SEMANTICS

Critical concepts in linguistics

Edited by Javier Gutiérrez-Rexach

Volume VI

Discourse and dynamics

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2003
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Routledge is an imprint of the Taylor & Francis Group

Typeset in Times by RefineCatch Limited, Bungay, Suffolk
Printed and bound in Great Britain by
MPG Books, Bodmin

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record for this book has been requested

ISBN 0-415-26632-7 (Set)
ISBN 0-415-26638-6 (Volume VI)

Publisher's note
References within each chapter are as they appear in the original complete work

CONTENTS

VOLUME VI DISCOURSE AND DYNAMICS

3 Definite descriptions governed by salience are discussed in my *Counterfactuals* (Blackwell, 1973), pp. 111–117: and in James McCawley, 'Presupposition and Discourse Structure', in David Denneen and Choon-Kyu Oh, eds., *Syntax and Semantics*, Vol. 11 (Academic Press, 1979). A similar treatment of demonstratives is found in Isard, *op. cit.*

Manfred Pinkal, 'How to Refer with Vague Descriptions' (presented at the Konstanz colloquium on semantics, September 1978), notes a further complication: if some highly salient things are borderline cases of *F*-hood, degree of *F*-hood and salience may trade off.

Indefinite descriptions that pave the way for referring expressions are discussed in Charles Chastain, 'Reference and Context', *Minnesota Studies in the Philosophy of Science* 7 (1975), 194–269, and in Saul Kripke, 'Speaker's Reference and Semantic Reference', *Midwest Studies in Philosophy* **2** (1977), 255–276. [Reprinted in this volume, Chapter 5]

4 See Charles Fillmore, 'How to Know Whether You're Coming or Going', in Karl Hyldgaard-Jensen, ed., *Linguistik 1971* (Athenäum-Verlag, 1972), and 'Pragmatics and the Description of Discourse', in Siegfried J. Schmidt, ed., *Pragmatik/ Pragmatics II* (Wilhelm Fink Verlag, 1976).

5 See the treatment of vagueness in my 'General Semantics', *Synthese* **22** (1970), 18–67. For arguments that hardly anything is flat or certain, see Peter Unger, *Ignorance* (Oxford University Press, 1975), pp. 65–68. For another example of accommodating shifts in resolution of vagueness, see the discussion of back-tracking counterfactuals in my 'Counterfactual Dependence and Time's Arrow', *Nous* **13** (1979).

6 See Angelika Kratzer, 'What "Must" and "Can" Must and Can Mean', *Linguistics and Philosophy* **1** (1977), 337–355. The accessibility semantics considered here is equivalent to a slightly restricted form of Kratzer's semantics for relative modality.

Knowledge and irrelevant possibilities of error are discussed in Alvin I. Goldman, 'Discrimination and Perceptual Knowledge', *Journal of Philosophy* **73** (1976), 771–791.

7 See J. L. Austin, 'Performative Utterances', in his *Philosophical Papers* (Oxford University Press, 1961), for the original discussion of performatives. For treatments along the lines here preferred, see E. J. Lemmon, 'On Sentences Verifiable by Their Use', *Analysis* **22** (1962), 86–89; Ingemar Hedenius, 'Performatives', *Theoria* **29** (1963), 1–22; and Lennart Aqvist, *Performatives and Verifiability by the Use of Language* (Filosofiska Studier, Uppsala University, 1972). Isard (*op. cit.*) suggests as I do that performative utterances are akin to other utterances that 'change the context'.

98

A THEORY OF INFORMATICS

Enric Vallduvi

Source: *The Informational Component*, PhD diss., University of Pennsylvania, 1990; Ann Arbor, MI: Garland Publishing Co., 1992, pp. 66–94.

In § 2.2 information packaging was characterized as a non-logico-semantic type of sentence 'meaning' concerned with the retrieval of information and its entry into the hearer's knowledge-store. Information packaging was defined as in (57) (=(11)):

(57) INFORMATION PACKAGING: A small set of instructions with which the hearer is instructed by the speaker to retrieve the information carried by the sentence and enter it into her/his knowledge-store.

where information is defined as that part of the propositional content which constitutes a contribution of knowledge to the hearer's knowledge-store. These instructions are meant to optimize the update of the hearer's knowledge-store by singling out the informative part of the sentence and articulating the ground in such a way as to indicate how this information fits the hearer's knowledge-store.

It was already pointed out in Ch. 2 that information packaging is sufficiently distinct from other types of pragmatic understanding to grant it autonomous status and, despite its close coexistence with logico-semantic meaning within the interpretive end of language, they also remain independent.[48] Therefore, it was suggested that the generation and interpretation of information packaging must be dealt with in an autonomous module, i.e. INFORMATICS.

This chapter provides an account of the exact nature of the role of informatics in the larger language apparatus. If building a coherent and comprehensive theory of informatics independent of these other interpretive modules proves feasible, it constitutes indirect validation of the autonomy hypothesis assumed. This account is based on the informational primitives discussed in § 3.2.2. It describes the means by which this primitives interact

to yield the information-packaging intructions and the exact interpretation of these instructions.

4.1 The role of informatics

4.1.1 The knowledge-store

The definition of information-packaging presented in Ch. 2 presupposes the existence of a hearer's knowledge-store with a given structure. In order to be able to describe the role of information-packaging in language understanding, it will be necessary first to discuss how the hearer's knowledge-store might be structured. To this purpose, the file metaphor in Heim's File Change Semantics (Heim 1983) will be adapted.

Heim views discourse referents as file cards in a file. Before the beginning of a discourse the hearer has a file with zero file cards (F_0). As the discourse progresses file cards are added and updated. For instance, after the utterance of (58), the hearer puts the cards (59) in her/his file (ignoring the speaker's own file card):

(58) Pat told me a weird story today.

(59)

addr.: pat	addr.: story(z)
told z to speaker(pat)	weird(z) told by pat(z)

In Heim's terms, the hearer has gone from an F_0 to a given F_1. As the discourse continues, every utterance will cause a change of files from F_n to F_{n+1}. Given the following three-utterance continuation to the discourse started with (58), the hearer's F_4 will contain the file cards and the entries in (61) (ignoring the file card for 'a story'):

(60) She saw this man with a broccoli stalk.
Well, the guy starts munching on it,
and, lo and behold, he's arrested right away.

(61)

addr.: pat	addr.: man(x)	addr.: br. stalk(y)
told z to speaker(pat) saw x(pat)	seen by pat(x) has $y(x)$ munch on $y(x)$ arrested(x)	had by $x(y)$ munched on by $x(y)$

File F_4 has three cards which list all the attributes and relations between them specified by the propositional content of the sentences in (58) and (60). The file change from F_1 to F_4 involved the update of Pat's file card and the addition of cards for the man and the broccoli stalk and their update.

Heim's file, as pointed out in § 2.3.3, is akin to the notion of discourse model viewed as a mental representation of the entities involved in a discourse and their attributes and the links between them. For our purposes here, however, Heim's file metaphor will be applied to the hearer's knowledge-store. The knowledge-store is taken to be a large file with a number of file cards or ADDRESSES. Each address denotes an entity and under each address there are a number of entries specifying attributes and relations pertaining to that entity. Unlike in Heim's, there is no file F_0 before the beginning of a discourse, since the hearer's knowledge is not null at the start of an interaction, i.e., at the start of a discourse the knowledge-store contains addresses denoting hearer-old discourse-new ('unused') entities. So, for instance, in (58) the hearer would not add a file card 'Pat' because it would already be there from previous shared knowledge.

How is the knowledge-store modified and updated? The referential status of entities (cf. 2.3.2) plays a crucial role in this process. According to Heim and others, an indefinite NP will cause the hearer to start a new file card or create a new address and a definite NP will indicate that an already-existing address must be activated. In example (60) above, the indefinites *this man* and *a broccoli stalk* instruct the hearer to create a new address and the pronominal forms *she*, *he*, and *it* indicate that the address they denote has been already created.[49] The relations and attributes that make up the propositional content of the utterances in (60) are then entered under each one of the addresses involved in those utterances. In (61), the address for 'man' contains the knowledge that he was munching on a broccoli stalk and the address for 'broccoli stalk' contains the knowledge that it was being munched on by the man. This reflects the fact that, after the discourse in (60), one knows *a*) about the broccoli stalk that it was munched on, and *b*) about the man that he was munching on a broccoli stalk.

The role of referential status marking in language understanding, then, is the management of addresses in the hearer's knowledge-store: creation and activation of addresses or file cards. This is not only the view in Heim 1983 and Prince 1988b, for instance, but it agrees with the observation made repeatedly in the literature (cf. § 2.3.2 above) that referential status is a property of NPs and/or referents independent of the sentential context in which they occur.

This schematic view of the knowledge-store will be enough to allow a description of the role of information packaging. Information packaging is concerned precisely with the update of the knowledge-store only with respect to the entry of information. In other words, while referential status marking is responsible for the creation of new addresses or activation of existing

addresses, information packaging is responsible for the actual update of these addresses. It indicates what part of the utterance constitutes information and, furthermore, it shows where this information goes and how it fits under a particular address. This will be discussed in what follows.

4.1.2 Redundancy in the entry of data

In spite of its advantages and intuitive appeal, the file metaphor has an important drawback: its inefficiency. This inefficiency is due to two characteristics of the file metaphor: *a*) it does not take into account the knowledge already existent in the file, and *b*) it requires multiple entry of the same propositional content. Let us discuss them in this order.

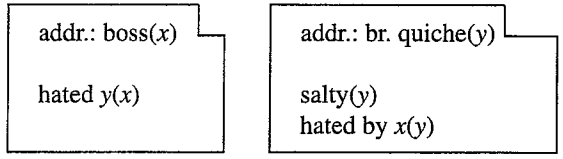
Preexistent knowledge

Suppose a speaker has just uttered sentence (62)a with the propositional content in (62)b.

(62) a. The boss hated the salty broccoli quiche.
b. hated (the salty broccoli quiche, the boss)

Referential-status marking indicates, in this case, that no addresses need be created, only that two already-existing addresses, 'the boss' and 'the salty broccoli quiche', must be activated. Then the hearer must perform her/his filing task and must enter this data into her/his knowledge-store. After this task is performed, a partial view of the hearer's knowledge-store may be represented as in (63):

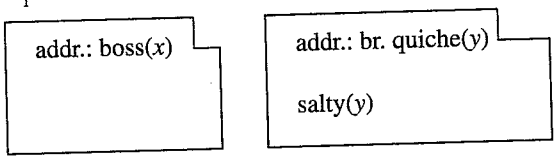
(63)



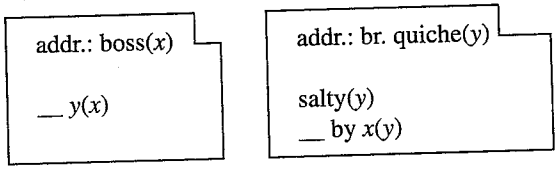
There are two addresses, one for 'the boss' and one for 'the broccoli quiche', and under each of them the relevant attributes, i.e. that the quiche was salty, and the relations that hold between the two addresses, i.e. that the boss hated the quiche and that the quiche was hated by the boss, are recorded.

Returning to the two hypothetical hearers H_1 and H_2 introduced in § 2.2, let us again assume that H_1 knows nothing about the existence of a connection between the boss and the broccoli quiche and that H_2 knows of the existence of such a connection but does not know which one it is. The following is what their knowledge-store might look like before the proposition in (62) is communicated, (64) for H_1 and (65) for H_2:

(64) H_1:



(65) H_2:



The difference in what H_1 and H_2 know is represented in the difference between the content of the file cards in (64) and (65). In (64) there is no information about the connection between the boss and broccoli, but in the case of (65), both the knowledge of the connection and the fact that the nature of this connection is unknown must be represented. This is done by means of the blank (__).

Within the file metaphor, however, there is no way to take this difference into account. Propositional content is entered in the same way independent of what the hearer may or may not already know. This process is not very efficient given that some of the entry of data in the case of H_2, for instance, is totally redundant and unnecessary. What is needed to avoid the redundant entry of data is the notion of information. A view of data entry in terms of propositional content does not allow us to take into account the distinction between H_1's knowledge-store and H_2's knowledge-store. Information does, since it is defined precisely with respect to the hearer's knowledge-store. As noted above, it is information that is responsible for the different structural encodings of the proposition in (62) when addressing H_1, shown in (66)a, or when addressing H_2, shown in (66)b:

(66) a. The boss hated [$_F$ the salty broccoli QUICHE].
b. The boss HATED the salty broccoli quiche.

Speakers are sensitive to the different make-ups of different hearers' knowledge-stores and package the proposition they want to communicate in different packaging instructions accordingly. Even though the propositional content of (66)a-b is the same, the information they carry is different. The marking of information allows for a more efficient process of data entry, since the hearer need not re-record knowledge s/he already has.

Multiple entry

In the discussion around the discourse in (60) it was pointed out that in Heim's approach—and indeed in any other discourse model account—the propositional content encoded in a given sentence is recorded on every file card whose referent is evoked in the discourse. Under the address for 'man', for instance, it was entered that he was munching on a broccoli stalk, and under 'broccoli stalk' it was entered that it was being munched on by a man. This means that the *same* knowledge is entered into the hearer's knowledge-store twice or more depending on how many participants are involved in a given utterance. The updating of the hearer's knowledge-store after sentence (67),

(67) The boss gave Mary a broccoli stalk.

would require that the same propositional content be entered three times, one for each address involved in the sentence.[50] As noted, this seems to be a necessary process to guarantee that all the knowledge one has about a given entity is available when evoking that entity.

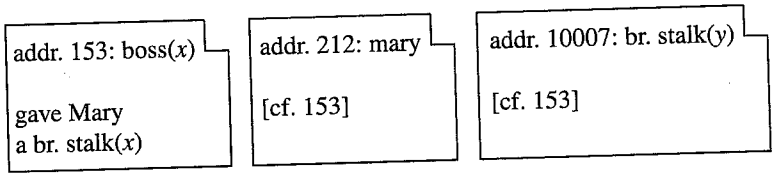
There is, however, another way to achieve the same results. The traditional file metaphor is based on a now 'old-fashioned' filing system. Let us consider the following example: in 1950 Smith's Body Shop from Anytown, a new customer, purchased 25 alternators from Jones Auto Parts. Mrs. Jones kept a customer file, where she kept track of their purchases, a stock file, where she kept track of the inventory, and a third file containing file cards for each city in her business area, where she listed all customers from that city. After Smith's purchase Mrs. Jones had to update three file cards. She had to create a file card for Smith's Body Shop and enter that he had purchased 25 alternators, update the 'alternator' card by registering Smith's purchase and subtracting 25 alternators from the stock, and update the 'Anytown' card by adding Smith's Body Shop to the list of customers in Anytown. This process is parallel to the process assumed to take place in the entry of data into the hearer's knowledge-store from the perspective of a traditional file metaphor.

Suppose now the same purchase takes place in 1990. Mrs. Jones abandoned her files and acquired a computer where she still keeps track of the same data. After Smith's purchase, Mrs. Jones enters the following data into her database: Smith's Body Shop from Anytown, 25 alternators. She only enters the data once, probably under the 'Smith's Body Shop' card, but if she wants to check her stock, she can call up the 'alternator' card and will find that the sale of 25 alternators to Smith's Body Shop has been registered there too, and if she wants to see which customers she has in Anytown, she may call up the 'Anytown' card and Smith's Body Shop will be there.

This, of course, is how any current database or hypercard program works. It is clearly less costly and it achieves the same results. This hypercard

efficiency can be applied to the file-structured knowledge-store quite straightforwardly. Let us consider sentence (67) again. The updating of the hearer's knowledge-store need not require that the data be entered three times. It may be entered only under, let us say, 'the boss' and is then made available somehow to the other two address corresponding to the other participants in the sentence. This availability may be achieved in two ways. The data may be entered under the other addresses at a later stage, or the other addresses may get a crossreference index corresponding to the address 'John' through which the relevant knowledge stored under 'John' becomes accessible. The second option is chosen for expository purposes, but the implications of each option are left unexplored:

(68)



The information-packaging articulation of the sentence is also responsible for the entry of data in this hypercard fashion, thus avoiding the redundancy that would otherwise arise. As noted, a packaging instruction does not merely signal what the information of the sentence is, it further specifies *how* this information contributes to the hearer's knowledge. This is the role played by the ground, with specific subtasks for both the link and the tail. First, given a certain information, where in her/his knowledge-store must the hearer enter it? Taking the discussion above into account, it is proposed that the particular task of the link as an address-pointer is to 'point to' the address in the hearer's knowledge-store under which information is entered. In example (67) *the boss* is the link. This means that the address 'the boss' in the hearer's knowledge-store is designated as the address under which the oncoming information of the sentence must be entered. This avoids the need for multiple entry under several addresses, thus improving in efficiency the necessary task of knowledge-store update.

Second, there are cases where knowledge communicated to the hearer in a given proposition partially overlaps with knowledge already present in the hearer's knowledge-store. Without the notion of information, the entry of propositional content becomes redundant in this respect as well. The ground, and specifically the tail, allows the hearer to relate the information in the sentence with the relevant knowledge already present in her/his knowledge-store in the way to be detailed below. This potential second source of inefficiency is avoided as well, thanks to information packaging.

4.1.3 Information packaging as data entry

Information packaging must be viewed then as being responsible for the entry of data into the hearer's knowledge-store. If propositional content were entered into the hearer's knowledge-store without any further articulation, as in the traditional file metaphor, the result would be a very inefficient system of data-entry. The role of information packaging is to achieve an efficient and nonredundant update of the hearer's knowledge-store. By representing not only propositional content but also information, natural language improves the overall efficiency of the process of communication.

The role of informatics as a linguistic component, then, is to generate and interpret packaging instructions. Structuring sentences into information-packaging instructions is as fundamental a part of language production as encoding propositional content in sentence structure, and interpreting these instructions is as important as decoding propositional content. Information is superimposed on or represented in parallel with propositional content. The interpretation of both types of 'meaning' must remain independent. The propositional content of a sentence represents knowledge that speakers have and wish to communicate to hearers. This knowledge speakers have is, in principle, definable independently of the hearer that will benefit from receiving it. Information is a reflection of the speaker's assumptions about the hearer's knowledge-store. When speakers wish to communicate a proposition they take into account how much of the knowledge represented in that proposition will actually contribute something to the hearer's knowledge-store. Information packaging is crucially affected by the linguistic context by virtue of this dependence on the speaker's assumptions about the hearer's knowledge and attentional state, but propositional content is constent across speaker–hearer interactions.

Information is a notion relevant for the *entry* of data into the hearer's knowledge-store What is *entered* in the knowledge-store is information; what *is* in the knowledge-store is knowledge. In other words, the knowledge-store contains knowledge but receives information. The speaker's knowledge must be 'squeezed' into information in order to be transferred to the hearer's knowledge-store. Squeezing propositions into instructions increases the efficiency of the transfer by avoiding all redundancy in the process of data entry into the hearer's knowledge-store. Informatics, then, is not concerned with the interpretation or the representation of meaning in the sentential structure, but with the interpretation and representation of information for the purposes of data entry into the knowledge-store.

Having stated what the exact role of informatics is within the larger task of language production and understanding, let us now proceed to a description of the manner in which this role is carried out by discussing the elements of the theory.

4.2 Elements of the theory

4.2.1 Primitives

The primitives of the theory and their hierarchical configuration were introduced and discussed in § 3.2.2. It is the trinomial hierarchical articulation in (69) (=(44)):

(69) S={FOCUS, GROUND}
GROUND={LINK, TAIL}

This articulation is based on the insights of the proposals previously made in the literature. In particular, it reflects both the focus-ground split (i.e. focus-presupposition, focus/open-proposition, or focus-background) and the fact that within the ground there generally is a 'special' topiclike element, the link, which appears in sentence-initial position (analogous, with the provisos noted in Ch. 3, to the topic in the topic-comment framework or the theme in the theme-rheme approach).

These primitives and their arrangement in a given sentence make up the information-packaging instruction represented by that sentence. The INSTRUCTION is the central object in the informatics, much in the same way the proposition is the central object in the semantics. In order to represent these packaging instructions, we will need to introduce some notation. The notation used may resemble on some occasions the language used in representing logico-semantic meaning. It must be emphasized, though, that the instructions spelled out in this manner evidently do not represent logico-semantic meaning but information packaging. Any similarity between the two is, therefore, purely notational.

Before starting the discussion of the different instructions that can be construed with these atomic primitives, one notational detail must be introduced. Information packaging, as noted, is concerned with the representation of information and the directions needed to enter that information into the hearer's knowledge-store. It was also noted that every uttered sentence must provide some information since otherwise there is no raison d'être for the sentence to exist in normal communication.

Let us represent the (variable amount of) information that all sentences must provide as in (70):

(70) Φ [*information*]

The symbol Φ (capital phi) stands for an informational one-place operator. It will be called 'focus operator'. Everything within its scope (the clause) is informative, i.e., the scope of Φ constitutes all the information provided by the sentence. Given that packaging instructions are a speaker-designed

information-retrieval mechanism for the hearer, the operator Φ may be read procedurally as 'retrieve', and the instruction in (70) as 'retrieve *information*', whatever the value of this information may be.

The structure in (70) is the simplest of information-packaging instructions, where an entire sentence is information, corresponding to the all-focus sentences seen in § 3.2.3. Many times, however, not all of the sentence is information: a ground is licensed to indicate how the information must be entered into the hearer's knowledge-store. Sentences with a ground correspond to the link-focus, link-focus-tail, and focus-tail structures discussed in § 3.2.3. Let us now turn to the structure of these more complex instructions.

4.2.2 Packaging instructions and interpretation

The information carried by the sentence is encoded by the focus constituent, even though its informational force is a relational property crucially dependent on its relation with the ground. This means that the focus will always be within the scope of Φ, but, as will be seen in a moment, the focus, strictly speaking, is not *all* there can be within the scope of Φ. The ground, in contrast, is not information, its only informational force being to permit the appropriate entry of information into the hearer's knowledge-store. Therefore, it cannot appear within the scope of Φ.

Let us start by looking at the informational value of links. Links, as part of the ground, play the anchoring role in the instruction. It was noted in § 3.2.2 that the link was an address pointer of sorts. In the previous section, it was concluded that links point toward the address in the hearer's knowledge-store under which the information of the sentence must be entered. It instructs hearers to 'go to' the address it denotes in their knowledge-store and then enter the information provided by the sentence under that address. Links, therefore, are informationally interpreted as a bipartite element: the instruction 'go to' and the address hearers are instructed to go to.

This interpretation may be represented with a quantifier-like element that will be called Λ (capital lambda): Λx is read 'go to x' and the address denoted by the link constitutes the range of x. The representation of the link is, then, as in (71)a (for an address α), while (71)b is the representation of the link expression *the boss* in the link-focus sentence *The boss* CALLED, which is read 'go to address "the boss" ':

(71) a. Λx, $x = \alpha$.
b. Λx, x = the boss.

The fact that the information of the sentence is to be entered under the address denoted by the link is represented by the latter taking scope over this information. The informational task of the link—pointing to a given address—is independent of the referential status of the NP that acts as link.

A definite NP instructs the hearer to activate a preexisting address whether the NP is a link or not. But only if it is a link will the hearer be instructed to go to that address and enter information under it. Links tend to be definite NPs because, as a default, the speaker cannot point to an address in the hearer's knowledge-store if that address is not there already. It is not impossible, however, to instruct the hearer to create an address and to point to that address simultaneously. This is the case in sentences where the link is an indefinite NP. The exact way in which the link takes scope over the information of the sentence will be discussed further after the notation for focus is introduced.

The focus, which must be entirely within the scope of Φ, will be represented by simply writing it in boldface within that scope. The actual formalization of an all-focus sentence, therefore, will not be as in (70) above, but as in (72)a. A particular instantiation of this instruction, for the all-focus sentence *[$_F$ The* BOSS *called]* is (72)b:

(72) a. Φ[focus]
b. Φ[the boss called]

Taking this into account, a link-focus sentence may be represented in the following manner. The link's relation to the information of the sentence is represented through a quantifier-variable structure, i.e. the quantifier Λ binds a variable in the clause, as shown in (73)a. In (73)b, this abstract instruction is illustrated with the particular instruction for the link-focus sentence *The boss* CALLED:

(73) a. Λx, $x = \alpha$ [Φ[x **focus**]]
b. Λx, x = the boss [Φ[x **called**]]

The instruction in (73) is the instruction encoded in all link-focus sentences. It combines the interpretation of the link with the interpretation of the focus operator. From a speaker's point of view, it may be procedurally read as follows: 'I instruct you to go to address α in your knowledge-store and then retrieve the information of the sentence by adding focus under *α*.' Or, in the case of (73)b, 'I instruct you to go to the address "the boss" in your knowledge-store and then retrieve the information of the sentence by adding under "the boss" that he called.' It is clear now why it was noted that the focus is not *all* that stays within the scope of Φ; the variable bound by the link-operator is also found within that scope. As will be seen below, the fact that this variable bound by the link remains within the scope of Φ is actually quite important for the interpretation of information packaging.

Finally, the notation for the tail must be introduced. The tail, being part of the ground, must escape the scope of Φ. This may be represented by abstracting the tail away from the clause, leaving a variable behind within the scope

of Φ. This 'abstraction' will be represented by means of a lambda-like construct. The *graph* λ will be borrowed from the lambda-calculus used in logical semantics. Again, it must be pointed out that all that is borrowed is the symbol, but none of the semantics of the lambda-calculus. The notation for tails is illustrated in (74) for a given tail *β*:

(74) $\lambda x[\Phi[\textbf{focus}\ x]]\ (\beta)$

The tail's task, as noted above, is to further specify how the information must be entered under a given address. If the ground contains a tail it means that the information of the sentence cannot be simply *added* under the address denoted by the link. Instead, it indicates that part of the proposition communicated is knowledge already contained under that address and that the information of the sentence must be construed in some way with that knowledge instead of merely added.

But how exactly does the presence of a tail affect the nature of the packaging instruction? The operator Φ instructs the hearer to retrieve the information contained in its scope, but in the discussion on the representation of link-focus sentences, the operator Φ was read as 'retrieve information by adding focus'. This is because the operator Φ actually comes in two brands: 'retrieve information by adding focus' and 'retrieve information by substituting focus for the blank in the ground "link__tail" (which is already under the address denoted by the link)'. These may be abbreviated as 'retrieve-add' and 'retrieve-substitute'. What the presence of the tail does in a given packaging instruction is alter the nature of Φ, turning it from a 'retrieve-add' to a 'retrieve-substitute'.

Let us see this by comparing the link-focus sentence in (75) with the link-focus-tail one in (76). The (b) examples are the representations of the packaging instruction they encode (instruction (76) represents both the link-operator and the tail in one instruction):

(75) a. The boss [$_F$ hates BROCCOLI].
b. $\Lambda x_1, x_1$ = the boss $[\Phi[x_1\ \textbf{hates broccoli}]]$
(76) a. The boss HATES broccoli.
b. $\Lambda x_1, x_1$ = the boss $[\lambda x_2[\Phi[x_1\ \textbf{hates}\ x_2]]$ (broccoli)

Both (75) and (76) have the same link. The representation of the link is, therefore, identical in both sentences, and can be read as 'I instruct you to go to the address "the boss" and then . . .' Once the address denoted by the link is 'gone to', the hearer is in a position to retrieve the information provided by the sentence, at least in the case of (75). In the case of (76), the hearer needs to know more about the way in which the information must be entered.

This difference is reflected in the distinction between 'retrieve-add' and 'retrieve-substitute'. In (75), the instruction continues '. . . and then retrieve

the information of the sentence by adding under "the boss" that he hates broccoli'. In the tailful (76), the instruction continues '. . . and then retrieve the information of the sentence by substituting *hates* for the blank in *he __ broccoli*, which is already under "the boss"'. In other words, in (75) the information 'hates broccoli' is merely added under 'the boss', while in (76) the information 'hates' is taken to fill the gap in the knowledge already existent under 'the boss', since under this address there is already an entry for '__ broccoli'. This filling of the gap is represented by the tail's triggering of the retrieve-substitute brand of Φ. The tail, then, in the fashion discussed above, prevents the hearer from redundantly treating part of the proposition communicated by the hearer as information. A tailful instruction directs the hearer to some entry under a given address and indicates that the focus completes or alters in some way that entry.

The previous discussion shows how informational primitives interact and combine to form packaging instructions. The number of possible packaging instructions totals four, corresponding to the four informational articulations discussed in § 3.2.3. The following is a list of the four possible informational articulations of a sentence and the packaging instructions associated with them. Their interpretation is discussed in detail immediately below:

1. Link-focus: $\Lambda x_1, x_1 = a[\Phi[x_1\ \textbf{focus}]]$
2. All-focus: $\Phi[\textbf{focus}]$
3. Link-focus-tail: $\Lambda x_1, x_1 = a[\lambda x_2[\Phi[x_1\ \textbf{focus}\ x_2]]\ (\beta)$
4. Focus-tail: $\lambda x_2[\Phi[\textbf{focus}\ x_2]\ (\beta)$

The basic constant informational operator is the one-place Φ, which takes scope over the clause. In the informatics, every sentence is interpreted as having the minimal structure 'Φ[*information*]', i.e., it is interpreted as 'retrieve the information carried by the sentence'. This is a mere representation of the fact that all sentences must carry information. More complex instructions are derived from this basic skeleton. Their purpose, as noted above, is to assure a nonredundant entry of information into the hearer's knowledge-store.

All-focus sentences represent the simplest of instruction, their structure being identical to the basic skeleton 'Φ[*information*]'. The information in the sentence is carried by the entire structure, i.e., the focus is the entire sentence. This was represented as in (77):

(77) $\Phi[\textbf{focus}]$

The following sentences are the examples of all-focus structure that were seen in § 3.2.3, and (c) represents the instruction each sentence encodes):

(78) a. [$_F$ The BOSS called].
b. [$_F$ Ha trucat l'AMO].
c. Φ[**the boss called**]

(79) a. Waiter! [$_F$ There's a fly in my cream of broccoli soup]!
b. Cambrer! [$_F$ Tinc una mosca a la crema de bròquil]!
c. Φ[**(there's) a fly in my cream of broccoli soup**]

(80) a. What doesn't the boss like?
BROCCOLI.
b. Què no li agrada, a l'amo?
el BRÒQUIL.
c. Φ[**broccoli**]

Their interpretation is as follows. Sentence (78), already seen in (72), is interpreted in the following terms (from a hearer's point of view): 'I am instructed to retrieve the information of the sentence by adding to my knowledge-store that the boss called'. Similarly, in (79), the waiter informationally interprets the sentence as 'I am instructed to retrieve the information of the sentence by adding to my knowledge-store that there is a fly in the speaker's soup'. Basically, what this instruction says is that the entire propositional content is a contribution of knowledge to the knowledge-store, i.e., that the information of the sentence is equivalent to its propositional content. Sentence (80) is of a different kind and will be discussed in a moment.

The linkless sentences just discussed are peculiar in that no particular address in the hearer's knowledge-store is specified for the subsequent entry of information. These all-focus sentences have been described as portraying a state of affairs or event (cf. e.g. Sasse 1987), i.e., contrary to link-focus sentences, all-focus sentences are not 'about' a specific entity. This intuition is captured by having the information of the sentence be entered under a temporary or situation address. The contents of this situation address are not meant to be permanent the way the content of regular addresses is. Before the contents of the situation address are deleted they are transferred to the addresses of the entities that participate in the discourse by means of the mechanisms described in the previous sections.

The link indicates that information must be entered under the address denoted by the link. Links are interpreted as 'go to address x (the range of x being determined by the denotation of the link phrase) and under x Φ'. The following are the interpretations for some of the sentences in §3.2.3 (the example *The boss* CALLED has been discussed already):

(81) a. The boss [$_F$ visited a broccoli plantation in COLOMBIA].
b. L'amo [$_F$ va visitar una plantació de bròquil a COLÒMBIA].
c. Λx_1, x_1 = the boss [Φ[x_1 **visited a broccoli plantation in Col.**]]

(82) a. The boss$_1$ [$_F$ I wouldn't BOTHER t_1].

b. L'amo$_1$ [$_F$ no l'EMPRENYARIA t_1].
c. Λx_1, x_1 = the boss [Φ[**I wouldn't bother** x_1]]

A hearer interprets the instruction encoded in (81) as 'I am instructed to go to the address "the boss" and then retrieve the information of the sentence by adding under "the boss" that he visited a broccoli plantation in Colombia', and the one encoded in (82) as 'I am instructed to go to the address "the boss" and then retrieve the information of the sentence by adding under "the boss" that the speaker wouldn't bother him'.[51]

The discussion around example (76) showed already how tailful sentences are interpreted, but they will be reviewed here as well for the sake of completion. While the presence of a link indicates that information must be entered under a given address, the presence of a tail indicates that, under the address denoted by the link, the information fills a gap in some partial entry instead of being a mere addition. This property of tails was captured by saying that it alters the nature of Φ from a 'retrieve-add' to a 'retrieve-substitute'. Let us illustrate this with the following examples:

(83) a. The boss HATES broccoli.
b. L'amo l'ODIA, el bròquil.
c. Λx_1, x_1 = the boss [λx_2[Φ[x_1 **hates** x_2]] (*broccoli*)

(84) a. I can't believe this! The boss is going crazy!
BROCCOLI, he wants now.
b. No m'ho hagués cregut mai! L'amo està ben boig!
BRÒQUIL, vol ara.
c. λx_1[Φ[he x_1 **broccoli**] (wants)

Sentence (83), as mentioned, is interpreted as 'I am instructed to go to the address "the boss" and then retrieve the information of the sentence by substituting *hates* for the blank in *the boss__broccoli* which is already under "the boss"'. 'Hates' is not merely added, but substituted for the gap in the entry '__ broccoli(x)' under boss(x). This is what distinguishes (83) from its corresponding tailless sentence.

Example (84), in contrast, has no link, since it is assumed that the hearer is already at the address under which information must be entered. As noted, links are pointers. They are only necessary if the hearer needs to *go* to a given address to enter the information of the sentence under that address. At the time a sentence S_n is uttered, the hearer is located at a given address a, under which s/he was entering the information of sentence S_{n-1}. If S_n is a linkful sentence, the hearer is instructed to move to another address b before proceeding to enter the information carried by S_n. However, if the information in S_n is to be entered under a, there is no need to instruct the hearer to move to a different address, since s/he is currently at a already. Therefore, the presence of a link is unnecessary in S_n, which will be a linkless sentence.

Therefore, the hearer would interpret the instruction of sentence (84) as 'I am instructed to retrieve the information of the sentence by substituting broccoli for the blank in *he wants*__ under the current address.' Example (80), the all-focus sentence fragment seen above, is parallel to (84) in that a link is rendered unnecessary. But, in addition, (80) requires no tail either, since (the speaker assumes) the hearer needs no ground at all to enter the information of the sentence appropriately. The instruction encoded in (80) is then 'I am instructed to retrieve the information of the sentence (fragment) by adding **broccoli** under the current address'.[52]

The instruction interpretations discussed in this section have been spelled out in a rather cumbersome way. For convenience, the following shorthand notation will also be used:

1. All-focus (Φ_A): RETRIEVE-ADD(**focus**)
2. Link-focus (Λ, Φ_A): GO-TO(link), RETRIEVE-ADD(**focus**)
3. Link-focus-tail (Λ, Φ_S): GO-TO(link), RETRIEVE-SUBSTITUTE(**focus**)
4. Focus-tail (Φ_S): RETRIEVE-SUBSTITUTE(**focus**)

The account of informatics presented in this chapter explains how from the informational articulation of the sentence, encoded by syntactic and prosodic means, one may derive the packaging instructions that indicate to hearers what the information of the sentence is. The four instructions proposed are derived in a systematic way from the informational primitives of the sentence. Moreover, these four instructions seem to cover most, if not all, the informational articulations described in the literature. In the remainder of this chapter, some further features of this theory of informatics will be discussed.

4.3 Features of the theory

4.3.1 Motivation

The account of informatics presented in the previous section is empirically based on the informational primitives identified in the literature as revised above in § 3.2. It is further validated by its wide coverage, attained with only a very small set of instructions. Nevertheless, there are other motivations, both empirical and conceptual, for the particular representations chosen. These further motivations are discussed in this section.

One of the features of the theory is the representation of links as informational quantifiers in quantifier-variable structures. There is an additional clear empirical motivation for viewing links as quantifier-variable structures. Links, given their informational task, are inherently sentence-initial. Now, if the link phrase is a complement or an adjunct—or a subject in a VS language—it must move from its postverbal thematic position to the

sentence-initial slot leaving a gap behind. The result is clearly a surface configuration where the link c-commands the clause and binds its trace. This syntactic configuration, is matched, in the informatics, by a quantifier-variable structure where the link-operator takes scope over the clause and binds a variable in it.

A putative exception to this syntactic configuration is the case of the subject in languages with basic SV order. Subjects in English, for instance, tend to be interpreted as links, given their default existential force (cf. Horn 1989), but there is no movement and trace-binding structure to represent it. In the last few years, however, a number of proposals for both Romance and Germanic languages have appeared that suggest that the surface sentence-initial position of the subject is a derived one.[53] If these proposals are correct, even preverbal subjects would fit the pattern. In any event, the quantifier-variable structure in the informatics matches a generalized XP_1-t_1 structure in the syntax.

In packaging instructions, tails, like links, are removed from the scope of Φ to reflect the fact that they are part of the ground and, therefore, not informative. This characteristic of tails was formalized by abstracting the tail phrase from the scope of Φ. It is interesting that in several languages, including Catalan, French and Italian, the phrases that make up the tail are found in a derived position as well. Thus, in these languages, tail phrases are removed from the clause by means of a right-detachment, as in Catalan (85) (=(4) in Ch. 1) and Italian (86) (the (b) sentences are the corresponding canonicals):

(85) a. L'amo l_1'ODIA t_1, el bròquil$_1$.
the.boss *obj* 3s-hate the broccoli
'The boss HATES broccoli.'
b. L'amo odia el bròquil.

(86) a. Il capo li_1 ODIA t_1, i broccoli$_1$.
the boss *obj* 3s-hate the broccoli
'The boss HATES brocccoli.'
b. Il capo odia i broccoli.

The presence of the clitic object pronominals (l_1 and li_1) in the (a) sentences reveals that the tail phrases, coindexed with these clitics, are not in their base positions, since cooccurrence of clitic and argument is otherwise illicit. In this respect, Catalan and Italian differ from English.

As a result of detaching both links and tails from the clause, the core clause is left, at the surface, containing only the focus of the sentence. In other words, these languages seem to reflect information packaging in a much more salient way than, for example, English, involving not only prosody but the syntax in the process. The informational representation of sentences proposed in the previous chapter reflects closely these structural

operations performed in Catalan and Italian, with respect to both links and tails.

A different kind of empirical motivation comes from the fact that with the representation proposed above, one is able to reflect the relational nature of informational properties without any need for stipulation. As pointed out in § 4.2.2, the focus is not all there is within the scope of the focus operator Φ. This is indeed the case in all-focus sentences, but as soon as there is a ground, one or more variables are found with the focus within the clause, bound by either the link-operator or the tail's λ-like operator. Let us illustrate this with the by now familiar instruction in (87):

(87) a. The boss HATES broccoli.
b. L'amo l_1'ODIA t_1, el bròquil$_1$.
c. Λx_1, x_1 = the boss [λx_2[Φ[x_1 **hates** x_2]] (*broccoli*)

The presence of these variables within the scope of Φ, i.e. within the information of the sentence, is not accidental. The clause portion in the instruction (87)c,

(88) $\phi[x_1$ **hates** $x_2]$

reflects the fact that the information to be retrieved by the hearer is not just '**hates**' in isolation but 'x_1 **hates** x_2', where the values of the variables are fixed by the operators that bind them. In other words, the information of the sentence is 'hates' but only when interpreted with respect to the ground 'the boss is in some relation with broccoli'.

This crucial relational nature of focus has been defended by many authors in the literature and is the gist of Prince's (1981, 1986) account of focus, in which the 'new information' carried by the sentence is the fact that the focus instantiates the variable in the open-proposition (see § 3.1.4 above). The representation proposed for tailful sentences adopts Prince's insight straightforwardly.

There is also an important conceptual motivation behind this representation. If the view of information packaging argued for in this study is correct, the purpose behind information-packaging instructions is to optimize the entry of information in the hearer's knowledge-store. They single out what part of the sentence makes a contribution to the hearer's knowledge-store and indicate where and how the hearer should enter that contribution. Information therefore, is central to the packaging instruction. As noted, all sentences must carry information since, otherwise, there is, in principle, no informational reason for the utterance to exist.

Most semantic and syntactic analyses of focus (cf. Chs. 6 and 7) put forward representations of focus-background structure where the focus is the element that is, in some way or another (as a quantifier-like element or a

λ-abstracted term), abstracted away from the sentence to a peripheral position. Such an approach does not reflect the core status of focus as the informational motivation of the sentence. The representation proposed here captures this conceptual point by taking the all-focus sentences as basic and have Φ take scope over the clause. The cases where there is a ground are, in some sense, derived from this basic all-focus structure by abstracting away the ground phrases, so that they can perform their task as data-entry instructions. In other words, a link and a tail exist only when they are needed to make sure that information is entered appropriately.

Taking this stance, we automatically obtain an efficient handling of the all-focus and the link-focus sentences, where most of the overt material is focal. In representations of focus that take the ground as basic and have the focus raise to a peripheral position, the incorporation of all-focus and link-focus cases is problematic.[54]

4.3.2 Some facts captured

There are a number of additional facts discussed in the pragmatic literature that the account as presented above did not incorporate directly but that are nevertheless captured in an indirect way. One is the feeling of 'aboutness' that has inspired the notion of topic. Contrary to the view of the topic-comment approach, aboutness here is not a causal correlate of linkhood, but just a consequence of the fact that the information carried by the sentence is entered under the address denoted by the link. This is why speakers intuitively feel that, in some sense, the sentence is about the link and not about the other entities involved in the sentence, whose addresses are updated only via crossreferencing. Two other facts that are also incorporated into our analysis are contrastiveness and felicity conditions in topicalization. Discussion of these two issues, which was postponed above, is undertaken in this section.

Contrast

Contrast is a discourse notion which is found pervasively in the literature. Even though it is considered a primitive in, for example, Kuno 1972 and Chafe 1976, it has no place in our informational articulation. Contrast, however, is not a unified phenomenon in that there exist two distinct types. These two types of contrast are evident in examples (89) and (90):

(89) Broccoli I like,
but pork rinds I hate.
(90) She gave a SHIRT to Harry, not a tuxedo.

There is a feeling of contrast both between the link phrases in (89) and in the

focus of (90). Chafe seems to conflate the two in his 'foci of contrast', but, as Szabolcsi (1981:518) observes, these two types of contrast are different in nature. We may label them link-contrast and focus-contrast.

Prince (1984:220), writing about felicity conditions for topicalization, argues that link-contrast is not a primitive but a derived notion. In her account of topicalization the topicalized phrase must represent an entity already evoked in the discourse, or else an entity standing in a salient set relation to another evoked discourse entity. The feeling of contrast that is obtained in most topicalizations is just an artifact of the set understanding that licenses the construction. Prince's formulation is later modified in Ward 1985 and Ward & Prince 1986, as will be seen in the next section, but they all share the view that contrastiveness is derivable as a 'side-effect' of the actual reason for topicalization.

In fact, link-contrast is also a derived notion from the perspective of aboutness. If a sentence is understood as being about a topic, then it may be understood that it is not about another topic, given the right opposition exists between both topics. In our system, link-contrast is derived in a similar way. If in (89), for instance, the hearer is told that s/he has to go to the address 'broccoli' and enter the information of the sentence there, then it must mean that another (related) address like 'pork rinds' should not be 'gone to', and that, therefore, the information must not be entered under it. Link-contrast can be derived, then, the way it was derived in previous accounts.

Ward 1985, in order to account for the contrastive or set feeling of foci (focus-contrast), proposes that foci must belong to a relevant scale. Ward argues that the variable in the open-proposition must be on a scale and that the focus that instantiates the variable represents a value on that scale. Ward's proposal is valid for standard focus/open-proposition structures, but if one expands the notion of focus to include the focus in cases of all-focus and link-focus sentences, his approach loses some appeal, since it would be hard to find a scale for, let us say, all-focus sentences. In our theory, focus-contrast is also a derived notion. As Ward points out, contrast occurs, for the most part, in the focus/open-proposition sentences, i.e. our link-focus-tail and focus-tail structures. In these tailful structures the focus operator Φ is interpreted not as RETRIEVE-ADD but as RETRIEVE-SUBSTITUTE, i.e. retrieve the information of the sentence by substituting 'focus' for the blank in the relevant ground frame. The blank in the relevant ground frame may be a real *blank*, as is in the examples discussed so far or the context in (91), where the utterer of (91)b assumes his/her hearer has the entry '__ broccoli(x)' under the boss.

(91) a. S_1 So we gave him [=the boss] this huge bouquet of broccoli for his birthday and it looked like he was very happy with it.
b. S_2 I don't get it. The boss HATES broccoli.

But the blank need not be a real blank. It may be some preexistent element that the hearer is instructed to substitute in an entry under a given address. For example, in (90), *a shirt* is meant as a substitute for *a tuxedo*, i.e., the speaker assumes that the hearer has the entry 'gave a tuxedo to Harry(x)' under the current address and indicates with his/her informational encoding that 'a tuxedo' must be removed and substituted by 'a shirt'. What is substituted here is not a real blank but the entry segment 'a tuxedo', so that after (90) is uttered the entry under the current address is not 'gave a tuxedo to harry(x)' but 'gave a shirt to Harry(x)'. It is clear that in this case there is an inherent contrast between the two argument of the focus operator 'substitute x for y', one of which is the focus. In example (90), the contrastive feeling is provided by the operation of substitution carried out from 'a tuxedo' to 'a shirt'. The blank notation, then, is used as shorthand for any element that gets substituted in the manner described, be it a real blank or a preexistent elements that needs substitution.

Felicity in topicalization

When the incompleteness of coverage in the focus-background articulation was discussed in § 3.1.4, it was mentioned that there was a sound proposal to cover the gap in the work of Prince 1981, Ward 1985, and Ward & Prince 1986. Ward & Prince 1986, for instance, argue convincingly that the correct generalization capturing what felicity conditions must be met for topicalization to be licit is as follows:

> Discourse Condition on Preposing in Topicalization:
> The entity represented by the preposed constituent must be related, via a salient partially ordered set relation [poset], to one or more entities already evoked in the discourse model. (1986:4)

This criterion seems to account for all the data at hand, but there is one objection that may be raised: that this condition on preposing is a necessary condition but not a sufficient one. In other words, there may be NPs that encode entities that are related via a poset to another entity in the discourse model which nevertheless appear in situ and not in a preposed slot. An example is (92), from Ward (1985:ex.109):

(92) Colonel Bykov had delivered to Chambers in Washington six Bokhara rugs which he directed Chambers to present as gifts from him and the Soviet Government to the members of the ring who had been most cooperative. *One of these rugs Chambers delivered to Harry Dexter White.* Another he gave to . . .

Here it is not only the preposed phrase *one of these rugs* which is in a poset relation with an entity in the previous discourse (set-subset), but also

Chambers (identity), and *Harry Dexter White* (set-subset; HDW is a member of the ring).

There is one way, however, in which Prince and Ward's insight with respect to preposing can be adapted into the approach presented here. Given the interpretation of links as denoting an address in the knowledge-store and an instruction to go to that address, Ward & Prince's Discourse Condition can be reinterpreted as a constraint in the mutual accessibility of these addresses. In other words, hearers cannot jump from one address to the other unless those two addresses are related via a poset relation. The ban on the preposing of constituents that denote addresses that fail to be in a poset relation with some already-evoked address is, then, a reflection of the fact that the address the hearer is instructed to go to is not accessible from the address s/he is at at the time of utterance. Only addresses that are in a poset relation with the current address are accessible. This observation on the poset relation condition on the mutual accessibility of addresses is, of course, made with a speculative slant. It remains for a more general theory of cognition to determine whether this observation is a valuable one or not.

4.3.3 Pronouns

It was mentioned in passing § 4.2.2 above that the role of pronouns in our information-packaging instructions needed further discussion. This section is devoted to this.

The informative part of the sentence—the scope of Φ—crucially includes the variables bound by the link and the tail, if there is a ground at all. These bound variables permit the interpretation of focus as a relational notion. The question that may arise now is whether there may be free variables as well. The answer to this question is yes: pronominal forms, for the most part, enter the informational structure of the sentence as free variables under the scope of Φ.

The pronouns that participate in the packaging instruction as free variables are the so-called 'weak pronouns' (cf. Rigau 1986). In English, weak and strong pronouns are not phonologically distinct, but in many languages, including some Germanic and Romance varieties, there exist sets of both weak and strong pronouns. Weak pronouns are always unstressed—this applies to English as well—and generally cliticize onto other sentence elements or may even be phonetically null. Strong pronouns are always stressed and have fuller phonetical shape than their weak counterparts.

Let us compare a packaging instruction with free variables ((93)a) to a packaging instruction with bound variables ((93)b):

(93) a. $\Phi[x_1$ **focus** $x_2]$
b. $\Lambda x_1,\ x_1 = a,\ \lambda x_2[\Phi[x_1$ **focus** $x_2]]\ (\beta)$

In this abstract representation it may be observed that while the instruction with the bound variables, (b), is a tripartite link-focus-tail sentence, the instruction with the free variables, (a), is a simple all-focus sentence. This means that the (b) sentence is interpreted as GO-TO plus RETRIEVE-SUBSTITUTE while (a) is merely a RETRIEVE-ADD. The following are particular instantiations of the instructions in (93):

(94) a. The boss HATES broccoli.
b. L'amo l_1'ODIA t_1, el bròquil.
(95) a. He HATES it.
b. *pro* l'ODIA.

The information-packaging instruction encoded in (94) is 'I am instructed to go to the address "the boss" and then retrieve the information of the sentence by substituting **hates** for the blank in *he__broccoli* under "the boss"', but the information packaging of (95) is just 'I am instructed to retrieve the information of the sentence by adding x_1 **hates** x_2 under the current address.' In other words, in (95) there is no need for an address or for further specification on how to enter the sentence under that address. The speaker assumes that the hearer is already at the correct address and, further, that he does not possess any of the knowledge encoded in that proposition. In contrast, in (94) the speaker assumes the hearer needs to be told how the sentence contributes to her/his knowledge-store. To make this distinction between (94) and (95) may seem, at first blush, quite counterintuitive. After all, the two sentences have a parallel syntactic structure and parallel prosodic contour. But while this is true, it is also true that, somehow, there is a difference in markedness between (94) and (95): the former is clearly marked compared to its canonical *The boss likes* BROCCOLI, but the latter is obviously unmarked when compared to *He hates* IT, which is on the verge of ungrammaticality. This difference in markedness matches the intuition that while *broccoli* feels like a tail in (94), *it* does not feel like a tail in (95). In fact, a comparison of the Catalan sentences (the (b) sentences) in (94) and (95) suggests that these two postfocal elements in English are informationally distinct. The object *el bròquil*, equivalent to English postfocal *broccoli* is encoded in the syntax typical of tails in that language (cf. § 4.3.1), but the object clitic equivalent to English *it* is not.

One could actually view sentences like (95) as equivalent to sentence fragments as the one discussed in § 4.2.2. That is, sentence (95) is informationally parallel to an all-focus sentence and the presence of the pronouns is due to independent syntactic motivations, namely, the θ-criterion and the idiosyncratic requirement in some languages that all arguments of the verb be phonologically spelled out. In other words, in languages where zero-anaphora is permitted the verb would be the only overt element in the phrase, as is partially the case in Catalan, where the subject is omitted. The

prediction here is that sentence (95), repeated here in context in (96)b, is informationally equivalent to sentence (97)b (=(53) in § 3.2.3)

(96) a. How does the boss feel about broccoli?
b. He HATES it. (cf. *HATES)
(97) a. What doesn't the boss like?
b. BROCCOLI.

This equivalence, at least at the intuitive level, seems to be correct. In both (96) and (97) the link is missing. As noted, this is because the information must be recorded under the address the hearer is currently at. Links denote an address, but they also instruct the hearer to go to that address. If the address relevant for information entry is the current one, there cannot be a link in the sentence. Sentence (96)b, however, is a tailless sentence as well. The hearer is instructed to add 'he hates it' under the address s/he is currently at. Contrary to the use of a tailful structure, which is designed by the speaker to avoid redundant data entry in the hearer's knowledge-store, the use of (96)b indicates that the speaker does not assume that the hearer has any of the knowledge encoded in the proposition communicated, e.g. the knowledge that 'he is in some relation to it'. Therefore, the presence of the pronoun in (96)b must be due to noninformational requirements. It is true that the address denoted by *it* is already not only hearer-old but also active, which allows the correct interpretation of the anaphoric form. But, as noted in the discussion of referential status, that process is independent of information packaging. In other words, the variable in (96)b is free as far as the packaging instruction is concerned. It remains for a theory of reference to say how the free variable is referentially identified.

However, in natural language there are strong pronouns as well. In our representation, due to the formal characteristics of strong pronouns pointed out above, they are treated like regular lexical material. Strong pronouns, within a given sentence, may be part of the focus, or even make up the focus by themselves, as pointed out in § 2.3.2 and further illustrated by the following example:

(98) S_1: Good morning. I am here to see Mrs. Bush again.
S_2: Sure, Mr. Smith. Let's see . . . One of her assistants will be with you in a second.
S_1: Could I see HER today? I'm always talking to her assistants.

And they can be part of the ground as well, as illustrated by the link pronominal phrase in (99):

(99) a. Him I don't want.
b. A ell_1 *pro* no el_1 vull.

In this example the hearer is instructed to go to the address denoted by *him* and enter information there. Independently, *him* is marked for referential status as being an already-activated address, thus allowing the hearer to know which address it denotes. This identification, as noted, is the responsibility of reference resolution and not of information packaging.

Notes

48 As noted in Ch. 1 this belief is not universally held, and proposals that try to reduce information packaging to logico-semantic meaning exist, as well as proposals that, despite recognizing a difference between the two, require the presence of informational elements in the logico-semantic representation. These proposals are addressed in Ch. 7.

49 Actually, to be exact, there is a further difference between definites in general and pronouns in particular. Both denote preexistent addresses but differ in that pronouns denote salient preexistent addresses (cf. Chafe 1976, Prince 1988b) and other definites nonsalient ones. In other words, definites trigger an activation of a dormant preexistent address. Pronouns simply indicate that their referent is in activation at the time of utterance.

50 Independently, the bearer is instructed here, via referential-status marking, to activate the preexistent addresses 'the boss' and 'Mary' and create a new address for 'a broccoli stalk'.

51 In the sentence representing this instruction there is a subject pronoun *I* which is not part of the focus, but still appears within the scope of Φ. The role of pronouns in information packaging will be discussed in § 4.3.3 below.

52 What gets added is not actually 'broccoli' but 'x_1 x_y broccoli', where the free variables are independently identifiable. See § 4.3.3 and § 5.3 for discussion.

53 Cf. Fukui & Speas 1986, Kroch, Heycock & Santorini 1988, Bonet 1989, Fernández-Soriano 1989, Santorini 1989. We will return to this issue in Chs. 5 and 6. All preverbal subjects in Catalan are interpreted as links, but that is not the case in English. Most are, but some are not, as in the case of *[$_F$ The BOSS called]* (cf. the 'ambiguity' of subjects with regard to their existential force or lack thereof; Horn 1989 suggests that the former are topics and the latter are not).

54 See, for example, Rochemont's (1986) development of an LF rule of Focus Raising within the Government & Binding Theory of syntax, where all-focus sentences are focus-raised in their entirety leaving an empty clause behind (cf. Ch. 6).

References

Bonet, E., 1989. Postverbal subjects in Catalan. MIT. (Unpublished MS.).

Chafe, W.L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li, ed. *Subject and topic*. NY: Academic Press. Pp. 25–55.

Fernández-Soriano, O. 1989. Strong pronouns in null-subject languages and the Avoid Pronoun Principle. Paper presented at the First Student Conference in Linguistics. Cambridge: February 4–5.

Fukui, N. & M. Speas. 1986. Specifiers and projection. *MIT Working Papers in Linguistics* 8.128–72.

Heim, I. 1983. File Change Semantics and the familiarity theory of definiteness. In R. Bäuerle et al., eds. *Meaning, use, and interpretation of language*. Berlin/New York: de Gruyter. Pp. 164–189.

Horn, L. 1989. *A natural history of negation*. Chicago: University of Chicago Press.
Kroch, A., B. Santorini & C. Heycock. 1988. Bare infinitives and external arguments. *NELS* 18.271–85.
Kuno, S. 1972. Functional sentence perspective. *Linguistic Inquiry* 3.269–320.
Prince, E.F. 1981a. Topicalization, Focus-Movement, and Yiddish-Movement: A pragmatic differentiation. *BLS* 7.249–64.
Prince, E.F. 1981b. Toward a taxonomy of given/new information. In P. Cole, ed. *Radical Pragmatics*. New York: Academic Press. Pp. 223–255.
Prince, E.F. 1984. Topicalization and Left-dislocation: A functional analysis. In S.J. White & V. Teller, eds. *Discourses in reading and linguistics. Annals of the New York Academy of Sciences*. vol. 433. Pp. 213–25.
Prince, E.F. 1986. On the syntactic marking of presupposed open propositions. *Parasession papers. CLS* 22.208–22.
Prince, E.F. 1988b. The ZPG letter: Subjects, definiteness, and information-status. To appear in S.A. Thompson and W. Mann, eds. *Discourse description: Diverse analyses of a fund-raising text*. Amsterdam/Philadelphia: John Benjamins.
Rochemont, M. 1986. *Focus in Generative Grammar*. Amsterdam/Philadelphia: John Benjamins.
Santorini, B. 1989. *The generalization of the verb-second constraint in the history of Yiddish*. University of Pennsylvania Ph.D. dissertation.
Sasse, H.-J. 1987. The thetic/categorical distinction revisited. *Linguistics* 25.511– 80.
Szabolcsi, A. 1981. The semantics of Topic-Focus Articulation. In T.M.V. Jannsen & M.B.J. Stokhoj, eds. *Formal methods in the study of language*. Amsterdam: Mathematisch Centrum. Pp. 513–40.
Ward, G. 1985. *The semantics and pragmatics of preposing*. University of Pennsylvania Ph.D. dissertation.
Ward, G. & E.F. Prince. 1986. On the topicalization of indefinite NPs. Paper presented at the 1986 Annual Meeting of the Linguistic Society of America. New York.