Incorporation as Unification

Ágnes Bende-Farkas*
IMS, Stuttgart University
e-mail: agnes@ims.uni-stuttgart.de

Summary: This paper proposes an analysis of a particular Definiteness Effect construction, that involving the light verb *have* as found in e.g. *John has a sister*. The analysis is based on the notion of term unification, which has played an important role in computer science and computational linguistics but, to my knowledge, hardly within formal semantics. According to term unification, the verb *have* and ist object phrase both introduce a predicational "term" consisting of a predicate and its arguments. Combining verb and object phrase involves unifying these two terms. When unification fails for some reason, the combination is unacceptable. The analysis is capable of extension to other Definiteness Effect constructions, including those involving "light" verbs in Hungarian.

1 Incorporation

In Van Geenhoven (1996) the author proposes the following treatment of "incorporation constructions" – constructions in which a verb and a nominal form one unit, with the effect that there is something of the kind denoted by the nominal which satisfies the relevant argument position of the verb. Such constructions are found in West-Greenland Eskimo, which van Geenhoven discusses, and many other languages besides. Incorporating verbs, according to van Geenhoven, have entries of the following forms:

(1) *a.* $TVs: \quad \lambda P.\lambda x.\exists y.[V(y,x) \wedge P(y)]$ *b.* $IVs: \quad \lambda P.\exists x.[V(x) \wedge P(x)]$

Here P is a variable of type $\langle e,t\rangle$, which gets instantiated by a common noun phrase (i.e. a predicate of type $\langle e,t\rangle$ when the verb is combined with a noun). The form of the entry is then responsible for the existential interpretation of the combination.

*The ideas underlying this paper are the result of close cooperation with Hans Kamp, to whom special thanks are due. This work has also benefited from discussions with Peter Krause, Barbara H. Partee, Mats Rooth and Ede Zimmermann, whose suggestions are gratefully acknowledged with the usual disclaimers.

Van Geenhoven proposes that an analogous analysis is also applicable to certain to "Definiteness Effect" constructions, which involve an NP position that may be occupied by indefinite NPs but not by definite or quantificational NPs, and in particular to constructions involving the English "weak" verb *have* and its counterparts in many other languages. Typical instances of weak *have* are found with object phrases whose head nouns are relational, such as *wife*, as in *Clyde has a wife*. *Have* as it is used in this sentence is called weak, because the actual relation that the sentence conveys between subject and object is not expressed by the verb itself but rather by the object noun. That such uses of *have* are subject to the Definiteness Effect can be seen by comparing the continuations (2a) and (2b).

(2) Clyde married Bertha. (a) So now he has a wife.
(b) So now he has her. (??)

(2a) can be paraphrased as *Now he is married*, but (2b) cannnot be paraphrased as *Now he is married to her.*

In order to apply the analysis in (1) to weak *have*, indefinites like *a wife* have to be taken to contribute to the verb-object construction a nominal (i.e., in this case, just *wife*). So, it is necessary to argue for some kind of type coercion here Partee (1987)). This is a complication, moreover, which equally arises when this strategy is applied to other Definiteness Effect constructions which accept full indefinite NPs, rather than, or as well as, bare nominals. It is clear, however, that in general this won't be enough to arrive at a fully satisfactory analysis. For there are significant differences between the various incorporation and Definiteness Effect contructions across languages. In particular, some such constructions allow the verb's object to act as antecedent to subsequent anaphoric pronouns (West Greenland Eskimo, Mohawk), whereas others do not (Hindi is one of these, see e.g. Dayal(1998)). Especially intriguing in this regard are the so-called "light" verbs of Hungarian (Szabolcsi (1986)). This is a substantial class of verbs which all show the Definiteness Effect, and which can be combined both with bare nominals and with full indefinite NPs. Pronominalization is possible when the object is a full NP, but not when it is a bare nominal:

(3) *a.* János macskát talált. ???Fekete volt.
John cat-Acc found. ???Black was
"There was a cat-finding event by John. ?It was black."
b. János talált egy macskát. Fekete volt.
John found one cat-Acc. Black was.
"John found a cat. It was black."

Evidently, a differentiated analysis of the different types of Definiteness Effect constructions is needed in order to explain these further differences. In the present paper an analysis for English weak *have* is proposed, which can be extended to account naturally for some of those differences, including the puzzle presented by Hungarian light verbs. The proposal borrows from Partee (1999a) and earlier work by the same author, but adds some elements of its own.

2 Unification

According to this proposal, then, the light verb introduces a higher order (relational) discourse referent. Provisionally this is taken to be of a pronominal character, needing to be matched to/unified with a property (with a relation in the case of *have*). This pronominal character distinguishes *have* from *own*, which can be seen from the difference between *John has a sister* and *John owns a sister*.[1]

Have is assigned the following entry in a space-saving linear notation, where DRSes are ordered pairs of a universe (set) and a condition set: $\langle U_K, Con_K \rangle$.

(4) $\langle \{s, \alpha, \beta\ R\}, \{R(s, \alpha, \beta)\} \rangle$

In (4), $R(s, \alpha, \beta)$ means that s is a state of affairs to the effect that α stands in the relation R to β.

The 'pronominal' status of R is to be understood in the following way: the condition "$R(s, \alpha, \beta)$" is to match a parallel condition supplied by the object NP. We assume, along with other work on bottom-up, compositional DRT (Kamp–Reyle t.a.) that the indefinite contributes a representation of the form (5):

(5) $\langle \{s', \gamma\}, \{s' \colon \langle \{\delta\}, \{Sister(\delta, \gamma)\} \rangle\} \rangle$

(5) says that s' is the state of γ being a sister of δ. δ is an implicit argument of *Sister*, because it is usually inaccessible to pronominal anaphora on its own, like the implicit arguments of compounds (Heim (1982)).

Combining (4) with (5) involves matching of the conditions $R(s, \alpha, \beta)$ $Sister(s', \delta, \gamma)$. This is understood in the sense of term unification. The two 'terms' unify if (i) $R(^{-1}) = Sister$, and (ii) the arguments match pairwise: $s = s', \alpha = \delta, \beta = \gamma$. (ii) is familiar from standard DRT, but (i) requires elucidation. In a simplified version of the analysis, R is an unconstrained 2-place relation variable. According to the present version, *have* might not be said to be "weightless" anymore. So, $R = Sister$ amounts simply to matching subject and object.

Applying the required (well-known) DRT-rules, we get the following DRS for the sentence *John has a sister*:

(6) $\langle \{n\ s\ s'\ \alpha\ \beta\ R\ \gamma\}, \{n \subseteq s', John(\alpha), R(s, \alpha, \beta), R = Sister, \beta = \gamma\} \rangle$

This is still not the entire picture, however, because *have* is not completely weightless. That it has some contribution of its own, and does constrain to some extent the set of possible relations, is shown by the oddity of *John has a divorcée*, intended to convey that John is divorced. The set of possible relations that can be matched to *have* is

[1] In this paper we concentrate on the basic, "light" meaning of *have*. We are aware of the differences between "light" *have* and "small clause" *have*, such as *John has his coat under the bed*, or between *have* of possession and cases like *John has Mary's cats now*, meaning that Mary's cats are with John for the time being.

some subset GEN of all possible 2-place relations. According to this version of the theory, *John has a sister* is acceptable because $Sister \in GEN$.[2]

Non-relational nouns such as *car* are assumed to be coerced into relational nouns, as suggested for genitive constructions in Jensen–Vikner (1994) and Partee (1999b). But then this means, as it should, that (part of) the meanings of their relational variants (say, Car' in the case of *car*) are in the set GEN. They can be given entries like the following:

(7) *a. a cat-of:*
b. $\langle\{s,x\},\{s{:}\,\langle\{\alpha\},\{cat(x),R_G(\alpha,x)\}\rangle\}\rangle$
c. $\langle\{s,x\},\{s{:}\,\langle\{\alpha\},\{\textit{cat-of}(\alpha,x)\}\rangle\}\rangle$

The common trait of (7b–c) is the introduction of a second (implicit) argument (α), with which the relevant discourse referent stands in some (GEN-)relation. (7c) incorporates the relevant relation in its meaning, whereas (7b) is more explicit. R_G in (7b) is one of the relations typically expressed by the possessor relation. R_G is not pronominal, unlike of *have* itself.

We take GEN to be closed under union or summation. This is motivated by examples like the following:

(8) John has a sister and two cats

The conjoined NP is taken to introduce a collective discourse referent X which is the sum of the two conjuncts. Likewise, the relation contributed by X is the smallest relation in GEN that comprises $Sister$ and *Cat(s)-of*.

3 The Definiteness Effect with Have

On this approach, the object of *have* predicts the right kind of properties for the Definiteness Effect.

Like NPs in *there*-insertion contexts, the object of *have* is to introduce a new discourse referent (McNally (1998) a.o.). This is ensured by the requirement (implicit in unification) that the relevant discourse referent should be unbound, where binding includes anaphoric binding and anchoring. Discourse-new definites may be admitted as objects of *have*, precisely because they are not bound to previous context:

(9) John has the smartest cat I know

[2] *Have* may not simply be conflated with the contribution of the Genitive, at least not in English. The set of relations expressible with light *have* seems to be a proper subset of those expressible by means of the Genitive, as seen from the following contrasts. (*Hundertwasser's house* is taken to mean a/the house designed by H.)
(*i*) *a. #Hundertwasser has a house in Stuttgart*
b. Yesterday we visited Hunderwasser's house in Stuttgart
(*ii*) *a. Virginia Woolf has a manuscript/#a hat in this museum*
b. This museum keeps Virginia Woolf's manuscripts/hats

Indefinites containing a presupposition, as in *John has his cat*, are excluded for the same reason as anaphoric definites.

Strong "quantifying" NPs are excluded from the object position of *have* for two (closely related) reasons. First, *have* provides a predicational frame, so to speak, which expects a term-type object. A quantificational NP like *every sister* simply does not "fit" into that frame, and there is no access to the free variable required by *have*.

Second, if one starts out from the quantifier, and tries to introduce *have* in its nuclear scope, several things can happen. Even if we take (10a) as grammatical, and do want to get something, the best is an uninformative reading.

(10) *a.* ∗John has every sister:
b. $\langle\{j, n\}, \{\forall x.\langle\{s, x\}\{n \subseteq s, s{:}\,\langle\{\beta\}, \{x\mathit{Sister\text{-}of}\,\beta\}\rangle\rangle,$
$\langle\{s', R, \alpha, \}, \{n \subseteq s', R \in GEN, s{:}\,jR\alpha\}\rangle\}\rangle$

If we assume that unification can be performed for (10b), this yields the tautology reading familiar from Barwise and Cooper: *every sister of John is a sister of John*. But unification should not be possible in this case: *have* is taken to expect a unifier in its local context.[3]

For the same reason, that is, because *have* expects a local unifier, only narrowest scope is possible under operators, or in attitude contexts, in sentences like *John thinks that Mary has a sister*, or *Whenever we visited him, John had a cat*.

Mutatis mutandis, the same holds for object NPs with constituent negation: within the scope of negation, *have* and its object are in each other's local contexts. So, a sentence like *John has no sisters* will be paraphrased as *It is not the case that John has a sister*.

4 Hungarian Light Verbs

Hungarian light verbs are amenable to the same kind of analysis that has been proposed for English *have*. One important difference with English *have* is that typically, Hungarian light verbs provide more content.

For instance, the lexical entry for the verb *talál* 'find' is said to contain an event description (there is a finding event by the agent), and a consequent state. It is the description of the consequent state that contains a higher order discourse referent awaiting unification. This is based on the intuition that the consequent states of these verbs are not hard-wired in the lexicon, so to speak, but are contributed to a large extent by the Theme argument, iff it has the required properties.

(11) *a.* *talál* 'find':
b. $\langle\{ec, \alpha\}, \{ec{:}\,\langle\{e, s, P, \beta\}, \{e{:}\,\alpha find\ \beta, s{:}\,P(\beta)\}\rangle\}\rangle$

[3] For reasons of space, a full account of this will be provided in the longer version of this paper.

(11b) contains a higher order discourse referent P which needs to be unified with the property contributed by the internal argument. In the case of (3b), for instance, P is unified with *cat*. The first-order discourse referent β for the internal argument is introduced internally to the event description. This way, β on its own is inaccessible to pronominal anaphora.

Hungarian bare nominals are not taken to introduce discourse referents, at least, not accessible ones. Hence, in the case of (3a), the internal argument β remains unmatched and thus inaccessible. If the object of *talál* is a full NP, as in (3b), it introduces a discourse referent at the appropriate DRS level, which is then accessible for anaphora. Thus the difference between (3a) and (3b) is seen to follow from the the different semantics for bare nominals and full NPs, eliminating the need to stipulate an ambiguity for the incorporating structure itself.

References

[1] Werner Abraham and Sjaak de Meÿ, editors. *Topic, Focus and Configurationality.* John Benjamins, Amsterdam, 1986.

[2] Mark Baker. *Incorporation: A Theory of Grammatical Function Changing.* University of Chicago Press, Chicago, 1988.

[3] Ann Copestake, Dan Flickinger, and Ivan Sag. Minimal Recursion Semantics: An Introduction, 1997.

[4] Per Anker Jensen and Carl Vikner. Lexical Knowledge and the Semantic Analysis of Danish Genitive Constructions. In S.L. Hansen and H. Wegener, editors, *Topics in Knowledge-based NLP Systems*, pages 37–55. Samfundslitteratur, Copenhagen, 1994.

[5] Hans Kamp and Uwe Reyle. *From Discourse to Logic.* Kluwer, Dordrecht, 1993.

[6] Louise McNally. Existential Sentences without Existential Quantification. *Linguistics and Philosophy*, pages 353–392, 1998.

[7] Barbara H. Partee. Weak NPs in HAVE-Sentences. In Jelle Gerbrandy, Maarten Marx, Maarten de Rijke, and Yde Venema, editors, *Liber Amicorum for Johan van Benthem.* Amsterdam, t.a.

[8] Anna Szabolcsi. From the Definiteness Effect to Lexical Integrity. In Abraham and de Meÿ [1].

[9] Veerle van Geenhoven. *Semantic Incorporation and Indefinite Descriptions.* PhD thesis, Tübingen University, Tübingen, 1996.