text{DEF}) = +$$

$$(19) \quad \text{engem} \quad (\uparrow\text{PRED}) = \text{'pro'}$$

$$(\uparrow\text{PERS}) = 1$$

$$(\uparrow\text{NUM}) = \text{SG}$$

Although this complicates the distribution of [DEF], it simplifies the grammar in another respect, by making the verb form sensitive to only one factor (formal definiteness) rather than two (definiteness and person). Hence the final stage has no person restriction equation at all:

(20) **Stage 5: Elimination of person restriction**

$$V_{aff} \quad (\uparrow\text{OBJ}) = \downarrow$$

$$(\downarrow\text{PRED}) = \text{'pro'}$$

$$(\uparrow\text{TOPIC}) = \downarrow \quad (\uparrow\text{OBJ DEF}) =_c +$$

$$(\downarrow\text{INDEX-PERS}) = 3$$

$$(\downarrow\text{INDEX-NUM}) = n \in \{\text{SG, DU, PL}\}$$

This is our proposed representation for modern Hungarian.

Before moving on, we must mention an alternative hypothesis regarding the development of the Hungarian objective conjugation, put forth by Dalrymple and Nikolaeva (to appear). They suggest that there is an earlier stage at which the objective conjugation is used for objects of all persons, as in Northern Ostyak, and then its usage is narrowed to third person objects. Assuming that the objective conjugation suffixes are derived from third person pronouns, this narrowing process involves loss of the pronoun-derived person specification, followed by a later step at which it is added back in. Our derivation has fewer steps, and we prefer it on these grounds.

To summarize, we posit that third person object pronouns were incorporated into verbs, and that feature loss is a mechanism of historical change that is active in the history of Uralic objective conjugations. These assumptions explain the lack of object person agreement in Northern Ostyak, the person restriction in Eastern Ostyak and Samoyedic languages, and the person restriction and lack of object number agreement in Hungarian.

4 Why were only third person pronouns incorporated?

A key assumption underlying our analysis of Uralic is that only third person pronouns were incorporated. This has allowed us to relate the third person restriction in Hungarian to the third person restriction in Eastern Ostyak and Samoyedic languages. Simply put, the objective conjugations are limited to the third person because it is only third person pronouns that were ever incorporated.

This raises the question of why incorporation was originally limited to third person pronouns. One possible explanation builds on the assumption that agreement derives from pronoun doubling constructions, like (21) (the *NP-detachment hypothesis*).

(21) I like him, that guy.

If the doubled NP were first or second person, it would be a pronoun, because first and second person analytic forms are always pronouns. Pronoun doubling of another pronoun may be possible in certain contexts, but it performs a more limited range of functions, as evidenced by the oddness of the following examples:

(22) #I like him, him.

(23) #I like you, you.

As Ariel (1999; 249, fn. 8) points out, “there is hardly any use for first/second persons in NP detachment constructions (cf. **I, I will do it*).”

As a general theory of how pronouns become agreement markers, the NP-detachment view is a bit too simple, because there are quite a few languages in which there is incorporation of first and second person pronouns, to the exclusion of third person pronouns. Incorporation of first and second person pronouns can also be shown to precede incorporation of third person pronouns in some cases (see Ariel 1999 and references cited therein). This motivates Ariel (1999) to posit an alternative hypothesis regarding the process by which pronouns become agreement markers, based on *Accessibility Theory*: high accessibility leads to reduction of a pronoun, and first and second person pronouns are more highly accessible than third person pronouns. It may be, however, that the NP-detachment hypothesis is correct for some cases; indeed, Fuss (2005) argues that the set of processes that can lead from pronoun to agreement marker is heterogeneous. Cases such as ours, in which only third person pronouns incorporate (and the external NP is moreover topical), conform precisely to the predictions of the NP-detachment hypothesis.

To summarize, our explanation for the person restriction in Hungarian relies on the assumption that only third person pronouns were incorporated at an earlier stage of the language. We further speculate that incorporated pronouns may have been restricted to third person because those pronouns participated in NP-detachment constructions. In such constructions, the doubled NP is usually lexical, hence third person, so only third person pronouns incorporated.

5 Other accounts of the third person restriction

The third person restriction on the objective conjugation in Hungarian has intrigued many scholars of Hungarian, and several other explanations have been proposed.

Bartos (2001) suggests an account in terms of ergativity, noting that it is “reminiscent of an ergative-type split ... first and second person pronouns follow a

nominative-accusative pattern, as opposed to third person ones engaging in an ergative-absolutive pattern” (p. 322). There is, however, no evidence of ergativity in Hungarian; pronouns of all persons can be marked with the accusative marker *-t*, and the subject appears in nominative (unmarked) case regardless of the person of the object.

É. Kiss (2005) proposes that the person restriction can be understood using the notion of an inverse system. She posits the following animacy hierarchy for Hungarian:

$$1\text{SG} > 1\text{PL}/2 > 3$$

along with the following constraint:

(24) *Inverse agreement constraint* (É. Kiss 2005)

An object agreeing with a verb must be lower in the animacy hierarchy than the subject agreeing with the same verb, unless the subject represents the lowest level of the animacy hierarchy.

Kiss’s proposal ties together two puzzling facts of Hungarian: the person restriction, and the fact that a special suffix is used with first person singular subjects and second person objects (*-lak/-lek*):

(25) Szeret-lek
 love-1SG.OBJ:2
 ‘I love you’

When the subject is first person singular and the object is second person, the subject outranks the object on the person hierarchy, so object agreement is allowed, according to Kiss’s theory; this accounts for the existence of *-lak/-lek*. On our view, the existence of *-lak/-lek* is independent from the fact that first and second person objects do not count as definite; these two facts must be stipulated individually.

However, we are skeptical of Kiss’s proposal for several reasons. First, in true inverse systems, the use of inverse morphology is determined by the relative rank of subject and object on an animacy hierarchy, but in Hungarian, the presence of object agreement is for the most part conditioned *only* by the features of the object. The rank of the subject does not affect the possibility of object agreement whenever the object is third person, that is, whenever the objective conjugation is used. (This is captured by Kiss’s exception clause “unless the subject represents the lowest level of the animacy hierarchy”, along with the additional implicit condition that in that case, the object may be no higher than the subject.) The use of the objective conjugation is not conditioned by the subject’s person value or indeed any other properties of the subject; only *-lak/-lek* is sensitive to both subject and object.

Furthermore, morphological marking (by an inverse marker) normally appears when the alignment between the animacy hierarchy and the grammatical relation

hierarchy is *non*-harmonic, that is, when there is a *violation* of the alignment constraint. In the case of Hungarian, the morphologically marked case occurs when the constraint is *satisfied*. This makes the putative inverse system in Hungarian a rather atypical exemplar of its category.

Moreover, we know of no other inverse agreement systems among the Uralic languages; Kiss's argument for its plausibility comes from the existence of inverse agreement systems in Chukchee, Koryak and Kamchadal (Comrie 1980), which are not genetically related to Hungarian, although "it does not seem implausible to hypothesize that Proto-Uralic, a distant ancestor of Hungarian, and Proto-Chukotko-Kamchatkan could belong to the same Sprachbund" (É. Kiss 2005; p. 115). However, if geographic influence on Proto-Uralic were responsible for the modern-day properties of Hungarian then we would expect at least some other Uralic languages to show traces of it.

A third explanation for the person restriction in Hungarian was given by Comrie (1977; 10), who claims that because first and second person pronouns are inherently definite, there is no need to mark them explicitly. Dalrymple and Nikolaeva (to appear) criticize this theory as follows:³

This explanation is based on the premise that the primary function of object agreement in Eastern Uralic is the marking of definiteness. However, in all Eastern Uralic languages except modern Hungarian, information structure and not definiteness plays the primary role in patterns of object agreement. The Hungarian situation thus is likely to be secondary, as is also confirmed by the Old Hungarian data.

Dalrymple and Nikolaeva propose instead that it is their typical information structure status, rather than their inherent definiteness, which sets first and second person pronouns apart here. They claim that "the Samoyedic languages (Nenets, Selkup and Nganasan) and Old Hungarian have grammaticalised the tendency for first and second person pronouns to be likely primary topics and unlikely secondary topics." The *primary topic* is the topic that is most topical; a *secondary topic* is a topic that is less topical than the primary topic. Dalrymple and Nikolaeva assume that the subject position tends to align with the primary topic and that the object position tends to align with the secondary topic, following Nikolaeva (2001). Since first and second person pronouns tend to be primary rather than secondary topics, they will tend not to align with the object position. There is a problem with Dalrymple and Nikolaeva's explanation as well, however: the same reasoning could be used to explain the opposite distribution. It is generally the rare case that is more marked, rather than the other way around (Dixon 1994; 85ff).

From the above discussion, it should be clear that the person restriction is a phenomenon that has been tricky to account for. The feature loss view provides a

³The Old Hungarian data to which they refer consists of examples pointed out by Marcantonio (1985) in which the objective conjugation is used with an object that is not marked for definiteness, suggesting that topicality, rather than definiteness, was the factor governing the use of the objective conjugation at that stage.

simple and well-motivated explanation for this notorious puzzle. Another advantage of the feature loss view is that it simultaneously accounts for the lack of object person agreement in Northern Ostyak, to which we will now return.

6 Person agreement: structural versus historical accounts

Recall Baker's claim above that the facts of Northern Ostyak are difficult to account for under the Index/Concord (*dualist*) view of agreement, rightly pointing out that the endings in question should not be analyzed as Concord targets. What we have shown is that these facts are not problematic for the dualist view if Index agreement is characterized historically, rather than as a feature bundle that necessarily includes person. On the view being advanced here, the morphological endings in Northern Ostyak are not Concord targets, but rather Index targets whose person specification has been lost. This implies that sensitivity to the person feature is not a definitional feature of Index agreement; rather, Index agreement should be defined as agreement that historically derives from incorporation of personal pronouns.

An alternative hypothesis comes from Baker (2008). Baker has a *monist* view of agreement, according to which verbal and adjectival agreement are of the same ilk. To account for the special properties of the person feature, Baker proposes a universal structural condition on person agreement:

(26) **Structural Condition on Person Agreement (SCOPA)**

A functional category F can bear the features +1 or +2 if and only if a projection of F merges with an NP that has that feature, and F is taken as the label for the resulting phrase.

This amounts to saying that a target head can agree in person with a first or second person controller only if the controller is the complement or the specifier of the target head. It explains why person agreement is absent from predicative adjectives that lack a structural subject position, as in the French example (1), where a copula is needed to project the subject position.

It is important for Baker's monist approach that this condition apply not only to adjectives but also to verbs. In support of his theory, Baker cites languages where verb agreement normally includes the person feature, but that feature is lacking in special constructions in which the controller's structural relation to the target does not comply with the SCOPA. For example, agreement between a tensed verb and a nominative object in Icelandic is ruled out by the SCOPA, assuming that nominative objects are not in the specifier of the projection corresponding to the verb, as Baker does. And as the SCOPA predicts, Icelandic verbs do not agree in first or second person with nominative objects. While nominative objects can be third person, triggering verb agreement in number ((27)), nominative objects cannot be first or second person ((28) and (29)):

- (27) a. Henni leiddust Peir
her.DAT was.bored.by-3PL they.NOM
‘She was bored with them’
- b. Henni leið-ist bókin sín
her.DAT was.bored.by-3SG book self’s
‘She finds her own book boring’
- (28) *Henni leidd-umst við
her.DAT was.bored.by-1SG we.NOM
‘She was bored with us.’
- (29) *Ég veit að honum lík-ið þið
I know that him.DAT like-2PL you.NOM.PL
‘I know that he likes you all.’

The case of Icelandic is slightly different from the case of predicate adjectives, however. With predicate adjectives, the target shows the same default person value across all controllers, while agreeing in number (this is *partial agreement*). In the case of Icelandic, the SCOPA-violating configuration with first and second person controllers is *ineffable*. Baker’s (2008) theory does not specify which of these two outcomes will occur when the SCOPA is violated, although Baker (to appear) proposes that the variation is related to case assignment.

The dualist hypothesis, in contrast, is capable of predicting when partial agreement will occur. The dualist hypothesis does not allow for the type of partial agreement where the agreement features encoded by a target form can be selectively ignored depending on the syntactic structure. On the dualist view, ϕ -features are bundled on Index and Concord target morphemes, and agree in an ‘all-or-none’ fashion (c.f. Chomsky 2001, Béjar 2008). It allows for person insensitivity only when person distinctions are entirely absent from the target morpheme’s paradigm. Two kinds of target morpheme can lack person: (i) Concord morphemes (for example, those on French adjectives); (ii) Index morphemes from which person distinctions have been historically lost (as in Northern Ostyak object agreement). Hence the dualist hypothesis predicts:

- (30) If a given target morpheme reflects the person feature of the controller in any controller position, then that target morpheme will reflect the controller’s person in every controller position.

The ineffability cases do not violate the prediction in (30), and are thus consistent with the dualist hypothesis, because the third person endings in (27) can be seen as displaying full number and person agreement with their third person controllers. Partial agreement, where person agreement morphology is available but the target morpheme agrees only in number, violates (30), because this constitutes a dissociation of the features that are hypothesized to be bundled together. According to (30),

Language	Target	Trigger	Pers. Morph.?	Result
Swahili, etc.	A	subjects	no	partial
English ‘dialect’	V	Spec, CP	yes	both
Passamaquoddy	V	Spec, CP	yes	ineffable
Icelandic	V	nom. obj.	yes	ineffable
Gujarati	Aux	acc. obj.	yes	none
Chicasaw	V	nom. obj.	yes	ineffable
Nahuatl, etc.	V	2nd obj.	yes	ineffable
Lokaa	V	obj. of gerund	yes	ineffable
Basque	V	controlled subj.	yes	ineffable
Passamaquoddy	V	embedded subj.	yes	ineffable
Northern Ostyak	V	acc. obj.	no	partial
Sakha	V	Spec, CP	yes	partial

Table 4: SCOPA-violating configurations

partial agreement should be possible only when person agreement morphology is absent from the relevant paradigm.

Table 4 lists the SCOPA-violating agreement configurations that Baker (2008, to appear) discusses, and classifies each according to the category of the agreeing head (“Target”), the location of the controller (“Trigger”), whether the agreeing head can inflect for person (“Pers(on) Morph(ology)?”), and whether the SCOPA violation results in ineffability or partial agreement (“Result”). (In the case of the English ‘dialect’, there is partial agreement for some person/number values of the controller and ineffability for other ones, which is why it is assigned the value ‘both’; in Gujarati, the auxiliary verb shows neither number nor person agreement with the target, so it is assigned ‘none’.) The crucial columns are the last two: Baker’s theory allows for languages in which person morphology is available but there is partial agreement, but the dualist theory predicts that this should not happen.

In all but one case, the prediction in (30) is straightforwardly met. For every line with a “yes” in the “Pers(on) Morph(ology)?” column, “partial” is *not* the value of the “Result” column— except for Sakha (Turkic). The case of Sakha requires closer inspection.

In Sakha ECM constructions, the predicate can lack the usual person/number inflection, registering plural number with the *-LAR* suffix, regardless of person (Baker to appear). Person agreement in these constructions is optional; example (31) shows the ECM construction with full person/number agreement; (32) shows the crucial variant with reduced agreement.

- (31) Min ehigi/ehigi-ni бүгүн kyaj-yax-xyt dien erem-mit-im.
I you/you-ACC today win-fut-2PL COMP hope-PAST-1SG
‘I hoped that you would win today.’

- (32) Min ehigi-ni/*ehigi bugün kyaj-yax-tar-a dien erem-mit-im.
 I you-ACC/you today win-past-PL-3 COMP hope-PAST-1SG
 ‘I hoped that you would win today.’

It is not clear whether this is a real counterexample to the prediction in (30). The first question is whether the plural suffix *-Lar* (of which *-tar* in (32) is an allomorph) is in the same person/number paradigm as the first and second person morphemes. The suffix *-Lar* has a number of properties distinguishing it from the other agreement affixes. It marks plural on nouns, suggesting it is a Concord target. It is optional, in the sense that an unmarked element has general number, i.e. either singular or plural (Vinokurova 2005; p. 144); on the basis of such optionality, among other reasons, Hahm (2010) argues that the cognate morpheme in Turkish is not agreement at all, but rather an indicator of semantic plurality. Also, *-Lar* has a historical source distinct from that of person/number inflection (Hahm 2010, citing Adamović 1985; p. 27 and Good and Yu 2005). We tentatively suggest the plural suffix is either not agreement at all, or else a Concord target morpheme that does not derive from pronoun incorporation. The suffix *-a*, glossed as third person, may be an Index target, perhaps a default form arising in the absence of person agreement. But this would not be partial Index agreement, but rather the absence of any Index agreement at all—in which case *ehigi-ni* ‘you-ACC’ is not really a controller. Because number and person are encoded in separate morphemes, they may in principle have different controllers, or one may have a controller while the other lacks one. A clear counterexample to the dualist hypothesis would involve a single morpheme that simultaneously encodes number and person. A more detailed study of Sakha and its ECM construction is needed before we know definitively whether it falsifies the predictions of the dualist hypothesis.

Although Sakha requires further investigation, all of the cases listed in Table 4 corroborate the prediction that partial agreement is possible only when the target affix fails to encode person morphologically. Thus, the dualist view is capable of predicting when partial agreement, as opposed to ineffability, will occur. In this respect, our dualist theory is stronger than Baker’s (2008) theory.⁴

7 Conclusion

The special distribution of person agreement can be largely explained as a consequence of its historical origin in incorporated pronouns. Synchronically, this is reflected in the distinction between Index targets and Concord targets; only the former are historically derived from pronouns, so only the former signal the person of their controller. Exceptions to that generalization arise when the person feature is lost over time, as we suggested for the historical development of object agreement in the Uralic languages. Synchronically, this means that Index targets do not

⁴Baker (to appear) does offer a way of explaining this variation based on case assignment, and further research is needed to distinguish between these two views.

always constrain the person feature of their controller. This view accounts for the main features of the Uralic objective conjugations, including the mysterious third person restriction in Hungarian.

We cannot offer any explanation for why person was lost and number was not in Northern Ostyak, but feature loss does not seem to be a process restricted to person; Toivonen's study of Finnish possessive affixes showed that feature loss can target both person and number. A typological investigation of feature loss and its constraints is currently being undertaken by the first author (Coppock 2010).

The dualist view of agreement, according to which Index and Concord are distinct feature bundles, leads to the prediction that person agreement can be dissociated from number agreement only when the feature specification has been lost from the target affix. In contrast to Baker's (2008, to appear) theory, it predicts that the presence of person agreement should not depend on the structural relationship between the controller and the target. We have shown preliminary typological evidence that this prediction is borne out.

References

- Adamović, M. (1985). *Konjugationsgeschichte der türkischen Sprache*. E. J. Brill, Leiden.
- Ariel, M. (1999). The development of person agreement markers: From pronouns to higher accessibility markers. In Barlow, M. and Kemmer, S., editors, *Usage-based Models of Language*, pages 197–260. CSLI Publications, Stanford.
- Baker, M. C. (2008). *The Syntax of Agreement and Concord*. Cambridge University Press, Cambridge.
- Baker, M. C. (to appear). When agreement is for number and gender but not person. *Natural Language and Linguistic Theory*.
- Bartos, H. (2001). Object agreement in Hungarian: A case for minimalism. In Alexandrova, G. M. and Arnaudova, O., editors, *The Minimalist Parameter: Selected Papers from the Open Linguistics Forum, Ottawa, 21-23 March 1997*, volume 192 of *Current Issues in Linguistic Theory*, pages 311–324. John Benjamins, Amsterdam.
- Béjar, S. (2008). Conditions on phi-agree. In Harbour, D., Adger, D., and Béjar, S., editors, *Phi Theory: Phi-Features across Modules and Interfaces*. Oxford University Press, Oxford.
- Bresnan, J. and Mchombo, S. (1987). Topic, pronoun, and agreement in Chicheŵa. *Language*, 63:741–782.
- Butt, M. (2007). The role of pronominal suffixes in Punjabi. In Zaenen, A., Simpson, J., and King, T. H., editors, *Architectures, Rules and Preferences*, pages 341–368. CSLI Publications, Stanford.

- Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M., editor, *Ken Hale: A Life in Language*, pages 1–52. MIT Press, Cambridge, MA.
- Comrie, B. (1977). Subjects and direct objects in Uralic languages: A functional explanation of case-marking systems. *Études Finno-Ourgriennes*, 12:5–17.
- Comrie, B. (1980). Inverse verb forms in Siberia: Evidence from Chukchee, Koryak and Kamchadal. *Folia Linguistica Hungarica*, 1:61–74.
- Coppock, E. (2010). Pathways from pronoun to agreement and their destinations. Post-doctoral research project funded by the Swedish Research Council.
- Coppock, E. and Wechsler, S. (under revision). The objective conjugation in Hungarian: Agreement without phi-features. University of Texas at Austin.
- Corbett, G. (2006). *Agreement*. Cambridge University Press, Cambridge.
- Dalrymple, M. and Nikolaeva, I. (to appear). Objects and information structure. Unpublished book manuscript.
- den Dikken, M. (2006). When Hungarians agree (to disagree): The fine art of ‘phi’ and ‘art’. CUNY Graduate Center.
- Dixon, R. M. W. (1994). *Ergativity*. Number 69 in Cambridge Studies in Linguistics. Cambridge University Press.
- É. Kiss, K. (1987). *Configurationality in Hungarian*. Akadémiai Kiado.
- É. Kiss, K. (2005). The inverse agreement constraint in Hungarian: A relic of a Uralic-Siberian Sprachbund? In *A Festschrift for Henk van Riemsdijk*. John Benjamins, Amsterdam.
- Fuss, E. (2005). *The Rise of Agreement: A Formal Approach to the Syntax and Grammaticalization of Verbal Inflection*. John Benjamins, Amsterdam.
- Givón, T. (1976). Topic, pronoun and grammatical agreement. In Li, C. N., editor, *Subject and Topic*, pages 149–188. Academic Press, New York.
- Good, J. and Yu, A. C. L. (2005). Morphosyntax of two Turkish subject pronominal paradigms. In Heggie, L. and Ordóñez, F., editors, *Clitic and Affix Combinations: Theoretical Perspectives*, pages 315–342. John Benjamins, Amsterdam.
- Gulya, J. (1966). *Eastern Ostyak Chrestomathy*. University of Indiana Press, Bloomington.
- Hahm, H.-J. (2010). *A Cross-linguistic Study of Syntactic and Semantic Agreement: Polite Plural Pronouns and Other Issues*. PhD thesis, University of Texas at Austin.

- Hajdú, P. (1968). *Chrestomathia Samojedica*. Tankönyvkiadó, Budapest.
- Helimski, E. A. (1982). *Vengersko-samodijiskije lexičeskije paraleli*. Nauka, Moscow.
- Honti, L. (1984). *Chrestomathia Ostiacica*. Tankönyvkiadó, Budapest.
- Kálman, B. (1965). *Vogul Chresomathy*, volume 46 of *Indiana University Uralic and Altaic Series*. Indiana University Press, Bloomington.
- Keresztes, L. (1989). *Chrestomathia Mordvinica*. Tankönyvkiadó, Budapest.
- Kortvély, E. (2005). Verb conjugation in Tundra Nenets. *Studia Uralo-Altaica*, 46.
- Lehmann, C. (1988). On the function of agreement. In Barlow, M. and Ferguson, C. A., editors, *Agreement in Natural Language*. CSLI Publications, Stanford.
- Marcantonio, A. (1985). On the definite vs. indefinite conjugation in Hungarian: A typological and diachronic analysis. *Acta Linguistica Scientiarum Hungaricae*, 35:267–298.
- Nikolaeva, I. (1999). Object agreement, grammatical relations, and information structure. *Studies in Language*, 23:331–376.
- Nikolaeva, I. (2001). Secondary topic as a relation in information structure. *Linguistics*, 39:1–41.
- Stassen, L. (1997). *Intransitive Predication*. Oxford Studies in Typology and Linguistic Theory. Clarendon Press, Oxford.
- Szabolcsi, A. (1994). The noun phrase. In Kiefer, F. and É. Kiss, K., editors, *The Syntactic Structure of Hungarian*, volume 27, pages 179–274. Academic Press, New York.
- Toivonen, I. (2001). Language change, lexical features, and Finnish possessors. In Butt, M. and King, T. H., editors, *Time Over Matter*, pages 209–225. CSLI Publications, Stanford.
- Vinokurova, N. (2005). *Lexical categories and argument structure: a study with reference to Sakha*. PhD thesis, University of Utrecht.
- Wechsler, S. (to appear). Mixed agreement, the person feature, and the Index/Concord distinction. *Natural Language and Linguistic Theory*.
- Wechsler, S. and Zlatić, L. (2003). *The Many Faces of Agreement: Morphology, Syntax, Semantics, and Discourse Factors in Serbo-Croatian Agreement*. CSLI Publications, Stanford.