

Number neutrality and anaphoric update of pseudo-incorporated nominals in Persian (and weak definites in English)*

Manfred Krifka
*ZAS Berlin &
Humboldt Universität zu Berlin*

Fereshteh Modarresi
ZAS Berlin

Abstract Persian shows differential object marking. We argue that objects without the object marker *rā* are pseudo-incorporated in the verbal predicate; specifically, existential closure over the vP binds the event variable, and nominals in the vP are interpreted as dependent definites with respect to the event. This results in an apparent number neutrality and a maximality interpretation of pronouns typical of the E-type strategy. The semantic contribution of nominals is modeled in DRT. We also argue that weak definites in English are interpreted similar to pseudo-incorporated nominals in Persian, and offer an explanation why they are restricted to readings referring to institutionalized activities.

Keywords: incorporation, discourse referents, Persian, differential object marking, existential closure, E-type pronouns, DRT, weak definites

1 Introduction

In this article we discuss the semantics of pseudo-incorporated object nominals (PINs) in Persian. They form a closer syntactic tie with the verb than regular objects (Massam 2001 for Niuean; Farkas & de Swart 2003 for Hungarian, Massam 2009, Borik & Gehrke 2015). While PINs are not morphologically incorporated, (they can be syntactically expanded), they are syntactically less elaborate than regular objects. In Niuean and Hungarian, they lack articles, and hence have been analyzed as NPs, not DPs (we will speak of “nominals”). PINs are typically realized adjacent to the verbal head, but can also be moved to focus positions. As for their interpretation, PINs are generally number-neutral and non-specific.

* Thanks to the Deutsche Forschungsgemeinschaft (DFG) for financial support through Sonderforschungsbereich 632 and the project ANAPIN, as well as the Bundesministerium für Bildung und Forschung (BMBF) for support of ZAS. We also thank the audiences at presentations at ZAS, the Research Institute for Linguistics at the Hungarian Academy of Sciences, the Queen Mary University London, and the anonymous reviewers of SALT 26 for valuable comments.

In Persian, an object can be realized with an object marker *rā* or without it, a phenomenon known as differential object marking. Without *rā*, a singular bare object is interpreted as indefinite and number-neutral interpretation; with *rā*, as singular and typically definite, Hence objects without *rā* marker can be analyzed as PINs (cf. also Dayal 2011, 2015 for Hindi; Öztürk 2005 for Turkish; Baker 2014 for Tamil, as structurally similar languages).

- (1) a. *Leili sib-rā khærid.* b. *Leili sib khærid.*
 Leili apple-OM bought.3SG Leili apple bought.3SG
 ‘Leili bought the apple.’ ‘Leili bought an apple / apples.’

Another property of PINs is that they are problematic antecedents. Anaphoric uptake is a controversial topic for morphological incorporation; there is evidence that it is possible in at least some languages (van Geenhoven 2008 for West Greenlandic, Mithun 2010 for Kapampangan). For PINs, anaphoric uptake is sometimes possible (Asudeh & Mikkelsen 2000 for Danish, Massam 2001 for Niuean, and Dayal 2011 for Hindi). For PINs, anaphoric uptake is sometimes possible, but restricted, cf. Asudeh & Mikkelsen 2000 for Danish; Massam 2001 for Niuean; Dayal 2011 for Hindi. The most thorough discussion of this issue is by Farkas & de Swart (2003) for Hungarian. According to their discussion, PINs are neither discourse transparent nor discourse opaque, but what they call “discourse translucent”. Farkas & de Swart find that anaphoric uptake is possible for some speakers with null anaphora, as in (2).

- (2) *János_i beteg_j vizsgált a rendelőben.*
 Janos_i patient.ACC_j examine.PAST the office.in
 ‘Janos_i patient_j-examined in the office.’

Ø_i túl súlyosnak találta ??ő_{t_j} / Ø_j és beutaltatta Ø_j a kórházba.
 pro_i too severe.DAT find he_j.ACC pro_j and intern.CAUSE.PAST pro_j the hospital.in
 ‘He_i found him_j too sick and sent him_j to hospital.’

However, Yanovich (2008) reports that overt pronouns are possible, as in (3):

- (3) *A bátyám házat_i vett a múlt héten. Egész vagyont adott érte_i.*
 ‘The brother house_i-bought last week. He spent a fortune for it_i.’

We will develop a semantic representation of PINs within the framework of Discourse Representation Theory (DRT) that captures their non-specificity, their number neutrality, and their restricted anaphoric potential. It makes a new prediction, a maximality effect in their interpretation. We will first turn to two existing proposals that attempt to explain the anaphoric potential of PINs.

2 Discourse translucency: Farkas & de Swart (2003)

Farkas & de Swart (2003) analyze PINs in DRT (Kamp & Reyle 1993; Kamp, Reyle & van Genabith 2011). In order to simplify the discussion, we will present their theory with Persian data. We will represent discourse representation structures (DRSs) as “flat” structures; $[x_1 x_2 \mid \Phi(x_1), \Psi(x_1, x_2)]$ stands for a DRS with two discourse referents (DRs) x_1, x_2 that satisfy the conditions $\Phi(x_1)$ and $\Psi(x_1, x_2)$. As usual, a DRS K is interpreted with respect to (wrt) a model $\langle A, F \rangle$, where A is the universe of discourse, and F is a function that maps the constants of the DRS language to entities, sets or tuples constructed from A . The DRS K is true wrt a model $\langle A, F \rangle$ iff there is a mapping g from the DRs of K into A such that each condition of K is verified by g wrt $\langle A, F \rangle$. For example, $[x_1 x_2 \mid \Phi(x_1), \Psi(x_1, x_2)]$ is true wrt $\langle A, F \rangle$ iff there is a g such that $g(x_1) \in A$, $g(x_2) \in A$, $g(x_1) \in F(\Phi)$, and $\langle g(x_1), g(x_2) \rangle \in F(\Psi)$. The condition $x = \alpha$ is verified iff $g(x) = F(\alpha)$.

The following examples illustrate the interpretation of a sentence with a regular indefinite marked by the singular indefinite article *yek* (4), and the corresponding pseudo-incorporating structure (5), following Farkas & de Swart (2003). K_0 is the empty input DRS that is updated (+) by a clause.

- (4) $K_0 + [Leili \text{ [[yek sib] khærid}]]$
 $= [x_1 x_2 \mid x_1 = \text{LEILI, APPLE}(x_2), \text{BUY}(x_1, x_2)]$, two DRs introduced: x_1, x_2
- (5) $K_0 + [Leili \text{ [sib khærid}]]$
 $= [x_1 \mid x_1 = \text{LEILI, APPLE}(y), \text{BUY}(x_1, y)]$ just one DR introduced: x_1
 $= K_1$ y : thematic argument

While regular indefinites introduce a DR, PINs don't; they are represented by a free variable y , a “thematic argument”. Lacking a determiner that expresses number information, thematic arguments are number neutral. This leads to the question how they are interpreted – more specifically, how conditions with thematic arguments are verified. Farkas & de Swart (2003) propose that an assignment g verifies a DRS-condition $\Phi(\alpha_1, \dots, \alpha_n)$ wrt a model $\langle A, F \rangle$ iff there is some sequence $\langle a_1, \dots, a_n \rangle$, with $a_1, \dots, a_n \in A$, such that $\langle a_1, \dots, a_n \rangle \in F(\Phi)$, and if α_i is a DR, $a_i = g(\alpha_i)$. If α_i is a thematic argument, there is no restriction. Hence, free variables are interpreted existentially. For example, an assignment g verifies the condition $\text{APPLE}(y)$ wrt $\langle A, F \rangle$ iff there is an a , $a \in A$, such that $a \in F(\text{APPLE})$, and it verifies the condition $\text{BUY}(x_1, y)$ iff there is an a , $a \in A$ such that $\langle g(x_1), a \rangle \in F(\text{BUY})$.

This explains why thematic arguments are problematic antecedents: They are not represented by DRs. But it also raises the question why they can be taken up at all. For this, Farkas & de Swart propose a rather complex rule. It can be justified that the rule is complex, as the complexity arguably corresponds to the additional effort needed to refer back to PINs. But the specific rule that they propose does not work for technical reasons, as has been argued for by Yanovich (2008).

Here are the details. Farkas & de Swart (2003) suggest that when a suitable DR cannot be found for a non-overt pronominal in a DRS K , a new DR x_j is introduced with a condition of the form $x_j \approx y_i$, where y_i is a thematic argument that is part of a preceding condition $\Phi(\dots, y_i, \dots)$ in K or in a DRS that is superordinate to K . An assignment g verifies the condition $x_j \approx y_i$ if g maps y_i onto an individual a_i that is the i -th element of an n -tuple $\langle \dots, a_i, \dots \rangle$ that verifies the condition $\Phi(\dots, y_i, \dots)$. This is illustrated in (6), a continuation of (5):

- (6) $K_1 + [Majnoon\ khord = \emptyset]$
 $= [x_1 \quad | \quad x_1 = LEILI, APPLE(x_2), BUY(x_1, y_2)$
 $\quad x_3\ x_4 \quad | \quad x_3 = MAJNOON, x_4 \approx y_2, EAT(x_3, x_4)]$
- true wrt an assignment g and a model $\langle A, F \rangle$
 iff a. $g(x_1) = F(LEILI)$
 b. there is an a_2 such that $a_2 \in A$ with $a_2 \in F(APPLE)$
 c. there is a sequence $\langle a_1, a_2 \rangle \in AxA$ with $g(x_1) = a_1$ and $\langle a_1, a_2 \rangle \in F(BUY)$
 d. $g(x_3) = F(MAJNOON)$
 e. $g(x_4) = a_2$,
 f. $\langle g(x_3), g(x_4) \rangle \in F(EAT)$

One problem of this interpretation was pointed out by Yanovich (2008). As a_2 is bound independently by two existential quantifications in (b) and (c), the conditions in (6) would be satisfied if there is an apple, and there is something that Majnoon bought (not necessarily an apple!). Another problem is the following: As a_2 is bound by existential quantifiers, it is not possible to refer back to the specific value of a_2 , as (e) attempts to do. We need a representation that binds a_2 across the conditions (b), (c), and relates the new DR x_4 to that entity. Furthermore, it is unclear why PINs are number neutral in the first place: They may lack a determiner, but they do have the morphological feature singular.

3 Number-neutral discourse referents: Modarresi 2015

Farkas & de Swart (2003) also cannot explain why covert, but not overt pronominals should be able to access thematic arguments – if their observation is correct, cf. (3). Modarresi (2015) points out that covert pronouns are special insofar as they, being covert, cannot express number. Hence they are suited to pick up number-neutral PINs as antecedents. Modarresi makes use of standard version of DRT by Kamp & Reyle (1993), who distinguish between singular, plural, and number-neutral DRs, and argues that PINs introduce number-neutral DRs (rendered by ξ_i).

- (7) *Leili porteghal khærid. Majnoon khord-∅ /-ʔesh/ -ʔeshoon.*
 Leili orange bought.3SG Majnoon ate-PRO/-it/-them
 ‘Leili bought orange(s). Majnoon ate it / them.’

$$\begin{array}{l|l} [x_1 \ \xi_2 & | \ x_1 = \text{LEILI, ORANGE/S}(\xi_2), \text{BUY}(x_1, \xi_2) \\ x_3 & | \ x_3 = \text{MAJNOON, ATE}(x_3, \xi_2)] \ \xi_2: \text{number-neutral DR} \end{array}$$

If world knowledge suggests an atomic / non-atomic interpretation of number-neutral DR, anaphoric reference with singular / plural overt pronouns is possible. For example, (8a) and (b) are fine as people that buy an apartment normally buy just one, whereas people that buy a carrot buy more than one.

- (8) a. *Leili apartman_i khærid. Gheimæt-esh_i bala bood.*
 Leili apartment bought.3SG. price-its high was.3SG
 ‘Leili bought apartment(s). Its price was high.’
- b. *Leili hævij_i khærid. Majnoon khord-eshoon_i.*
 Leili carrot bought.3SG. Majnoon ate-them.
 ‘Leili bought carrot(s). Majnoon ate them.’

Modarresi’s account suggests an inherent connection between the two main properties of PINs: Their number-neutral interpretation, and their reduced ability to serve as antecedent for anaphoric elements. However, it does not explain why PINs are number neutral. Also, it suggests that anaphoric uptake with covert pronouns, or with overt pronouns in cases where world knowledge suggests an atomic or a non-atomic entity, should be straightforward. But anaphoric uptake is sub-optimal; a speaker that intends to continue talking about the entities in question would rather not introduce them by PINs.

4 E-Type pronouns, maximality, and abstraction & summation

We will suggest a new type of analysis for PINs that explains their number neutrality and restricted anaphoric potential. Anaphora to PINs will be analyzed as E-type pronouns (Evans 1980), i.e. as pronouns with quantifier antecedents that do not c-command them. A standard example is (9):

- (9) *Few congressmen_i admire Kennedy, and they_i are very junior.*
 ‘There are (only) few congressmen that admire Kennedy,
 and the congressmen that admire Kennedy are very junior.’

E-type pronouns come with a signature effect: As the paraphrase *the congressmen that admire Kennedy* in (9) suggests, they are interpreted maximally. Heim (1990) discusses contrasts like (10a,b):¹

1 A better minimal pair would be *Three wine glasses broke last night. They had been expensive* and (10b). However, numerals trigger scalar implicatures, here ‘exactly three...’, and then *they* would also refer to the sum of all wine glasses that broke. Another candidate would be (10a) and *At least one wine glass ... Now the continuation with the plural pronoun They... appears less felicitous, due to a syntactic clash with the singular of at least one wine glass.*

- (10) a. *A wine glass broke last night. It was very expensive.*
 (o.k. if several wine glasses broke last night and only one was expensive.)
 b. *At least three wine glasses broke last night. They were very expensive.*
 (all the wine glasses that broke last night were very expensive).

E-type pronouns are often seen as involving a descriptive theory of pronouns (cf. Neale 1990, Heim 1990, Elbourne 2005), but this should not be taken as their defining property, just as a particular implementation (cf. Nouwen *subm.*). What distinguishes them from regular indefinites is that their antecedents are not supposed to introduce DRs with an unlimited life span. Yet anaphoric reference is possible, as (9) and (10b) show. In DRT, such anaphoric uptake has been treated by a special operation over DRSs (Kamp & Reyle 1993, Hardt 2003), as in (11):

- (11) *John beats most donkeys he owns. They complain.*
 $[x_1 \mid x_1 = \text{JOHN}, [x_2 \mid \text{DONKEY}(x_2), \text{OWN}(x_1, x_2)] \langle \text{MOST } x_2 \rangle [\mid \text{BEAT}(x_1, x_2)]$
 $\xi_3 \mid \xi_3 = \Sigma x_2 [x_2 \mid \text{DONKEY}(x_2), \text{OWN}(x_1, x_2), \text{BEAT}(x_1, x_2)]]$

The first clause introduces a condition stating that most ways to extend the assignment g to include the DR x_2 such that the DRS $[x_2 \mid \text{DONKEY}(x_2), \text{OWN}(x_1, x_2)]$ is true are also ways that make the DRS $[\mid \text{BEAT}(x_1, x_2)]$ true. In a subsequent sentence, an anaphoric expression like *they* or *the donkeys* can be used to “synthesize” a discourse referent “out of what the DRS of the first sentence provides” (Kamp, Reyle & van Genabith 2011). In (11) a number-neutral DR ξ_3 is constructed that is identified with the sum of entities for which it holds that they are donkeys that John owns that he beats.

The general rule involves the presence of a condition like $K' \ Q \ K''$, the abstraction of a DR of this condition, the summation over the values of this abstraction, and the identification with a new number-neutral DR, as in (12):

- (12) If K contains a duplex condition $K' \ Q \ K''$, form the union $K''' = K' \cup K''$, choose a DR x from K''' , add a new DR ξ to K , add the condition $\xi = \Sigma x \ K'''$ where $\Sigma x \ K'''$ is interpreted wrt an assignment g and a model $\langle A, F \rangle$ as the the sum of all $a \in A$ such that there is an extension g' of g with $g'(x) = a$ that verifies K''' wrt $\langle A, F \rangle$.

The maximality effect arises due to the summation operation $\Sigma x \ K'''$ in this rule. These are the representations we get for (10a) vs. (b):

- (13) a. $[x_1 \mid \text{WINEGLASS}(x_1), \text{BROKE}(x_1), \text{EXPENSIVE}(x_1)]$
 b. $[\mid [x_1 \mid \text{WINEGLASS}(x_1)] \langle \geq 3 \ x_1 \rangle [\mid \text{BROKE}(x_1)]$
 $\xi_2 \mid \xi_2 = \Sigma x_1 [x_1 \mid \text{WINEGLASS}(x_1) \langle \geq 3 \ x_1 \rangle [\mid \text{BROKE}(x_1)], \text{EXPENSIVE}(\xi_2)]]$

An assignment g verifies the condition $K' \langle \geq n \ x \rangle \ K''$ wrt a model $\langle A, F \rangle$ iff g can be extended to g' in n -many ways such that $g'(x) \in A$, and g verifies K' and K''

that it maps x_2 to an object such that it verifies the conditions – it does allow for there being more than one such extension. As existential quantification EC does not have alternatives like numerals, it does not trigger the scalar implicature that there is just one such extension. Furthermore, anaphoric reference is achieved via the established abstraction and summation rule. Following Modarresi (2015), this explains why number-neutral covert pronominals are particularly suited for anaphoric reference to PINs, and why, depending on world knowledge, overt singular and plural pronouns can be used as well, as in (8a), or (8b).

The analysis also makes a new prediction, namely that we should observe maximality in the anaphoric uptake of PINs. This is indeed the case. (17) is pragmatically odd, as *digari* ‘another’ would have to refer to the sum of the houses that Ali has, which does not allow for additional houses that Ali has. However, cases in which the first clause introduces a DR with a *yek*-marked indefinite as in (18) are not quite perfect either, as they implicate that that Ali has not more than one house (*yek* is more similar to English *one* than to the indefinite article, *a*).

(17) *Ali khaneh darad. # Khane-ye-digari ham dard ke ejareh mideh.*
 Ali house have.3SG house-LINKER-other also has that rent give.3SG
 ‘Ali has house(s). He also has another house that he rents out.’

(18) *Ali yek khaneh darad. (#) Khane-ye-digari ham dard ke ejareh mideh.*
 Ali a house have.3SG house-LINKER-other also has that rent give.3SG
 ‘Ali has a house. He also has another house that he rents out.’

In conclusion, the proposed analysis of PINs allows us to explain a number of phenomena: The fact that PINs are syntactically close to the verbal predicate, hence subject to existential closure over the vP, the apparent number neutrality of the interpretation, the somewhat more difficult anaphoric uptake, the preference for covert anaphora, and the maximality effect in case of anaphoric uptakes.

6 PINs as event-dependent definites

Object PINs in Persian stand in minimal opposition with object nominals that are marked by the accusative postposition *rā*. (1a) suggests that *rā* is a definiteness marker, which has often been assumed (e.g. Ghomeshi 2003). However, the object marker also occurs with the indefinite markers *yek* and *-i*, leading to a specific interpretation, as in (19a,b) (Karimi 2003), and with indefinites in the restrictor of generic quantifiers, as in (19c,d) (cf. Dabir-Moghaddam 1992, Krifka 2001).

(19) a. *yek film-rā didaem.* b. *film-i-rā didaem.*
 a movie-OM see.1SG movie-IDEF-OM see.1SG
 ‘I saw one of the movies’ ‘I saw a certain movie.’

- c. *serke shir-rā mi-borræd.* d. *kowboye-ha tanbako-rā mi-javand.*
vinegar milk-OM dur-curdle.3SG cowboy-PL tobacco-OM DUR-chew.3PL
‘Vinegar curdles milk.’ ‘Cowboys CHEW tobacco.’

Modarresi (2014) proposed that *rā*-marking indicates scrambling of the object nominal out of the vP (cf. also Browning & Karimi 1994, Karimi 2003). In the current setting, a *rā*-marked object will escape existential closure over the vP. In case there is a generic quantifier, it is interpreted in the restrictor, as in (19c,d).

The definite interpretation of bare nominals as in (1a) is surprising: Following our analysis of such nominals as indefinites in Section 5 we should expect an indefinite interpretation of *rā*-marked bare nominals. We only predict that *rā*-marked bare nominals lose their number neutrality and have a singular interpretation, that is, that they can be taken up by singular pronouns, that this uptake is easy, and that no maximality effect can be detected. But this is not what we find.

For this reason, we would like to explore the option that bare nominals always have a **definite** interpretation. But as this definiteness is not detectable inside the vP, it must be dependent on some other variable. We suggest that this is a Davidsonian event argument that undergoes existential closure, which indirectly binds any DR dependent on it. That is, we revise our analysis of (15) as follows:

- (20) $K_0 + [Leili_1 EC [{}_{vP} t_1 sib\ khærid]]$
 $= [x_1 \mid x_1 = LEILI, \exists [e_3\ x_2 \mid x_2 = APPLE-OF(e_3), BUY(x_1, x_2, e_3)]]$
 $= K_1$

We assume an event DR e_3 introduced by the event argument of the verb. The bare noun *sib* is interpreted as a function that identifies the unique apple of this event and introduces a DR x_2 for it. The nominal has to be interpreted within existential closure, as it is dependent on the event variable, explaining the apparent indefiniteness. The resulting interpretation allows for more than one apple being bought by Leili, as there could be multiple buying events, which explains the apparent number-neutrality. Anaphoric uptake is achieved by abstraction and summation, cf. (21).

- (21) $K_1 + [Majnoon\ khord-\emptyset]$
 $= [x_1 \mid x_1 = LEILI, \exists [x_2\ e_3 \mid x_2 = APPLE-OF(e_3), BUY(x_1, x_2, e_3)]]$
 $\xi_5\ x_4 \mid x_4 = MAJNOON, \xi_5 = \Sigma x_2 [x_2\ e_3 \mid x_2 = APPLE-OF(e_3), BUY(x_1, x_2)],$
 $EAT(x_4, \xi_5)]$

When bare nominals are *rā*-marked and scramble out of the vP, they cannot be related to the event of the verbal predicate. Under the assumption of a uniform interpretation of bare nominals as dependent definites, there must be another DR or entity that they can take as argument. This is illustrated in (22).

- (22) *tooye sæbæd yek sib væ yek golabi bood. Leili sib-rā bærdašt.*
 in basket an apple and a pear was.3SG Leili apple-OM took.3SG
 ‘There was apple and a pear in the basket. Leili took the apple.’

$[X_1 \xi_2 x_3 x_4 X_5 \mid \text{BASKET}(X_1), \text{APPLE}(X_2), \text{PEAR}(X_3), X_4=X_2 \oplus X_3, \text{IN}(X_1, X_4),$
 $x_6 x_7 \mid x_6=\text{LEILI}, x_7=\text{APPLE-OF}(X_4), \exists[e_8 \mid \text{TAKE}(x_6, x_7, e_8)]]$

The first clause introduces a plural DR X_4 . The scrambled bare nominal in the second clause is a functional definite that identifies the unique apple of X_4 , thus establishing coreference with the DR x_3 . We can assume the same mechanism in cases like (23), in which the functional definite is applied to a single apple.

- (23) *Yek sib too sæbæd bood. Leili sib-rā bærdašt.*
 an apple in basket was.3SG Leili apple-OM took.3SG
 ‘There was an apple in the basket. Leili took the apple.’

$[X_1 x_2 \mid \text{BASKET}(X_1), \text{APPLE}(X_2), \text{IN}(X_1, X_2),$
 $x_3 x_4 \mid x_3=\text{LEILI}, x_4=\text{APPLE-OF}(X_2), \exists[e_5 \mid \text{TAKE}(x_3, x_4, e_5)]]$

The analysis of bare nominals as dependent definites leads to a uniform interpretation of such expressions inside and outside the scope of vP. It also predicts that in case the bare nominal should be related to a previously established DR, it should scramble, as this prevents the event-dependent interpretation.

The current theory relates the interpretational properties of bare nominals to scrambling out of the vP, hence should also apply to subjects, which are generated within vP. Subjects do not have a morphosyntactic marker of scrambling, but may form a phonological phrase with the verb, in which case they receive accent and are interpreted as indefinite (cf. (24b), Modarresi 2014).

- (24) a. *Ketāb oftad.* b. ***Ketāb*** *oftad.* boldface: accent
 book fell.3SG book fell.3SG
 ‘The book fell.’ ‘Some book(s) fell.’

We also predict the correct interpretation of PINs in generic sentences. (19c,d) showed that bare object nominals with *rā*-marking are interpreted in the restrictor of the generic operator. This is consistent with standard assumptions of the semantic partition under a generic operator (cf. Diesing 1992, Krifka e.a. 1995):

- (25) GEN [*serke*₁ *shir-rā*₂ [_{vP} *t*₁ *t*₂ *mi-borræd*]]

$[\mid [X_1 x_2 e_3 \mid x_1 = \text{VINEGAR-OF}(e_3), x_2 = \text{MILK-OF}(e_3), \text{IN}(X_1, X_2, e_3)]$
 $\langle \text{GEN } e_3 \rangle [e_4 \mid e_4 \subseteq e_3, \text{CURDLE}(X_1, X_2, e_4)]$

We conclude that a uniform interpretation of bare nominals in Persian as dependent definites is tenable and in fact well motivated. In Section 9 we will argue that explicit definiteness marking actually shows up in related cases in English.

7 Overtly marked indefinites in Persian

We now turn to nominals that are marked as indefinite. If scrambling of objects must be indicated by *rā*, then a *yek*-marked nominal without *rā* as in (4) is interpreted within the vP, resulting in interpretation (26a).

- (26) $K_0 + [Leili_1 EC [_{VP} t_1 [yek\ sib] khærid]]$
 a. $[x_1 | x_1 = LEILI, \exists[e_2\ x_3 | APPLE(x_3), \#(x_3)=1, BUY(x_1, x_3, e_2)]]$
 b. $[x_1\ x_3 | x_1 = LEILI, APPLE(x_3), \#(x_3)=1, \exists[e_2 | BUY(x_1, x_2, e_3)]]$

While the DR x_3 is not dependent on the event, it is introduced under existential closure. Consequently, we predict number-neutrality and a maximality interpretation of anaphoric uptake. But this is not what we find. We propose that this is because (26a) is semantically indistinguishable from the PIN interpretation (20). The indefinite marker *yek* contributes the number information $\#(x_3)=1$, but has no effect on the overall interpretation. Hence this interpretation is blocked by the expression involving a PIN, (20).

But indefinites can also have a wide-scope interpretation independent of their syntactic position, as argued for in Kamp & Reyle (1993, 3.7.3). This would lead to the interpretation (26b), without scrambling of the indefinite nominal outside of the vP (van Geenhoven 1998, Geurts 2000).

To be sure, *yek*-marked indefinites can be *rā*-marked, which indicates scrambling outside of the vP according to our analysis. But *rā*-marking of *yek*-marked indefinites is disfavored. A plausible reason for this is that the wide-scope interpretation can be achieved without scrambling, and so scrambling is blocked. We find it for instance when the indefinite nominal should take wide scope with respect to another quantifier, as in (27), where it c-commands the other quantifier:

- (27) *yek ketab-rā har daneshjoo-i bayad be-khoonad*
 a book-OM each student-I must SUBJUNCTIVE-read.3SG
 ‘There is a book that each of the students must read.’

There is another way of expressing indefiniteness, by *i*-marking, which also can be combined with *rā*-marking:

- (28) a. $[Mæn_1 [_{VP} t_1 roobah-i didæm]]$ b. $[Mæn_1 roobah-i-rā_2 [_{VP} t_1 t_2 didæm]]$
 I fox-I saw.1SG I fox-I-OM saw.1SG
 ‘I saw a fox’ (not: foxes) ‘I saw a certain fox.’

The suffix *-i* satisfies several functions, including individuation of objects (cf. Daniel 2009). We propose that *i*-marking conveys a selection of an individual out of a kind, sum, or a plurality of entities. This may be specified by a relative clause, but also by an antecedent. It should be modeled by choice functions, which map the set denoted by the nominal to an element of that set. There are dif-

ferent ways of spelling this out – choice functions may be bound by general existential closure, and hence be higher-order DRs (Reinhart 1997), or they may be contextual variables (Kratzer 1998). In (29), f is a choice function, hence $f(\text{APPLE})$ is mapped to an apple; following Reinhart 1997, f would be introduced as a DR.

$$(29) \quad K_0 + [Leili_1 EC_2 [{}_{VP} t_1 sib-i kharid]] \\ = [x_1 (f) \mid x_1 = \text{LEILI}, \exists [e_2 x_3 \mid x_3 = f(\text{APPLE}), \text{EAT}(x_1, x_3, e_2)]]$$

Being identified by a choice function, the reference of *sib-i* is not dependent on the event variable of the vP existential closure. Consequently, (29) does not result in a number-neutral interpretation, and anaphoric uptake will not exhibit the maximality effect. Anaphoric uptake would still be mediated by abstraction and summation; this predicts that anaphoric uptake of a scrambled *i*-marked noun, as in *Leili sib-i-rā kharid*, would be easier than with non-scrambled nouns. It is also possible to combine *yek* with *-i*, as in *Leili yek sib-i kharid*, which also appears to improve the possibility of anaphoric uptake.

8 Plurals and collectives

As we have seen, PINs have a number-neutral interpretation, in spite of being semantically singular, due to the existential closure over the event argument of the verbal predicate. This explains why plural marking of PINs is avoided: it would be semantically superfluous. Existential closure $\exists [e \ x \mid \dots]$ has to be read as: ‘there is at least one e, x such that...’, which naturally allows for the existence of more than one event e and object x .

However, plural marking of PINs does occur, but then does not just indicate a multitude of objects, but a multitude of events. Modarresi (2014) analyzes nominal plural marking as an exponent of plural marking over the vP; she analyzes *ketāb-ha khand-ad* ‘book-PL read-3SG’ as PL(book read). Alternatively, we can assume that under nominal plural marking an interpretation is enforced that involves a sum of spatiotemporally distinct events that stands in relation to a sum of entities. This is illustrated in (30), where E_2 is a sum of reading events and X_3 are the books of these reading events, where the plural marker on the nominal indicates that different books were read at different parts of the reading event.

$$(30) \quad Maryam_1 EC [{}_{VP} t_1 ketāb-ha khand-ad]. \\ \text{Maryam} \quad \text{book-PL} \quad \text{read-3SG} \\ \text{‘Maryam has read (many) books at different occasions.’} \\ [x_1 \mid \text{MARYAM} = x_1, \exists [E_2 X_3 \mid \text{BOOKS-OF}(E_2) = X_3, \text{READ}(x_1, X_3, E_2)]]$$

We now turn to collective predicates. Dayal (2011, 2015) discusses such cases in Hindi and finds that bare nominals are fine with predicates meaning ‘collect’,

which supports a number-neutral interpretation of the PIN, as number-neutrality allows for plural reference. This is similar to Persian:

- (31) *Ali tambr jam-mi-konad*,
 Ali stamp collect-DURATIVE-do.3SG
 ‘Ali collects stamps’

This appears to be a problem for the current analysis: The representation would include a part $\exists [x_3 e_3 \mid \text{STAMP-OF}(e_2) = x_3, \text{COLLECT}(x, x_3, e_2)]$, where x_3 refers to the unique stamp of a collecting event e_2 , which appears to be at odds with the collective meaning of ‘collect’. However, *collect* can be interpreted as ‘adding an entity to a collection’, and then this interpretation is plausible.

If we consider collective verbs meaning ‘compare’ or ‘connect’, an interpretation along the lines of ‘adding an entity to a comparison’ seems implausible, as any case of comparison involves at least two objects. Interestingly, these predicates do not allow for singular PINs in Persian, as predicted. However, the current theory also predicts that plural PINs should be possible. Typically, these plural PINs also are *rā*-marked, as in (32), but there are also cases without *rā*, as in (33). Importantly, singular objects are not possible for these verbs, which is against an analysis of a basic number-neutral interpretation of PINs.

- (32) *Maryam gheimat-ha-rā moghayese mi-konad*
 Maryam price-PL-OM compare DURATIVE-do.3SG
 ‘Maryam compares prices’

- (33) *che doost-ha be ham moarrefi-na-kardam*
 what friend-ha to each.other introduce-not-did.3SG
 ‘which friends did I not introduce to each other?’

9 Weak definites in English

We have analyzed bare nominals in Persian as dependent definites that receive an apparently indefinite interpretation when occurring within the vP. We would like to propose that in English, we have a similar range of interpretations of definite nominals, where the definite interpretation is visible by the definite article. The use of definite DPs to refer to entities given by the background knowledge of speaker and hearer, in the situation in which the conversation takes place, or in the preceding text is well known. But we find an apparently indefinite use with so-called weak definites (Poesio 1994, Carlson 2006, Schwarz 2014), as in (31a,b). In this interpretation, the accident victims might have been taken to different hospitals, and John and Mary might have read different newspapers.

- (34) a. *Every accident victim was taken to the hospital.*
 b. *John and Mary read the newspaper.*

Weak definites have been analyzed as kind-referring expressions, e.g. in Carlson & Sussman 2005, Aguilar-Guevara & Zwarts 2010 and Schwarz 2014. We would rather suggest that they are interpreted as dependent definites, like PINs in Persian. That is, we propose the following analysis:

- (35) *Mary took John to the hospital.*
 $[x_1 x_2 \mid x_1 = \text{MARY}, x_2 = \text{JOHN},$
 $\exists[e_3 x_4 \mid x_4 = \text{HOSPITAL-OF}(e_3), \text{TAKE-TO}(x_1, x_2, x_4, e_3)]]]$

We predict anaphoric uptake of weak definites to be possible, but less straightforward than uptake of wide-scope definites, as this requires abstraction and summation. This is corroborated by Aguilar-Guevara & Zwarts (2010), who find that while weak definites can be antecedents to anaphora, they are rather picked up by full definite noun phrases than by pronominals. Also, we predict a maximality effect of anaphoric uptake of weak definites, which can be observed in (36):

- (36) *Every victim was taken to the hospital. They declared a state of emergency.*
 (*they*: the hospitals were victims were taken too)
 $[\quad \mid [x_1 \mid \text{VICTIM}(x_1)] \Rightarrow \exists[e_2 x_3 \mid x_3 = \text{HOSPITAL-OF}(e_2), \text{TAKEN-TO}(x_1, x_3, e_4)]$
 $X_4 \mid X_4 = \Sigma_{x_3}[x_1 e_2 x_3 \mid \text{VICTIM}(x_1), \text{HOSPITAL-OF}(e_2), \text{TAKEN-TO}(x_1, x_3, e_4)],$
 $\exists[e_5 \mid \text{DECLARE-EMERGENCY}(X_4, e_5)]]]$

An important difference between PINs in Persian and weak definites in English is that the latter are restricted to cases referring to an “institutionalized” activity (Klein e.a. 2013). For example, while the object in (31a) has a weak definite reading, as to take someone to a hospital for treatment is an institutionalized activity, (34a) and (b) do not have this interpretation, as taking someone to an arena is not an institutionalized activity, and foreign dignitaries on a state visit most likely want to inspect a hospital, and not get treated there.

- (37) a. *The victims were taken to the arena.*
 b. *The foreign dignitaries on a state visit were taken to the hospital.*
 (38) a. *John and Mary burned the newspaper.*
 b. *John and Mary read the book.*

Similarly, the object in (31b) has a weak definite reading, as getting informed about recent events by reading a newspaper is an institutionalized activity. (35a) does not have an institutionalized interpretation, and neither does (35b), as ‘acquire information by reading a book’, presumably because books lack the ephemeral qualities of newspapers as sources of information.

The restriction to institutionalized readings for English weak definites can be motivated as follows. Consider the meaning of *take to the hospital*, rendered as a DRS with a weak definite, and subject and object still unspecified, as in (39).

- (39) *take to the hospital*
 $[e_1 x_2 \mid x_2 = \text{HOSPITAL-OF}(e_1), \text{TAKE}(x, y, x_2, e_2)]$

This applies to events e_1 in which x takes y to x_2 , where e_1 is restricted to events for which a hospital is defined, by $x_2 = \text{HOSPITAL-OF}(e_1)$. The natural class of events for which a hospital is defined, and the agent of the event takes the theme to that hospital, are those for which the theme is taken to the hospital for treatment. After all, to apply medical treatment to persons is part of the qualia structure of the noun *hospital* (cf. Pustejovsky 1995 for the notion of qualia). In contrast, in the case of *take to the arena*, such a natural class is not defined.

It is certainly plausible that when a combination of a verb and a noun refers to an institutionalized interpretation, then the verb applies to events e for which there is an entity x intrinsically related to e . Hence we should expect to find weak definites, and PINs in general, in expressions that refer to institutionalized readings. However, this does not explain why, in English, weak definites appear to be restricted to such institutionalized readings, in contrast to bare nouns in Persian (which, as we argued, are also dependent definites).

One important feature that distinguishes the two languages is that Persian explicitly marks whether an object is interpreted inside or outside existential closure over the vP, by the object marker *rā*. English lacks such marking. If English definites could be interpreted as freely as Persian bare nominals as vP-internal or vP-external, this would result in a high amount of structural ambiguities. Hence the vP-internal interpretation of definites is generally disfavored except in cases of institutionalized readings, where they are specifically licensed by conventionalized interpretations. This does not say that the internal reading of *read the newspaper* cannot be derived compositionally, only that this interpretation is structurally disfavored and surfaces only if the corresponding meaning is established in the lexicon, here as ‘gathering information about recent events by reading a newspaper’.

This predicts that languages that have a more perspicuous marking for the vP-internal vs. external interpretation of dependent definites make use of this distinction to a larger degree. An interesting case is German, which allows for scrambling but does not mark object scrambling as clearly with morphosyntactic means as Persian with *rā* marking. But German has a subclass of non-anaphoric definites, used mainly for entities given in the background or associative anaphora, including dependent definites (cf. Schwarz 2013, Frey 2015); they differentiate between the strong definites as *in das Hospital bringen* and weak definites as *in-s Hospital bringen* ‘take to the hospital’. This clearer distinction might allow for a higher occurrence of vP-internal definites in German. On the other hand, English

allows for bare nominals as in *go to school* (cf. Stvan 2009) which also may be interpreted as dependent definites without the use of explicit definite articles. Obviously, the use that languages make of the option of vP-internal, dependent definites, and how this interacts with other available options of (in)definiteness and scope marking, has to be explored more systematically.

References

- Aguilar-Guevara, Ana & Joost Zwarts. 2010. Weak definites and reference to kinds. In Nan Li & David Lutz (eds.), *Semantics and Linguistic Theory (SALT) 20*, 1-15. Ithaca, NY: CLC.
- Asher, Nicholas & Alex Lascarides. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.
- Asudeh, Ash & Line Mikkelsen. 2000. Incorporation in Danish: Implications for interfaces. In Ronnie Cann e.a. (eds.), *Grammatical Interfaces in HPSG*, 1-15. Stanford: CSLI Publications.
- Baker, Mark. 2014. Pseudo noun incorporation as covert noun incorporation: Linearization and crosslinguistic variation. *Language and Linguistics* 15: 5–46.
- Borik, Olga & Berit Gehrke, (eds.), *The Syntax and Semantics of Pseudo-Incorporation*. Leiden: Brill.
- Borik, Olga & Berit Gehrke. 2015. An introduction of the syntax and semantics of pseudo-incorporation. In Olga Borik & Berit Gehrke (eds.), 1-46.
- Browning, Maggie A. & Ezat Karimi. 1994. Scrambling to object positions in Persian. In Norbert Corver & Henk van Riemsdijk (eds.), *Studies on Scrambling*, 61-100. Berlin: Walter de Gruyter.
- Carlson, Greg N. 2006. The meaningful bounds of incorporation. In Svetlana Vegeler & Liliane Tasmowski, (eds.), *Non-Definiteness and pPurity*, 35-60. Amsterdam: John Benjamins.
- Carlson, Gregory & Rachel Sussman. 2005. Seemingly indefinite definites. In Stephan Kepser & Marga Reis, (eds.), *Linguistic Evidence*, 26-30. Berlin: Mouton de Gruyter.
- Dabir-Moghaddam, Mohammad. 1992. On the (in)dependence of syntax and pragmatics: Evidence from the postposition *-ra* in Persian. In Dieter Stein (ed.), *Cooperating with Written Texts*, 546-573. Berlin: Mouton de Gruyter.
- Dayal, Veneeta. 2011. Hindi pseudo-incorporation. *Natural Language and Linguistic Theory* 29, 123-167.
- Dayal, Veneeta. 2015. Incorporation: Morphosyntactic vs. semantic considerations. In Olga Borik & Berit Gehrke, (eds.), 189-221.
- Diesing, Molly. 1992. Bare plural subjects and the derivation of logical representations. *Linguistic Inquiry* 23: 353-380.

- Elbourne, Paul. 2005. *Situations and Individuals*. MIT Press.
- Evans, Gareth. 1980. Pronouns. *Linguistic Inquiry* 11, 337-362.
- Farkas, Donka F. & Henriette de Swart. 2003. *The Semantics of Incorporation*. Stanford: CSLI Publications.
- Frey, Werner. 2015. NP-Incorporation in German. In Borik, Olga & Berit Gehrke, (eds.), 225-261.
- van Geenhoven, Veerle. 1998. *Semantic Incorporation and Indefinite Descriptions: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. Stanford: CSLI Press.
- Geurts, Bart. 2000. Buoyancy and strength. *Journal of Semantics* 17, 315-333.
- Ghomeshi, Jila. 2003. Plural marking, indefiniteness and the noun phrase. *Studia Linguistica* 57, 47-74.
- Hardt, Daniel. 2003. Plural pronouns with quantified antecedents. In Matthias Weisgerber (ed.), *Proceeding of Sinn und Bedeutung*, 149-158. Arbeitspapiere des SFB 114, Universität Konstanz.
- Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13, 137-177.
- Kamp, Hans & Uwe Reyle. 1993. *From Discourse to Logic*. Dordrecht: Kluwer.
- Kamp, Hans, Uwe Reyle & Josef Van Genabith. 2011. Discourse Representation Theory. In Guenther, Franz & Dov M. Gabbay (eds.), *Handbook of Philosophical Logic*, 125-394. Berlin: Springer.
- Karimi, Simin (2003). On object positions, specificity and scrambling in Persian. In Simin Karimi (ed.), *Word Order and Scrambling*, 91-124. Oxford: Blackwell.
- Klein, Natalie, et al. 2013. Experimental investigations of weak definites and weak indefinite noun phrases. *Cognition* 128: 187-213.
- Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide-scope indefinites? In: Susan Rothstein (ed.), *Events and Grammar*, 163-196. Dordrecht, Boston, London: Kluwer Academic Publishers.
- Krifka, Manfred. 1995. Focus and the interpretation of generic sentences. In Greg N. Carlson & Francis J. Pelletier (eds.), *The Generic Book*, 238-264. Chicago: University of Chicago Press.
- Krifka, Manfred. 2001. Non-novel indefinites in adverbial quantification. In Cleo Condoravdi & Gerard Renandel de Lavalette (eds.), *Logical Perspectives on Language and Information*, 1-40. Stanford: CSLI Press, 1-40.
- Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19: 153-197.
- Massam, Diane. 2009. Noun incorporation: Essentials and extensions. *Language and Linguistics Compass* 3: 1076-1096.

- Mithun, Marianne. 2010. Constraints on compounding and incorporation. In: Vogel, Irene & Sergio Scalise (eds.), *Compounding*. Amsterdam: John Benjamins, 37-56.
- Modarresi, Fereshteh. 2014. *Bare nouns in Persian: Interpretation, Grammar, and Prosody*. Doctoral dissertation. Humboldt Universität zu Berlin.
- Modarresi, Fereshteh. 2015. Discourse properties of bare noun objects. In Olga Borik & Berit Gehrke (eds.), 189-221.
- Neale, S. 1990. *Descriptions*. Cambridge, MA: MIT Press.
- Nouwen, Rick, (subm.). E-type pronouns: congressmen, sheep and paychecks. *The Semantics Companion*, Wiley.
- Öztürk, Balkız. 2005. *Case, Referentiality and Phrase Structure*. Amsterdam/Philadelphia: John Benjamins.
- Paul, Daniel. 2009. The individuating function of the Persian “indefinite suffix”. In Simin Karimi e.a. (eds.), *Aspects of Iranian Linguistics*, 309-328. Cambridge Scholars Publishing.
- Poesio, Massimo. 1994. Weak definites. In Mandy Harvey & Lynn Santelmann (eds.), *Semantics and Linguistic Theory (SALT) 4*, 282-299. Ithaca, N.Y.: CLC Publications.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA.: MIT Press.
- Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20, 335–397.
- Schwarz, Florian. 2013. Two types of definites cross-linguistically. *Language and Linguistic Compass* 7: 534-559.
- Schwarz, Florian. 2014. How weak and how definite are weak indefinites? In Ana Aguilar-Guevara e.a. (eds.), *Weak Referentiality*. John Benjamins.
- Stvan, Laurel Smith. 2009. Semantic incorporation as an account for some bare singular count noun uses in English. *Lingua* 119: 314-333.
- Yanovich, Igor. 2008. Incorporated nominals as antecedents of anaphora, or How to save the thematic arguments theory. *University of Pennsylvania Working Papers in Linguistics* 14: 367-379.

Manfred Krifka
 ZAS &
 Humboldt-Universität zu Berlin
 Unter den Linden 9
 D-10099 Berlin, Germany
 krifka@zas.gwz-berlin.de

Fereshteh Modarresi
 Zentrum für Allgemeine Sprachwissenschaft
 (ZAS)
 Schützenstraße 18
 D-10117 Berlin, Germany
 fereshteh.modaresi@gmail.com