Anna Volkova* and Eric Reuland

Reflexivity without reflexives?

Abstract: What prevents pronominals from being locally bound? Does this a) reflect an intrinsic property of pronominals (Chomsky 1981), is it b) a relative (economy) effect, that only shows up where there is a more dedicated competitor (see from different perspectives, Safir 2004, Boeckx, Hornstein and Nunes 2007, Levinson 2000), does it c) have a semantic basis as in Schlenker (2005), or does it d) follow from general conditions on agree based chains, and reflexive predicates (Reuland 2011a)? To help resolve this issue, we investigate Khanty (Uralic, spoken in Northwest Siberia), a language that is reported to allow locally bound pronominals (Nikolaeva 1995), and assess whether it in fact does have them, and, which factors come into play when local binding obtains.

Keywords: pronominals, reflexivity, binding

1 Introduction

In her very informative description of the Uralic language Khanty, Nikolaeva (1995, 1999a) observes that Khanty has no dedicated reflexive pronouns; instead, personal pronouns are used, as illustrated in (1):

(1) Utłtiteχo$_i$ łuveł$_{i/k}$ išək-s-əłłe.
    teacher   he.ACC   praise-PST-SG.3SG
    The teacher praised him(self).

The object pronoun *łuveł* in (1) can be either bound by the subject of its clause, or receive a value from the discourse. This observation is intriguing from a typological perspective, since it clearly differentiates Khanty from other Uralic languages (see Volkova 2014). It is also intriguing from a theoretical perspective, as it goes against a standard principle of the canonical binding theory (CBT), and also

*Corresponding author: Anna Volkova: National Research University Higher School of Economics. Utrecht institute of Linguistics OTS, Universiteit Utrecht. E-mail: a.volkova@uu.nl
Eric Reuland: Utrecht institute of Linguistics OTS, Universiteit Utrecht. E-mail: e.reuland@uu.nl

poses a challenge for other theories, such as Reflexivity Theory (Reinhart and Reuland 1993, Reuland 2011a). Condition B of the CBT requires pronominals to be free in their local domain (Chomsky 1981), which would rule out a reading with the subject as the antecedent of the pronoun:

(2) Binding condition B
    A pronominal must be free in its local domain[1].

The challenge to Reflexivity theory is that the predicate *praise* in (1) allows a reflexive interpretation, which, prima facie, is not licensed, contrary to what the theory requires.

This puts Khanty right in the middle of a theoretical debate. What is the status of condition B, or conditions on reflexivity? Are there limits to cross-linguistic variation? And if so, what are they? Are there commonalities in the restrictions languages impose on binding relations, and if so, what do they come from? What do the Khanty facts imply for the choice between theories?

Khanty is not the first language that has been found to pose a challenge to extant approaches to binding. In a range of Germanic and Romance languages one finds local binding of 1st and 2nd person pronouns, and, as already discussed in Everaert (1986), Frisian, like Old English (see Van Gelderen 2000), allows locally bound 3rd person pronominals. Also in French, one finds bound pronominals that might qualify as local (Pica 1984, 1986, Zribi-Hertz 1995), as well as in Haitian Creole (Zribi-Hertz and Glaude 2007).

Despite such prima facie exceptions, there is a common view among language typologists that the use of special strategies is the 'norm' (Moyse-Faurie 2008, Heine and Miyashita 2008). Language after language appears to use special strategies to express reflexivity (see Faltz 1985, Lust et al. 2000, Genušiene 1987, Schladt 2000 for overviews), with a rough division into nominal and verbal strategies (although it is not always trivial to tell them apart). One finds SELF-reflexives, as in English *himself*, Dutch *zichzelf*, Norwegian *seg selv*, Icelandic *sjalfan sig*, etc. One finds body-part reflexives as in Georgian *tav tavis*, Basque *bere burua*, Indonesian *dirinya sendiri*, Yoruba *ara rè*, and many other languages, or doubling of a pronominal expression as in Malayalam *taan tanne*, or Tsaxur *wuž-e: wuž*. One also finds 'reflexive clitics', as in Romance (Italian, French, etc.) or Slavic (Serbo-Croatian, Czech, Polish).

Many languages have more than one way of expressing reflexivity. With some verbs English allows a bare form. In Dutch and Scandinavian one also finds a

---

1 Roughly, the domain of its nearest subject.

simplex reflexive (*zich*, *seg*, etc.), alongside the complex one. In Russian some verbs allow reflexivization with the affix *sja*, and the tonic reflexive *sebja*. One also finds two strategies in a language like Sakha (Vinokurova 2005). A well-known example of a language with verbal reflexives is Kannada (Lidz 1995). In an overview Franssen (2010) shows that around 45% of the Australian languages derive reflexives by some form of verbal affixation.

Given this, it seems highly unlikely that the need to use special strategies is accidental, a position that is shared by most approaches to binding. The question is, then, why the use of special strategies would be the norm, and how to deal with the diversity. Are there non-trivial language universals in this domain, and if so, what would hide them from our view where their manifestation is less apparent?

In recent years this goal of identifying language universals has been challenged by researchers from the cognitive sciences, for instance by Evans and Levinson (2009, E&L). For Evans and Levinson, language is a 'bio-cultural hybrid', and languages can vary as much as human cultures can. The extent of variation among languages is such that the search for universals is misguided. For Levinson (2000), the employment of special strategies in binding is just a tendency, reflecting pragmatic preferences, where some languages simply have not yet developed the tools to express these preferences.[2]

It is not the place here to provide a general response (see Reuland and Everaert 2010, and many contributions in the same Lingua volume for detailed responses). What is important, though, is to acknowledge the challenge posed by linguistic diversity and take it seriously.

Clearly, many properties of languages are idiosyncratic, as for instance the arbitrary form-meaning combinations one finds in the lexicon. So, there is no deep reason why the conventional English word for table would be the loan *table* rather than *stool*, or its Russian equivalent would be *stol,* and not just a loan, or would not be from a totally different root. However, if we find properties that language after language appears to share, irrespective of how distant they are in terms of historical relation, these are in need of explanation. Quite plausibly, such universal properties provide a window into the nature of the language faculty. Hence, the quest for language universals is an important part of linguistic theory.

So, the fundamental question one has to face with respect to the challenge posed by E&L, is whether the cross-linguistic variation in binding within the local domain is of the same nature as the variation in the words for table, or chair,

---

2 Unlikely, though, this is given the common time scale of language development.

namely just a reflection of the fundamental arbitrariness of form-meaning map-
pings, or whether there are significant commonalities reflecting fundamental
principles of grammar.

There is only one way to resolve this question. Namely to work on a case by
case basis. To carry out detailed investigation of languages with patterns that
appear to defy the possibility of generalization. To determine whether the 'excep-
tions' are real, or whether appearances are deceptive and there are factors that
initially escaped observation.

In order to do so effectively we have to use articulate hypotheses to guide the
investigation. So, we have to face three related questions. One, is there a *reason*
to expect that the employment of special strategies to express reflexivity is uni-
versal? Second, is this in fact *true*? And, thirdly, how precisely do these strategies
work (and why is their presence not always apparent)?

Khanty is the main subject of the present article. Starting out from the obser-
vation in (1), we address the challenge it poses, but we also place its properties in
a broader context.

In doing so, we will follow the heuristics proposed in the minimalist program
(Chomsky 1995 and subsequent work). In order to achieve explanation we should
start from as few assumptions as possible. If the assumptions we start out with
are too poor, we will find this out in the course of empirical investigation. If we
start out with an overly rich system, we run the risk of never finding out that a
simpler system would have done the job as well. Thus, we set out to explain the
properties of languages in terms of what it minimally takes to have a language
and basic properties of the human cognitive system ('third factors', Chomsky
2008). This strategy we will apply to binding, taking the position that binding-
specific properties should be posited only as a last resort.

2 The issue

In order to be able to answer our questions, we will start being precise about the
notion of reflexivity.[3] Consider, then some binary predicate head P, e.g. *defend*. In
its transitive use – *Jack defends Jill* – *defend* assigns an agent role to its subject
Jack and a patient role to its object *Jill*. In its reflexive use – *Jack defended himself*
– both roles end up on the same argument, here Jack (with some proviso to be
discussed later). So, in the case of *Jack defended himself* we have a reflexive use of

---

3 Staying close to its formal definition. Formally we can say that a binary relation R on a set S is
reflexive iff for every element x of S, xRx.

the predicate *defend*.[4] Compare this to the predicate *wash*, and take it again in its transitive use – *Jack washed Jill.* For this predicate we have two options for a reflexive use, namely *Jack washed himself*, and *Jack washed*. In both cases one participant plays two roles, the washer and the washee.

Thus, the basic notion of reflexivity we have is the following: *A predicate is reflexive iff one semantic argument bears two of the predicate's semantic roles.* Clearly, this excludes any configuration from qualifying as reflexive where what is semantically the same argument is assigned roles from two different predicates. Consider, for instance the case of possessor binding as in Russian *Masha ljubit svoju koshku* 'Masha loves her cat'. Here there is a dependency between *Masha* and *svoju*, and traditionally expressions such as *svoj* are indeed called possessive reflexives. However, obviously, the sentence does not denote a reflexive loving-relation, and no reflexive predicate is formed.[5]

Naturally, whether or not this is the right way to cut the pie is an empirical matter. From a cross-linguistic perspective, however, one must conclude the linguistic means involved in the representation of these two types of dependencies are rather different. For instance, there are many languages that dissociate them, requiring dedicated reflexives for locally bound objects, but allowing simple pronominals as possessors (as in English). Consequently, also the type of explanation they require must differ. In this article we will focus on reflexivity, but where relevant we will address other issues of binding as well.

3 Representing reflexivity

One general way of bringing about that two semantic roles of a predicate end up on the same semantic argument is by identifying two co-argument expressions each bearing a semantic role. This is an effect of binding. There has been considerable discussion in the literature of the notion of binding. In much of the literature one finds a definition based on Chomsky (1981):

(3) *a* binds b iff *a* and *b* are coindexed and *a* c-commands *b*

---

4 Strictly speaking we must distinguish between a predicative head in the lexicon, and its use with arguments in a sentence. This distinction is reflected in the definition of a predicate in Reinhart and Reuland (1993): A [. .] predicate formed of (a head) P, is P and all it arguments [. . .]. Where no misunderstanding arises we will simply use the term *predicate P* both for the head P and for the *predicate formed of the head P.*

5 It is important to note that the definition of reflexivity we are using is neutral as to the specific manner in which it is expressed. More specifically it covers both 'nominal' and 'verbal' reflexivization. We are grateful to one reviewer for making us clarify this point.

However, as already shown in Reinhart (1983) the notion of co-indexing is problematic (see Reuland (2011b) for a recent overview of the problems). Chomsky (1995) concludes that indices are not possible morpho-syntactic objects, and therefore they violate the 'inclusiveness condition, which limits syntactic computations to morpho-syntactic objects' in the numeration. Hence co-indexing is not available as a means to encode dependencies. This necessitates a reassessment of the binding theory, since even the definition of binding in Chomsky (1981) is no longer applicable. In view of this, Reinhart (2006) proposes the following *definition* of binding, reducing 'linguistic' binding to the logical notion of binding we need anyway:

(4) a. *a* A-binds *b*, iff *a* is the sister of a λ-predicate whose operator binds *b*.
     b. a (λx (P (x . . . . . . x)))

So, in (4), *a* A-binds *x*, and more concretely, binding in the case of (5a) is represented as in (5b), or the long-distance binding in (6a) as in (6b):

(5) a. Masha ljubit svoju koshku 'Masha loves her cat'
     b. Masha (λx (loves (x [x's cat])))

(6) a. Julia nadeetsja chto John ljubit ee 'Julia hopes that John loves her'
     b. Julia (λx (x hopes [that John loves x]))

At this point we need not go into the technicalities of the mapping. The intuition behind (5) and (6) is that anaphors such as *svoj* or pronominals such as *ee* or *her* are translated as variables in logical syntax. The subject (*Masha* or *Julia*) undergoes quantifier raising, leaving a trace, that in turn is translated as a variable, which is identified with the variable translating the anaphors/pronouns, making binding in the sense of (4) possible (see, for instance Heim and Kratzer (1998), or Büring (2005) for formal details.) The result is that the roles of being the experiencer of the love and being the possessor of the cat or that of being the experiencer of the hope and the source of the love are being borne by the same arguments. Note, that as will turn out to be important later, binding dependencies may differ in their syntactic status. The bound interpretation of *ee* in (6a) is possible, but not necessary, since *ee* may also refer to an individual in discourse. Hence, the dependency is not syntactically encoded, but just represented at the level of logical form (logical syntax in the sense of Reinhart 2006). In cases of anaphor binding, however, binding is enforced, and represented syntactically, as we will see in more detail in section 5.

Taking this as our cue it would seem that the most straightforward representation of a reflexive predicate would be as in (7), where both the subject argument and the object argument are bound by the same λ-operator.

(7) DP (λx (V x x))

Furthermore, given that pronominals can be translated as variables, one would expect that a common source of (7) would be (8a), instantiated by sentences as in (8b), etc., with the DP binding the pronominal, indicated by italics:

(8) a. *DP* Verb *Pronominal*
     b. *Julia defended *her*
     c. *Viktor nenavidel *ego*   'Viktor hated him'
     d. *Frans kende *hem*   'Frans knew him'

As we already noted, though, this is not how it generally works. As a rule one doesn't find the configuration of a subject a verb and a simple pronominal object as in (8a) with a reflexive interpretation. Hence, some factor must prevent 'brute force reflexivization' (BFR), that is, one argument binding a co-argument without further ado. The standard interpretation rules we need independently, must be prevented from mapping the expressions (8b–d) onto the representation in (7). This was the role of condition B of the CBT.[6]

However, as we saw, Condition B as stated in (2), cannot be maintained. It is empirically inadequate as a universal, and it relies on co-indexing, which is theoretically untenable. Thus, we must replace it by principles that are independently motivated, and are ideally also able to capture the cross-linguistic variation. The task, then, is two-fold. The theoretical task is to replace the notion of co-indexing with mechanisms that are compatible with minimalist guidelines,

---

6 An anonymous reviewer wonders why the discussion is mostly restricted to subject-verb-object configurations. This is a consequence of the well-known obliqueness hierarchy in natural language, see already Keenan and Comrie (1977) for a hierarchy in relativization positions. Testelets and Toldova (1998) based on a survey of anaphora in East Caucasian languages proposed a hierarchy of the positions a reflexive bound by the (matrix) subject can occupy in a sentence. The hierarchy presented in (i) predicts that if a language uses dedicated reflexives for indirect objects it uses them for direct objects as well, but not vice versa.

(i) DO_{2place} > $DO_{ditrans}$ > Coargument > Non-coarg. > $DP_{non-fin}$ > DP_{fin}

Thus, the first case represents the core case of reflexivization. The further removed the more interference of other factors is to be expected. See, for instance the analysis of local binding in French, as discussed in Reuland (2011a).

the empirical task is to determine the factors that enter into the variation, and to accommodate them in the simplest possible manner.

4 Deconstructing binding

For a starter it is important to see that the conditions on binding – whatever underlies them – do not constitute a unitary phenomenon. This is straightforwardly demonstrated by the contrast between Dutch and Frisian (but recapitulated in many languages). The relevant pattern is illustrated in (9) and (10). Dutch has two anaphor types, simplex anaphors (SE-anaphors, *zich*), and complex anaphors (*zichzelf*). The element *zich* is specified for person, but not for number and gender. The choice of *zich* versus *zichzelf* depends on the type of predicate: agent-theme verbs generally allow *zich*, other verbs such as subject experiencer verbs require *zichzelf*. ECM subjects can be *zich*. 3rd person pronominals cannot be locally bound.

(9) Dutch
    a. *Winnie* waste *zich/*hem*   SE-anaphor
        Winnie washed
    b. *Winnie* bewonderde *zich*(zelf)/*hem(zelf)*   *complex anaphor*
        Winnie admired himself
    c. *Winnie* voelde [*zich/*hem* wegglijden]   SE-anaphor
        Winnie felt [SE slip away]

Comparing this pattern to Frisian we see that one factor is constant, the other changes:

(10) Frisian
    a. *Winnie* waske *him*   pronominal instead of SE-anaphor
        Winnie washed
    b. *Winnie* bewûndere *him*(sels)*   *complex anaphor*
        Winnie admired himself
    c. *Winnie* fielde [*him* fuortglieden]   pronominal instead of SE-anaphor
        Winnie felt [SE slip away]

These paradigms show that there is a distinction between: i. properties of predicates and their effects, and ii. properties of 'pronouns' and their effects. Frisian requires a complex anaphor just where Dutch does, but it has a bound pronominal just where Dutch has a SE-anaphor. The fact that one factor varies, whereas

another one does not, is enough to establish the point that the conditions on binding don't constitute a unitary phenomenon. Consequently, we have to look for an explanatory model that captures the distinction. We will first discuss the SE-anaphor/pronominal contrast in section 5. We will see that it reflects the way in which anaphoric dependencies are represented in syntax. In section 6 we will discuss the relevant properties of predicates.

5 Encoding binding in syntax

SE-anaphors are essentially pronominals with a deficient specification for φ-features. *Zich* (and its Scandinavian counterparts *seg*, and *sig*) are all specified for 3rd person, but have unvalued number and gender features.[7] Two questions come up: i. How is the dependency in (9a,c) encoded? ii. Why is a 3rd person pronoun instead of *zich* ruled out in Dutch, but not in Frisian?[8]

5.1 The formation of feature chains

First consider the issue of encoding (note that these questions are extensively discussed in Reuland (2011a: Chapter 5) to which the reader is referred for details). Under minimal assumptions, syntax has precisely one way of representing identity, namely by the *y is a copy of x*-relation. This is the relation that underlies movement as in *The men were admired (the men)*; this relation also underlies Agree, which copies feature values. That is, Agree copies features from a source and uses these to fill/overwrite features in the matrix of a (partially) empty target. This is technically implemented in Pesetsky and Torrego's (P&T, 2007) theory of feature chains. Agree results in literal identification of features (in P&T's terms, turning 'occurrences' into 'instances').[9]

---

7 It is crucial to distinguish between 'having a feature that is unvalued' and various forms of syncretism. See Reuland (2011a: 137–138) for detailed discussion.

8 There is a further question: Why does *zich* have to have an antecedent? In fact, nothing intrinsic requires an element with a deficient φ-feature specification to acquire a full specification. It is just the syntactic environment that mechanically enforces the chain (or blocks it, as is illustrated by the role of subjunctive in licensing Icelandic logophoric *sig*, see Reuland and Sigurjonsdottir 1997, Reuland 2011a).

9 The notion of identity at stake is illustrated by difference between *The men were admired (the men)*, and *The men admired the men*, where the two occurrences of *the men* are easily interpreted as referring to different groups of men in the latter case, but not in the former.

Some of the relevant 'abstract' properties of the configuration in which Agree can encode a binding dependency by creating a feature chain are indicated in (11), with the subject/external argument (EA) as the prospective binder in the low position before movement to the T-domain:

(11) [Tns [EA [v* [V [SE . . . .]]]]]

The process of chain formation is set in motion by the structural Case feature, since structural Case is an independent syntactic 'trigger' in this environment. The chain is formed by check/agree via a series of probe-goal relations, essentially based on structural Case being analyzed as uninterpretable Tense. Thus, the φ-feature dependency gets a free rider on the Case dependency. The dependencies are summarized in (12), with the subject EA providing the required valued and interpretable instance of [φ] (where *u* stands for unvalued and *val* for valued) (see Reuland 2011a for more detail):

(12) $[\text{Tns}_{u\varphi} \, [\text{SE}_{u\varphi} \, [\text{EA}_{\textbf{val}\varphi} \, [v^{\star}_{\; u\varphi} \, [\, V \, (\text{SE}_{u\varphi}) \, . \, . \, . \, . \,]]]]]$

In P&T's approach the exchange of values in the formation of a feature chain unifies the features it contains. By valuation, feature values are copied/ overwritten, hence all the tokens of φ in (12) share instances of one feature. Just as with copies in dislocation structures, copying/overwriting of feature values encodes identity. Hence, we do have a syntactic representation of the binding relation.

Why is this allowed with SE-anaphors and not with pronominals? This is captured by a general *condition on chains*. Overwriting features amounts to deleting the features that are overwritten. Overwriting is subject to a general principle of recoverability of deletions (PRD, Chomsky 1995). Informally, a occurrence of a feature cannot be overwritten by another occurrence of that feature unless this doesn't limit interpretive options. A SE-anaphor and its antecedent only share interpretive constants (category, person). But the pronominal *he* in let's say *Mark hoped he would win* may be interpreted as Mark, but also as someone else. This entails that *Mark* and *he* cannot become members of the same syntactic chain (although they can become co-valued by an interpretive process). Consequently, with SE in the position of (12) true identity of feature bundles can be effected; with a 3rd person pronominal in the position of SE in (12) no chain can be formed.[10]

---

10 This reflects a crucial contrast with 1st and 2nd person pronouns. Their interpretation is constant (modulo proxies), within one reportive context, hence in their case chain formation is not blocked by the PRD. This explains why so many languages do allow local binding of 1st and 2nd

This answers the question of why the pronominal cannot enter a chain. It leaves open, however, why this option cannot be bypassed, by simply having the pronominal bound in logical form by the procedure sketched in section 3. That is, why don't we have (13b) alongside (13a), with *hem* semantically bound by *Winnie*, as in (13c)?

(13) a. *Winnie* voelde [*zich* wegglijden]
       Winnie felt [SE slip away]
     b. *Winnie voelde [*hem* wegglijden]
       Winnie felt [him slip away]

The answer resides in economy. The envisaged interpretation is (14):

(14) Winnie (λx. (x felt [x slip away]))

Deriving this representation from (13b) by encoding the dependency in the syntax is blocked as a violation the PRD. PRD is a fundamental principle of grammar. Any 'attempt' to violate it – i.e. creating a configuration that would lead to its violation – entails that the derivation is cancelled (Chomsky 1995). If a derivation is cancelled all other options to continue this derivation are blocked. More specifically, the possibility to bypass the prohibition of chain formation by semantic binding is blocked: 'Rejection is final' (Reuland 2011a). Note, that we don't have to compare the *zich*- and the *hem*-options, invoking global economy. The *hem*-option is just blocked in its own right. This leads to the next question: how to account for the Dutch-Frisian contrast?

5.2 The Dutch-Frisian contrast

The main question from our present perspective is the following: Given this fundamental impossibility of deriving (14) from (13b), why is it possible to derive (14) from the Frisian (10c), repeated here as (15)?

(15) *Winnie* fielde [*him* fuortglieden]
     Winnie felt [SE slip away]

---

person pronouns (see Reuland 2001, 2011a for discussion). Note furthermore that this strict identity between chain members has the interesting consequence that the interpretation should be uniform in all positions. We will return to this issue in section 6.

The reason resides in a minor parametric difference: Frisian allows licensing of object pronominals with non-structural Case (Hoekstra 1994). Such a Case is not T-related in the sense of P&T. Consequently, there is a derivation in which *him* in (15) is not visible as a goal for probing. There is no configuration that would lead to a violation of the PRD, hence the derivation is not cancelled. Therefore, semantic binding of *him* by *Winnie* is not blocked (see Reuland 2011a for extensive discussion).[11]

Note, that nothing in this account hinges on any stipulated property or principle that is specific to binding. Binding by chains and the restrictions it is subject to are just a by-product of general principles of grammar, and their interaction with language specific morpho-syntactic properties.[12] This explanation of local binding in Frisian is illustrative of what we may expect cross-linguistically: Minor differences in morpho-syntax, may lead to striking differences at the macro-level.

6 Reflexivity

6.1 Approaches to reflexivity

As we saw in the previous section, part of the answer to the question of why brute force reflexivization is excluded resides in the process of chain formation. However, this cannot be all. As the paradigms in (9) and (10) illustrate, what Dutch and Frisian share is that reflexive subject experiencer predicates require a complex reflexive.[13] In this they are not alone. We see this in other languages as diverse as Russian, and Sakha.[14] Other languages require a complex reflexive across the board, or have a verbal strategy alongside a nominal strategy (see, for in-

---

11 Note that Van Gelderen (2000) shows that Old English, another language with locally bound pronominals, lacked structural accusative Case.

The analysis of Frisian given here has been criticized by Rooryck and Vanden Wyngaerd (2011). However, this criticism ignores the difference between features being unvalued and the effects of syncretism (discussed in Reuland 2011a: Ch. 5), hence misses the point.

12 That no dedicated binding theory is needed is argued in a range of contributions by Reuland from (1995) onwards, lastly in (Reuland 2011a). A similar claim is made in Rooryck and Vanden Wyngaerd (2011). However, their approach does not succeed in avoiding a violation of the inclusiveness condition, witness the *-annotation of feature values resulting from Agree.

13 Thus, unlike what is occasionally claimed in the typological literature (e.g. Schladt 2000), Frisian is not a language without a reflexive.

14 A notable exception are languages with reflexive clitics (Romance, Slavic). For reasons of space, we refer to Marelj and Reuland (forthcoming) for explanation.

stance, Lust et al. 2000, for an overview of reflexivity markings in a range of languages spoken in India, and Schadler 2014 for discussion of yet a different range of languages). The question is why. There is extensive discussion in the literature of what makes reflexive predicates special.

Farmer & Harnish (1987: 557) argue that there is a *Disjointness presumption* on arguments (P4, in their numbering):

(16) P4 The arguments of a predicate are intended to be disjoint, unless marked otherwise.

Thus, *Alice* washed *her* violates P4, requiring the form *Alice* washed *herself* instead. However, it is unclear why this presumption would hold. Why would the co-argument domain be special? Why wouldn't such a presumption hold for larger chunks? Why wouldn't special marking of *her* be necessary in **Alice** saw *that nothing would save **her***?

Levinson (2000), builds on Farmer and Harnish, and includes the contrast between verbs allowing a simplex reflexive and those requiring a complex reflexive in the discussion. In his view the notion 'Prototypicality of action' is central, as also proposed in König and Siemund (2000).

Certain verbs denote actions that are prototypically '*other-directed*'. If such verbs are used as reflexives this must be specifically encoded (e.g. intensification with *self*, e.g.). If a prototypically 's*elf-directed*' verb is so used, no special marking is needed (although neither analysis distinguishes this issue from the question of the distribution of pronominals versus anaphors).

Prototypicality of action, however, does not provide the answer. Among the verbs most clearly resisting *self*-less reflexivization in Dutch, Scandinavian, etc. are subject experiencer verbs with meanings such as *admire* or *know*, which do not denote actions at all. There is no independent test for self- or other-directedness (or +/−*natural reflexivity*) as a relevant property of verbs. Is *ontwapenen* 'disarm' prototypically other-directed or self-directed? No way to tell, but it is content with a SE-anaphor. Why would *branden*, which is highly unpleasant if it happens to you be intrinsically self-directed, as the SE-anaphor in *Jan brandde zich* 'Jan burnt himself' would seem to require in this line of analysis? Why would *bewonderen* 'admire' convenient as self-admiration is, be other-directed as the obligatoriness of a SELF-anaphor would seem to indicate?

For many languages it is impossible to make the case that the marked form used for reflexives has properties of an intensifier, focus marker, etc. Russian uses *sebja*, which doesn't have any such properties. Many languages use bodypart expressions such as 'his head', 'his body' to express reflexivity. There is no non-arbitrary correlation between the meaning of such bodypart expressions and

intensification. There is a language, Zande, which puts the reflexive argument in a PP (Schladt 2000). It is unclear what in such a case the intensification would reside in. The conclusion must be that an approach along these lines cannot tell us why reflexivity would require special means. That's why one has to look for firmer ground to explain the special status of reflexivity.

6.2 Deriving the need for licensing: IDI

As we will see, the special status of reflexivity need not be stipulated. Rather it follows from a general property of grammatical computations, in fact of computations in general. In order to see this, let's go back to the representation of reflexivity in (7), repeated here:

(7) DP (λx (V x, x))

A crucial aspect of (7) is that we see two identical objects in a local domain: the two occurrences of the variable x. The mapping of (8a) onto (7) presupposes that the two occurrences of x in (7) can be kept apart. However, that is not a foregone conclusion. In fact, it is well known that the grammatical system has trouble handling occurrences of identicals in a local domain. This is reflected in the Obligatory Contour Principle of phonology (Leben 1973), the distinctness condition on linearization (Richards 2002) and the anti-locality condition on movement (Abels 2003). This is a consequence of a trivial fact: the *Indistinguishability of Indistinguishables* (IDI). In general we can only keep track of different occurrences of identical objects if they occur in a space with sufficient structure to keep them apart. Note, that this is a *general truth*, not specific to grammar. In (7) it is the – implicit – coordinate system of a piece of paper that allows us to distinguish the two occurrences of x. However, what if we have a space without order? Trivially, in a space without order the expression *x x* cannot be distinguished from the expression *x*. The question is, then, what implications this has for the linguistic computational system.

An influential idea since Kayne (1994) and Chomsky (1995) is that order is a property of the expression side of the grammar, Phonological Form (PF). Not of the grammar per se, and not of the interface with the interpretive system (the C-I interface). Grammar, however, does have hierarchy. So, in (8a), repeated here as (17a), we minimally have the structure in (17b):

(17) a. DP V Pron
      b. [$_{VP}$ DP [$_{V'}$ V Pron]]

Making the translation explicit, (17b) might be taken to yield (18):

(18) DP (λx ([$_{VP}$ x [$_{V'}$ V x]]))

A structure as in (18) presupposes that syntactic structure is preserved at the C-I interface. However, such a supposition is unwarranted. Just as in the mapping from syntax to PF only that structure is preserved that is phonologically interpreted, in the mapping from syntax to the C-I system only that structure should be preserved that is semantically interpreted. As Chomsky (1995) notes, only terms, i.e. heads and maximal projections are interpreted; intermediate projections are not. So, in fact, the V$'$ in (17b) is not interpreted, and therefore not represented at the interface level. This entails that at the C-I interface there is not sufficient structure in the co-argument domain to distinguish two occurrences of the argument variable. Thus, upon translation at the interface (17b) is not rendered as (18), and subsequently, (7), but as in (19). The interpretative system cannot see two argument variables, but only one as indicated in (19). Consequently, the verb does not express a relation, but a property:

(19) DP (λx ([$_{VP}$ V x]))[15]

Since the verb has two theta-roles to assign, the transition to a representation in which there is only one argument variable creates an indeterminacy. This is why this derivation is illicit (see Reuland 2011a for more discussion).

6.3 Strategies for licensing reflexivity

Logically, there are three ways to prevent IDI from causing reflexive relations to be inexpressible.

6.3.1 Protection

One strategy is based on *protection*. This involves a derivation from a numeration including material that keeps the arguments distinct. Such material is for instance a SELF-type element, a bodypart expression, or the doubling of a

---

15 This does not change if one adopts a more articulate structure with a v★ and a V, since the V will have to head-move onto the v★, making its original projection transparent.

pronominal element, all strategies we saw realized. The effect of such a protection strategy is the formation of a structure as in (20a), which is interpreted as in (20b) (see Reuland and Winter 2009 for its precise semantics):

(20) a. . . . . [V x [x Morph]]
　　　b. λx (V x f(x)), where f maps x onto an element that can stand proxy for ||x||.

So, going back to (9b) and (10b), the interpretation of *Winnie bewonderde zich* or *Winnie bewûndere him*, would lead to (19), where no relation is expressed. The alternatives *Winnie bewonderde zichzelf* and *Winnie bewûndere himself* will be mapped onto (21), assuming the SELF-element makes an independent contribution to the interpretation:

(21) a. Winnie λx (V x [SELF x])
　　　b. Winnie λx (V x f(x))

Here, crucially, the arguments are kept distinct, and the relational character of the verb preserved. Thus, f(x) should denote an individual that can stand proxy for Winnie, including Winnie herself. As we all know, since Jackendoff (1992), this is in fact true. *Himself* can not only have the value Winnie, but also Winnie's picture, statue, etc. This is, then, not just a quirk, but an intrinsic feature of the licensing strategy.[16] At the same time, this analysis explains why so many languages use body-part reflexives to express reflexive relations. Body-part expressions provide very natural proxies, enabling us to have our cake and eat it: express reflexive relations (or rather a suitable approximation thereof) while keeping the arguments distinct. Doubled pronominals serve the same function.

　　　Thus, one crucial factor ruling out local binding of a pronominal, or a simplex anaphor is the effect of IDI. A simplex pronominal carries no protection of the relevant variable. No protection is needed, however, in environments where binding does not cause a reflexive predicate to be formed, since binder and bindee are not semantic co-arguments. This happens, for instance, with locative PPs (as in *Alice* put the bottle next to *her*), or when a matrix subject binds an ECM subject. In such environments simple pronouns or simplex anaphors (depending on the conditions on chain formation) can be found.

---

16 Expressing *near-reflexivity* to use a term coined by Lidz (2001), elaborating a distinction discussed earlier in Reuland (1995), see further Reuland (2001).

6.3.2 Bundling

The other strategy involves an operation on *argument structure*. As follows from the discussion in section 2, any other process ensuring that two roles of a predicate end up on one argument will do as well to represent reflexivity. The Theta System (TS) of Reinhart (2002), Marelj (2004) and Reinhart & Siloni (2005) does precisely this. TS provides a general theory of operations on argument structure, capturing well-known alternations such as that between *John opened the door – The door opened*, or *John worried about his health – His health worried John*. One of the operations TS provides is *Bundling of θ-roles*. The Bundling operation reduces the internal argument of a two-place predicate and bundles the internal role (theme) and the external role (agent) into a composite agent-theme role. Thus, the bundling operation provides a theoretical basis for the valence reduction effects that so often show up in reflexives cross-linguistically.

(22)　Bundling:
　　　　a.　$V_{acc} (\theta_1, \theta_2) \rightarrow R_s(V) (\theta_{1,2})$
　　　　　　(where $\theta_{1,2}$ stands for the Bundling of θ_1 and θ_2)
　　　　b.　$V [Agent]_1 [Theme]_2 \rightarrow V[Agent\text{-}Theme]_1$
　　　　c.　Restriction: Agent-Theme verbs

As stated in (22c), Reinhart and Siloni note an intriguing cross-linguistic generalization, namely that bundling in the lexicon is restricted to (subsets of) Agent-Theme verbs. Although they don't provide an explanation for this restriction it has a clear advantage over earlier attempts to characterize the class of lexically reflexives discussed in section 6.1: *It can be assessed independently*. Also descriptively, it appears to be more adequate than characterizations in terms of (self/other)-directedness or 'naturalness'.[17]

　　Valence reduction may also affect the Case assigning properties of the predicate. In Reinhart and Siloni's system, the reduction operation applies to transitive entries that assign accusative Case.[18] Reflexivization is parameterized in the following respect: Languages vary as to whether valence reduction also eliminates

---

17 In Reinhart and Siloni's system theta-roles are coded in 2 binary features. Agent is [+cause, +mentally involved], theme is [−cause, −mentally involved]. These are the two most distinct roles. Reuland (2011a) suggests that it is this maximal distinctness that allows them to be bundled.
18 See Marelj and Reuland (forthcoming) for discussion of how this view can be reconciled with insights from Distributed Morphology.

the accusative (e.g. English, Hebrew), as illustrated by English *John washed*, or leaves a Case residue that still has to be checked, for instance by a nominal affix as in Russian *Ivan pomylsja*, or by an expletive such as Dutch *zich*. That is, according to TS, *zich* in *Winnie waste zich* is only there to check the residual case left by the Bundling operation. It has no independent argument status. Thus, to make this explicit, at the interface *Ivan pomylsja*, or *John washed* have the structure in (23):

(23) DP (λx ([$_{VP}$ V$_{[Agent-Theme]}$ x])

One of the facts supporting the idea that Bundling reduces the internal argument of a two-place predicate is the wax museum context of Jackendoff (1992), mentioned earlier. As Jackendoff notes, *himself* in (24a) can be interpreted either as the person Ringo or as Ringo's wax statue. Crucially, however, in (24b) a statue interpretation is not available.

(24) a. {Upon a visit to Mme Tussaud wax museum,} *Ringo washed himself*. (OKRingo, OKRingo's statue)
   b. {Upon a visit to Mme Tussaud wax museum,} *Ringo washed*. (OKRingo, *Ringo's statue)

The simplest account is to say that there is no object projected in this case. However, even if one would wish to argue for a null-object (which then would have a special and restricted distribution) it cannot be a semantic argument. Interestingly, the same contrast is found in Dutch between reflexives with *zichzelf* and *zich* (Reuland 2001), and see Reuland and Winter (2009), and Reuland (2011a) for formal details and an explicit semantics.[19] In Russian the same contrast obtains between reflexives with *sebja* and *sja*.

Another test involves object comparison from Zec (1985), see also Dimitriadis and Que (2009), illustrated in (25):

---

[19] An anonymous reviewer remarks that the absence of proxy-readings does not necessarily indicate the absence of an argument, referring to Giorgi (2007). Giorgi discusses cases of long-distance anaphora (LDA) where indeed a proxy-reading is absent. Giorgi explains this in terms of a theory where LDA involves unsaturated argument positions. To us it seems that a discussion of whether an unsaturated argument is or isn't an argument is rather moot. But note, that although the execution is different, an approach in terms of chain formation achieves the same effect, since as observed in fn. 11, chain members are semantically uniform (see Reuland 2011a for an analysis of LDA in Germanic in terms of chain formation). This is precisely what is captured by the semantic interpretation of feature chains with *zich* by Reuland & Winter.

(25) a. Bill washes himself more often than John.
     b. Bill washes more often than John.

The sentence in (25a) has two readings. One is that Bill washes Bill more often than John washes John; the other one is that Bill washes Bill more often than Bill washes John. In (25b), however, object comparison is impossible. It doesn't allow the reading that Bill washes Bill more often than Bill washes John. Since in (25a) object comparison is available the predicate must have an object argument to enter the comparison. Again, the simplest answer to the question of why object comparison is impossible in (25b) is that there is no object argument to start with. The same pattern obtains in Dutch, similarly arguing against the argument status of *zich*.

6.3.3 Separation

Logically one may conceive of yet another strategy to license reflexivity, namely guarantee that at the level where IDI applies, the two variables are not on one grid. This is in fact the strategy applied in Zande, where in reflexives the object argument is realized in a PP (Schladt 2000, see also Schadler 2014). This is also the case where French allows locally bound pronominals, as in *Jean est fier de lui* 'Jean is proud of himself', where *Jean* and *lui* are not on the same grid, unlike in its Dutch counterpart (see Reuland 2006, 2011a for discussion). This option plays a role in our subsequent discussion of datives in Khanty.

6.4 Towards resolving the puzzle

All this paves the way for the discussion of reflexivity in Tegi Khanty. Can we understand the anaphora facts in this language on the basis of the minimal assumptions sketched, and without recourse to unmotivated stipulations?

7 Introducing Tegi Khanty

7.1 Sociolinguistic background

Khanty is a language of the Uralic language family spoken by some 9,500 people (2010 census). According to the Ethnologue (Lewis et al. 2013), the ethnic

population totals 28,700 spread up over several thousand square kilometers in Northwest Siberia, Russia. Living in remote communities, the Khanty have developed a dialectal continuum, the further ends of which diverge greatly both in grammar and lexicon and are mutually incomprehensible (Nikolaeva 1999a). The Khanty language is listed as definitely endangered by UNESCO (Moseley 2010). The variety reported below is spoken in the village of Tegi on the river Malaya Ob' in the Beryozovsky District in the north-western part of the Khanty-Mansi Autonomous Region.[20]

The Tegi variety can be considered severely endangered: it is spoken only by the older generation, who rarely use it among themselves. Most of our native speakers switched to Russian at home when raising children, which resulted in the next generation being unable to speak the language, though some understand it. This development can at least partly be explained by a certain pressure from the school system: the education in primary school is almost entirely in Russian so the children have to speak Russian by the age of seven when they go to school.[21] Another contributing factor is that migration of Khanty speakers even over small distances to other villages or towns where some other variety of Khanty is spoken, makes them face a relatively high degree of unacceptance of their dialect, which also, reportedly, forces them to switch to Russian.

All the data presented in this paper was compiled during a fieldwork trip to Tegi in July 2012. As part of the fieldwork party of Lomonosov Moscow State University[22] we did elicitation sessions with ten native speakers of Khanty using Russian as the language of communication. We also studied the collection of texts recorded from the native speakers by other members of the field party. Some follow-up questions were addressed via email. All examples below were cross-checked with at least two native speakers of Tegi Khanty.

---

20 In what comes next we will use the term Khanty as referring to the Tegi Khanty variety if not indicated otherwise.

21 They have a few hours of Khanty every week, but using text books in a standardized variety, that is rather different from Tegi Kanty.

7.2 Grammar sketch

Khanty has a basic SOV structure coupled with a relatively free word order (Nikolaeva 1999a). The sentence structure is highly dependent on the structure of the discourse and passive is widely used to preserve the topic of the discourse.

The nominal system has three morphological cases – Nominative, Dative, and Locative. The direct object is encoded with the unmarked case (the same form as Nominative) and differential indexing on the verb with the choice between subject and subject-object agreement. Personal pronouns distinguish three morphological cases – Nominative, Dative and Accusative. The language has three numbers: singular, dual and plural. The pronominal system has three persons: 1st, 2nd and 3rd.

In the verbal domain there are two synthetic tenses – past and nonpast, and an analytical future tense created with the auxiliary verb *pitti* 'start'. The language has a rich system of aspectual markings and affixes reflecting changes in argument structure such as causativisation, and detransitivisation.

What turns out to be crucial for resolving our puzzle is that Khanty distinguishes two types of verbal agreement: obligatory subject agreement and optional object agreement (see the next subsection, 7.3). Objects associated with agreement on the verb usually have a special status with respect to the information structure of the sentence and differ in their syntactic behavior from objects that are not associated with agreement. Marked objects tend to be associated with old information, secondary topic, while unmarked objects tend to express new information (see Nikolaeva 1999a, Nikolaeva 1999b, Nikolaeva 2001 for detailed discussion). According to Nikolaeva (1999a), objects triggering agreement are characterized by a relatively free position in the sentence, while objects that do not trigger agreement are usually situated in the immediately preverbal position. Further, Nikolaeva (1999a) claims that objects triggering agreement are realized in a VP-external position.

Under favourable discourse conditions the language easily allows subject drop (26) (see Nikolaeva 1999a).

(26) Xoleivt   aŋk-em   ropota-ja   an   mant-ł.   N'avrem piła
     tomorrow   mother-1SG   work-LOC   NEG   go-NPST.3SG child   with
     jomtepit-ł.
     play-NPST.3SG
     Tomorrow my mother won't go to work. (She) will play with the child.

Example (26) shows that, once introduced, the subject can be omitted in the following sentences under the topic continuance condition.

Khanty also allows object drop, as illustrated in (27):

(27) *One autumn day a man went to the forest and suddenly saw a bear in the lake*
caught up in the ice.
S'ałta   keł-ən   metš-əs-łe,   vuti-šek   tał-s-əłłe,
then   rope-LOC   tie-PST-SG.3SG   bank-COMP   drag-PST-SG.3SG
χun-əł-a   nuχ   perət-s-əłłe.
belly-3SG-DAT   up   turn-PST-SG.3SG
Then (he) tied (it) with a rope, dragged (it) to the bank of the lake and turned
(it) over.

Once the object has been introduced in the discourse it can be dropped, but, cru-cially, this is possible only if the verb carries *subject-object* agreement. The verbs in (27) *metšəsłe*, *tałsəłłe*, and *perətsəłłe* agree with the subject of the sentence in person and in number and with the object in number: the marker – *(əł)łe* ex-presses that the subject is 3rd person singular, while the object is singular. With this configuration of markers on the verb the object itself can be omitted. Let's therefore look at the agreement system in more detail.

7.3 Subject agreement versus object agreement

Both types of agreement are illustrated in the two forms of (28):

(28) Utłtiteχo   poχ-leŋki   išək-s-əłłe /   išək-s.
teacher   boy-DIM   praise-PST-SG.3SG /   praise-PST.3SG
The teacher praised the boy.

The verb agrees with the subject in number and person and with the object in number. The relevant dependencies are indicated in (29b):

(29) a. Učitel-t$_i$ łiveł$_{i/k}$ išək-s-əł-əł.
teacher-PL   they.ACC   praise-PST-PL-3PL
The teachers praised them(selves).
  b. Subject$_{[Pers;Num]}$ Object$_{[Pers;Num]}$ Verb-T-O$_{NUM}$-S$_{PERS;NUM}$

The morphological realization of these forms of agreement is given in tables 1 and 2.

Table 1: Subject agreement in Tegi Khanty, past tense

Person	Number					
	SG		DU		PL	
1	išak-s	-əm	išak-s	-əmən	išak-s	-uv
2	išak-s	-ən	išak-s	-ətən	išak-s	-əti
3	išak-s		išak-s	-əŋən	išak-s	-ət

Table 2: Subject-object agreement in Tegi Khanty, past tense

Subject		Object					
		SG			PL		
SG	1	išak-s-	Ø	-em	išak-s-	əł	-am
	2	išak-s-	Ø	-en	išak-s-	əł	-ən
	3	išak-s-	(əł)łe		išak-s-	əłłe	
DU	1	išak-s-	Ø	-emən	išak-s-	əł	-amən
	2	išak-s-	Ø	-əłən	išak-s-	əł	-łən
	3	išak-s-	Ø	-əŋən	išak-s-	əł	-łən
PL	1	išak-s-	Ø	-ev	išak-s-	əł	-əv
	2	išak-s-	Ø	-əłən	išak-s-	əł	-łən
	3	išak-s-	Ø	-əł	išak-s-	əł	-əł

This overview of the use of subject and object agreement is necessary for the further discussion, because, as we will see in section 8.3, object agreement plays an essential role in obtaining a locally bound reading of the 3rd person pronoun *łuv*. As a preliminary, let us first briefly discuss, where object agreement (henceforth ObjAgr) is situated in the syntactic representation. A possible choice would be to identify ObjAgr with AgrO (Mahajan 1990), introduced in the Minimalist program (Chomsky 1992) as a locus of object agreement under the Split Infl Hypothesis (Pollock 1989).

Note, that Chomsky (1995: ch. 4) argued that conceptually AgrO is suspect, as Agr-projections do not receive an interpretation at the interface with the interpretation system. Rather its effect should be replaced by checking with the outer member of a two-layered VP-shell for active transitive verbs, as proposed by Hale & Keyser (1991) (see Ura (2003) for discussion of Agr-based and Agr-less Case theories). A conceptual argument doesn't carry any force, however, in the case of an agreement that is overtly realized, as we see here in Khanty. In any case, following the traditional understanding of AgrO it has to be situated lower than T.

A potential problem with an analysis along this line is that it does not reflect the order of morphemes on the verb. ObjAgr follows the tense marker in Khanty, which is always situated immediately after the verb stem, followed in turn by the subject agreement marker, as indicated in (29b). Given Baker's Mirror Principle (Baker 1985) the surface morpheme order is derived via successive-cyclic head-movement. If so, the functional projections containing subject and object agreement should be higher on the spine than T. Thus, ObjAgr should be positioned above T, but below the functional projection containing the subject agreement. This state of affairs is illustrated in (30), which shows the two possible positions for ObjAgr: The lower position βP (30b) would reflect the canonical low AgrO position; the higher position of βP in (30c) is in accordance with the Mirror Principle.

(30) a. Učitel-t$_i$ łiveł$_k$ išək-s-əł- əł.
       teacher-PL   they.ACC   praise-PST-PL-3PL
       The teachers praised them.

   b.

c.

```
                              γP
                    ┌──────────┴──────────┐
                 SAgr                    βP(II) ?
                   |              ┌────────┴────────┐
                -ət₃PL          OAgr               αP
                                  |         ┌───────┴───────┐
                               -ət-PL      TM              vP
                                            |        ┌──────┴──────┐
                                         -s-PST     DP            v′
                                                     |       ┌─────┴─────┐
                                               učiteltNOM; PL  v        VP
                                                'teachers'          ┌────┴────┐
                                                                    V        DP
                                                                    |         |
                                                                 išək-     tiveł ACC; PL
                                                                 'praise'   'they'
```

γ - subject agreement
β - object agreement
α, TM - tense marker
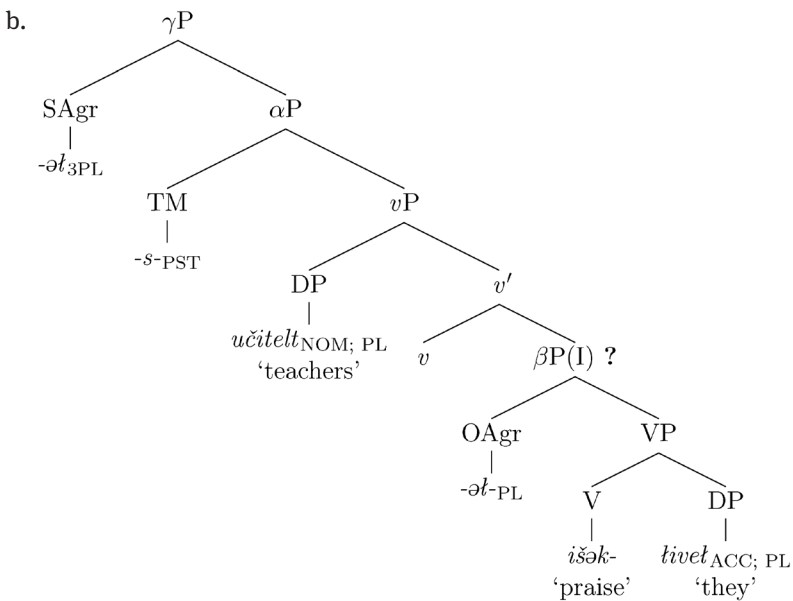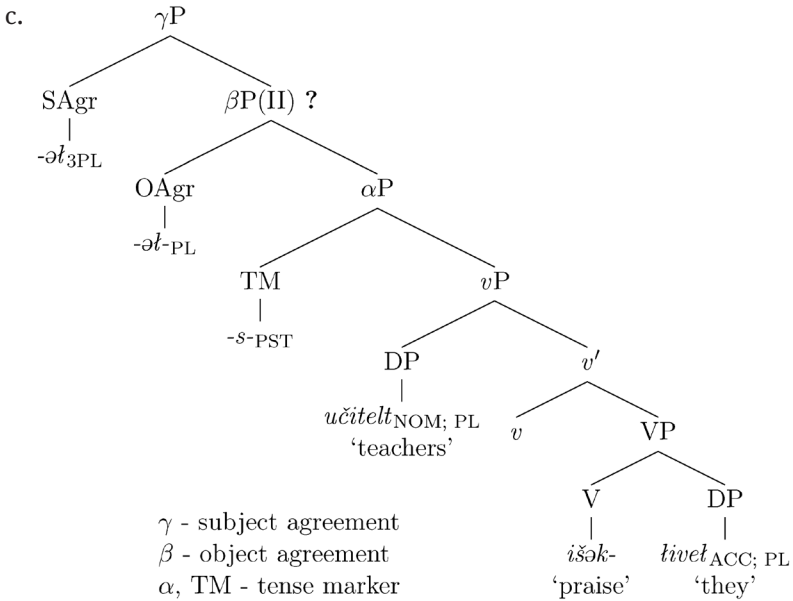
Crucial is, that either way ObjAgr is positioned below subject agreement, and hence it intervenes between subject agreement and the VP (with the locus of the direct object's first Merge). As we will see, this is enough for the purposes of the present article.[23] In order to derive the surface order of the overt arguments, some standard subsequent movements will be assumed to take place. Hence the subject will end up in Spec γP, and the object into Spec βP (= Spec ObjAgr), deriving the structure Nikolaeva proposes.

Having introduced the relevant essentials of the Khanty grammatical system, and specifically, its agreement system, we will now proceed with the reflexivization strategies in Khanty.

8 Reflexivization strategies

As we saw in section 6, reflexivity must be licensed, either by bundling of thematic roles or by protection. Let us explore, then, what strategies are employed by Khanty. In the first part of this section we will discuss a verbal strategy, in the second part the nominal strategy.

---

23 In fact, both could even be correct if one were to assume an independent movement of ObjAgr from the one position to the other. We are not aware of convincing evidence for such a movement, though.

8.1 Reflexivization with *-ij(ł)*

It is uncontroversial that Tegi Khanty has at least one specific reflexivization strategy, namely by using the suffix *-ij(ł)*, as in the following examples.

(31) a. *l'oχətti* 'wash' – *l'oχətijłti* 'wash oneself'
    b. Łuv   l'oχət-ij-əł.
       S/he   wash-DETR-NPST.3SG
       He washes.

(32) a. *eŋχəsti* 'undress' – *eŋχəsijłti* 'undress oneself'
    b. Poχ   eŋχəs-ij-əs.
       boy   undress-DETR-PST.3SG
       The boy took his clothes off.

If the suffix is present, no direct object argument can be realized (33b).

(33) a. As'i   n'avrem   l'oχət-əł.
       father   child   wash-NPST.3SG
       The father washes the child.
    b. *Pet'a   n'avrem   l'oχət-ij-s.
       Petja   child   wash-DETR-PST.3SG
       Int: Petja washes the child.

The use of the suffix *-ij(ł)-* is subject to limitations. Crucially, it can be used only with a limited amount of verbs: it is available with some agent-theme verbs like *wash*, *defend*, but not with subject-experiencer verbs like *know, remember, trust* – cf. (34). Thus, Khanty patterns with a range of languages studied so far, and is in accordance with Reinhart and Siloni's generalization.[24]

---

24 An anonymous reviewer takes issue with this line of argument, arguing that it is not surprising that the suffix *-ij(ł)-* does not attach to verbs such as *nuomti* 'remember', since no one would ever want to say (34), and similarly, that it is not surprising that there is no special nominal reflexivization strategy in Khanty, since the verbal one suffices for what one would want to say. This criticism, however, does not take into account that 'verbal' and 'nominal' reflexivization strategies co-occur cross-linguistically very generally, and are accompanied by the same restriction and contrast we discuss here. There is no reason to assume that remembering oneself is any stranger in Khanty than in English, Dutch, or than to love oneself in Sakha. It is crucial to see that the use of natural language is not limited to what is 'usual', or 'normal'. Rather the power of natural language is that it can be used to talk about what is abnormal, or impossible. And interestingly enough even if one talks about the impossible, this does not by itself affect the grammar.

(34) a. *nuomti* 'remember' – *nuomtiti
    b. Pet'a-jən[25]   ɫuv   ɫuveɫ   nuom-ɫ-əɫɫe.
       Petja-2SG   he   he.ACC   remember-NPST-SG.3SG
       Petja remembers himself.

The reciprocal interpretation that cross-linguistically often patterns with reflexivization, is not available.

(35) N'avrem-ət   l'oχət-i-s-ət.
    Child-PL   wash-DETR-PST-3PL
    The children washed (each of them – himself / *each other).

Given our discussion so far, the most straightforward interpretation of the pattern is that *-ij(ɫ)* licenses a Bundling operation. Bundling should create a composite θ-role, and its result should display both the properties of the agent and the theme roles. The following modification test with the adverb *tăɫaŋteɫn* 'completely' shows the availability of the theme role properties (see Dimitriadis and Everaert 2012 for discussion).

(36) Pet'a-jən   *(n'an' pul)   tăɫaŋteɫn   joχi   ɫe-s(-ɫe).
    Petja-2SG bread   piece   as.a.whole down   eat-PST(-SG).3SG
    He ate the piece of bread completely.

According to Dimitriadis and Everaert (2012), adverbs like *painfully* or *completely*, target the explicit theme/patient role. Example (36) shows a two-place predicate *ɫeti* 'eat' modified with the adverb *tăɫaŋteɫn* 'completely, as a whole'. If the adverb is present, the internal argument of *ɫeti* 'eat' cannot be omitted. A similar effect is present in (37a). In (37b) the verb does not have an internal argument, however, modification with the adverb *tăɫaŋteɫn* 'completely, as a whole' is still

---

As a final note, we were struck by the fact that our informants were in fact very well able to reflect on their language and our questions, and did not hesitate to correct us, if our sentences were unacceptable since they violated conventions. A striking example was that they corrected us, when we asked sentences, that contained a phrase with *ubit' medvedja* 'kill a bear'. One should not say *kill a bear*, we were informed, only *dobyt'* 'obtain' one, with a detailed explanation of why this is so in their culture.

25 Here the subject *Pet'a* carries a 2nd person singular possessive suffix. This suffix is used with proper names occupying subject and sometimes object positions as a discourse marker to indicate some contextual relation to the listener.

possible, which shows that the only argument of the verb in (37b) preserves properties of the theme.

(37) a. Pet'a-jən   tăłaŋtełn   noχ   l'oχət-s-əłłe.
     Petja-2SG   as.a.whole   up   wash-PST-SG.3SG
     (He) washed Petja completely.
   b. Van'a-jən   tăłaŋtełn   l'oχət-ij-s.
     Vanja-2SG   as.a.whole   wash-DETR-PST.3SG
     Vanja washed himself completely.

Although the theme role is there, there is no evidence for an independent argument bearing it. This is shown by the distribution of proxy-readings in (38).

(38) a. Gorbachev   muzej-a   joχt-əs   i   łuv   pam'atnik
     Gorbachev   museum-DAT   come-PST.3SG   and   he   monument
     s'ijal-s.   ?Łuv   l'oχət-ij-s.
     see-PST.3SG   he   wash-DETR-PST.3SG
     Gorbachev came to the museum and saw a monument to himself. He washed.
   b. Gorbachev   muzej-a   joχt-əs   i   łuv   pam'atnik
     Gorbachev   museum-DAT   come-PST.3SG   and   he   monument
     s'ijal-s.   Łuv   łuveł   l'oχət-s-əłłe.
     see-PST.3SG   he   he.ACC   wash-PST-SG.3SG
     Gorbachev came to the museum and saw a monument to himself. He washed himself / the monument.

In (38a) with the verb form *l'oχətti* 'wash' and no overt object the only reading available is that Gorbachev washed himself, and not his statue/picture/etc. This is just like in its English counterpart. Its counterpart with the pronominal in (38b) allows both readings. The simplest interpretation of these facts is that the suffix *-ij(ł)* indeed marks Bundling and reduction of the object argument.[26] Thus, one of

26 An anonymous reviewer comments that the absence of proxy-readings by itself does not show that no argument can be projected, given the observation in Giorgi (2007) that LDA's in Italian or Dutch don't allow proxy-readings either. This is correct, see also fn 11 and 20. The difference, however, is that the existence of Italian *sè* and Dutch *zich* is independently motivated. If one were to postulate a null 'reflexive' object in English or Khanty for these cases, it would require highly idiosyncratic properties, and a special, highly restricted distribution. It seems to us that given Occam's razor there is little reason to pursue this line. But, if one were to postulate a null-element here it would require a specific licenser (*-ij(ł)*) and should not qualify as an independent argument.

the ways to satisfy the IDI in Khanty is with the help of the suffix *-ij(ł)*, yielding Bundling and reduction of the internal argument.

8.2 Locally bound *łuveł*

Coming back to the starting point of this article, the puzzle we face is that, according to Nikolaeva (1995), Khanty has no dedicated reflexive pronouns; instead, personal pronouns are used. It was illustrated by the example in (1), repeated here:

(1) Utłtiteχo$_i$ łuveł$_{i/k}$ išǝk-s-ǝłłe.
     teacher   he.ACC   praise-PST-SG.3SG
     The teacher praised him(self).

This is what we found as well. The question is, then, how Khanty can use its pronouns to express reflexivity? Just by 'brute force' binding – that is, binding without any further licensing (which would make it a true exception to the idea that reflexivity must be licensed), or does it have structural properties that independently license reflexivity?

Strictly speaking, however, (1) is not enough to show that we have binding – hence reflexivity in the sense defined. Since the subject is referential, we still must check that the dependency is indeed binding, and not simply co-reference, which is much less restricted. To check that we have binding we replaced the referential subject by a quantificational subject as in (39):

(39) Nemχojat$_i$ łuveł$_{i/k}$ ănt   išǝk-ł-ǝłłe.
      no.one   he.ACC   NEG   praise-NPST-SG.3SG
      No one praises himself / him.

Here, the same pattern obtains. To make sure, however, that elements such as *nemχojat* indeed are truly quantificational, like their counterparts in English, we put them through the Heim test (Heim 1982), which predicts a contrast between the a. and the b. case in (40).

(40) a. Poχ-leŋki   vuonłtǝ-s-le   ar.   Ari-ti   (łuv)   pit-ǝł?
          boy-DIM   learn-PST-SG.3SG   song   sing-INF   he   start-NPST.3SG
          The boy learnt the song. Will he sing?
      b. Nemχojat   ar   ănt   vuonłt-ǝs.   *Ari-ti   łuv   pit-ǝł?
          no.one   song   NEG   learn-PST.3SG   sing-INF   he   start-NPST.3SG
          Int.: *No one learnt the song. Will he sing?*

In (40) *poχleŋki* 'boy' is a referential expression, hence the 3rd person pronoun *łuv* 'he' can refer back to it across the sentential boundary. In contrast, *nemχojat* 'no one' is not an individual denoting expression, and therefore *łuv* 'he' cannot co-refer with it. Binding is impossible across a sentential boundary due to the absence of c-command (Reinhart 1983). This yields an illicit mini-discourse for this intended interpretation.[27,28]

One might consider the possibility that the pronoun *łuv* would be special in some sense. For instance that it could be logophoric. However, the language presents no evidence for that: there is no other pronominal paradigm, and *łuv* does not show the properties of logophors. It is not restricted to reportive contexts, and its antecedent is not limited to the individual whose words or thoughts are transmitted in the reported context in which the pronoun occurs (see Clements 1975, Sells 1987).[29]

(41) Maša$_m$ jast-əs   Ivan-a$_i$ što   łuv$_{i/m}$ Boris   išək-ł-əłłe.
     Masha   say-PST.3SG   Ivan-DAT   that   s/he   Boris   praise-NPST-SG.3SG
     Masha says to Ivan that s/he praises Boris.

Example (41) shows the 3rd person singular pronoun *łuv* used in a reportive context. It can take as an antecedent either argument of the main clause, specifically, the indirect object, which would not be the option if it were logophoric.

For the considerations given so far, the facts in (1) (and 39) raise issues on two different levels: they could potentially constitute a violation of IDI or of constraints on chain-formation.

Hence, to explain the way the Khanty anaphoric system functions we need to establish that no chain is formed between the subject and the locally bound third person pronominal, and that the bound variable is protected.

This is where the Agreement system comes in. If one takes a closer look at the reflexivity pattern in Khanty, it turns out that the bound reading of the pronoun *łuveł* is possible only if there is *object agreement* on the verb. Example (42b) shows

---

27 In the variety of Khanty described by Nikolaeva (1999a) quantifiers cannot be used in the subject position.

28 An anonymous reviewer remarks that this sentence is rather weird, and wonders what is the relevance of such a result. The relevance is just that we want to make sure that in this respect the Khanty quantifier behaves like its counterpart in English, where the equivalent of (40b) is equally odd.

29 That is, it needs neither be a SOURCE, nor a SELF, nor a PIVOT in the sense of Sells (1987), as is also demonstrated by (41).

that with just subject agreement on the verb, the 3rd person pronoun *ɬuveɬ*, as the internal argument of the verb, can be interpreted only disjointly. If the verb carries both subject and object agreement, as in (42c)–(42d), the pronoun *ɬuveɬ* is ambiguous between a bound and a disjoint interpretation.

(42) a. Utɬtiteχo$_i$ ɬuveɬ$_j$ išək-s.
      teacher   he.ACC   praise-PST.3SG
      The teacher praised him.

 b. *Utɬtiteχo$_i$ ɬuveɬ$_i$ išək-s.
      teacher   he.ACC   praise-PST.3SG
      Int.: The teacher praised himself.

 c. Utɬtiteχo$_i$ ɬuveɬ$_j$ išək-s-əɬɬe.
      teacher   he.ACC   praise-PST-SG.3SG
      The teacher praised him.

 d. Utɬtiteχo$_i$ ɬuveɬ$_i$ išək-s-əɬɬe.
      teacher   he.ACC   praise-PST-SG.3SG
      The teacher praised himself.

Why is (42b) illicit? Given the considerations introduced in sections 5, and 6, (42b) violates both IDI and conditions on chain formation. Hence its ill-formedness is simply to be expected. The question is then, what the role is of ObjAgr in (42d), where it apparently licenses the reflexive interpretation (just as in (1), and (39), sentences that also have ObjAgr).

We will now show how ObjAgr in Khanty provides protection for the variable, satisfying IDI. Briefly, it adds complexity, and thus helps keeping separate the two argument variables. And, as a side effect, it also prevents an ill-formed chain from being formed.

As a first step, recall from section 7.2, that object agreement on the verb facilitates object drop. However, under this type of object drop no reflexive interpretation is available (also noted in Nikolaeva 1999a).

(43) *One autumn day a man went to the forest and suddenly saw a bear in the lake caught up in the ice.*
      S'aɬta   keɬ-ən   metš-əs-ɬe,   vuti-šek   taɬ-s-əɬɬe,
      then   rope-LOC   tie-PST-SG.3SG   bank-COMP   drag-PST-SG.3SG
      χun-əɬ-a   nuχ   perət-s-əɬɬe.   = (27)
      belly-3SG-DAT   up   turn-PST-SG.3SG
      Then (he) tied (it/*himself) with a rope, dragged (it/*himself) to the bank of the lake and turned (it/*himself) over.

Thus, example (43) shows that object agreement licenses object drop, and that, in isolation as it is here, it does not allow a reflexive interpretation. This fact confirms that there is no brute force reflexivization in Tegi Khanty.

Hence, in order to license a reflexive interpretation in Khanty, two conditions must be met. First, there must be object agreement on the verb, and second, there must be an overt pronominal present. The simplest way to bring these factors together is as follows:

(44) *Complexity in Tegi Khanty*
  – Object agreement licenses a null object pronoun.
  – The overt *łuveł* forms a constituent with the null object.

For current purposes we will assume that *łuveł* is adjoined as in (45) but nothing hinges on it. Potentially it could also occupy the specifier position preceding the null pronoun. Crucial is that there is complexity.

If *łuveł* is adjoined, the structure of (42d) under its reflexive interpretation is (45), with Ø licensed by object agreement. Under this analysis the accusative marking on *łuveł* comes from the agreement between the head and the adjunct.

(45) Utłtiteχo$_i$ [Ø łuveł$_i$] išək-s-əlle.
  teacher   Ø he.ACC praise-PST-SG.3SG
  The teacher praised himself.

The availability of this option is supported by the fact that *łuv* can be used as an intensifier, although not all speakers accept this usage (note that as an intensifier it should be stressable, hence in this capacity it cannot be null).

(46) Jełp škola puš-s-ə(t) Komarova łuv joχt-əs.
  new school open-PST-3PL Komarova he come-PST.3SG
  {LC: Komarova is the governor of the Khanty-Mansi Autonomous Region.}
  For the opening of the new school Komarova herself came.

Irrespective of whether one adopts the specifier or the adjunct option – or either, since they are fully compatible – this analysis comes entirely for free. Very general considerations, and the cross-linguistic patterns require complexity. Khanty has a null pronominal, and independent means to create complexity. Hence, as far as IDI is concerned no mystery remains.

There remains the question of chain formation. But note, that under the account as it has developed, this question does not involve *łuveł* but rather the

null-pronoun in the head position. We know it must be licensed by ObjAgr, hence it cannot be the case that the constituent containing it is opaque.[30] In fact, it must be part of a chain with ObjAgr.

In an interesting respect, then, Khanty is reminiscent of Frisian (as discussed in section 5). Similarly to Frisian, in Khanty the conditions for chain formation between the pronoun Ø and the subject are not met, though for reasons that are different at the 'micro-level'. Informally speaking, the overtly expressed object agreement intervenes between subject agreement and the Ø pronoun, thereby preventing a configuration for chain formation with the subject, see (30) and (47c). At the stage where γ comes into play, object agreement (β) will already have checked any syntactic property of the pronoun (e.g. structural Case) that would make it visible for further probing. Hence the pronoun Ø and the subject do not give rise to a cancelled derivation.

It is important to distinguish issues of chain formation and the effects of IDI. In (47a)[31] object agreement licenses a null object pronoun. If the object were locally bound as in (47b), ObjAgr would prevent chain formation between the subject and the object, but the structure would still violate IDI, as nothing protects the second occurrence of the variable (see (47c) for an explicit structure).

(47)   a.   Učitel-t$_i$ Ø$_k$ išək-s-əł- əł.
           teacher-PL   they.ACC   praise-PST-PL-3PL
           The teachers praised them.
     b.   *Učitel-t$_i$ Ø$_i$ išək-s-əł- əł.
           teacher-PL   they.ACC   praise-PST-PL-3PL
           Int.: The teachers praised themselves.
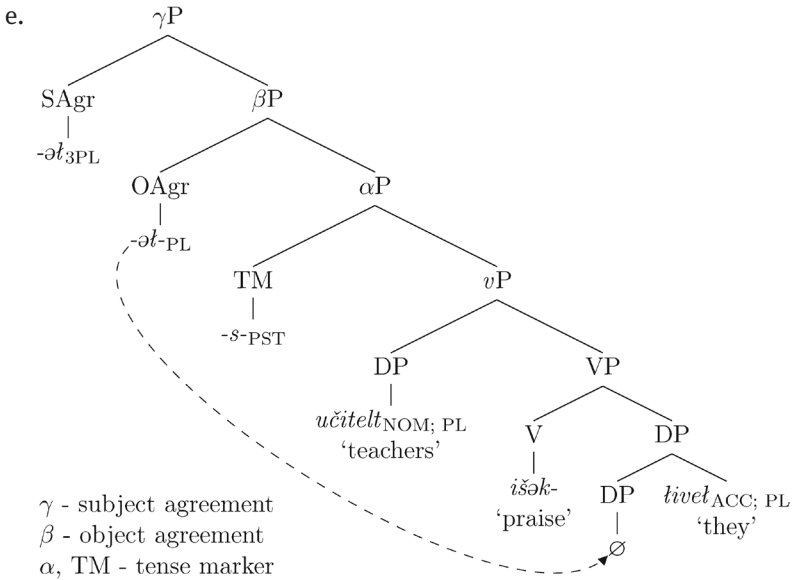
---

30 Just in case a reader wonders about English *himself*, note that here *him* is not in the head position.
31 Examples in (47) are constructed for ease of exposition based on (30) and the assumption that object agreement facilitates object drop as exemplified by (27).

c.

γ - subject agreement
β - object agreement
α, TM - tense marker
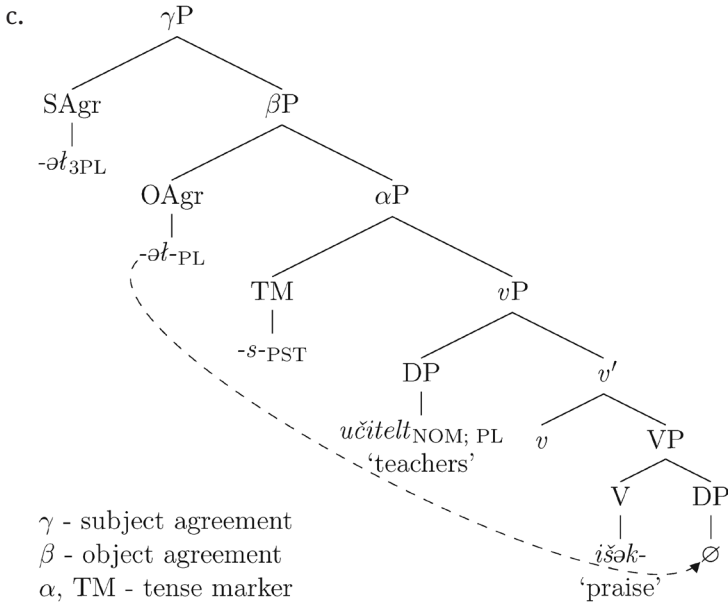
Consider next the case of an overt pronominal. The expression with a locally bound 3rd person plural pronoun *łiveł* in (48) is ambiguous between two syntactic structures – (48b) with bare *łiveł* and (48c) with a null object pronoun licensed by object agreement and *łiveł* adjoined to it. The structure in (48b) – also shown as a tree in (48d) – would violate IDI. Only the complex structure in (48c) also shown as a tree in (48e) satisfies IDI. The object can be variable bound at the C-I interface, given the protection scenario we sketched. Chain formation between *łiveł* and the subject is prevented as well.

(48) a. Učitel-t$_i$ łiveł$_i$ išək-s-əł- əł.
     teacher-PL   they.ACC   praise-PST-PL-3PL
     The teachers praised themselves.
    b. Učitel-t$_i$ łiveł$_i$ išək-s-əł- əł.
    c. Učitel-t$_i$ [Ø   łiveł]$_i$ išək-s-əł- əł

d.

```
                          γP
                 ┌─────────┴─────────┐
              SAgr                   βP
                │          ┌──────────┴──────────┐
              -ət₃PL    OAgr                     αP
                         │         ┌──────────────┴──────────┐
                       -ət-PL    TM                          vP
                                  │            ┌──────────────┴─────┐
                               -s-PST         DP                   v′
                                               │          ┌──────────┴──────┐
                                        učitelt_NOM; PL    v              VP
                                           'teachers'            ┌─────────┴────────┐
                                                                 V               DP
                                                                 │                │
                                                              išək-         tivet_ACC; PL
                                                             'praise'           'they'
```

e.

```
                          γP
                 ┌─────────┴─────────┐
              SAgr                   βP
                │          ┌──────────┴──────────┐
              -ət₃PL    OAgr                     αP
                         │         ┌──────────────┴──────────┐
                       -ət-PL    TM                          vP
                                  │            ┌──────────────┴─────┐
                               -s-PST         DP                   VP
                                               │          ┌─────────┴──────┐
                                        učitelt_NOM; PL    V              DP
                                           'teachers'      │        ┌──────┴──────┐
                                                        išək-      DP        tivet_ACC; PL
                                                       'praise'    │            'they'
                                                                   ∅
```

γ - subject agreement
β - object agreement
α, TM - tense marker
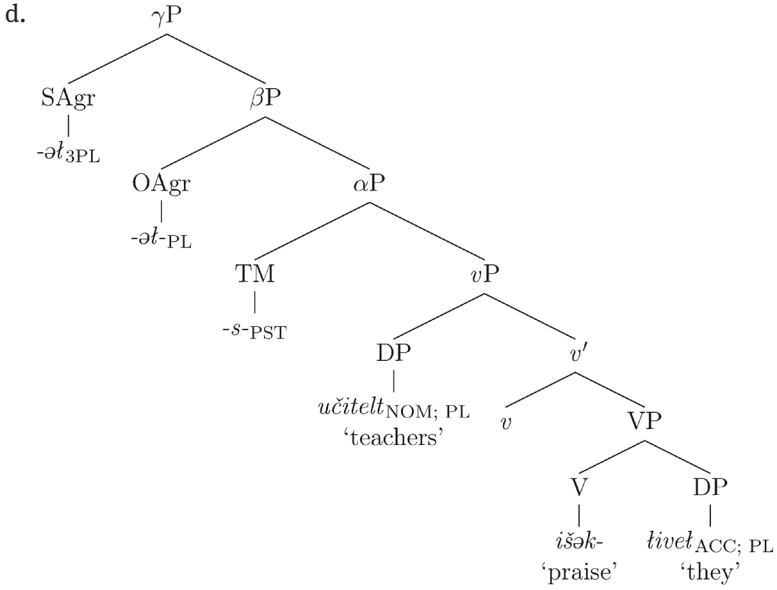
As to *łuveł* itself in (48e), as an adjunct it is not visible to A-chain forma-
tion, so no chain condition violation will ensue. As already noted, its accusative
case follows from fact that quite generally adjuncts of this type require case
matching.[32]

Next, note that *łuveł* also allows non-local binding, or can be free, just like its
English counterparts. The simplest analysis is that like any other argument DP it
can just occur with object agreement without a null pronominal being projected.
In that case it just behaves like one would expect a pronominal to behave. (Note
that Nikolaeva argues that Khanty is not a pronominal argument language in the
sense of Jelinek 1984).

The pronoun *łuv* can be used in postpositional phrases (49). In this context
it can be bound by the subject or take a discourse antecedent. In the former
case the postposition *ɔłŋałn* 'about' protects the variable and prevents chain
formation.[33]

(49) Pet'a-jən$_i$ Vas'a-jən-a$_j$ potərt-əs   łuv$_{i/j/k}$ ɔłŋ-ał-n.
     Petja-2SG   Vasja-2SG-DAT   say-PST.3SG   he   about-3SG-LOC
     Petja told Vasja about him(self).

The last issue we would like to discuss in this context are dative objects. Ex-
ample (50) shows the pronominal *łuveł* as a dative object of the verb *evəłti* 'trust':
it allows both a bound and a disjoint reading. How is the variable protected in this
case and could there be a chain formation violation?

Note, first of all that cross-linguistically dative objects show a variable be-
havior, depending on differences in case licensing and their relation to argument
structure (as is also expressed by the hierarchy in Testelets and Toldova 1998).
Even in English certain types of dative objects allow local binding of pronominals
(Webelhuth and Dannenberg 2006, see also Reuland 2011a: 382 n. 10). As illus-
trated in (50), dative in Khanty is expressed by the morpheme *-a* appended to the
accusative form *łuveł*, with the structure indicated in (50b). Here the Dative *-a*
morpheme is realized on top of the accusative form *łuveł*, and *-a* is in a position
comparable to that of a preposition. The simplest assumption is that *-a* licenses
*łuveł*'s case. If so, the latter is invisible for probing from the outside, and chain

---

32 Note that under the alternative analysis with *łuveł* as a specifier, *łuveł* will escape chain for-
mation by the locality principle regulating the distribution of dedicated possessive reflexives
(see Reuland 2011a: 5.6.4 for discussion).

33 We have no examples illustrating the behavior of *łuv* in DPs and nominalizations, but we
expect that it would not differ much from its use in PostPs.

condition effects are not expected. Whether IDI effects are to be expected depends on the precise relation of the dative object to the verbal grid, and specifically the role of -*a*. Here we discuss two possible factors, that each would be sufficient, but don't exclude each other.

If we assume for Khanty that the head *a* defines a projection that is separate from the main verb, the variable will not be on the same grid as its binder, creating *separation* (see the discussion in section 6.3.3). Furthermore, as is shown by the surface form, *łuveł* subsequently moves to the left of -*a*. We have insufficient evidence at this point to determine the nature of this position. However, if it is an A′-position rather than an A-position, this would also suffice to create complexity. (But note that separation is enough for the result to be derived.)

(50) a. Ivan$_i$ łuveła$_{i/k}$ evəlt-ł.
      Ivan   he.DAT   trust-NPST.3SG
      Ivan trusts himself / him.

   b.
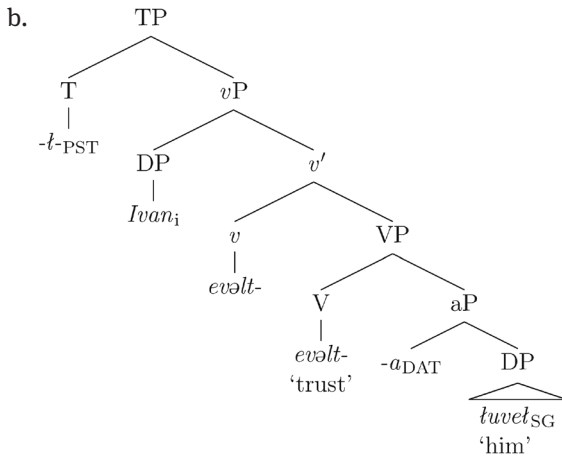
To derive the surface order a further leftward movement of subject and object will have to be assumed, as is standard.[34]

---

34 Again, this analysis of Khanty datives introduces no novel assumptions that are specific to binding. It does rely on two assumptions concerning the status of *a*. One is that it is a case licenser, which seems unavoidable. The other that it creates separation. This is an empirical assumption, but one that is entirely within the realm of options that is attested cross-linguistically. We leave it to the reader to assess how other approaches would deal with these facts.

8.3 Overt doubling: *łuv łuveł*

The covert doubling analysis just sketched finds independent support in the fact that Tegi Khanty speakers also use another reflexive strategy previously un-attested for Khanty – they may double the pronominal, for instance *łuv* for third person singular, to create an anaphor *łuv łuveł*:

(51) Maša   jast-əs   Ivan-a   što   łuv   łuveł   išək-ł-əłłe.
     Masha   say-PST.3SG   Ivan-DAT   that he   he.ACC   praise-NPST-SG.3SG
     Masha said to Ivan that he praises himself.

Not all speakers agree on the obligatoriness of the reflexive interpretation in (51), rather suggesting an interpretation depending on the context. The confounding factor here is the possibility of having a null subject, which allows for two different syntactic configurations for the sentence: 1. where *łuv* is the subject of the embedded clause, and *łuveł* is the object; 2. where the subject is null, and *łuv łuveł* is the object.

Hence, although interesting, this fact is not conclusive. However, in the following configuration this ambiguity is eliminated, and here *łuv łuveł* reflexivizes the predicate even in a pragmatically unfavorable context:

(52) a. Vas'a-jən$_v$ Maša-jən$_m$ par-s-əłłe   łuveł$_{v/??m}$ łap-ti.
        Vasja-2SG   Masha-2SG   ask-PST-SG.3SG   he.ACC   feed-INF
        Vasja asked Masha to feed him.
     b. ?Vas'a-jən$_v$ Maša-jən$_m$ par-səłłe
        Vasja-2SG   Masha-2SG   ask-PST-SG.3SG
        [łuv   łuveł]$_{m/*v}$ łapti.
        he   he.ACC   feed-INF
        Vasja asked Masha to feed herself.

While in (52a) for pragmatic reasons the most plausible interpretation for *łuveł* is the subject of the matrix clause, in (52b) the doubled pronoun *łuv łuveł* can be interpreted only locally, which, as the speaker noted, makes the sentence "sound funny" (in Russian and presumably in Khanty *feed oneself* can be used only in the sense 'earn enough money to buy food for oneself', but not in the sense 'consume food').

The anaphor *łuv łuveł* can occupy not only co-argument positions, as in (52), but also adjunct positions in postpositional phrases.

(53) a. Nemχojat   ɫuv   ɫuv   vŏrŋ-aɫ-a   ănt   potərt-əs.
      no.one   he   he   with-3SG-DAT   NEG   say-PST.3SG
      No one said anything about himself.

    b. Maša-jən$_m$ op-əɫ$_s$ piɫn   ɫuv   ɫuv$_{m/?*s}$ ɔɫŋ-aɫ-n
      Masha-2SG   sister-3SG   with   he   he   about-3SG-LOC
      potərt-əs.
      say-PST.3SG
      Masha talked with her sister about herself.

We, therefore, conclude, that *ɫuv ɫuveɫ* is obligatorily locally bound. Speakers use the doubled form *ɫuv ɫuveɫ* as the locally bound object of subject experiencer verbs (54), although it is not required. With agent-theme verbs *ɫuv ɫuveɫ* is also allowed, as in (52b).

(54) a. Maša-jən$_m$ Pet'a-jən-a$_p$ n'ɔt-s.   Pa   Pet'a-jən-a
      Masha-2SG   Petja-2SG-DAT   help-PST.3SG   and   Petja-2SG-DAT
      n'ɔt-əm   veraɫn   ɫuv   ɫuveɫa$_{m/*p}$ evəɫt-ɫ.
      help-PFT.PART   after   he   he.DAT   trust-NPST.3SG
      Masha helped Petja. After (she) helped Petja, she trusts herself.

    b. Maša-jən$_m$ Pet'a-jən-a$_p$ n'ɔt-s.   Pa   Pet'a-jən-a
      Masha-2SG   Petja-2SG-DAT   help-PST.3SG   and   Petja-2SG-DAT
      n'ɔt-əm   veraɫn   ɫuveɫa$_{p/?m}$ evəɫt-ɫ.
      help-PFT.PART   after   he.DAT   trust-NPST.3SG
      Masha helped Petja. After (she) helped Petja, she trusts him.

How can *ɫuv ɫuveɫ* be obligatorily locally bound in the relevant contexts? The simplest answer is to pursue a parallel with Tsaxur suggested by Toldova (p.c.): in case of *ɫuv ɫuveɫ* the first element has the case of antecedent and the second one the local case. The pronoun *ɫuv* needs to get its case licensed. The nearest licenser that is compatible is AgrS, consequently the feature sharing creates a dependency with via the subject via AgrS, which enforces the dependency.[35]

---

[35] Theoretically we need no more assumptions anyway than we need for Tsaxur (see, e.g. Toldova 1999). Both *ɫuv*, and *ɫuveɫ* have an unvalued Case feature, hence are possible goals. Intuitively, what we need is that both are in a position accessible for probing. Furthermore, as soon as one has been probed and valued by ObjAgr the other has to wait until the next probe has been merged. After valuation ObjAgr is inert, hence does not intervene. Informally, then, the technical means for such a derivation are available. For a detailed derivation more details of the structure of the Khanty TP would have to be determined than is presently the case. What we can conclude, though, is that such a strategy that involves case matching requires a role of Agree in the

This strategy has not been reported before in Khanty. This can be due to the fact that the form was thought to be not really noteworthy, and its interpretation escaped attention. It could also be that it reflects a recent development following the changes in the language under the pressure of Russian. Further instances of such pressure are the (partial) loss of its properties by the object triggering object agreement (cf. Nikolaeva 1999a) and the newly developed ability for the non-specific quantified expressions to occur as subjects, which is ungrammatical for the variety of the language described by Nikolaeva (1999a). As object agreement is crucial to licensing reflexivity with a locally bound pronominal, its disruption would lead to the development of the new strategy.

Summing up, like other languages investigated, despite initial appearances, Tegi Khanty is no exception to the generalization that reflexivity must be licensed. It avoids BFR, and employs at least two 'classic' reflexive strategies: a valence reduction operation accompanied by bundling of θ-roles and protecting a variable by means of doubling a pronominal.[36]

But, importantly, we also showed that, as a language, Tegi Khanty despite its initially puzzling property of allowing locally bound pronominals and the fact that it is spoken by people with an intriguing, though rapidly vanishing culture, is no more exotic than our neighbor Frisian.

9 Consequences

There is a number of competing accounts of Binding Theory. A number of them explore the idea of competition from various angles. Levinson (2000) suggests that pronominals can be bound if there is no more pragmatically dedicated competitor. Rooryck and Vanden Wyngaerd (2011, R&W) in a proposal to account for the local binding of 1st and 2nd person pronouns in Germanic and Romance, and 3rd person pronouns in Frisian, make use of an APBE principle:

---

syntactic encoding of binding. So it does restrict the choice of possible theories. We may leave a more detailed analysis for further research.

36 Other initially problematic languages are Zhuang, analyzed in Schadler (2014), and Fijian, described by by Dixon (1988), and later by Park (2012), and Schadler (2014). We refer to these works for details. In all these cases we found that appearances were deceptive, and that an in-depth analysis allowed an explanation of initially puzzling patterns. Haitian Creole, and similar languages are on our to-do list. But there are reasons to believe that the principles we argue for also apply to these languages.

(55) Absence of Principle B effects (APBE)
Pronouns behave like anaphors when a dedicated class of reflexive pro-
nouns is lacking. (p. 19).

Such competition based approaches would not help clarify the situation with
Khanty pronominals, and the necessity of object agreement to license a reflexive
interpretation would stay unaccounted for (unless our analysis would be essen-
tially incorporated). R&W's approach is Agree-based, though, hence it would
be conceptually compatible with the existence of Case-based dependencies. As
noted in fn. 13, their *-notation violates inclusiveness, though, which is again
problematic.

A competition theory of derived complementarity (Safir 2004) also faces
problems. Safir (2004) proposes a scale of relative degree of dependency for mor-
phological forms and a form to interpretation principle (FTIP), which "determines
whether or not a dependent interpretation is supported by a given form in syntac-
tic context" (p. 102).

(56) Generalization: Between any two anaphors, the more referentially specified
one is more dependent, whereas among non-anaphors, the more referen-
tially specified one is less dependent (p. 86) which for Germanic yields the
following scale:
SIG-SELF ≫ pronoun-SELF ≫ SIG ≫ pronoun ≫ r-expression

(57) FTIP (simplified): If X c-commands position Y, and *z* is the lexical form or
string that fills Y, and it is not the most dependent form available in Y, then
Y cannot be interpreted as identity dependent on X.[37]

While Safir's approach fares well for languages with several dependent forms in
complementary distribution, it does not provide the tools to analyze languages
with no visible competition, or where 'specialized' and less 'specialized' forms
coexist in a binding environment (like *ɫuv ɫuveɫ* and *ɫuveɫ*). Moreover, in its (2004)-
form it does not reduce the notion of a dependency to more basic properties
(hence does not comply with inclusiveness).

---

37 FTIP
If a) X c-commands position Y,
b) z is the lexical form or string that fills Y,
c) w is a single form more dependent than z,
d) both w and z could support the same identity-dependent interpretation if Y were
exhaustively dependent on X, then (the referential value for) Y cannot be interpreted as identity
dependent on X.

Hornstein (2000), and Boeckx, Hornstein and Nunes (2007) base their take on binding on movement-chains. According to these authors, an anaphor is the morphological offspring of a copy of the antecedent, which in cases of local binding can surface as a reflexive (see the discussion in Drummond et al. 2011). Much like Reuland (2011a) this approach reduces binding to a special case of another kind of grammatical relation, namely movement. However, it is largely ignorant of morphology: "the form of the anaphor plays no real role in the interpretation afforded" (Drummond et al. 2011: 399) and the theory does not take into account the various morphological realizations reflexivity can take across languages like verbal reflexives. Thus it has little to say about the Khanty facts we discussed.

Schlenker (2005) proposes a reduction of the binding conditions to semantic principles. Such a reduction would potentially be very interesting. One of the main principles of his theory is Non-Redundancy, which together with his further specific assumptions about the interpretation function is intended to capture the Condition B effects.

(58) Non-Redundancy:
    No object may occur twice in the same sequence of evaluation of a predicate.

But, the reduction to conditions on evaluation sequences embodies the claim that condition B effects have to be purely semantic; this, in turn, predicts that there cannot be languages with locally bound pronominals, which is contrary to fact, and provides no 'handle' on the cross-linguistic variation observed. Finally, the condition of Non-redundancy on evaluation sequences is not independently motivated.

10 Conclusions

In this paper we provided an overview of general principles applying to binding, leading to a puzzle about binding in Tegi Khanty. We provided a sketch of the grammar of Khanty, and offered an account for the use of locally bound personal pronouns it exhibits. The key factor here is the presence of object agreement, which is instrumental in creating the complexity needed to satisfy one of the key principles we argued for, IDI. In addition, the presence of ObjAgr prevents violating conditions on chain formation. We also described a reflexive strategy that was previously unattested for Khanty, and which involves doubling of the personal pronoun *łuv łuveł*. This strategy may be specific to the Tegi Khanty variety.

In the account presented, only general principles underlying grammatical processes were used. So, the analysis comes for *free*, modulo some specific properties of Khanty, such as the existence of object agreement. On a more general level, we showed that contra what is posited by advocates of approaches like E&L there are important universals to be discovered within the linguistic diversity. Thus, Khanty provides support for a general research strategy that doesn't take linguistic phenomena at face value, but only draws conclusions on the basis of detailed investigation. We contrasted the approach we presented with a number of competing binding theories, showing that currently we see no alternatives that can account for cases like Khanty in a parsimonious manner.

This paper is based on the discussion in chapter 2 of Anna Volkova's dissertation (Volkova 2014). The names of the authors appear in some alphabet's alphabetical order.

References

Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Storrs, CT: University of Connecticut PhD thesis.

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3). 373–415.

Boeckx, Cedric, Norbert Hornstein & Jairo Nunes. 2007. Overt copies in reflexive and control structures: A movement analysis. In Conroy, Anastasia, Chunyuan Jing, Chizuru Nakao & Eri Takahashi (eds.), *University of Maryland Working Papers in Linguistics* (UMWPiL) 15. 1–46.

Büring, Daniel. 2005. *Binding theory*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1981. *Government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Otero, Carlos & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Clements, George N. 1975. The logophoric pronoun in Ewe: Its role in discourse. *Journal of West African Languages* 10. 141–177.

Dimitriadis, Alexis & Martin Everaert. 25-05-2012. Characterizing reflexivization: semantic and syntactic perspectives. In *Debrecen Workshop on Argument Structure*. University of Debrecen, Hungary.

Dimitriadis, Alexis & Min Que. 2009. The doubly marked reflexive in Chinese. In Sobha L. Devi, Antonio Branco & Ruslan Mitkov (eds.), *Anaphora processing and applications*, 80–90. Berlin: Springer.

Dixon, Robert Malcolm Ward. 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press. Oxford: Oxford University Press.

Drummond, Alex, Dave Kush & Norbert Hornstein. 2011. Minimalist construal: Two approaches to A and B. In Boeckx, Cedric (ed.), *The Oxford Handbook of Linguistic Minimalism*, 396–426. Oxford: Oxford University Press.

Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5). 429–492.

Everaert, Martin. 1986. *The syntax of reflexivization*. Dordrecht: Foris.

Faltz, Leonard. 1985. *Reflexivization: a study on universal syntax*. New York and London: Garland Publishing.

Farmer, Ann & Robert Harnish. 1987. Communicative reference with pronouns. In Verschueren, Jef & Marcella Bertuccelli-Papi (eds.), *The pragmatic perspective*, 547–565. Amsterdam: Benjamins.

Franssen, Floris. 2010. Australian and Austronesian anaphora: An archival approach. Utrecht: Utrecht University MA thesis.

Gelderen, Elly van. 2000. *A history of English reflexive pronouns: person, self, and interpretability*. Amsterdam: Benjamins.

Geniušiene, Emma. 1987. *The typology of reflexives*. Berlin: Mouton de Gruyter.

Giorgi, Alessandra. 2007. On the nature of long distance anaphors. *Linguistic Inquiry* 38(2). 321–342.

Hale, Ken & Jay Keyser. 1991. On the syntax of argument structure. *Lexical Project Working Papers*, 34. Cambridge, MA: MIT Center for Cognitive Science.

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Amherst, MA: University of Massachusetts PhD thesis.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

Heine, Bernd & Hiroyuki Miyashita. 2008. The intersection between reflexives and reciprocals: A grammaticalization perspective. In König, Ekkehard & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 169–225. Berlin: Mouton.

Hoekstra, Jarich. 1994. Pronouns and case: on the distribution of Frisian *harren* and *se* 'them'. *Leuvense bijdragen* 83(1). 47–65.

Hornstein, Norbert. 2000. *Move! A minimalist theory of construal*. Oxford: Blackwell.

Jackendoff, Ray. 1992. Mme. Tussaud meets the binding theory. *Natural Language and Linguistic Theory* 10. 1–32.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2(1). 39–76.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Keenan, Edward & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

König, Ekkehard and Peter Siemund. 2000. Intensifiers and reflexives: A typological perspective. In Frajzyngier, Zygmunt & Traci Curl (eds.), *Reflexives: Forms and functions,* 41–74. Amsterdam: Benjamins.

Leben, William. 1973. Suprasegmental phonology. Cambridge, MA: MIT dissertation. Distributed by Indiana University Linguistics Club.

Levinson, Stephen. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the World*, 17th edn. Dallas, TX: SIL International. Online version: http://www.ethnologue.com.

Lidz, Jeffrey. 1995. Morphological reflexive marking: Evidence from Kannada. *Linguistic Inquiry* 26. 705–710.

Lidz, Jeffrey. 2001. Condition R. *Linguistic Inquiry* 32. 123–140.

Lust, Barbara, Kashi Wali, James W. Gair & Karumuri V. Subbarao (eds.). 2000. *Lexical anaphors and pronouns in selected South Asian languages: A principled typology*. Berlin: Mouton de Gruyter.

Mahajan, Anoop K. 1990. *The A/A-bar distinction and movement theory*. Cambridge, MA: MIT dissertation.

Marelj, Marijana. 2004. *Middles and argument structure across languages*. Utrecht: LOT International Dissertation Series.

Marelj, Marijana & Eric Reuland. Forthcoming. Clitics and reflexives – reducing the lexicon-syntax parameter. In Everaert, Martin, Marijana Marelj, Eric Reuland & Tal Siloni (eds.), *Concepts, syntax, and their interface: The Theta-system*. Cambridge, MA: MIT Press.

Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*, 3rd edn. Paris: UNESCO Publishing. Online version: http://www.unesco.org/culture/en/endangeredlanguages/atlas.

Moyse-Faurie, Claire. 2008. Constructions expressing middle, reflexive and reciprocal situations in some Oceanic languages. In König, Ekkehard & Volker Gast (eds.), *Reciprocals and reflexives: Theoretical and typological explorations*, 105–168. Berlin: Mouton.

Nikolaeva, Irina. 1995. *Obdorskij dialekt chantyjskogo jazyka (The Obdorsk dialect of Ostyak – in Russian)*. Hamburg: Mitteilungen der Societas Uralo-Altaica.

Nikolaeva, Irina. 1999a. *Ostyak*. Munich: Lincom Europa.

Nikolaeva, Irina. 1999b. Object agreement, grammatical relations, and information structure. *Studies in Language* 23(2). 331–376.

Nikolaeva, Irina. 2001. Secondary topic as a relation in information structure. *Linguistics* 39(1). 1–49.

Park, Karen. 2012. *The selective properties of verbs in reflexive constructions*. Oxford: University of Oxford doctoral dissertation.

Pesetsky, David & Esther Torrego. 2007. The Syntax of valuation and the interpretability of features. In Karimi, Simin, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation,* 262–294. Amsterdam: Benjamins.

Pica, Pierre. 1984. Liage et contiguïté. In Milner, Jean-Claude (ed.), *Recherches sur l'anaphore*, 119–164. Paris: CNRS.

Pica, Pierre. 1986. Subject, tense and truth. In Guéron, Jacqueline, Hans-Georg Obenauer & Jean-Yves Pollock (eds.), *Grammatical Representation*, 259–291. Dordrecht: Foris.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.

Reinhart, Tanya. 2002. The theta system: an overview. *Theoretical Linguistics*, 28. 229–290.

Reinhart, Tanya. 2006. *Interface strategies: Reference set computation*. Cambridge, MA: MIT Press.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24. 657–720.

Reinhart, Tanya & Tal Siloni. 2005. The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry* 36(3). 389–436.

Reuland, Eric. 1995. Primitives of binding. In GLOW Newsletter 1995. Paper presented at GLOW 18, Tromsø, Norway.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry*, 32(3). 439–492.

Reuland, Eric. 2006. Gradedness: Interpretive dependencies and beyond. In Fanselow, Gisbert, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in Grammar: Generative Perspectives,* 45–69. Oxford: Oxford University Press.

Reuland, Eric. 2011a. *Anaphora and language design*. Cambridge, MA: MIT Press.

Reuland, Eric. 2011b. Syntax and interpretation systems: How is their labour divided? In Boeckx, Cedric (ed.), *The Oxford Handbook of Linguistic Minimalism*, 377–395. Oxford: Oxford University Press.

Reuland, Eric & Martin Everaert. 2010. Evans and Levinson's cabinet of curiosities: Should we pay the fee? *Lingua* 120(12). 2713–2716.

Reuland, Eric & Sigríður Sigurjónsdóttir. 1997. Long distance binding in Icelandic: Syntax or discourse? In Bennis, Hans, Pierre Pica & Johan Rooryck (eds.), *Atomism in binding*, 323–334. Dordrecht: Foris.

Reuland, Eric & Yoad Winter. 2009. *Binding without identity: Towards a unified semantics for bound and exempt anaphors*. In Sobha L. Devi, Antonio Branco & Ruslan Mitkov (eds.), *Anaphora processing and applications,* 69–79. Berlin: Springer.

Richards, Norvin. 2002. A distinctness condition on linearization. *West Coast Conference on Formal Linguistics (WCCFL)* 20. 470–483.

Rooryck, Johan & Guido Vanden Wyngaerd. 2011. *Dissolving binding theory*. Oxford: Oxford University Press.

Safir, Ken. 2004. *The Syntax of anaphora*. Oxford: Oxford University Press.

Schadler, Dagmar. 2014. *Reflexivity: Licensing or enforcing*. Utrecht: LOT International Dissertation Series.

Schladt, Mathias. 2000. The typology and grammaticalization of reflexives. In Frajzyngier, Zygmunt & Traci Curl (eds.), *Reflexives: Forms and functions*, 103–124. Amsterdam: Benjamins.

Schlenker, Philippe. 2005. Non-redundancy: Towards a semantic reinterpretation of binding theory. *Natural Language Semantics*, 13(1). 1–92.

Sells, Peter. 1987. Aspects of logophoricity. *Linguistic Inquiry* 18. 445–479.

Testelets, Yakov & Svetlana Toldova. 1998. Refleksivnyje mestoimenija v dagestanskix jazykax i tipologija refleksiva (Reflexive pronouns in the Daghestanian languages and the typology of reflexive – in Russian). *Voprosy Jazykoznania* (4). 35–57.

Toldova, Svetlana. 1999. Mestoimennye sredstva podderžanija referencii (Pronominal means of reference tracking). In Kibrik, Alexander & Yakov Testelets (eds.), *Elementy caxurskogo jazyka v tipologičeskom osveščenii (Studies in Tsakhur: a typological perspective – in Russian)*, 629–674. Moscow: Nasledie.

Ura, Hiroyuki. 2003. Case. In Baltin, Mark & Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, 334–373. Oxford: Blackwell.

Vinokurova, Nadezhda. 2005. *Lexical categories and argument structure: A study with reference to Sakha*. Utrecht: LOT International Dissertation Series.

Volkova, Anna. 2014. Licensing reflexivity: Unity and variation among selected Uralic languages. Utrecht: LOT International Dissertation Series.

Webelhuth, Gert & Clare J. Dannenberg. 2006. Southern American personal datives: The theoretical significance of dialectal variation. *American Speech* 81(1). 31–55.

Zec, Draga. 1985. Objects in Serbo-Croatian. *Berkeley Linguistics Society (BLS) 11*. 358–371.

Zribi-Hertz, Anne. 1995. Emphatic or reflexive? On the endophoric character of French *lui-même* and similar complex pronouns. *Journal of linguistics* 31. 333–374.

Zribi-Hertz, Anne & Herby Glaude. 2007. La réflexivité en haïtien: réexamen et comparaison. In Gadelii, Karl & Anne Zribi-Hertz (eds.), *Grammaires créoles et grammaire comparative*, 151–182. Saint-Denis: Press Universitaires de Vincennes.