

# *University of Pennsylvania Working Papers in Linguistics*

---

*Volume 14, Issue 1*

2008

*Article 28*

PROCEEDINGS OF THE 31ST ANNUAL PENN LINGUISTICS  
COLLOQUIUM

---

## Incorporated nominals as antecedents for anaphora, or How to save the thematic arguments theory

Igor Yanovich\*

\*Moscow State University

Copyright ©2008 by the authors.  
<http://repository.upenn.edu/pwpl>

# Incorporated nominals as antecedents for anaphora, or How to save the thematic arguments theory

Igor Yanovich

## **Abstract**

The paper discusses the incorporation facts, mainly from Hungarian and Hindi, and the extended version of DRT developed by Farkas and de Swart (2003) to capture the properties of incorporated nominals across languages. It is shown that though the framework as defined by Farkas and de Swart does not derive all predictions it was aimed to derive by the authors, it can be easily extended into a version which does predict the relevant data. However, after examining more closely the data from Hungarian-type incorporation languages, I put forward a preliminary hypothesis alternative to Farkas and de Swart's proposal, namely, that the possibility of anaphora to singular incorporated nominals is not a parameter of a language, but arises in languages like Hungarian if and only if the context makes it clear that the maximal entity denoted by the incorporated nominal is at most atomic. It is left for future empirical research to find out whether this hypothesis can actually account for the whole range of relevant data.

# Incorporated Nominals as Antecedents for Anaphora, or How to Save the Thematic Arguments Theory

Igor Yanovich\*

## 1 Introduction

It is well known that Incorporated Nominals (INs) may differ across languages with respect to their discourse transparency, that is, their ability to be antecedents for pronouns in subsequent sentences. Moreover, there may be differences as to this ability within a single language with respect to INs of different morphological number, in languages where both morphologically singular and morphologically plural INs are allowed.

For example, West Greenlandic INs are discourse transparent, cf. (1), while in Hindi and Hungarian plural INs are discourse transparent and singular ones are not; see (2) and (3). Farkas and de Swart (2003) (F&dS, henceforth) also argue that Hungarian INs may not antecede overt pronouns, but may antecede covert ones, (4), and thus that there is one more dimension of possible differences.

(1) *West Greenlandic (from van Geenhoven 1998:187):*

Aani qimmi-qar-p-u-q.  
Aani.ABS dog-have-IND-[-tr]-3Sg.  
Miki-mik ati-qar-p-u-q.  
M.-inst name-have-IND-[-tr]-3Sg  
'Aani has a dog<sub>i</sub>. It<sub>i</sub> is called Miki.'

(2) *Hindi (from Dayal 1999):*

a. anu kitaab paRh-rahii-hai. \*vo bahut acchii hai  
Anu book read-PR-PROG It very good be-PR  
'Anu is reading a book<sub>i</sub>. It<sub>i</sub> is very good.'  
b. anu apne bete ke liye laRkiyaaN dekh rahii hai.  
OK vo unkaa swabhaav jaanna caahtii hai.

---

\*This paper has benefited from the discussion at the 31<sup>st</sup> PLC, as well as from the valuable comments of Chris Barker, Anna Szabolcsi, and two anonymous reviewers. Many thanks to Olga Lukhminskaya for her help with constructing Hungarian examples and to Anna Szabolcsi for her judgments and, again, for the help with constructing examples in Hungarian. Of course, all remaining mistakes are my own. I am grateful for the partial support of this work by the National Science Foundation under Grant No. BCS-0418311 to B.H. Partee and V. Borschev.

‘Anu is seeing girls<sub>1</sub> for her son.  
She wants to find out their<sub>1</sub> temperament.’

(3) *Hungarian (from Farkas and de Swart 2003:135):*

- a. János<sub>1</sub> beteget<sub>2</sub> vizsgált a rendelőben  
Janos patient.Acc examined the office.in  
?? Ø<sub>1</sub> Túl súlyosnak találta őt<sub>2</sub>  
he too severe.Dat found he.Acc  
és beutaltatta Ø<sub>2</sub> a kórházba.  
and intern.Cause.Past him the hospital.in  
‘Janos<sub>1</sub> patient<sub>2</sub>-examined in the office.  
He<sub>1</sub> found him<sub>2</sub> too sick and sent him<sub>2</sub> to hospital.’

- b. János<sub>1</sub> beteget<sub>2</sub> vizsgált a rendelőben  
Janos patient.Pl.Acc examined the office.in  
OK Ø<sub>1</sub> Túl súlyosnak találta őket<sub>2</sub>  
he too severe.Dat found they.Acc  
és beutaltatta Ø<sub>2</sub> a kórházba.  
and intern.Cause.Past them the hospital.in  
‘Janos<sub>1</sub> patients<sub>2</sub>-examined in the office.  
He<sub>1</sub> found them<sub>2</sub> too sick and sent them<sub>2</sub> to hospital.’

(4) *Hungarian (from Farkas and de Swart 2003:135-136):*

János<sub>1</sub> beteget<sub>2</sub> vizsgált a rendelőben  
Janos patient.Acc examined the office.in  
‘Janos<sub>1</sub> patient<sub>2</sub>-examined in the office.’

a. A singular IN binding an overt pronoun:

?? Ø<sub>1</sub> Túl súlyosnak találta őt<sub>2</sub>  
he too severe.Dat found he.Acc  
és beutaltatta Ø<sub>2</sub> a kórházba.  
and intern.Cause.Past him the hospital.in  
‘He<sub>1</sub> found him<sub>2</sub> too sick and sent him<sub>2</sub> to hospital.’

b. A singular IN binding a covert pronoun:

OK Ø<sub>1</sub> Túl súlyosnak találta Ø<sub>2</sub>  
he too severe.Dat found he.Acc  
és beutaltatta Ø<sub>2</sub> a kórházba.  
and intern.Cause.Past him the hospital.in  
‘He<sub>1</sub> found him<sub>2</sub> too sick and sent him<sub>2</sub> to hospital.’

Farkas and de Swart set up a Discourse Representation Theory-based (DRT) framework to capture these data. They introduce a new kind of variable—thematic arguments—whose properties should help capture the properties of INs. The thematic arguments extension to classical DRT is described in Section 2. However, as I show in Section 3, this extension in fact does not work as intended because of two technical problems. In the same section, I describe the improvements which allow the theory to actually work. In Section 4, I discuss F&dS’s analysis of the distinction between singular and plural INs with respect to anaphora, cite old and new data on the subject, and argue that F&dS’s analysis is based on an empirical generalization which makes too sharp a distinction between singular and plural INs. In Section 5, I propose a new analysis of the data from Section 4 which allows to explain the generalization that plural INs usually allow and singular INs usually do not allow anaphora, as well as the exceptions to this generalization. Whether this new analysis is in fact viable is subject to future empirical research, but at least it allows for a natural account of the known set of data and makes interesting predictions about incorporated mass nouns, which have not received much attention in previous studies of incorporation.

## 2 The Thematic Arguments Framework

The widely accepted informal view on nominal incorporation states that an IN does not provide a full-fledged verbal argument, but instead forms a complex predicate with the verb, like the English “berry” in the compound “berry-picking”. The nominal does not occupy a semantic argument slot, but rather adds a restriction on possible objects of picking (cf. Dayal 2003, Chung and Ladusaw 2003, a.o.). The incorporated nominal usually does not constitute a full-fledged DP, and its interpretational import is different from the import of such a DP. Rather than introducing a discourse referent (or something like that) and filling the argument slot, the IN serves as a semantic adjunct. There are many conceivable ways to capture this basic intuition, some of them are present in the literature. One such theory is that of Farkas and de Swart (2003), to be discussed below.

F&dS build their theory as an extension to DRT (Kamp and Reyle 1993, a.o.) First of all, F&dS introduce a new type of variable, in addition to the familiar discourse referents: thematic arguments. Their notational convention requires to use  $u$ ,  $v$ ,  $t$  for discourse referents and  $x$ ,  $y$ ,  $z$  for thematic arguments. Both discourse referents and thematic arguments are essentially logical variables ranging over individuals. The difference between the two classes of variables lies in DRS-construction (Discourse Representation Structure) and interpretation rules: the rules treat them differently.

First, there is a difference concerning how the two types of variables are introduced into DRSs. In short, “normal”, full-fledged argument DPs introduce “normal” discourse referents, while “unusual” incorporated nominals introduce “unusual” thematic arguments. The actual story is a bit more complex, however. F&dS argue that lexical items, when they come from the lexicon, carry with them DRS-conditions in which arguments are represented by thematic arguments. For example, the noun “student” carries a condition **student(x)** (remember that we use x-z for thematic arguments, not discourse referents), and the verb “leave” carries a condition **leave(x)**. When these lexical items are processed (or, more precisely, when the corresponding part of the syntactic structure is processed) during DRS-construction, the processing rules, in the simplest case, replace thematic arguments with discourse referents. For the thematic arguments of common nouns, this job is performed by the rules associated with determiners, and for the thematic arguments of verbs, by the rules for combining the verb with its syntactic arguments. What is important is that only discourse referents may be listed in the universe of a DRS. Thematic referents are “less independent”, and they do not appear in the universe.

I will illustrate how this system works with the example of processing the sentence “A student leaves” and refer the reader to Farkas and de Swart (2003) for more technical details:<sup>1</sup>

(5) a. Syntactic representation:

{ : [ [DP a [NP student] [VP leaves]] ] }

b. The common noun is processed:

{ : [ [DP a [NP **student(x)**] [VP leaves]] ] }

c. The article introduces a new discourse referent and “binds” the thematic argument of the common noun to it (Determiner Instantiation, in F&dS’s terms):

{ **u** : [ [DP **u** [NP **student(u)**] [VP leaves]] ] }

d. The verb is processed:

{ **u** : [ [DP **u** [NP **student(u)**] [VP **leave(x)**] ] ] }

e. The verb combines with its argument (Argument Instantiation, in F&dS’s terms):

{ **u** : [ [DP **u** [NP **student(u)**] [VP **leave(u)**] ] ] }<sup>2</sup>

<sup>1</sup>I use the following conventions for representing DRSs: {... : ...} corresponds to a box, the universe of the DRS being described to the left of the “:”, and the conditions—to the right of it. [] stands for familiar syntactic constituency. Words (parts of the syntactic structure) are in normal font, and meanings are in bold.

<sup>2</sup>Farkas and de Swart (2003) do not discuss explicitly the deletion of syntactic

So in the “normal” course of events, DRS conditions enter into the DRS along with thematic arguments, but these arguments are replaced with discourse referents, or instantiated, during the derivation.

For noun phrases, the instantiation of the thematic argument is performed by the determiner. However, since there is no D in incorporated nominals, they are left with thematic arguments in the conditions they introduced. By the time such a nominal must combine with the verb, it does not have a corresponding discourse referent, and thus a different mode of combination must be used. While full-fledged DPs combine with the verb via Argument Instantiation, a rule replacing the relevant thematic argument of the verb with the discourse referent introduced by the DP, incorporated nominals combine with the verb via Unification:

- (6) *Unification*. Replace the relevant thematic argument  $y$  of a verbal predicate with the thematic argument  $z$  contributed by a nominal argument of the verb. (Farkas and de Swart’s 2003 (83), p. 65)

Here is the result of Unification for a Hungarian incorporation example:

- (7) János betegeget vizsgált.  
 Janos patient- examined.

|  |
|--|
| u  |
| Janos(u)<br>patient(x)<br>examined(x, u) |

Of course, such a DRS may not be interpreted with the standard verifying rules, because the standard rules may not deal with the thematic arguments. F&S formulate the relevant verification rule as follows:

- (8) A function  $f$  verifies a condition of the form  $P(\mathbf{a}_1, \dots, \mathbf{a}_n)$  relative to a model  $M$  iff there is a sequence  $\langle e_1, \dots, e_n \rangle \in E^n$ , such that  $\langle e_1, \dots, e_n \rangle \in I(P)$ , and if  $\mathbf{a}_i$  is a discourse referent,  $e_i = f(\mathbf{a}_i)$ , and if  $\mathbf{a}_i$  is a thematic argument,  $e_i$  is some element in  $E$ .  
 (Farkas and de Swart’s 2003 (82), p. 63)

---

structure from the DRS, and we follow them. It would not be hard to formulate the construction rules more accurately, so that they would put the conditions directly in the DRS and delete the relevant parts of the structure after they are processed.

Thus, embedding functions do not store the values for thematic arguments as they do for discourse referents: it suffices to have an individual which satisfies a condition for that condition to be verified. After it has been verified, we do not keep track of this individual. Note that this is different from how discourse referents are treated in DRT: the verifying values for discourse referents are remembered by the embedding functions, and that ensures that all instances of the same discourse referent get the same value. The verification rule in (8) does not enforce the same for thematic arguments.

If so, how then do F&dS capture anaphoric references to incorporated nominals? They formulate a special connective for such anaphoric dependencies,  $\approx$ , which is different from the usual  $=$  used for anaphoric conditions in the standard variant of the theory. The first argument of the connective  $\approx$  must be a discourse referent, and the second a thematic argument (hence it is syntactically different from the “ $=$ ”, both arguments of which are discourse referents). The part of the pronoun introduction rule that adds a condition with  $\approx$  is formulated as follows:

- (9) ... If an accessible and suitable discourse referent  $\mathbf{u}$  cannot be found, add a condition of the form  $\mathbf{v} \approx \mathbf{x}_i$ , where  $\mathbf{x}_i$  is an accessible and suitable thematic argument that is part of a condition  $\mathbf{P}(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n)$  in  $\text{Con}_K$  or  $\text{Con}_{K'}$  of some  $K'$  that is superordinate to  $K$ .<sup>3</sup>

(Farkas and de Swart (2003), p. 144)

Farkas and de Swart formulate the following verification clause for conditions introduced by (9):

- (10) A function  $\mathbf{f}$  verifies a condition of the form  $\mathbf{v} \approx \mathbf{x}_i$ , where  $\mathbf{v}$  is a discourse referent and  $\mathbf{x}_i$  is an (uninstantiated) thematic argument that shows up in the  $i$ -th position of a predicative condition of the form  $\mathbf{P}(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n)$ , iff  $\mathbf{f}$  maps  $\mathbf{v}$  onto the individual  $e_i$  that is the  $i$ -th element of the  $n$ -tuple  $\langle e_1, \dots, e_n \rangle$  that verifies the condition  $\mathbf{P}(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n)$ .

(Farkas and de Swart (2003), p. 144)

---

<sup>3</sup>While F&dS do not define the notions of accessibility and suitability, the former does not seem to be needed, since they explicitly mention that the condition hosting a thematic referent to which the pronoun is bound may occur only in superordinate DRSs. As for the latter notion, it seems that F&dS presuppose that “suitable” means here either “satisfying sortal restrictions and the pronoun’s presuppositions”, or “such as intended by the speaker”, or both. Either way, it makes sense, so we will not discuss it further.

### 3 Tuning up the Thematic Arguments Framework

Although Farkas and de Swart (2003) presents intriguing data and intuitively clear informal analysis for it, unfortunately, their formal system fails to give the intended results. The problem is connected to thematic arguments—the new type of variable introduced by F&dS. The interpretation of thematic arguments is regulated by two verifying rules: (8) for regular conditions with thematic arguments and (10) for conditions of the form  $v \approx x_i$ . Both rules are insufficient for the purposes for which they are formulated: namely, (8) does not provide the intended interpretations for simple DRSs with incorporation, and (10) does not provide the intended interpretations for anaphora referring to thematic arguments.

- (7) János beteget vizsgált.  
 Janos patient- examined.

|  |
|--|
| u  |
| Janos(u)<br>patient(x)<br>examined(x, u) |

Consider the interpretation of (7), repeated here, under the rule in (8). (8) states that to verify a condition containing a thematic argument, we need only to find some individual satisfying the condition. Imagine a situation where there is a patient named John, and a nurse named Mary, and where János examined Mary, who is not a patient, and did not examine John. To satisfy the condition **patient(x)**, we need to find an individual who is a patient. There is such an individual, namely, John, so the condition is satisfied. To satisfy the condition **examined(x,u)**, we need to find some individual who was examined by **u** (János). Mary is such an individual, because she was examined by János, and thus this condition is too satisfied. So F&dS's rule predicts that (7) is true in this situation, while actually (7) is false.

The problem is clear: while it is intended that the two instances of **x** in (7) “refer” to the same individual, (8) does not require it. An anonymous reviewer has suggested that this problem may be dealt with by requiring that all conditions in a DRS be interpreted as conjoined; then the two instances of **x** in (7) would be in the same big formula, not in two different formulas, and both instances would have to refer to the same individual. This does not

seem like a good solution to me, though it certainly is a matter of personal taste rather than of empirical consequence. In standard DRT, DRS conditions are interpreted as though they were conjoined, but it is not the direct effect of conjoining them, but of using embedding functions to keep track of discourse referents. So why should we use literal conjunction after we introduce thematic arguments, if we have not done so when we had only discourse referents? Another problem with using literal conjunction is that it will not help once the problem with the anaphoric rule in (10) is considered.

An alternative solution is to allow embedding functions to keep track of thematic arguments too. In effect, thematic arguments will be listed in the universe part of a DRS just as discourse referents are. This change from the original F&dS proposal is not very serious: even under F&dS (2003), both are just different types of variables.

Another problem concerns the other rule for thematic argument interpretation—the one in (10). The rule states that the discourse referent introduced by the pronoun must be mapped to the individual which may be referred to by a thematic argument which serves as an antecedent. The nature of this problem is the same: since we do not keep track of what individuals thematic arguments refer to, we cannot ensure that the pronoun picks up the very same individual that was picked when we interpreted the condition with the thematic argument: the pronoun may refer to any individual satisfying the condition, not only to the exact individual we used to verify its truth.

While it was possible to solve the first problem by stipulating that all conditions in a DRS must be interpreted as conjoined, it will not help us to solve the second problem, because the anaphoric condition may be part of an embedded DRS, and the conditions from embedded DRSs may not be simply conjoined with conditions of a higher DRS: If they are just conjoined, then the discourse referents from the universes of the embedded DRSs will obtain text-level scope. However, if we allow embedding functions to “record” the values of thematic arguments, as suggested above, the problem disappears, and anaphoric reference to thematic arguments will be as simple as it is to discourse referents.

It is not that all the differences between the two types of variables disappear after such changes to the framework: there still remain several important differences concerning how thematic arguments and discourse referents are treated by the theory. It is important to note that the changes we have just introduced are of a technical nature: applied to F&dS’s framework, they allow it to actually derive the results the authors wanted to derive.

#### 4 Singular vs. Plural Incorporated Nominals

Farkas and de Swart argue that the differences between plural and singular INs in Hindi and Hungarian, shown in (2-3), should be explained as follows: both singular and plural INs introduce thematic arguments, and not discourse referents; however, the plural morphology carries a presupposition of existence of a discourse referent. Furthermore, they stipulate that presuppositions introduced by morphological markers must be accommodated as locally as possible. Therefore, plural INs have a corresponding discourse referent, and thus may be referred back to by a general anaphoric mechanism; singular INs, on the other hand, may be referred back too only with the help of the special mechanism associated with the  $\approx$  connective. Some languages lack this mechanism altogether, as Hindi, while such languages as Hungarian use it, but more restrictively than the usual anaphoric mechanism.

(4), according to F&dS, further demonstrates the need to have such a special mechanism specifically for singular, but not for plural INs: while a singular IN cannot antecede an overt pronoun in (4a), it can antecede the covert pronoun in (4b), which suggests that the DRT rules for this type of pronoun are different, and only the rule for the latter allows for  $\approx$  introduction.

However, there are reasons to suspect that there is no such distinction between covert and overt pronouns in Hungarian: if we replace the singular IN in (4a) with a full-fledged DP, the resulting example should be grammatical, according to F&dS's predictions; yet (11) is as bad as (4a):

- (11) <sup>??/\*</sup> János<sub>1</sub> vizsgált [egy beteg]<sub>2</sub> a rendelőben  
 Janos examined [INDEF patient].Acc the office.in  
 Ø<sub>1</sub> Túl súlyosnak találta őt<sub>2</sub>  
 he too severe.Dat found he.Acc  
 'Janos<sub>1</sub> patient<sub>2</sub>-examined in the office.  
 He<sub>1</sub> found him<sub>2</sub> too sick.'

As an anonymous reviewer pointed out to me, Hungarian generally does not allow overt pronouns in the direct object position, and this is probably the reason why these examples are bad. What is crucial is that the overt pronoun is equally bad with a singular IN and a full DP as an antecedent, and hence there is no need to have different rules for overt and covert pronouns in Hungarian: whatever rules out (4a) should also rule out (11).

As for the distinction between plural and singular INs, there are facts suggesting that the analysis proposed by F&dS makes this distinction too sharp—actually sharper than it is. While it is the general tendency both in

Hindi and Hungarian that plural INs easily antecede anaphoric pronouns and singular INs do not, there are important exceptions to this:

(12) *Dayal's (2003) example (43):*

anu-ne apne beTe ke-liye laRkii<sub>1</sub> cun lii.  
 Anu self's son for girl has-chosen  
 'Anu has chosen a bride for her son.'  
<sup>OK</sup> us-ne us-ko<sub>1</sub> ek sone-ka cen diyaa hai  
 she her one gold necklace has-given  
 'She has given her a golden necklace.'

Similar Danish facts were reported by Asudeh and Mikkelsen (2000):

(13) *Asudeh and Mikkelsen's (2000) example (12):*

a. Vita [købte 'hus<sub>1</sub>] sidste år. <sup>OK</sup> Det<sub>1</sub> ligger i Hals.  
 Vita bought house<sub>1</sub> last year It<sub>1</sub> lies in Hals  
 b. Mikkel [holdt 'forelæsning<sub>1</sub>]. \*Den<sub>1</sub> var spændende.  
 Mikkel held lecture<sub>1</sub> It<sub>1</sub> was interesting

Somewhat surprisingly, a Hungarian example parallel to (13a), with a singular IN, allows anaphora too (Anna Szabolcsi's judgment, p.c.):

(14) A bátyám házat<sub>1</sub> vett a múlt héten.  
 'The brother house<sub>1</sub>-bought last week'  
<sup>OK</sup> Egész vagyont adott érte<sub>1</sub>.  
 'He spent a fortune for it<sub>1</sub>.'

This suggests that F&dS's analysis too strongly differentiates between the semantics of singular and plural INs. While plural INs in languages like Hindi and Hungarian allow anaphora better than singular ones, the distinction is not as razor-sharp as F&dS describe it. While the data presented in this section is clearly just one of the steps to a better understanding and much empirical work is needed before any definite conclusions may be drawn, the empirical picture emerging so far may be summarized as follows:

- (15) a. Plural INs normally allow anaphora.  
 b. Singular INs usually do not allow anaphora.  
 c. However, there are certain contexts in which singular INs support anaphora much better than in other contexts.

The question of what those contexts in which singular INs allow anaphora have in common naturally suggests itself. Consider two pseudo-English

paraphrases of the examples above:

(12') Anu has bride-chosen for her son.

(14') Vita house-bought last week.

In both cases, the world knowledge about our culture supports the inference that there was, respectively, one bride and one house—we do not usually choose two brides for a single person at a time, and most people cannot afford to buy two houses in the same week. So in all of the examples in (12-14) that allow anaphora to singular INs there is at most one individual that may be referred to by the IN, according to our world knowledge.

I cannot see how Farkas & de Swart's theory would naturally account for this significant fact. In fact, the very existence of the "exceptions" in (12-14) is hard to explain within their theory, which states that plural INs are crucially different from singular INs in that they presuppose a discourse referent. In the next section, I propose a new analysis of the singular-plural IN distinction which is intended to capture not only the general tendency that plural INs allow and singular INs generally do not allow anaphoric reference to them, but also the apparent exceptions in (12-14).

## 5 Alternative Theory: Constructing Discourse Referents with the Help of the Context

I propose that thematic referents may never support anaphora directly; what seems to be anaphoric reference to a thematic referent is in fact reference to a discourse referent constructed using the context in which the thematic argument appears.

The classic variant of DRT described in Kamp & Reyle (1993) uses two procedures for the construction of a new discourse referent. The first one is Summation—a procedure that can produce a new plural referent summing up several already existing referents, needed for such examples as (16). The second operation is Abstraction, which creates a new referent referring to the sum of all individuals satisfying the conditions in both restrictor and nuclear scope of a tripartite quantificational structure. Abstraction is needed to account for discourses like (17), where "they" refers to the senators who came in.

(16) Lise<sub>i</sub> met Ann<sub>j</sub> yesterday. They<sub>i+j</sub> had not seen each other for years.

(17) Most senators came in. They were angry.

Note that Abstraction must produce the maximal possible individual, e.g., “they” in (16) cannot refer to a subset of those senators who came in. It can refer only to the maximal set.

I propose to introduce a third operation of this kind, **TA-Abstraction**, an analogue of Abstraction for contexts with thematic arguments. Here is the rule for this operation and the verification clause for conditions it produces:

- (18) Take a DRS  $K$  and turn it into  $K_1$ , adding to it a new discourse referent  $v$  and a condition of the form  $v = \Sigma x.K'$ , where  $K'$  is the copy of  $K$ .
- (19) Function  $f$  verifies  $v = \Sigma x.K'$  in  $M$  iff  $f(v) = \oplus \{a: a \in U_M \wedge (\text{function } f \cup \langle x, a \rangle \text{ verifies } K')\}$

Informally, (19) says that the new discourse referent  $v$  refers to a sum of all individuals  $a$  which can satisfy the conditions of  $K'$  when they are substituted for all the instances of the thematic argument  $x$ . In other words, this discourse referent refers to a maximal group of individuals that can be used to verify the conditions with the thematic argument  $x$  abstracted over.

The maximality of the constructed discourse referent helps to explain the differences in behavior between singular and plural INs and the “exceptions” in (12-14). As was shown by Farkas & de Swart (2003), plural INs denote non-atomic individuals, while morphologically singular INs are number-neutral, and there are both atomic and non-atomic objects in their denotation. Suppose we want to construct a new discourse referent from a context with a plural IN. To use an anaphoric pronoun we need to know the number of atomic individuals denoted by the newly introduced referent, because we have to choose an appropriate anaphoric pronoun. In the case of a plural IN, it poses no problem—this referent may not be atomic, because plural INs may not refer to atomic individuals, so anaphoric reference is in principle possible.

Now take the case of a singular IN. In general, we do not know if there is just one atomic individual satisfying the condition or there are more—or, even if the speaker knows it, the linguistic form she uses leaves that unspecified for the hearer. Thus it is not clear whether the new referent must be singular or plural, and I argue that this is what makes such anaphora bad. In principle, it is conceivable that the speaker might just generalize to the plural case, construct a plural individual and use a plural pronoun. But since plural regular DPs are not number-neutral, the speaker would cancel the number-neutrality in doing so, which is probably the reason why this option is actually not possible.

The seeming exceptions to this general rule in (12-14) receive a natural

explanation within the system just described. The problem with singular INs lies in their number-neutrality. However, in (12-14) world knowledge allows us to infer that the maximal individual referred to by the thematic argument is at most atomic. Hence the problem disappears, and the speaker may successfully construct a singular referent and use a singular pronoun.

While it would require a lot of empirical work to understand if our theory is actually viable, there are several interesting predictions it makes which point out novel kinds of data possibly relevant to the proper analysis of nominal incorporation. For example, our theory predicts that if a language uses only singular for mass nouns and mass nouns may be incorporated, then in such cases anaphora to singular INs should be possible: since the mass individual may not be plural, no conflict should arise. For now, I do not know if this prediction is borne out.

## References

- Asudeh, Ash, and Line Hove Mikkelsen. 2000. Incorporation in Danish: Implications for interfaces. In *Grammatical interfaces in HPSG*, ed. R. Cann, C. Grover, and P. Miller. Stanford, CA: CSLI Publications.
- Chung, Sandra, and William A. Ladusaw. 2003. *Restriction and Saturation*. Cambridge, MA: MIT Press.
- Dayal, Veneeta. 1999. Bare NPs, reference to kinds and incorporation. In *Proceedings of SALT IX*, 34–51. Ithaca, NY: Cornell University Press.
- Dayal, Veneeta. 2003. A semantics for pseudo-incorporation. Ms., Rutgers.
- Farkas, Donka, and Henriette de Swart. 2003. *The Semantics of Incorporation: From Argument Structure to Discourse Transparency*. Stanford, CA: CSLI Publications.
- van Geenhoven, Veerle. 1998. *Semantic Incorporation and Indefinite Description: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. CSLI, Stanford, Ca.
- Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic*. Dordrecht: Kluwer.

Chair of Theoretical and Computational Linguistics  
 Department of Philology  
 Moscow State University  
 1st building of humanities departments  
 GSP-2, Moscow, Russia, 119992  
 iyanovich@mail.ru