Noun Incorporation from a Semantic Point of View
Author(s): Veerle van Geenhoven
Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Place of Morphology in a Grammar (1992), pp. 453-466

Noun Incorporation from a Semantic Point of View *

Veerle van Geenhoven
Universität Stuttgart

0 Introduction

The fact that indefinite NP's are discourse transparent, as shown in (1), was one of the reasons for many semanticists (Heim (1982), Kamp (1981)) to believe that indefinite NP's are referential rather than quantificational expressions, as shown in (1)'s Discourse Representation Structure (DRS) under (2).

(1) Frederik owns a horse$_i$. It$_i$ is a Hanoveraner.

(2) $\exists x\, \exists y$ [horse(x) & own(f, x) & y = x & Hanoveraner(y)] [1]

Postal (1969) observes that words that are embedded into more complex words are not discourse transparent and that therefore incorporated nouns (INs) are anaphoric islands.

(3) Emily had bought fresh dog food$_i$, but it$_?$ didn't eat it$_i$.

The question mark index on the definite pronoun *it* in (3) indicates that *it* cannot be coindexed with any NP in this piece of discourse. Although *dog* appears to be the most obvious antecedent candidate of *it*, it is incorporated in the compound *dog food* and therefore it does not have an independent referential meaning which can be used as a potential antecedent for anaphoric expressions. This is made explicit by the unsolved equation in DRS (4).

(4) $\exists x\, \exists u\, \exists y$[**dog**-food(x) & buy(e, x) & u = **?** & y = x & ¬[eat(u, y)]]

From Sadock (1980) we learn that in West Greenlandic Eskimo personal suffixes on verbal affixes play much the same role as definite pronouns in English. His example (Sadock (1980): 311):

(5) Suulut timmisartuliorpoq. Suluusaqar**poq** aquuteqarl**lu**ni**lu**
Søren(ABS) airplane-made-INDIC-3sg wing-have-INDIC-**3sg**
rudder-have-INF-**4sg**-and.
'Søren made an airplane. It has wings and a rudder'.

According to Sadock the incorporated object 'airplane' is to be understood as an indefinite NP. A semantic representation of (5) along the lines of Discourse Representation Theory (DRT) would yield (6).

(6) $\exists x\, \exists Y\, \exists y\, \exists u\, \exists z$ [airplane(x) & make(s, x) & wing*(Y) & have(y,Y)
& y = x & rudder(z) & have(u,z) & u = x]

This paper is an attempt to capture the anaphoric potential of INs in some polysynthetic languages in the formal semantic framework of Discourse Representation Theory (DRT).[2] In this perspective, there are two questions that need to be answered. First, if word internal elements are discourse transparent, *what* is their semantic contribution, what do they *denote*? My answer to this question is that INs with anaphoric potential are referential expressions. On top of that, one has to think about *how* incorporated referents can be made *accessible* in the process of semantic construction. To make this second question more concrete: how do I get the semantic representation (6) out of (5)'s surface syntactic representation in the same way as I get (2) out of (1)? The view advocated in this paper is that apart from surface syntactic phrasal structures also word-level configurations have to be considered as triggering information for DRS construction.

Before I get to answering these two questions, I point at yet another issue related to discourse transparency. Discourse transparency has often been regarded as a valid argument in favor of a syntactic approach to noun incorporation (NI) in

polysynthetic languages (Sadock (1980), Baker (1988)). However, Ward and Sproat (1987) show that in English compounds word internal elements are often discourse transparent. Given that English compounding is commonly taken to be a lexical phenomenon, I will conclude that discourse transparency in itself is too weak to be a convincing argument in favor of a syntactic analysis of NI.

1 INs and Discourse Transparency: the Validity of the Argument

The fact that INs in polysynthetic languages have anaphoric potential, i.e. that they introduce discourse referents as in (5), is one of the main arguments to treat incorporation as a syntactic rather than a lexical phenomenon (Baker (1988), Sadock (1980), (1986)). In this section I want to question the validity of this argument, i.e. is the semantic transparency of word formation processes a convincing argument that these processes are syntactic?

In a recent paper, Ward, Sproat and McKoon (1991) discuss the anaphoric potential of nominal expressions that are embedded in English compounds. Although I am not convinced that all of their examples are acceptable, I borrow a subset of them to show that some word formation processes in English appear to be transparent as well (Examples from Ward, Sproat and McKoon (1991).

(7) Museum$_i$ visitors can see through its$_i$ big windows the 900-year-old Tower of London and the modern office blocks of the City financial district.

(8) At the same time as coffee beans were introduced, the Arabs made changes in the coffee$_i$ preparations, that greatly improved its$_i$ flavor.

(9) Bush$_i$ supporters would stay home figuring he$_i$'d already won.

Ward, Sproat and McKoon's account for the correct coindexing in (7), (8) and (9) is that the relation between the head and its left hand member in the respective compounds *museum visitor*, *coffee preparations* and *Bush supporters* is easily decomposable. A set of pragmatic principles - contrast, topicalization - is responsible for this decomposition process. Despite the fact that English compounds appear to be discourse transparent, this does not affect the standard view that English compounding is a lexical phenomenon rather than a syntactic, because the transparency is only partially present and the semantic relation in a compound between its head and its left hand member is often idiosyncratic.[3] I take this to be an indication that the discourse transparency of an IN is not one-to-one related to its being an instance of syntactic word-formation.

Nevertheless, the fact that the lexical semantics of English compounds is often transparent when it is used in a piece of discourse is important if one wants to make Ward, Sproat and McKoon's pragmatic principles visible in a formal semantic approach to discourse semantics, such as DRT. If the semantic decomposition of a compound is present in the lexicon, then its semantic representation will be inserted as a complex DRS condition into the DRS of the sentence of which it is used. I will illustrate this in giving the lexical semantic translation of the compound *museum visitor*.

(10) $\lambda\eta\, \exists x$ [museum(x) & visitor*(η, x)] [4]

In this DRS condition I have assigned the referent *x* existential force because one can infer from the extensionality of the verbal stem *visit-* that its object can get a de re reading if the context makes such a reading plausible.[5] When we now introduce (10) as a complex DRS condition into the DRS of (7) the referent of *museum*, i.e. *y*, will be in a position that is subordinated to the positions of the other referents. Given that in DRT accessibility is defined in terms of subordination, it becomes clear that *y* is still unaccessible as the antecedent of the pronoun *it*.

(11) ∃X ∃Y ∃u [∃y [museum (y) & visitor*(X,y)] & big-windows*(Y)
& u = ? & POSS(u, Y) & see-through(X,t,Y) ...]

Still a separate story has to be told to explain how *y* escapes its embedded position, because we do want to set *u* equal to it. The only way this explanation can go is to say that the referent *y* is accommodated to a level that is accessible for the referent *u* introduced by the pronoun *its*. Notice, that such an accommodation mechanism does not apply systematically, i.e. one would not want it to be at work for a DRS of (12).

(12) Museum visitors always wish that it? would be open on mondays too.

The fact that this accommodation mechanism does not apply systematically is the main reason why I have assigned a narrow scope reading to the lexically embedded *museum* in (10), since if I would have treated it as a free variable, sentences like (12) would automatically come out as semantically well-formed.

Notice also that without a separate accommodation mechanism any lexicalist approach to NI in polysynthetic languages (Di Sciullo and Williams (1987), Mithun (1984); (1986) and Rosen (1989)) is unable to treat discourse transparency in a straightforward way. I repeat (5) to illustrate into what DRS the lexical semantic decomposition of the complex words *timmisartulior-*, *suluusaqar-* and *aquuteqarl-* under (13) would be inserted.

(5) Suulut timmisartuliorpoq. Suluusaqar**poq** aquuteqarll**uni**lu
Søren(ABS) airplane-made-INDIC-3sg. wing-have-INDIC-**3sg**
rudder-have-INF-**4sg**-and.
'Søren made an airplane. It has wings and a rudder'.

(13) λη[∃u [airplane(u) & make(η,u)]]
λη[∃U [wing*(U) & have(η,U)]]
λη[∃u [rudder(u) & have(η,u)]]

Even with all of (13) available in the lexicon, the resulting DRS (14) shows that the airplane referent *u* is in a subordinated position and therefore unaccessible for anaphoric reference. This again follows from the simple reason that a lexicalist approach treats *timmisartulior-* as one lexical unit.

(14) ∃x ∃y [∃**u** [airplane(u) & make(s, u)] & ∃Y [wing*(Y) & have(x,Y)]
& x = ? & ∃z [rudder(z) & have(y,z)] & y = ?]

As for (11) we would have to stipulate a separate accommodation mechanism in the semantic component in order to let *u* escape its embedded position. The mechanism would yield the desired DRS (6) repeated here as (14').

(14') ∃x ∃Y ∃y ∃u ∃z [airplane(u) & make(s, u) & wing*(Y) & have(x,Y)
& x = u & rudder(z) & have(y,z) & y = u]

However, since in Greenlandic Eskimo INs appear to be systematically discourse transparent, the question arises whether accommodation really is the best explanation to represent this systematicity. In a syntactic treatment of NI one would not expect to encounter this problem, because there word formation is not lexically pre-processed but runs parallel with phrase construction. A syntactic approach makes it possible to keep DRS construction compositional and it would yield (14') without any further stipulation.

Another point of interest with regard to the notion of discourse transparency is that the object required by an anaphoric expression is not per definition an individual object. Nominal expressions can be discourse transparent in different ways. In example (1) the pronoun *it* requires an antecedent that denotes an individual.

(1) Frederik owns a horse. *It* is a Hanoveraner.

In (15) the definite pronoun *them* refers to a kind antecedent, and the CN-anaphor *some* in (16) takes a nominal sense as it antecedent.[6]

(15) If at least one chicken of Ottilie has laid an egg, she has had a nice breakfast. *They* are very good to eat. (Example from Kamp and Reyle (1991))

(16) Most books have been returned to the library. Unfortunately, *some* were damaged.

Also incorporated nominal expressions can establish a kind discourse referent.

(17) I'm a mystery-story buff, and I read a lot of *them*.

(18) I also have a paper on the interaction of child morphology with *their* phonological skills, ... (Examples from Ward e.a. (1991))

It has become a common assumption that word internal elements in English-like language name kinds (Sproat and Ward (1987), Hoeksema (1984)). There is also a point of general agreement among incorporationists who are working on polysynthetic languages, that in many of these languages INs often have a non-specific or generic meaning, or to put it slightly differently, that they name kinds as well. In what follows I will bring INs in nonpolysynthetic and polysynthetic languages under the same semantic type denominator.

2 The Semantic Contribution of Incorporated Nouns

2.1 Incorporated Nouns denote Properties

A standard work on reference to kinds is Carlson (1977). In his ontology kinds are objects just like individuals. In this paper, I take kinds to be properties, or in type-theoretic terms, intensions of sets. This view is also mentioned by Carlson as yet another non-quantificational way to look at kinds in a possible world semantic setting, but he raises two objections against it (cf. Carlson (1977): 161). Carlson argues that since in Montague Semantics all NPs are taken to denote sets of properties one would need an extra complication to turn a property into a set of properties. This first objection does not hold in the DRT framework in which I am presently at work: bare plurals are indefinite NPs and in DRT semantics these are referential expressions. His second objection is that in the "property view" the extension of a property differs when different extensional predicates are predicated of it.

(19) Dogs are barking at me.

(20) Dogs are mammals.

In (19) *dogs* reads as 'some dogs', whereas in (20) it reads as 'all dogs'. Also this is not a real objection to the view that kinds denote properties for I would not treat the predicates *are barking at me* and *are mammals* on a par. To call *are mammals* an "extensional" predicate is based on how our actual world happens to be, i.e. a world in which all instances of the kind dogs are mammals. The extensionality of *are barking at me* is related to what is going on at a particular time-location point in the actual world.[7]

Is there any evidence that kind naming INs denote properties? From the common belief that through the resolution of anaphors we can learn more about the semantic types of their antecedents, it seems plausible that INs do denote properties. I advocate the view that CN anaphors need a property as their antecedent and that the relation between such an anaphor and its antecedent is sense identity. To interpret (21) and (22) we are supposed to extract that sense from the context or, as in (23) and (24), from the preceding discourse.

(21) Isabel ate *two*.

(22) Benjamin needs *one* too.

(23) Jonas had five sardines for dinner and I ate *one*.
(24) Ten books have been returned to the library. *Five* got lost.

Indefinite CN anaphors have a double role in discourse semantics. Therefore, I propose that, on the one hand, they introduce a (set of) discourse referent(s) and, on the other, they introduce a DRS condition variable that holds of the latter. It is this variable that displays the anaphoric role of CN anaphors in that it has to be set equal to some other condition that has been introduced before (or that is known from the context as for (21) and (22)).

(25) $\exists X\ \exists Y\ \exists C\ [C = \text{sardine}^* \ \&\ \text{sardine}^*(X) \ \&\ |X| = 5 \ \&$ have-for-dinner(j,X) & $C(Y) \ \&\ |Y| = 1 \ \&\ \text{ate}(I,Y)]$

(26) $\exists X\ \exists u\ \exists Y\ \exists C\ [C = \text{book}^* \ \&\ \text{book}^*(X) \ \&\ |X| = 10 \ \&\ \text{library}(u) \ \&$ return-to(X,u) & $C(Y) \ \&\ |Y| = 5 \ \&\ \text{lost}(Y)]$

In (25) and (26) the CN anaphors *five* and *one* take up a property or a DRS condition in stead of a set, simply because such a set is neither available nor can its existence be presupposed.[8] Also in (27) and (28) the CN anaphor *some* and *one* can be read as 'some elephants' and 'one rabbit' respectively.

(27) Timu went on an elephant hunt. When he caught sight of *some*, he leveled his rifle to shoot them.
(28) Joe saw a fresh rabbit track, so he knew that there had to be *one* around.

That also here there is no actual (set of) individual referent(s) available in these cases is clear from the impossibility to use a definite pronoun in the respective second parts of these pieces of discourse.

(29) Timu went on an elephant hunt. When he caught sight of **them*, he leveled his rifle to shoot them.
(30) Joe was following a rabbit track, so he knew that **it* had to be around.[9]

The fact that properties or intensions of sets seem to escape lexical embedding whereas I claimed before that this is impossible for their extensions can be explained in a very straightforward way. Nouns stand to their semantic value, i.e. an intension, in the same way as proper names to the bearers of these names. We can therefore assume that every noun's denotation is rigid and that they are modally closed.[10]

The fact that an IN denotes a property relates directly to what I regard to be the meaning of an n-place predicate. Standard Montague semantics treats the meaning of a transitive predicate as a two-place relation between an individual and a quantifier. To account for the obligatorily referentially transparent reading of the object of an extensional transitive verb ('to find', 'to make', 'to kiss', ...) Montague postulates a necessary meaning shift that turns the quantifier denotation of its internal argument into an individual object denotation. Zimmermann (1991) argues against Montague's view and treats referentially opaque verbs as relations between an individual and a property. I take over Zimmermann's alternative and claim that every transitive predicate's basic meaning is a relation between an individual and a property. Again, the denotation of the object of an extensional predicate will, by postulation, be a (set of) individual object(s).[11] The denotation of its subject, i.e. the external argument, depends on whether the predicate is episodic or stative. In the former case, the subject is a (set of) individual(s), whereas in the latter it is a property.

Many arguments of Carlson's findings about kinds are based on the semantic behaviour of the bare plural construction in English. I said before that in DRT bare plurals are referential expressions or terms. From that perspective, it is rather straigthforward to regard all kind naming nominal expressions - the

indefinite NP, the bare plural, the IN - as denoting properties. In other words, INs can be the word-level counterpart of the indefinite NP construction. In nonpolysynthetic languages an IN can be the lexical word-level counterpart of indefinite NPs, in polysynthetic languages it can be its syntactic alternative. Evidence that supports this view is that in none of the literature about NI I know of, I encountered an example in which a quantificational NP was incorporated.[12]

2.2 An External Quantificational Source

Referential expressions get their quantificational force from an external source. In Carlson (1977) a distinction between two types of predicates, i.e. the individual-level and the stage-level predicate, accounts for the distinction between the 'genuine' *de dicto* or generic and the *de re* or existential reading that a bare plural in subject position can have. This is illustrated in (31) and (32).

(31) Horses are nice animals.
(32) Horses are galopping in Jim's meadow.

are nice animals is an individual-level predicate: it holds as a property of the kind 'horse', regardless whether some or even a lot of the instances of this kind are not so nice. The predicate *are galopping in Jim's meadow* is a stage-level predicate: it holds for some instances of the kind 'horse' at a particular time-place location. This explains why the bare plural in (32) gets an existential reading, whereas it gets a generic reading in (31).[13]

Also in object position, indefinites can have either a generic or an existential reading. This depends on the extensionality of the verbal predicate.

(33) Joachim wants a horse/horses.
(34) Rob bought a horse/horses.

Because *to want* is an intensional verb, the indefinite NP/bare plural in (33) has a de dicto or generic reading. This is not the case in (34). Here the indefinite NP/bare plural object gets an existential reading since the verb *to buy* forces its object to be extensional. Hence, *a horse/horses* in (30) denotes (an) instance(s) of the kind 'horse'.

I already pointed out that in English-like languages INs often have an infered existential reading (cf. *museum visitor* in (7)) which I take to come from an external source. In (7) the pronoun can be resolved with a linguistically expressed but incorporated antecedent only when the addressed subject, i.e. a museum visitor, really is in a museum (or reads a brochure about a museum of which she is a potential visitor) and we therefore are able to infer the existence of the latter from the context. In some polysynthetic languages we apparently have a systematic use of INs as referential expressions with an existential reading. According to Sadock, the reason that incorporated nominals in Greenlandic Eskimo can so easily introduce discourse topics is "that, in many cases, the language provides no nonincorporated form of equal or lesser complexity and idiomaticity" (cf. Sadock (1986): 25). I assume that like the bare plural construction and the indefinite NP in English, the existential reading of INs derives from an external source as well. The most obvious candidate is the kind of predicate in which the IN is embedded (Examples (19) and (27) from Sadock (1980)).

(35) Sapangarsivoq.
'He bought beads'
(36) Kusanartumik sapangarsivoq.
'He bought a beautiful bead'

In the Eskimo example (35) the noun *sapangaq* ('bead') is incorporated in an extensional transitive verbal affix *-gar-* ('to have'). Even if the meaning of an IN is

taken to be nonspecific - it could have been any set of beads - the transparency of the verb forces a de re reading of the IN. In (36) where the IN *sapangaq* is modified through an external modifier *kusanartumik* the former also gets an existential reading from the very same affix. An example in which the IN does not have to get an existential reading is (37) (Example (88) from Sadock (1985)).

(37) Kaali illoqarumavoq.
Kaali illu-qar-uma-voq.
Karl(ABS) house-have-want-INDIC/3s
'Karl wants to have a house'

Here the modal operator 'want', realized in Eskimo as the affix *-uma-*, has the property 'to have a house' in its scope. The IN *illu* ('a house') can have either an existential reading or a generic reading. Anaphoric reference with a definite pronoun to the modally embedded noun can only be resolved in the former case.

(38) Karl wants to have a house$_i$. It$_{i/?}$ is too expensive.

Apart from the incorporated object's referential transparency that is forced by an extensional transitive verb it would be interesting to see how other external sources - quantifiers, temporal and local adverbs, ... - are realized in Eskimo, that could give rise to an existential reading of an incorporated referential expression.

Apart from object incorporation, there are also languages in which we encounter subject incorporation. I borrow a Caddo example from Mithun (1984):

(39) Ná: kan-núh-'a'.
that water-run.out-will
'That water will run out.'

The incorporated subject *kan* ('water') is a referential expression that gets its existential reading from the temporal inflection on the predicate *nuh* ('run out'). This reading is also enforced by the 'stranded' demonstrative determiner *na* ('that').

Now that the semantic denotation of INs with anaphoric potential is said to be a referential expression, we can start to think about the question how to construct DRSs for incorporating configurations in polysynthetic languages. Although it was said that discourse transparency as such is not a valid argument in favor of the syntacticity of NI in polysynthetic languages, there appears to be a systematic referential transparency of objects incorporated in extensional transitive verbs, a systematicity that is missing in complex words in English-like language. I also showed that a lexicalist approach to NI is not able to cover this systematicity, unless a rather ad hoc accommodation mechanism is introduced.

3 DRS Construction of Object Incorporating Languages

Within the framework of DRT many attempts have been made to define construction procedures that yield semantic representations which have to be interpreted along the findings of DRT philosophy (cf. Kamp and Reyle (1991)). One way in which these algorithms can be set up is that they take surface syntactic structure configurations as their input from which stepwise DRSs are constructed. The DRS construction trigger in English indefinite NPs is the determiner *a(n)* for the singular case and the lack of a determiner for the plural one. When in a syntactic tree an occurrence of the determiner *a(n)* is encountered in an NP, a discourse referent will be introduced in the discourse universe and the nominal head of that NP will be translated into a DRS condition that holds of that referent.

In the foregoing I have argued that object INs in polysynthetic languages play much the same semantic role as indefinite NPs, in that they are referential expressions. We now need a structural trigger that yields the same result in DRS

construction as the indefinite determiner *a(n)* in my sketch above. The most obvious candidate for this task is the NI configuration itself and I therefore need a surface syntactic parse that does justice to that configuration. At this point the question arises in how far I can take advantage of existing syntactic approaches to NI. One would expect that surface syntactic parses of incorporating configurations based on the incorporation theories of Sadock (1985), (1991) and Baker (1988) can be used as the input of a DRS construction algorithm. However, I will show that this expectation is wrong and give an answer as to why it is wrong. I will then sketch how an alternative syntactic approach to object NI could be set up.

3.1 Existing Syntactic Analyses

3.1.1 Autolexical Syntax (Sadock (1985), (1991))

Sadock's (1985) auto-lexical model consists of a word-level and a phrase-level component that are set up in parallel as two autonomous combinatorial systems. In his view, morphological structure and syntactic structure are independent to the extent that the leaves of syntactic trees need not correspond to the roots of morphological trees (Sadock (1985): 387). An interface grammar has to link the two structural systems. From the perspective of the semanticist who wants to construct the semantic representation of NI configurations, it is somehow disappointing to read that of the two syntactic modules only the phrase-level syntax is claimed to be relevant for semantic interpretation. Hence, morphological information is doomed to be lost for interpretation.

The first thing we seem to lose is inflectional information. The double tree under (37) shows that the inflectional morpheme *-poq* is not present in the syntactic tree. It has been argued at length by Bach (1983) that inflection does have to be considered for interpretation. Putting the phonological aspect aside for now, Bach's grammar picture shows that semantic interpretation is not a privilege for phrasal syntax (Bach (1983): 65).

Word-grammar	*Phrase-grammar*
syntax	syntax
phonology	phonology
semantics	semantics

Sadock is aware of this and he claims that although inflectional affixes have no lexemic representation in the syntactic tree, the inflectional information of a word can be reflected by means of features on the syntactic tree (Sadock (1985): 403).

But in the (1985) autolexical model the syntactic tree doesn't involve morphological configurational information either. At the syntactic level an incorporated object argument gets the same representation as its non-incorporated version, i.e. one in which the argument is a maximal projection. Given that in an autolexical model only the syntactic trees are relevant for semantic interpretation we will get one and the same DRS for both configurations with incorporated arguments and their nonincorporated versions. However, many languages show meaning differences between incorporated structures and their nonincorporated counterparts, differences that are found in what they denote. Sadock (1980) himself gives us evidence for such a meaning difference in Greenlandic Eskimo (Examples (19) and (25) from Sadock (1980)).

(40) Sapangarsivoq.
'He bought beads'

(41) Sapanngamik pisivoq.
'He bought a bead'

(40) gives us the incorporated version of (41). Sadock notes that the incorporated versions of *sapangaq* ('bead') is always to be understood as an indefinite and that the incorporated version is also neutral with respect to number. Other denotational differences between incorporated and nonincorporated configurations we find in the Micronesian language Mokilese (Examples (3a) and (3b) from Mithun (1984)).

(42) Ngoah kohkoa oaring-kai.
I grind coconut-these
'I am grinding these coconuts'

(43) Ngoah ko oaring.
I grind coconut
'I am coconut-grinding'

Mithun notes that the incorporated coconuts in (43) are not refering to specific coconuts, but that they modify the activity of grinding. In Sadock's approach the incorporated and the non-incorporated versions of these Eskimo and Mokilesean examples will be assigned the same syntactic, downward tree structure. From his claim that only the latter delivers input to the semantic component, it follows that they will also be assigned the same truth conditions, which of course is an undesired result.

It is not clear to me whether Sadock wants his syntactic trees to be a S- or D-structural representation. What is systematically missing in the autolexical model is the surface syntactic parse of a sentence in which the incorporation configuration is present. The morphological trees stand all on their own, i.e. they are not interrelated. Hence, we would expect that the surface syntax is dekivered by the syntactic module, but this is neither the case. The part of the autolexical syntax model that I have left unmentioned till now is the interface grammar. In Sadock's (1991) model this interface component plays a central role in that it fixes the respective two-way connections between the syntactic, the morphological and the semantic component. The semantic component interacts with the morphological and the syntactic component in a similar but autonomous way. It is therefore unclear to me how the semantic representation that correlates with the morphological structures of the words of a particular sentence and the semantic representation that correlates with the syntactic tree of that sentence can be unified into one and the same semantic representation that contains the truth conditions of that sentence.

3.1.2 A Transformational Approach (Baker(1988))

Also Baker (1988) advocates the view that NI in polysynthetic languages is a syntactic phenomenon. Amongst his arguments in favor of a syntactic approach discourse transparency is a convincing argument that incorporation is a syntactic operation.[14] But, if we take a closer semantic look at Bakers's approach to incorporation it is clear that we encounter a situation that is very similar to the autolexical one that I have decribed above.

Baker explains incorporation in terms of thematic paraphrasability, which alludes to the fact that an IN can be paraphrased as a maximal projection. To illustrate his view I take one of his examples from Onondaga (Examples from Baker (1988)).

(44) Pet wa-ha-hwist-ahtu-t-a.
Pat PAST-3ms/3N-money-lost-CAUS-ASP
'Pat lost money.'

(45) Pet wa-ha-htu-t-a ne o-hwist-a.
Pat PAST-3ms/3N-lost-CAUS-ASP the PRE-money-SUF
'Pat lost the money.'

Example (44) contains a nominal expression *hwist* ('money') that is incorporated in the verbal expression *ahtu* ('to loose') of which it is the object argument. (45) contains the same expression but now it is realized as the nominal head of a free NP argument. It is also (45) that is the D-structural source of (44) since Baker defines NI as a structure dependent operation, i.e. a word-level version of move-α:

(46) [... [VP [V][*NP* [N]]] ...] ⇒ [... [VP [V [N_i V]][NP [t_i]]] ...]

Although he recognizes the discourse transparency of INs in polysynthetic languages, according to his view an IN N_i in (46) does not have an independent referential meaning, but receives that meanig from its D-structural *NP* source via the N-chain. This means that if we take a Bakerian surface structure level as the input for DRS construction we will have to reconstruct its D-structural representation in order to interprete that chain. But it is not difficult to see that we get again the wrong result in that the incorporated version, e.g. (44), and the nonincorporated one, e.g. (45), are then assigned the same truth conditions. The syntactic representation that is relevant for interpretation is, as in Sadock's approach, the same for both the incorporated version of a verbal argument and its nonincorporated counterpart.

Although Sadock's and Baker's views of grammar modelling are conceivably distinct, the identification of their respective weak points in the treatment of NI can be brought under the same denominator, which is the mistaken view that thematic paraphrasibility equals truthconditional paraphrasibility. At the grammar level that is relevant for interpretation, neither Sadock nor Baker distinguish between the incorporated and the nonincorporated realization of a nominal expression. In what follows I regard NI in polysynthetic languages as being a part of an independent word-formation operation that is syntactically real.

3.2 A Monostratally Syntactic Approach to NI

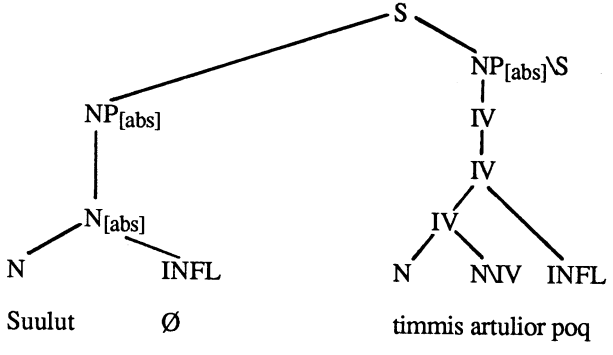
For my purposes, I will only borrow from the syntactic principles of Categorial Grammar and Categorial Morphology (Hoeksema 1984, Moortgat 1983). Traditional categorial systems mostly adapt a strict Montagovian framework for its interpretation. I'd rather translate the syntactic categories into building blocks for DRS construction. The reason that I borrow from CG syntax lies in the fact that this monostratal syntactic framework draws a useful distinction between linguistic operations and grammar rules. Dowty (1979) makes this distinction even more finegrained in that, on the one hand, operations are either word-level or phrase-level and, on the other, grammar rules are either syntactic or lexical. Given this crossclassification one is not forced to think of NI as being either a syntactic or a lexical phenomenon because word-level and phrase-level operations can show up both in syntactic and in lexical rules. In Hoeksema (1984) we find a thorough analysis in a Dowty-like CG framework of relational and synthetic compounds in English-like languages that involves incorporation. In these languages, NI is part of partially productive word-level operations and it therefore shows up in lexical rules. This is a purely language specific matter and it does not imply that word-level operations are lexical per se. To ascribe affixation and compounding in polysynthetic languages a syntactic status, we need syntactic rules that are defined in terms of word-level operations. Notice that the discussion as to whether NI is a lexical or a syntactic phenomenon ends up being a purely language specific issue.

A test whether a particular instance of NI is syntactic rather than lexical is to check whether it is systematically transparent or not. Notice also, that in this monostratally syntactic approach incorporation is a descriptive term which has no procedural meaning.

In the second section, I have pointed at a semantic parallelism between the English indefinite NP and INs. I have also suggested that in order to make this parallelism explicit in DRS construction a syntactic NI configuration itself would be assigned the same role as the indefinite NP configuration. Whereas in (1) the presence of the determiner *a* in the NP *a horse* triggers the introduction of a discourse referent, in (47) the fact that the N *airplane* is strictly adjacent to its head *made* triggers essentially the same thing. In this way, the lack of a syntactic determiner is intercepted: although INs in a polysynthetic language as Eskimo are word parts, their word part status is exactly what accounts for their systematic transparency.

(47) Suulut timmisartuliorpoq.
Søren(ABS) airplane-made-INDIC-3sg
'Søren made an airplane'

(47')

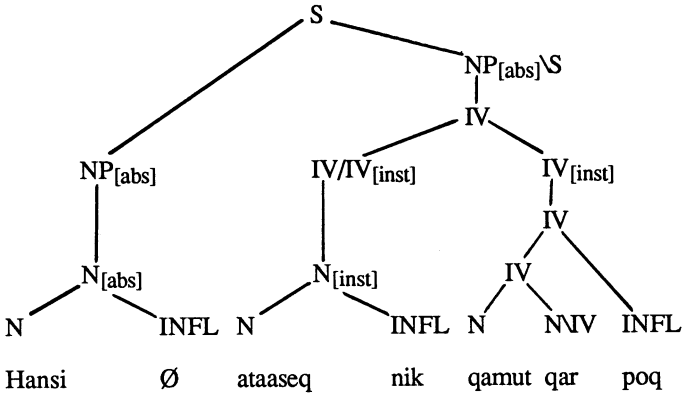


The syntactic tree (47') is the result of a bottom up parsing strategy along the following lines: every (complex) word is checked whether it appears in the lexicon. If so, the lexical semantic and syntactic information is brought into the parse. If not, the word is parsed on line. For that purpose there has to be a separate stem-and-affixes lexicon available that delivers the lexical semantic and combinatorial properties of its entries.

One last problem that I want to comment on is the treatment of external modifiers in Eskimo. I treat them as VP modifiers, but simultaneously I want to keep track of the fact that semantically they modify an IN. Following Gazdar (1982) I use categories labeled with features. The case feature [inst] on IV indicates the fact that every IV assigns to its nonincorporated "arguments" the instrumental case. Since *gar* becomes an IV after it has combined with the N *qamut* this feature shows up only in the phrasal domain of the tree.

(48) Hansi ataatsinik qamuteqarpoq.
Hansi-ø ataaseq-nik qamut-qar-poq
Hans-(ABS) one-INST/pl sled-have-INDIC/3s
'Hans has one sled'

(48')



The semantic contribution of *ataaseq-* manifests its status of a nominal modifier. It introduces a DRS condition variable *C* that has to be set equal to the condition introduced by the IN; it also introduces a set variable *Y* of which *C* holds (|Y| = 1) that has to be set equal to the set discourse referent introduced by the IN. All of this can be read of from DRS (49).

(49) $\exists X \exists Y \exists C [C = \text{sled}^* \& \text{sled}^*(X) \& C(Y) \& |Y| = 1 \& Y = X \& \text{have}(h,X)]$

4 Concluding remarks

In this paper I have argued that INs in some polysynthetic languages, among which is Greenlandic Eskimo, are referential expressions that get quantificational force from an external source. Moreover, from the perspective of DRS construction, incorporated referents can be made accessible in a systematic way only if the syntactic/semantic parse of complex words runs simultaneously with the syntactic/semantic parse of phrases. The hypothesis that INs are referential expressions is based on few examples and it needs a lot more further investigation to check whether INs in other but extensional contexts can still be treated as being semantically parallel to indefinites.

Another interesting point that falls out of this discussion is the observation that incorporation is part of a word-level operation that can show up both as a lexicalized and as a syntactic phenomenon. The question whether incorporation is lexical or syntactic is therefore reduced to a language specific question rather than a language universal one.

5 Footnotes

[*] The construction of this piece of discourse draws upon the accessibility of the following referents: Franz Beil, Hans Kamp, Manfred Krifka, Ede Zimmermann. All unresolved errors are my own.

[1] Because I have to save space, I will not use 'boxese' to write down Discourse Semantic Representations, but rather its predicate logic translation (cf. Kamp and Reyle (1991)). The DRSs are underspecified with respect to tense and agreement information.

[2] In Mithun (1984) we also find languages (e.g. Koryak, Huahtla Nahuatl) in which the incorporated noun itself plays the role of an anaphoric expression. I will not bring these cases into the present discussion.

[3] The compound *dog food* has at least two readings, i.e. 'food for dogs' and 'food made of dog meat'.

[4] The *-index on DRS conditions indicates that these conditions are applicable both to sets of discourse referents and to individual ones. The variable η ranges over sets and individuals.

[5] I will get to the importance of an external quantificational source for de re readings of INs further on in this paper.

[6] I borrow the term 'CN anaphor' from Nerbonne, Iida and Ladusaw (1990). These authors elaborate an analysis of the extensional cases of CN anaphora in Situation Semantics (see also Webber (1978)). In my view all cases of CN anaphora are cases of sense identity. Moreover, in what follows nominal senses and kinds will have the same semantic type.

[7] Carlson also distinguishes these two types of predicates as individual-level and stage-level predicates. *are mammals* is a special kind of individual-level predicate, in that it holds of all individuals of the kind of dogs without any exception. *are barking at me* is a stage-level predicate. I get to Carlson's distinction later on in the paper.

[8] Notice that in DRT numerals are treated as indefinite determiners and since these are not quantifiers they do not trigger the presupposition of a set that is quantified over.

[9] Given the fact that the relational noun *track* presupposes the existence of an agent who makes the track, some people don't have a problem with the use of *it* in (30). However, I would not regard that as a case of coreference between that pronoun and the incorporated *rabbit*, but rather explain it in terms of inferencing based on our world knowledge about tracks.

[10] I thank Ede Zimmermann for pointing this out to me.

[11] Notice that this view is related to an operation on n-place relations (n>1) formulated by Jacobson (1991), i.e. the z-operation, that turns every relation between n individuals into an n-place relation between an individual and n-1 properties. The need for this type shift follows from her variable-free semantic account of the Bach-Peeters sentences. Whereas I take the view that the individual object meaning of internal arguments is the "postulated" one, Jacobson takes it to be the basic one.

[12] Sadock reports the existence of incorporated quantifiers in Westgreenlandic Eskimo (Sadock (1991): 94-95). However, the apparent quantifiers in his examples are numerals and the weak "quantifier" *many*, and I consider neither of those to be true quantifiers.

[13] The distinction between these two sorts of predicates in particular has been criticized as being inadequate to explain particular generic readings (for discussion cf. Carlson (1988)). Although I am aware of the correctness of these criticisms, my point here is that any treatment of the bare plural has to make explicit the external source that it takes to be responsible for the bare plural's existential reading. For now, I take that source to be Carlson-like, stage-level predicates, since I only want to describe the *fact* that there is a distinction between the existential and the generic reading of an indefinite NP without making any real theoretical assumptions about its *source*.

[14] From the perspective of the theoretical approach in which Baker works out his syntactic analysis and in which the autonomy of syntax is a central concern, it is somehow strange that Baker takes discourse transparency, which is a semantic phenomenon, as a valid argument for his approach.

6 References

Bach, E. (1983) On the Relationship between Word-Grammar and Phrase-Grammar, in: *Natural Language and Linguistic Theory 1*, pp: 65-89.

Baker, M. (1988) *Incorporation*, The University of Chicago Press, Chicago-London.

Carlson, G.N. (1977) *Reference to Kinds in English*, PhD.Dissertation, University of Massachusetts, Amherst.

Carlson, G.N. (1988) The Semantic Composition of English Generic Sentences, in: G.Chierchia e.a. (eds) *Properties, Types and Meaning Vol II: Semantic Issues*, Kluwer, Dordrecht, pp. 167-193.

Di Sciullo, A.-M. and E.Williams (1988) *On the Definition of Word*, MIT Press Cambridge MA.

Dowty, D. (1979) *Word Meaning and Montague Grammar*, Reidel, Dordrecht.

Heim, I. (1982) *The Semantics of Definite and Indefinite Noun Phrases*, PhD.Dissertation, University of Massachusetts, Amherst.

Hoeksema, J. (1985) *Categorial Morphology*, Garland, New York.

Jacobson, P. (1991) Raising without Movement, to appear in R.Larson e.a. (eds.) *Control and Grammatical Theory*, Kluwer, Dordrecht.

Kamp, H. (1981) A Theory of Truth and Semantic Representation, in: J.Groenendijk e.a. (eds.) *Formal Methods in the Study of Language*, Amsterdam: Mathematisch Centrum, pp: 277-322.

Kamp, H. and U.Reyle (1991, to appear) *From Discourse to Logic*, Reidel, Dordrecht.

Mithun, M. (1984) The Evolution of Noun Incorporation, in *Language 60*, pp.847-894.

Nerbonne, J., M.Iida and W.Ladusaw (1990) Semantics of Common Noun Phrase Anaphora, in: A.Halpern (ed.) *Proceedings of the nineth West Coast Conference on Formal Linguistics*, pp.379-394.

Postal, P. (1969) Anaphoric Islands, in: Papers of the Chicago Linguistic Society 5, pp.205-239.

Rosen, S.T. (1989) Two Types of Noun Incorporation: a Lexical Analysis, in: *Language 65*, pp. 294-317.

Sadock, J. (1980) Noun Incorporation in Greenlandic, A Case of Syntactic Wordformation, in: *Language 56*, pp. 300-319.

Sadock, J. (1985) Autolexical Syntax: a Theory of Noun Incorporation and Similar Phenomena, in: NLLT 3, pp. 379-440.

Sadock, J. (1986) Some Notes on Noun Incorporation, in: *Language 62*, pp. 19-31.

Sadock, J. (1991) *Autolexical Syntax*, The University of Chicago Press, Chicago.

Sproat, R. and G.Ward (1987) Pragmatic Considerations in Anaphoric Island Phenomena, in: *Papers of the 23d regional meeting of the Chicago Linguistic Society 23*, pp. 321-335.

Ward, G., R.Sproat and G.McKoon (1991) A Pragmatic Analysis of so-called Anaphoric Islands, in: *Language 67*, pp. 439-473.

Webber, B. (1978) *A formal Approach to Discourse Anaphora*, PhD. Dissertation, BBN Report No. 3761.

Zimmermann, E. (1991) Do we bear attitudes towards quantifiers?, Paper presented at SALT 1, Cornell University, manuscript Universität Stuttgart.