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, we argue for an existential closure operator over the vP that binds the event variable of the verbal predicate, and that nominals within the vP are interpreted as dependent definites with respect to the event. We show that this results in an apparent number neutrality of such nominals, and a maximality interpretation of anaphoricuptakes typical of E-type pronouns. The semantic contribution of nominals within and outside of the vP is modeled in DRT. We also argue that weak definites in English make use of the same strategy as 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 less prominent 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 be moved to focus positions. As for their interpretation, PINs are generally number-neutral and non-specific.

∗ We would like to thank the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 632, and the Bundesministerium für Bildung und Forschung for financial support (F. Modarresi). We would also like to 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.

©2016 Krifka & Modarresi
In Persian, an object can be realized with or without an object marker rā, a phenomenon known as differential object marking. Without rā, a singular bare object achieves an apparently indefinite, number-neutral interpretation; with rā, it is interpreted 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 khārid.  
    Leili apple-OM bought.3SG  
    ‘Leili bought the apple.’  

(2) János, beteget, vizsgált a rendelőben.
    János, patient,ACC examine,PAST the office.IN
    ‘János, examined in the office.’

    Øi túl sulyosnak találta Øj és beutaltatta Øj a korházba.
    pro, too severe,DAT find he,ACC pro and intern,CAUSE,PAST pro the hospital.IN
    ‘He, found him, too sick and sent him, to hospital.’

However, Yanovich (2008) reports that anaphoric reference with overt pronouns is possible, as in (3):

(3) A bátyám házat, vett a múlt héten. Egész vagyont adott érte.
    ‘The brother house, bought last week. He spent a fortune for it.’

In this paper 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 the conditions under which anaphoric uptake is possible. 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 struc-
tures (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 \(<A, F>\), 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 \(<A, F>\) iff there is a mapping \(g\) from the DRSs of \(K\) into \(A\) such that each condition of \(K\) is verified by \(g\) wrt \(<A, F>\). For example, \([x_1 \ x_2 \mid \Phi(x_1), \Psi(x_1, x_2)]\) is true wrt \(<A, F>\) iff there is a \(g\) such that \(g(x_1) \in A, g(x_2) \in A, g(x_1) \in F(\Phi), \text{ and } g(x_1, g(x_2)) \in F(\Psi)\). 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.

\[(4) \quad K_0 + [\text{Leili} [\text{[yek sib] kh aired]]] = [x_1 \ x_2 \mid x_1 = \text{LEILI, APPLE}(x_2), \text{BUY}(x_1, x_2)], \quad \text{two DRs introduced: } x_1, x_2\]

\[(5) \quad K_0 + [\text{Leili} [\text{[sib kh aired]]}] = [x_1 \mid x_1 = \text{LEILI, APPLE}(y), \text{BUY}(x_1, y)] \quad \text{just one DR introduced: } x_1\]

\[= K_1 \quad y: \text{thematic argument}\]

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 a number, 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 \(\Phi(a_1, \ldots, a_n)\) wrt a model \(<A, F>\) iff there is some sequence \(<a_1, \ldots, a_n>\) with \(a_1, \ldots, a_n \in A\), such that \(<a_1, \ldots, a_n> \in F(\Phi)\), and if \(a_i\) is a DR, \(a_i = g(a_i)\). If \(a_i\) 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 \(<A, F>\) 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 \(<g(x_1), a> \in F(\text{BUY})\).

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. 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 $x_i$ is introduced with a condition of the form $x_i = y_i$, where $y_j$ is a thematic argument that is part of a preceding condition $\Phi(\ldots, y_j, \ldots)$ in K or in a DRS that is subordinate to K. An assignment $g$ verifies the condition $x_i = y_j$ if $g$ maps $y_j$ onto an individual $a_i$ that is the i-th element of an n-tuple $\langle \ldots, a_i, \ldots \rangle$ that verifies the condition $\Phi(\ldots, y_j, \ldots)$. This is illustrated in (6), a continuation of (5):

(6)  
$$K_1 + [\text{Majnoon khord}=\emptyset]$$
$$= [x_1 \mid x_1 = \text{Leili}, \text{Apple}(x_2), \text{Buy}(x_1, y_2)$$
$$x_3 x_4 \mid x_3 = \text{Majnoon}, x_4 = y_2, \text{Eat}(x_3, x_4)]$$

true wrt an assignment $g$ and a model $\langle A, F \rangle$

iff a. $g(x_1) = F(\text{Leili})$
    b. there is an $a_2$ such that $a_2 \in A$ with $a_2 \notin F(\text{Apple})$
    c. there is a sequence $\langle a_1, a_2 \rangle \in AxA$ with $g(x_1) = a_1$ and $\langle a_1, a_2 \rangle \in F(\text{Buy})$
    d. $g(x_3) = F(\text{Majnoon})$
    e. $g(x_4) = a_2$
    f. $\langle g(x_3), g(x_4) \rangle \in F(\text{Eat})$

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. Another problem is the explanation why PINs are number neutral in the first place: They may lack a determiner, but they 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 features like singular and plural number. Hence they are ideally suited to pick up number-neutral PINs as antecedents. Modarresi makes use of a feature in the 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 (here rendered by $\xi$).

(7)  
Leili por te ghal khÆrid. Majnoon khord-Ø /\text{z\textsuperscript{27}eshi} \text{z\textsuperscript{27}eshoon}.

Leili orange bought.3sg Majnoon ate-pros/-it/-them

‘Leili bought orange(s). Majnoon ate it / them.’
Pseudo-Incorporated Nominals

\[ x_1 \xi_2 \ | \ x_1 = \text{Leili, Orange}/s(\xi_2), \text{buy}(x_1, \xi_2) \]
\[ x_3 \ | \ x_3 = \text{Majnoon, ate}(x_3, \xi_2) \] \[ \xi_2: \text{number-neutral DR} \]

But if world knowledge suggests an atomic / non-atomic interpretation of number-neutral DR, anaphoric reference with singular / plural overt pronouns is possible. For example, (8)(a) 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)  
       Leili apartment bought.3SG. price-its high was.3SG  
       ‘Leili bought apartment(s). Its price was high.’
   
   b. Leili hævij, khærid. Majnoon khord-eshoon.  
       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 – this is just stipulated. 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. However, 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 both their number neutrality and their restricted ability to serve as antecedent for pronouns. We will analyze cases of anaphoric uptake as cases of E-type pronouns (Evans 1980, Nouwen subm.), that is, as pronouns with quantifier antecedents where the quantifier does not c-command the pronoun. A standard example is (9):

(9)  Few congressmen, admire Kennedy, and they, 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 in a maximal way. Heim (1990) discusses contrasts like the following:
(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 related to 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 sup-
posed to introduce DRSs with an unlimited life span. Yet anaphoric reference is
possible, as (9) and (10)(b) 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 assign-
ment 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 “synthe-
size” a discourse referent “out of what the DRS of the first sentence provides”
(Kamp, Reyle & van Genabith 2011). In (11) a number-neutral DR \( \xi_3 \) is con-
structed 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 \( K' \text{ Q K}'' \), the ab-
straction of a DR of this condition, the summation over the values of this abstrac-
tion, and the identification with a new number-neutral DR, as in (12):

(12) If a DRS \( K \) contains a duplex condition \( K' \text{ Q K}'' \),
    form the union \( K''' = K' \cup K'' \),
    choose a DR \( x \) from \( K''' \), add a new DR \( \xi \) 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 (10)(a) vs. (b):

---

1 An better minimal pair would be *Three wine glasses broke last night. They had been expensive* and (10)(b). However, numerals trigger a scalar implicature, here ‘exactly three wine glasses’, and then the pronoun *they* would also refer to the sum of all wine glasses that broke last night. Another candidate would be (10)(a) and *At least one wine glass broke last night...* In this case, however, the continuation with the plural pronoun *They...* appears less felicitous, due to a syn-
tactic clash with the singular of *at least one wine glass.*
Pseudo-Incorporated Nominals

(13) a. \([x_1 \mid \text{WINEGLASS}(x_1), \text{BROKE}(x_1) \mid \text{EXPENSIVE}(x_1)]\)

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

An assignment \(g\) verifies the condition \(K' \langle \geq n \rangle K''\) wrt a model \(\langle A, F \rangle\) iff \(g\) can be extended to \(g'\) in \(n\) distinct ways such that \(g'(x) \in A\), and \(g\) verifies \(K'\) and \(K''\) wrt \(\langle A, F \rangle\). While (13)(a) could be true in case at least three wine glasses broke and just one of them was expensive, (b) would be false.²

5 PINs as Nominals under Existential Closure

In this and the next section we will develop the essence of our proposal. We have seen that PINs cannot be analyzed as regular indefinites; this wrongly predicts that they can be taken up as easily as regular indefinites by anaphoric expressions. For this reason, we analyze structures with PINs as cases that are similar to duplex conditions; as a consequence, the anaphoric uptake of PINs requires a complex abstraction and summation operation (cf. Yanovich 2008 for this suggestion). But with PINs there is no overt quantifier, different from cases considered so far, (9) and (10)(b). We assume that there actually is a quantifier, introduced by the syntactic structure in which the PIN occurs – the existential closure operation associated with the level of the vP, as proposed by Diesing 1992.

Within DRT, one has to distinguish between two types of existential closure: General existential closure; we say that \(K\) is true wrt \(\langle A, F \rangle\) iff there is an assignment \(g\) from the DRs of \(K\) into \(A\) that verifies the conditions of \(K\) wrt \(\langle A, F \rangle\). And vP existential closure, expressed by its own existential quantifier, cf. (14).

(14) If \(K'\) is a DRS, then \(\exists K'\) is a DRT-condition.

An assignment \(g\) verifies \(\exists K'\) wrt a model \(\langle A, F \rangle\) iff

there is an extension \(g'\) of \(g\) such that \(g'\) maps the DRs of \(K'\) to \(A\),

and \(g'\) verifies the conditions of \(K'\) in \(\langle A, F \rangle\).

² We would like to remark here that the standard formulation of abstraction and summation as in (12) is preliminary as it stands. It assumes a kind of anaphoric uptake – reference to a condition \(K' Q K''\) – without actually modeling it. A better representation might be to assume that such conditions introduce a DR for the condition \(K'' = K' \cup K''\) that then can be picked up by the abstraction and summation rule. Such DRs for DRSs have been proposed in SDRT (Asher & Lascarides 2003).
We assume that $\exists$ scopes over vPs, leading to the following interpretation:

\[(15)\] \quad K_0 + [Leili, EC \ [x_1 \ sib\ khærild]] \quad \text{EC: existential closure}

\[= [x_1 \ | x_1 = \text{Leili}, \exists [x_2 \ |\ \text{APPLE}(x_2), \text{BUY}(x_1, x_2)]]\]

\[= K_1\]

The truth conditions of this DRS are the same as in (4), but the DR $x_2$ cannot be picked up because it occurs in a subordinated DRS, $\exists [x_2 \ | \ \ldots]$. It can be accessed only by the abstraction and summation rule, as in (16):

\[(16)\] \quad K_1 + [\text{Majnoon khorzd-}$\emptyset$]

\[= [x_1 \ | x_1 = \text{Leili}, \exists [x_2 \ |\ \text{APPLE}(x_2), \text{BUY}(x_1, x_2)]]

\[\xi_2 x_3 \ | x_3 = \text{Majnoon}, \xi_2 = \Sigma x_2 [x_2 \ |\ \text{APPLE}(x_2), \text{BUY}(x_1, x_2)], \text{EAT}(x_3, \xi_2)\]

We assume that existential condition $\exists K'$ makes $K'$ accessible for the abstraction and summation rule. The resulting interpretation is similar to Farkas & de Swart 2003, but does not suffer from its deficiencies. Furthermore, we do not need a new mechanism to explain anaphoric reference to PINs; it is a case of E-type anaphora to a quantified antecedent.

The analysis proposed here naturally explains number neutrality of PINs. The existential condition $\exists [x_2 \ | \ \ldots]$ requires that the assignment $g$ can be extended so 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 scalar alternatives like numerals, there is also no scalar implicature that there should be just one such extension.

Anaphoric reference is achieved via the abstraction and summation rule. Following Modarresi (2015), this explains why number-neutral covert pronouns 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 (8)(a), or (8)(b).

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).

\[(17)\] \quad \text{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.’

\[(18)\] \quad \text{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.’
Pseudo-Incorporated Nominals

In conclusion, the proposed analysis of PINs as involving a general existential closure over the vP 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. (1)(a) suggests that rā is a definiteness marker, which has often been assumed (e.g. Ghomešī 2003). However, the object marker also occurs with the indefinite markers yek and -i, leading to a specific interpretation, as in (19)(a,b) (Karimi 2003), and with indefinites in the restrictor of generic quantifiers, as in (19)(c,d) (cf. Dabir-Moghaddam 1992, Kripka 2001).

\[
(19) \quad \begin{align*}
\text{a. yek film-rā didām.} & \quad \text{b. film-i-rā didām.} \\
& \quad \text{movie-om see.1sg} \quad \text{movie-idef-om see.1sg} \\
& \quad \text{‘I saw one of the movies’} \quad \text{‘I saw a certain movie.’} \\
\text{c. serke shir-rā mi-borāed.} & \quad \text{d. kowboy-e-ha tanbako-rā mi-javand.} \\
& \quad \text{vinegar milk-om dur-curdlе.3sg} \quad \text{cowboy-pl tobacco-om dur-chew.3pl} \\
& \quad \text{‘Vinegar curdles milk.’} \quad \text{‘Cowboys CHEW tobacco.’}
\end{align*}
\]

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 (19)(c,d).

The definite interpretation of bare nominals as in (1)(a) 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 is bound by existential closure, which indirectly binds any DR dependent on it. That is, we revise our analysis of (15) as follows:
(20) $K_0 + [Leili, EC [\_p \text{ t}_1, \text{ sib k\text{"a}rid}]]$
\[= [x_1 \mid x_1 = \text{Leili}, \exists [e_3 \ x_2 \mid x_2 = \text{APPLE-OF}(e_3), \text{BUY}(x_1, x_2, e_3)]]\]
\[= K_1\]

We assume an event DR $e_3$ introduced by the event argument of the verb. The bare noun $\text{sib}$ 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 – hence the apparent indeterminateness. The resulting interpretation allows for more than one apple being bought by Leili, as there could be multiple buying events, each of which being related to a different unique apple. This explains the apparent number-neutrality. Anaphoric uptake is achieved by abstraction and summation, cf. (21). Given an assignment $g$ and a model $\langle A, F \rangle$, the summation $\Sigma x_2[e_3 | \ldots] \text{returns the sum of all elements } a \in A \text{ such that there is an extension of } g \text{ to } g' \text{ with } g'(x_2) = a$, where $g'$ verifies $[x_2 e_3 | \ldots]$ in $\langle A, F \rangle$.

(21) $K_1 + [\text{Majnoon khord-}O]$
\[= [x_1 \mid x_1 = \text{Leili}, \exists [x_2 e_3 \mid x_2 = \text{APPLE-OF}(e_3), \text{BUY}(x_1, x_2, e_3)]\]
\[\xi_5 x_4 \mid x_4 = \text{Majnoon}, \xi_5 = \Sigma x_2[x_2 e_3] x_2 = \text{APPLE-OF}(e_3), \text{BUY}(x_1, x_2)\]
\[\text{EAT}(x_4, \xi_5)\]

When bare nominals are $r\ddot{a}$-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 s\text{"a}b\text{"e}d ye\text{\text{"a}} sib v\text{"a} ye\text{\text{"a}} g\text{\text{"a}}l\text{\text{"a}} bi\text{\text{"a}} bood. Leili sib-r\ddot{a} b\text{\text{"a}}\text{\text{"a}}\text{\text{"a}}d\text{\text{"a}}\text{\text{"a}}s\text{\text{"a}}t.*
\begin{align*}
\text{in basket an apple and a pear was.3SG Leili apple-OM took.3SG} \\
\text{‘There was apple and a pear in the basket. Leili took the apple.’}
\end{align*}
\[\left[ x_1 \xi_2 x_3 x_4 x_5 ight] \mid \text{BASKET}(x_1), \text{APPLE}(x_2), \text{PEAR}(x_3), X_4 = x_2 \oplus x_3, \text{IN}(x_1 X_4), \]
\[x_6 x_7 \mid x_6 = \text{Leili}, x_7 = \text{APPLE-OF}(X_4), \exists [e_8 \mid \text{TAKE}(x_6 x_7, e_8)]\]

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 s\text{"a}b\text{"e}d bood. Leili sib-r\ddot{a} b\text{\text{"a}}\text{\text{"a}}\text{\text{"a}}d\text{\text{"a}}s\text{\text{"a}}t.*
\begin{align*}
\text{an apple in basket was.3SG Leili apple-OM took.3SG} \\
\text{‘There was an apple in the basket. Leili took the apple.’}
\end{align*}
\[\left[ x_1 x_2 \mid \text{BASKET}(x_1), \text{APPLE}(x_2), \text{IN}(x_1 X_2), \right.\]
\[x_3 x_4 \mid x_3 = \text{Leili}, x_4 = \text{APPLE-OF}(x_2), \exists [e_5 \mid \text{TAKE}(x_3 x_4, e_5)]\]

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. It is assumed that bare subjects in Persian are definite, but they also have an apparently indefinite interpretation (Modarresi 2014). Subjects form a phonological phrase with the verb when inside the vP, in which case they carry main accent. The apparently indefinite interpretation of the subject in (24)(b) arises if we assume a definite interpretation dependent on the event over the vP.

book fell.3SG book fell.3SG
‘The book fell.’ ‘Some book(s) fell.’

(19)(c,d) showed that in generic sentences, bare object nominals with rā-marking are interpreted in the restrictor of the generic operator. This is predicted under standard assumptions of the partition of semantic material under a generic operator (cf. Diesing 1992, Krifka e.a. 1995). For example, (19)(b) is interpreted as follows, where GEN introduces its own event variable e₃.

(25) GEN [serke₁, shir-rā₂ [v t₁ t₂ mi-borraed]]
    [ | [x₁ x₂ e₃ | x₁ = VINEGAR-OF(e₃), x₂ = MILK-OF(e₃), IN(x₁,x₂,e₃)]
     ⟨GEN e₃⟩ [e₄ | e₄ ⊆ e₃, CURDLE(x₁, x₂, e₄)]

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 an object nominal is always indicated by rā-marking, then a yek-marked nominal without rā as in (4) is interpreted within the vP. This would result in interpretation (26)(a).

(26) K₀ + [Leili₁ [v t₁ [yek sib] khaerid]]
a. [x₁ | x₁ = Leili, ∃[e₂ x₃ | APPLE(x₃), #(x₃)=1, BUY(x₁,x₃,e₂)]
b. [x₁ x₃ | x₁ = Leili, APPLE(x₃), #(x₃)=1, ∃[e₂ | BUY(x₁,x₂,e₃)]

While the DR x₃ is not dependent on the event, it is introduced under existential closure. Consequently, we predict number-neutrality and a maximality inter-
pretation of anaphoric uptake. But this is not what we find. This is because (26)(a)
is semantically indistinguishable from the PIN interpretation (20). The indefinite
marker ye:k contributes the number information #(x₃)=1, but has no effect on the
overall interpretation. Hence this interpretation is blocked by the PIN form (20).

But indefinites can have a wide-scope interpretation independent of their syn-
tactic position, as argued for in Kamp & Reyle (1993, 3.7.3). This would lead to
the interpretation (26)(b), without scrambling of the indefinite nominal outside of
the vP. One mechanism for this wide-scope interpretation is that the indefinite is
presuppositional (van Geenhoven 1998), or backgrounded (Geurts 2000). Now,
yek-marked indefinites can be rā-marked, which indicates scrambling outside of
the vP according to our analysis. But in cases like (26) this is disfavored, presum-
ably because the wide-scope interpretation can be achieved without scrambling,
and so scrambling is blocked. We find it only when the indefinite nominal should
take wide scope interpretation with respect to another operator, as in (27).

(27) ye:k kétar-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. [Māen₁ [v p₁ roobah-i didāem]]   b. [Māen₁ roobah-i-rā₂ [v p₁ t₂ didāem]]
    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 exis-
tential closure, and hence be higher-order DRs (Reinhart 1997), or they may be
contextual variables (Krätzer 1998). In (29), f is a choice function, hence f(APPLE)
is mapped to an apple; following Reinhart 1997, f would be introduced as a DR.

(29) K₀ + [Leili₁ EC₁ [v p₁ sib-i kharid]]
    = [ x₁ (f) | x₁=Leili₁, ∃[e₂ x₃ | x₃ = f(APPLE), EAT(x₁,x₃,e₂)]

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 re-
sult in a number-neutral interpretation, and anaphoric uptake will not show 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 that with non-scrambled nouns.
8 Plurals and Collectives

As we have seen, PINs have a number-neutral interpretation, in spite of being semantically singular. This explains why plural marking of PINs is avoided. However, it does occur, but then indicates a multitude of events. Modarresi (2014) has analyzed plural marking on the noun as a reflex of a verbal plural marking over the vP. Alternatively, we can assume that in this case, a reading is enforced which relates the plural object DR to a plural event DR, as in (30), where \( E_2 \) stands for a sum of spatiotemporally distinct events, and \( X_3 \) stands for a sum of books.

(30) \[ Maryam_1 \text{ EC} [vP t_1 ketāb-ha khand-ad]. \]
\[
\begin{align*}
\text{Maryam} & \quad \text{book-PL} \quad \text{read-3SG} \\
\text{‘Maryam has read (many) books at different occasions.’} \\
[x_1 | \exists [E_2 X_3 | \text{books}(X_3), \text{read}(x_1, X_3, E_2)]]
\end{align*}
\]

The current analysis predicts that singular PINs are avoided for collective predicates. Dayal (2011, 2015) discusses such cases. She finds that they are fine with predicates meaning ‘collect’, which is similar in Persian – cf. \( \text{Ali tambr jammi-konad} \), ‘Ali stamp collect-DURATIVE-do.3SG’, ‘Ali collects stamps’. Such predicates can be understood as referring to repeated events of adding a stamp to a collection, and hence are compatible with singular PINs. In contrast, verbs meaning ‘compare’ or ‘connect’, which require a sum-referencing object, do not allow for singular PINs. This is predicted by our analysis. However, it turns out that they also disfavor plural PINs but require \( rā \)-marked plurals referring to sum individuals that are given in the larger context, or in the restrictor of a generic quantifier, as in \( Maryam gheimat-ha-rā moghayese mi-konad ‘Maryam price-PL=OM compare DURATIVE-do.3SG’ \), ‘Maryam compares prices’. A possible reason is that plurals without \( rā \)-marking lead to a pluralization over events as in (30).

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 (31)(a,b). In this interpretation, the accident victims might have been taken to different hospitals, and John and Mary might have read different newspapers.

(31)  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:

(32)  *Mary took John to the hospital.*

\[ x_1 \ x_2 \ | \ x_1 = \text{MARY}, \ x_2 = \text{JOHN}, \]
\[ \exists [e_3 \ x_4 \ | \ x_4 = \text{HOSPITAL-OF}(e_3), \ \text{TAKE-TO}(x_1, x_2, x_4, e_3)] \]

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 pronouns. Also, we predict a maximality effect of anaphoric uptake of weak definites, which can be observed in (33):

(33)  *Every victim was taken to the hospital. They declared a state of emergency.*

\( \text{they: the hospitals were victims were taken too} \)

\[ [ x_1 \ | \ \text{VICTIM}(x_1) ] \Rightarrow \exists [e_2 \ x_3 \ | \ x_3 = \text{HOSPITAL-OF}(e_2), \ \text{TAKE-TO}(x_1, x_3, e_4)] \]
\[ X_4 = \sum x_3 \ [ x_1 \ e_2 \ x_3 \ | \ \text{VICTIM}(x_1), \ \text{HOSPITAL-OF}(e_2), \ \text{TAKE-TO}(x_1, x_3, e_4)], \]
\[ \exists [e_5 \ | \ \text{DECLARE-EMERGENCY}(X_4, e_5)] \]

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 (31)(a) has a weak definite reading, as to take someone to a hospital for treatment is an institutionalized activity, (34)(a) 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.

(34)  a. *The victims were taken to the arena.*
    b. *The foreign dignitaries on a state visit were taken to the hospital.*

(35)  a. *John and Mary burned the newspaper.*
    b. *John and Mary read the book.*

Similarly, the object in (31)(b) has a weak definite reading, as getting informed about recent events by reading a newspaper is an institutionalized activity. (35)(a) does not have an institutionalized interpretation, and neither does (35)(b),
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 (36).

\[(36) \text{take to the hospital} \]
\[ [e_1 x_2 \mid x_2 = \text{HOSPITAL-OF}(e_1), \, \text{TAKE}(x, y, x_2, e_2)] \]

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 licensed as being idioms. This does not entail that a the meaning verbal expression like *read the newspaper* cannot be derived compositionally, as *kick the bucket* – after all, it has a transparent derivation, $[e_1 x_2 \mid x_2 = \text{NEWSPAPER-OF}(e_1), \, \text{READ}(x, x_2, e_1)]$. But definites in English usually are interpreted outside the vP, hence the internal interpretation is an option only if such meanings are established in the lexicon, where they have an enriched meaning, here ‘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 distinc-
tion to a larger degree. An interesting case is German, which allows for scrambling but does not mark object scrambling as clearly as Persian. 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); they differentiate between the strong and the weak definite reading of in das Hospital bringen and 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 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

Pseudo-Incorporated Nominals


Schwarz, Florian. 2014. How weak and how definite are weak indefinites? In Ana Aguilar-Guevara e.a. (eds), *Weak referentiality.* John Benjamins


---

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