WPFG

A critique of FG-CNLU
Cees Hesp
Free University, Amsterdam

Chapter 1

Functional Grammar

1. Introduction

In this chapter we will give a short overview of the main tenets of the theory of Functional Grammar. For reasons of conciseness, we will largely restrict ourselves to the work of its founding father, Simon Dik.

2. Some methodological principles

Functional Grammar (FG) as advanced by Dik (1978a) and later publications is a general theory of the grammatical organization of natural languages. It approaches the subject from a functional perspective. In this perspective, language is first and foremost seen as a means of communication. FG, therefore, should be considered part of a wider, pragmatic theory of verbal interaction (Dik, 1978a:2).

In its explanations, however, FG does not only seek optimization with respect to standards of pragmatic adequacy, but also with respect to typological and psychological adequacy. In short, the rules and principles of FG should

a) reveal those properties of linguistic expressions that are relevant to the way they are actually used and relate those properties to rules governing verbal interaction;
b) be applicable to any type of language;
c) not be incompatible with what is known about the psychological mechanisms involved in natural language processing (e.g. Dik, 1978a: 6-8).

3. Scope and descriptive tools

The linguistic structures covered by FG are so-called independent linguistic expressions. These expressions are defined as "in no way dependent on their preceding or following context" (Dik, 1968: 164ff.; 1978a: 15). More concretely, FG concerns itself with sentences and in particular with clauses.

In FG, linguistic expressions are analysed into underlying predications. These predications consist of predicates into which terms have been inserted. Predicates designate properties of, or relations between, entities. Terms are expressions with referential potential, i.e. can be used to refer to entities in some world (Dik, 1978a: 15).

Predicates may either be basic or derived. Basic predicates are given in the lexicon. Derived predicates such as let go and be a teacher, on the other hand, are produced according to predicate formation rules. Lexical entries for predicates are given in the form of predicate-frames, specifying the following information (Dik, 1978a: 16):

(i)   their lexical form;
(ii)   their syntactic category (Noun, Verb or Adjective);
(iii)   the number of arguments they require;
(iv)   the selection restrictions on their arguments;

3

(v)   the semantic functions of their arguments.

An example of such a basic predicate-frame is given in (1).

(1) $give_V (x_1: human (x_1))_{Ag} (x_2)_{Go} (x_3: animate (x_3))_{Rec}$

The entry in (1) is to be read as:

> give is a V(erbal) predicate that takes three arguments. These arguments have the semantic functions of Ag(ent), Go(al) and Rec(ipient) respectively. The first argument is restricted to some human Agent, the third to some animate Recipient.

The selection restrictions themselves should again be thought of as having the basic properties of predicate-frames. For instance, "human" should properly be represented as in (2):

(2) $human_A (x_1)_\emptyset$

in which $\emptyset$ specifies "zero semantic function".
Terms, on the other hand, are formed according to term formation rules. The general schema is given in (3)

(3) $(op\text{-}x_i: pred\text{-}_1(x_i): ... : pred\text{-}_n(x_i))$

The variable op indicates one or more term operators. Each $pred\text{-}_j(x_i)$ indicates an "open predication" (possibly complex) having x_i as a free variable. The colon [:] indicates that information to the right restricts information to the left, and can be read as "such that". Example (4) may serve to clarify this:

(4) $(d1x_i: guitar_N (x_i))$

Representation (4) expresses that "the guitar" may be used to refer to any definite ("d") singular ("1") entity x_i such that x_i has the property of being a guitar.
When terms have been inserted into the argument slots of predicate-frames we have a nuclear predication (Dik, 1978a: 17). Such a nuclear predication is semantically interpreted as designating a (set of) State(s) of Affairs (SoA), or "what can be the case in some world" (Dik, 1978a: 32). Nuclear predications can be further extended by adding satellite terms. Satellites modify the SoA described, for instance by specifying the time or location (Dik, 1978a: 17, 25).

The predications (either nuclear or extended) that have so far been described may be expressed in different linguistic expressions. Two further levels of function assignment are responsible for this.
Terms, which already have the semantic functions given in the predicate-frame, may be assigned the syntactic functions of Subj(ect) or Obj(ect) (e.g. Dik, 1978a: 17). These functions are taken as specifying the perspective from which the SoA is described. The term with Subject function provides a first perspective on the SoA, and the term with Object function a second. Moreover, in many languages the term assigned Subject function will show concord with the verb as to person, number, etc. The

4

differential assignment of these functions is designed to account for such differences in actual expressions as between (5-8).

(5) John (AgSubj) gave the book (GoObj) to Mary (Rec)
(6) John (AgSubj) gave Mary (RecObj) the book (Go)
(7) The book (GoSubj) was given to Mary (RecObj) by John (Ag)
(8) Mary (RecSubj) was given the book (GoObj) by John (Ag)

Apart from the difference in syntactic function assignment, (5-8) are derived from the same underlying structure (9):

(9) Past give$_V$ (John)$_{Ag}$ (the book)$_{Go}$ (Mary)$_{Rec}$

Secondly, terms, which now have a semantic and (possibly) a syntactic function, may be assigned the pragmatic functions of Top(ic) or Foc(us) (e.g. Dik, 1978a: 130). The constituent with Topic function presents "the entity "about" which the predication predicates something in a given setting". The constituent with Focus function presents "the relatively most important or salient information with respect to the pragmatic knowledge of the speaker and hearer" (ibid.). Awaiting further research, Dik (1978a:144) assumes that there is just one Topic constituent per predication, and so too, it seems, for Focus. Example (10) may serve to illustrate the assignment of these pragmatic functions.

(10) Q: What about Rebecca?
      A: To her (RecTop) John (AgSubj) gave a diamond (GoFoc)

In a context in which the conversation has centred on the presents that John gave to various (known) people, and in which Q wants to know about Rebecca's present, A provides the answer as shown. From this example it will be clear that Topic function cannot universally be identified with Subject function, even though Topic function will frequently be assigned to Subject constituents (Dik, 1978a: 143).

When the three stages of function assignment have been gone through, and when the appropriate predication operators (e.g. DECL(arative) or INT(errogative)), predicate operators (e.g. Pres(ent), Past, Prog(ressive)) and term operators (e.g. d(efinite) and p(lural)) have been assigned, we have a fully specified underlying predication (FSUP). FSUPs are unordered with respect to the final order that constituents receive in actual utterances (e.g. Dik, 1978a: 170). The actual mappings of FSUPs onto linguistic expressions are taken care of by expression rules. These account for the following aspects of linguistic structure (Dik, 1978a: 157):

(i)   the form in which terms are realized, in particular by a) case marking and b) adpositions;
(ii)   the form in which the predicate is realized, in particular with respect to a) voice differences in the verb, b) auxiliary elements and c) agreement and cross-reference;
(iii) the order of constituents;
(iv)   stress and intonation.

As concerns the final order of constituents, three forces are considered to be at play in its determination (Dik, 1978a:174). First, there is a preference for having constituents with specific functional specifications

5

invariably in the same structural position. This aspect is accounted for by setting up one or more functional patterns. These patterns specify the positions that constituents of a given functional specification can take under certain conditions. A general schema for such functional patterns, with considerable cross-linguistic validity, is given in (11).

(11)   P2, P1 (V) S (V) O (V), P3

In this schema S and O indicate the pattern positions for Subject and Object constituents. Vs indicate the possible positions for Verbs, possibly divided into Vf(inite) and Vi(non-finite). Ps indicate "special positions". P2 is the position for left-dislocated constituents. It is typically used for extra-predicational constituents having the pragmatic function of Theme (Dik, 1978a: 175-176). Theme is an optional addition to the predication and presents a domain or universe of discourse with respect to which it is relevant to pronounce the following predication (Dik, 1978a: 130). P3 is the position for right-dislocated constituents and is typically used for extra-predicational constituents having the pragmatic function of Tail (Dik, 1978a:176). Tail is defined as an (optional) "afterthought" (Dik, 1978a:130). P1 indicates the initial position in the clause itself.

The second force at play is the preference for assigning specific positions to certain designated categories of constituents and to Topic or Focus constituents. This is accounted for by the inclusion of P positions, in particular P1, and by specifying rules and conditions under which these positions are filled. Designated P1 constituents in English are Wh-constituents, relative pronouns and subordinators (Dik, 1978a: 180).

The third determining force is the preferential left-to-right sequencing of constituents in increasing order of categorial complexity. This is accounted for by postulating a Language Independent Preferred Order of Constituents (LIPOC) (Dik, 1978a: 192). This preferred order is shown in (12).

(12)   PROcl < PRO < NP < NPP < V < NP < PNP < SUB

LIPOC is not held to apply to P positions (ibid.). It is intended to account, for instance, for the favourite position of pronouns as being before rather than after NPs. It also accounts for the favourite position of subordinate clauses as being after (almost) anything else.

Although much more could be said about the theory of FG, and although we will occasionally have to have a closer look at some of its details in the next chapters, we will end this overview by examining the way it organizes the lexicon.

4. The lexicon

Each basic predicate is given in the lexicon in the form of a predicate-frame that specifies the basic aspects of the constructions in which it can appear. Nothing, however, has been said so far about the meaning of these constructions. This is what we will turn to now.

To each of the lexical entries a meaning definition is added. These meaning definitions have the following general format:

(13) $B =_{df} A$

In this definition, B is some predicate-frame and A is some combination of semantically simpler predicate-frames (Dik, 1978a: 46). The underlying rationale of (13) consists of the following assumptions (Dik, 1978a: 46-47; 1978b: 21):

(A1) the defining predicates in meaning definitions are lexical items of the object language;

(A2) the definiens A of a meaning definition may not contain a proper subconfiguration B such that B itself constitutes the definiens of some other lexical item of the object language;

(A3) the structure of the definiens of meaning definitions is of the same formal type as the structures underlying linguistic expressions;

(A4) the definiens of a meaning definition is not directly accessible to the operation of syntactic rules;

(A5) in every language there is a set of (maximally) simple lexical items the meanings of which cannot be (further) defined by means of meaning definitions.

Assumptions (A1-A5) amount to the following:

a) The meanings of words are not defined in some abstract, language-independent code of thought but in the content words of the language itself (= A1).

b) The meanings of words are not decomposed directly into some set of (abstract) maximally simple or basic components but rather in a "stepwise" manner (= A2).

c) Both the definiendum and the definiens have the format of predicate-frames (= A3).

d) Meaning definitions do not appear in the structures underlying sentences but in the lexicon. They are used to interpret the underlying predications rather than to produce them (= A4).

e) The simple, undefinable, terms are semantic primitives (= A5).

Examples (14-15) and (16-18) may serve to illustrate this approach.

(14) $\text{kill}_V \, (x_1)_{Ag/Fo} \quad (x_2: \text{animate} \, (x_2))_{Go} \quad =_{df}$

$\text{cause}_V \, (x_1)_{Ag/Fo} \, (x_3: \{\text{die}_V \, (x_2)_{Proc}\} \, (x_3))_{Go}$

That is, to kill some being is to cause a process through which that being dies (Dik, 1978a: 49).

(15) $\text{die}_V \, (x_1)_{Proc} \quad =_{df} \text{come about}_V \, (x_2: \{\text{dead}_A \, (x_1)_\emptyset\} \, (x_2))_{Proc}$

That is, for someone to die is for a process to take place in which that someone comes to be dead (ibid.)

(16) $\text{bachelor}_N \, (x_1)_\emptyset \quad =_{df} \text{unmarried}_A \, (x_1: \text{man}(x_1))_\emptyset$

That is, for someone to be a bachelor is to be an unmarried man (Dik, 1978b: 24).

(17) $\text{man}_N \, (x_1)_\emptyset \quad =_{df} \text{male}_A \, (x_1: \text{person}(x_1): \text{adult}(x_1))_\emptyset$

That is, for someone to be a man is to be a male adult person (ibid.).

(18) $\text{person}_N (x_1)_\emptyset \quad =_{df} \text{human}_A (x_1: \text{being}(x_1))_\emptyset$

That is, for someone to be a person is to be a human being (ibid.).

From these examples we may (tentatively) conclude that forms such as "male", "human", "being" and "dead" should be considered semantic primitives of English.

5. Conclusion

This brief exposition of the theory of "Stepwise Lexical Decomposition" (Dik, 1978b) as incorporated in Functional Grammar (Dik, 1978a) ends our overview of the latter's main tenets. In the next chapter we will discuss Dik's proposals for a process model of the Natural Language User based on these principles.

Chapter 2

FG-CNLU

1. Introduction

Although Dik (1978a: 2) grants that a theory of language is not the same as a theory of language use, the strict distinction between competence and performance as advocated by Chomsky (1965: 4) has never been observed within Functional Grammar (FG). This is not only reflected in FG's fundamental view of language as a means of communication rather than as a formal system, but also in the programmatic emphasis on pragmatic and psychological adequacy. As a consequence, it has always been implicit in the FG literature that the knowledge which it specifies is actually used by the speaker/hearer.

Not until recently, however, has this claim been made explicit. Dik (1986b: 1) has stated the principle that FG should be devised in such a way that "FG should be a good candidate for providing a module for a naturalistic model of the Natural Language User (NLU)". In fact, a research program has been set up to see if it is possible to build a Computational Natural Language User (CNLU) based on FG principles (Dik, 1987b). This process model, called FG-CNLU, is "to simulate the actual, natural performance of (Human) NLU in normal communicative circumstances". It is to do so in a "psychologically adequate and realistic way" (Dik, 1987b: 4). As such, the building of a process model follows quite naturally from the methodological principles of FG. It is seen as a test for the overall adequacy of the theory of FG (Dik, 1986a: 1).

At the same time, however, the building of a process model serves an even more ambitious purpose. FG-CNLU is seen as an attempt at "the re-unification of grammar, logic and cognition" (Dik, 1986c: 18). It should provide "insight into how NLUs work" (Dik, 1987b: 4). It is no longer the case, therefore, that FG should solely (try to) be compatible with psychological theories of language processing. FG-CNLU itself has been proclaimed a psychological theory of language use. It will be clear that a number of additional assumptions were necessary to extend the domain(s) over which FG principles range for this to become possible.

In this chapter we will describe the recent proposals made by Dik (1986a, b, c, d; 1987a, b) for the building of a process model of the Natural Language User (NLU). We will first present the general outline of the model as sketched in Dik (1987b). This will allow us to refer to the different parts of the model. Moreover, it will enable us to divide the model into digestible chunks.

2. General outline of the model

A general outline of the model is given in figure 1 (opposite). Let us briefly elaborate on this schema. Arrows indicate processors and boxes represent data stores. The input to FG-CNLU is either spoken or written, and so is its output (A and B). Spoken input receives a phonetic representation through processor 1, and is stored in A*. Processor 2 provides a representation of written form for written input, and stores it in B*. Reverse mappings for output are processed by 1* and 2*, respectively. Processors 5 and 5* can map these representations onto one another. They are involved in writing down spoken information, or in reading written information out loud. Parsers 3 and 4 map the representations of A* and B* into fully specified underlying predications (FSUPs) and store them in C. Generators 3* and 4* map FSUPs onto phonetic representations and representations of written form, respectively. Processors 6 and 6* store UPs in and retrieve them from knowledge base D. (How FSUPs come to be UPs will be discussed below.) Processor 7 is used to infer pieces of information from the UPs stored in D and is called Functional Logic (FL). Processors 8 and 8*, finally, convert predications from one language into another by using a dictionary that contains L1-L2 (paraphrase) equivalences.

For ease of exposition, the general outline of the model as given in figure 1 may be simplified to figure 2 (opposite). (We will restrict ourselves to the apparatus needed for one language (L1), and assume that it is set to the "comprehension" mode.) The three main parts, which Dik attempts to re-unify (i.e grammar, logic and cognition), are now schematically represented as follows. (X) represents a language module that takes written or spoken language as input, and that delivers FG underlying predications as output. These outputs are stored in memory D by "gatekeeper" 6, and serve as input to "inference engine" 7. Language module X, thus, represents the "grammar" part. Inference engine 7, called Functional Logic, forms the "logic" part. And the sum total of what 7 can infer from the knowledge representations stored in D plus what 6 "knows" is already stored in D is the "cognition" part. Unification is achieved by using the same representational format, viz. the FG underlying predication, throughout.

We are now in the position to impose some structure on our further description and discussion of the model. We will first describe the underlying assumptions about the relation between language and thought as presented in the model. We will then describe the proposed contents for boxes X (the language module) and D (the knowledge base). Next, we will present the principles that govern their interaction, i.e. the operations carried out by processors 6 and 7.

3. Underlying assumptions

The process model advocated by Dik (1987b) utilizes the rules and principles of FG to characterize the linguistic knowledge of the speaker/hearer. It will be clear, however, that the building of a CNLU also calls for the characterization of other types of knowledge (e.g. referential knowledge, episodic knowledge and general knowledge of the world (see Dik, 1986a)), since all these types of knowledge are used in communicating with language. It will also be clear that from a design point of view the system should preferably have only one format for representing knowledge.
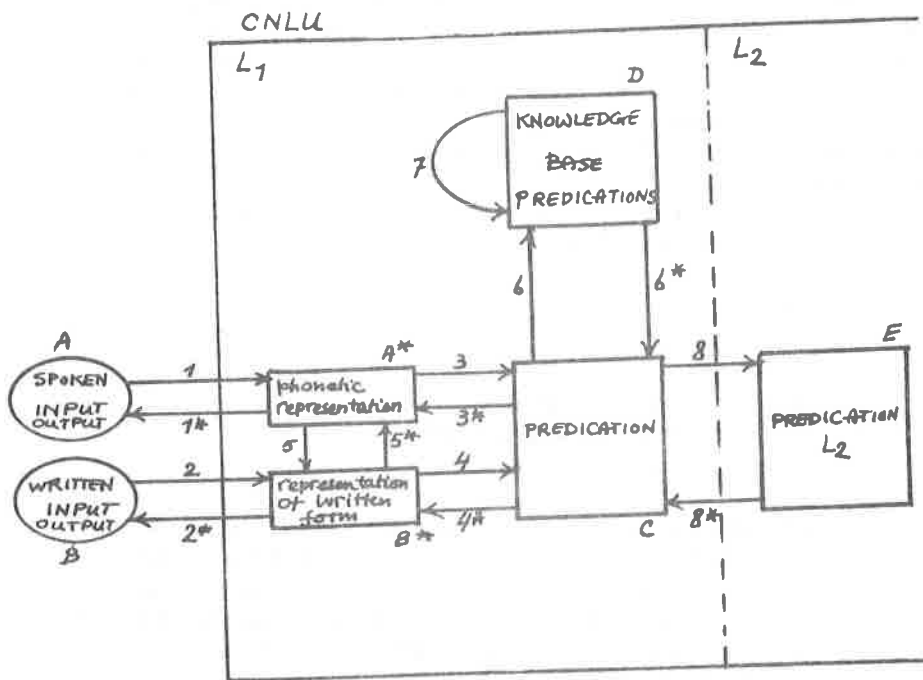
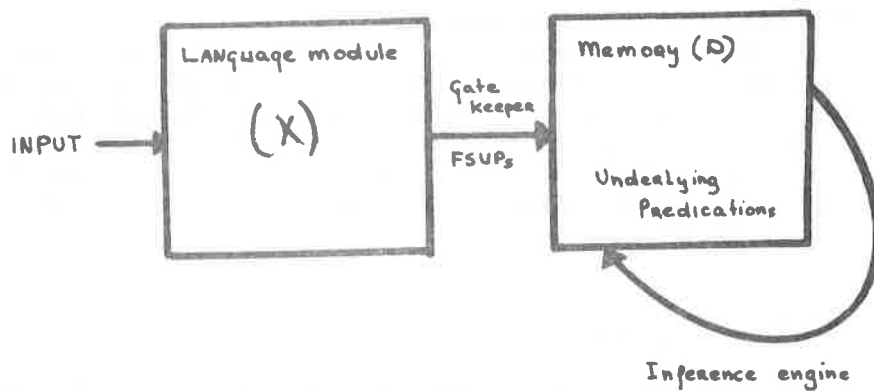Fig.1. Global structure of CNLU.

Figure 2. A simplified view of FG-CNLU

To arrive at a uniform representation of these various types of knowledge, Dik (1986a: 16) proposes the following. First, he strictly dichotomizes conceptual and perceptual knowledge by claiming that all knowledge is either one or the other. Second, he assumes that perceptual knowledge takes the form of images, i.e. "perceptual pictures of the entities and states of affairs they stand for". He further assumes that all other knowledge, including non-linguistic knowledge, is represented in verbal form. And finally, he suggests that verbal form is represented in FG format, viz. in underlying predications (UPs).

The FG predication, thus, plays a central role in the CNLU model. More specifically, it is used to represent (Dik, 1986b: 18):

(i)   the output of a parser;
(ii)   the input of a generator;
(iii)   the lexical definitions;
(iv)   the conceptual knowledge of the system;
(v)   the input and output of an inferencing calculus;
(vi)   the input and output of a translation module.

It will be obvious that this approach ties (conceptual) knowledge closely to the language in which it is expressed since FG predications largely consist of the actual content words of particular languages. A far-reaching consequence of this view is that Dik is forced to conclude that different languages do not share concepts. And he also concludes that there is no language-independent code of thought. It should be noted, however, that he sees this as an advantage of his theory, not as a disadvantage (1986a: 3-4).

4. The language module (Box X)

4.1. Data stores and representations

As one can judge from figures 1 and 2, FG-CNLU comprises 4 data stores per language that the system understands. Three of these data stores are internal to what we have called the language module (viz. A*, B* and C). These data stores and will be discussed here.

As yet, the only specifications that have been made about these data stores are the representational format of the contents of C and the general functions that these data stores are judged to have in language processing as a whole. No specifications are given for the periods of time over which these stores retain their contents, nor what their capacities are. In the following description of these data stores, therefore, we can only briefly relate the general ideas.

4.1.1. Phonetic representations (A*)

FG-CNLU so far has taken no particular stand with respect to the nature of the representations of phonetic form. FG-CNLU would be compatible, therefore, with either a phonemic or diphonic approach. The only requirement on the eventual representation is that processor 3 should be able to construct the predication underlying it, and that 3* should be able to convert a predication into it (Dik, 1987b: 9).

4.1.2. Representations of written form (B*)

For most practical purposes, according to Dik (1987b: 9), the written version of a text can at the same time be used for the internal representation within CNLU. A certain amount of pre-editing may be required, though (e.g. repairing damages, specifying intonation and stress contours).

4.1.3. The underlying predication (C)

Data store C contains the FSUPs of FG and needs no further comment, since these have been extensively discussed in the previous chapter.

4.2. Processors of the language module

As figures 1 and 2 show, there are ten processors internal to the language module. Here, too, very little can be said as yet about the actual working of these different processors, because development (of some of them, at least) is still in progress. We can only mention, therefore, some general principles, design objectives and suggestions that have been made for these processors.

4.2.1. Interactive interpretation

Interpretation is defined by Dik (1986a: 11) as "to attribute higher-level significance to lower-level data". There are four processors in the language module that fit this description, namely processors 1 and 3 for spoken input and processors 2 and 4 for written input (see figure 1). Other processors either convert higher-level data into lower-level data (processors 1*, 2*, 3* and 4*) or operate on data of equal status (processors 5 and 5*).

Interpretation, according to Dik (1986a: 12), should take place in a parallel, interactive top-down and bottom-up fashion. A rough outline of the parallel, two-way interactive interpretation model as proposed by Dik (1986a: 13) is given in figure 3 (overleaf). The bottom-up component consists of an Analyser that scans the data and that yields tentative analyses of these data. The top-down component consist of a Deducer which derives predictions about what these data might represent from the knowledge which the system already possesses. The outputs of the Analyser and Deducer are fed into one another so as to further guide the interpretation process. This process is monitored by a Matcher that decides at which point the prediction and analysis are sufficiently congruent to decide on an interpretation, using criteria of sufficient fit.

Given the above description, one might expect this interpretation model to be the major design objective for processors 1 to 4. In Dik (1987b: 17), however, this objective is only explicitly stated for processor 1. One thus might assume that processor 1 is subdivided into an Analyser, a Deducer and a Matcher, that it is linked up to a body of knowledge, and that it delivers an interpretation to the next processor in line, the parser (processor 3), in other words, that the internal organization of processor 1 is characterized by the interpretation model of figure 3. This, however, does not seem to be the (intended) case. Consider the following.

In describing the intended functioning of processor 1, Dik (ibid.) gives the following example:
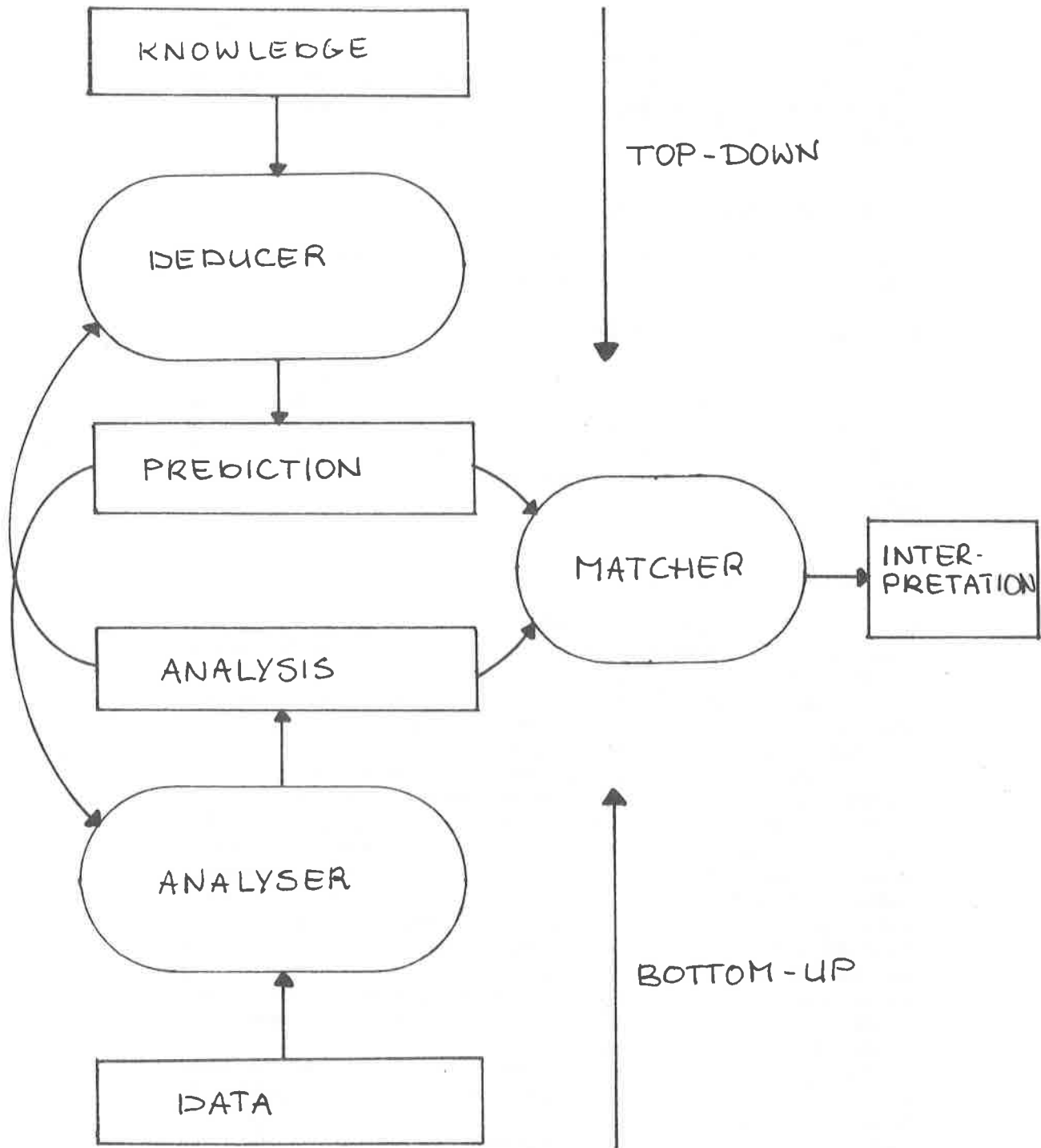
13

FIGURE 3. A TWO-WAY INTERACTIVE INTERPRETATION MODEL

(1) The man had a long scar on his cheek

He comments:

> Suppose that the spoken form of **scar** is pronounced in such a way as to be compatible with **scar**, **star** or **car**. The system should then nevertheless be able to arrive at the correct interpretation **scar**, on the basis of such higher-level knowledge as:
>
> a)   A man does not usually have a star or car on his cheek;
> b)   Stars cannot be said to be long;
> c)   A scar is something which can be said to be long and which one can have on one's cheek.

It will be clear, however, that much more is involved here then pure phonetic processing or "speech identification". In order to use the knowledge contained in a) to c), there must already have been a complete parse of sentence (1) for each of the proposed candidates. Each of these parses, moreover, must have been interpreted and checked against knowledge of the world. This cannot be the job of processor 1 alone, of course. We will assume, therefore, that the interpretation model as given in figure 3 should be taken to characterize the interpretation process as a whole. In other words, we will assume that the design objective holds for processors 1 to 4 in general.

A further proposal by Dik concerns processor 2*, but in an indirect way. In the wider system of FG-CNLU there will be a place where each content predicate and grammatical element of the language is stored together with its phonetic representation (Dik, 1987b: 9). This "place" is usually called the (mental) lexicon (see e.g. Garnham, 1985: 43). The nature of this phonetic representation would be such that it can be taken as a complex command which, when executed, will produce a token of the word in question, without going through the complex process of combining smaller elements (phonemes or even phonemic features) (Dik, ibid.). (The relevance of this proposal to our discussion, which as one will recall is limited to comprehension, will appear later on.)

Another suggestion concerns the principles along which a parser could be designed. Dik (1986c: 8) suggests that "a powerful parser could be built on the principle that

1)   first, the (main) predicate of a linguistic expression is identified,
2)   that then the predicate frame of that predicate is retrieved,
3)   and finally that the rich information contained in the predicate frame is used to identify and interpret the other constituents contained in the linguistic expression."

It will be clear that this view flows rather naturally from the theory of FG. A consequence is, however, that the parser cannot start processing until it has received the completed output from processor 1. That is, the parser cannot start parsing until all the words in the utterance have been phonetically processed. (This is also implicit in Dik's (1987b: 18) statement about the parser for written input. He suggests that the input to processor 4 consists of "a string of separated written word forms".)

5. Knowledge base (D)

The knowledge in knowledge base D has the form of UPs. The difference between the linguistically motivated knowledge representations (LMKRs) of D and the underlying predications of C, i.e. between UPs and FSUPs, is that UPs do not usually contain illocutionary operators (viz. DEC, INT or IMP) (Dik, 1987a: 21). These labels are not simply dropped, though. The part of FL (processor 7) that handles illocutionary logic takes the label DEC to mean that, if the speaker is sincere, the speaker believes that the predication is true, and that the speaker believes that the hearer may want to know the information in it (Dik, 1987a: 13). INT is taken to mean that the speaker should consult his/her knowledge base and provide an answer if possible (Dik, 1986c: 8). A similar interpretation is, probably, intended for IMP.

We may also gather from the examples that Dik gives of LMKRs that the syntactic and pragmatic functions of terms are not retained in the UPs of D. The knowledge representation of D will thus look like (2) (from Dik, 1987b: 20):
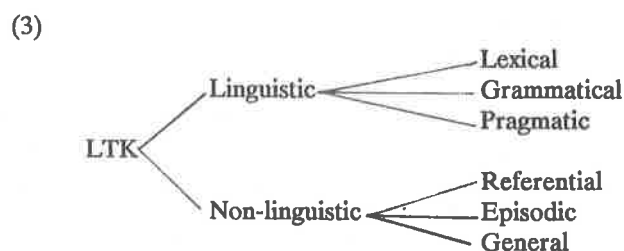
(2) Past kill$_V$ (Oswald)$_{Ag}$ (Kennedy)$_{Go}$

We will assume that syntactic and pragmatic function labels are somehow taken into account during the comprehension process; in other words, that the speaker does not assign these functions to no avail.

5.1. Knowledge typology of D

The knowledge stored in D can be divided into short-term knowledge and long-term knowledge on the basis of the following definitions (Dik, 1986a: 15):

A) short-term knowledge is knowledge which the system derives from what it perceives in the actual communicative situation;

B) long-term knowledge is knowledge which the system possesses independently of the particular situation.
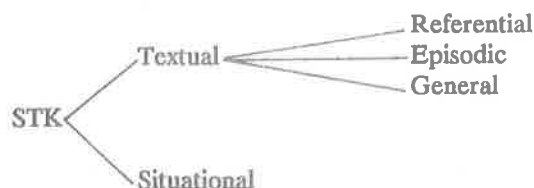
Long-term knowledge (LTK) is further divided into the following types on the basis of what it is about (and not on how it is represented, for the representational format of conceptual knowledge is assumed to be identical for each of these knowledge types):

(3)

```
                                       ┌─ Lexical
                    ┌─ Linguistic ─────┼─ Grammatical
                    │                  └─ Pragmatic
        LTK ────────┤
                    │                  ┌─ Referential
                    └─ Non-linguistic ─┼─ Episodic
                                       └─ General
```

And short-term knowledge (STK) into:

(4)

```
                                           ┌─ Referential
                          Textual ◄────────┼─ Episodic
                         ╱                 └─ General
                 STK ◄──
                         ╲
                          Situational
```

Linguistic (long-term) knowledge is divided into knowledge of the forms and meanings of lexical elements (=lexical knowledge), of how these may be combined into linguistic expressions (= grammatical knowledge), and of what conventions govern the use of these expressions (= pragmatic knowledge). Non-linguistic (long-term) knowledge, on the other hand, is divided into knowledge of entities (= referential knowledge), of States of Affairs (SoAs) in which these entities have been involved (= episodic knowledge), and of general rules and laws that are judged to apply to these entities and SoAs (= general knowledge).

Short-term knowledge is derived either from the verbal input (the ext) that the system receives (= textual knowledge) or from the situation in which it receives that input (= situational knowledge). Textual knowledge can again be divided into referential, episodic and textual knowledge.

Two further comments are necessary here. First, according to Dik (1986a: 15) this distinction between STK and LTK more or less coincides with the psychological distinction between Short-Term Memory (STM) and Long-Term Memory (LTM). (This claim will have some very important consequences for our later discussion of the model's psychological reality.) Second, the long-term knowledge of the system should be thought of as stored in thematically organized files, connected by multiple cross-references (Dik, 1987b: 21).

6. Processor 7 (Functional Logic)

Processor 7, called Functional Logic, is used to make inferences from the knowledge that is stored in knowledge base D. Since the knowledge in D has the form of predications, processor 7 takes predications as its arguments. Moreover, it is also required that its conclusions have the form of predications. Functional Logic, thus, has the same syntax as the language of underlying predications as determined by FG for a particular natural language. The syntactically well-formed logical expressions of FL are interpreted in terms of possible mental pictures.

(Mental) pictures, according to Dik (1987a: 9), have the following properties:

(A) Pictures consist of knowledge. Since knowledge is either perceptual or verbal, the language in which pictures are coded is partly identical to the language of UPs.

(B) Pictures are epistimological objects and, therefore, subject-dependent.

(C) Pictures are dynamic entities: they may be created, modified and

17

destroyed.

(D) Pictures are necessarily limited and finite: nobody has a complete picture of anything.

(E) Different pictures may at the same time be relevant to interpretation and reasoning.

Despite this reference to a "language" that consists partly of predications and partly of perceptual pictures, the subtheories of FL only concern linguistic material. The subtheories Dik (1987a: 12) recognizes are:

(a) Illocutionary Logic
which deals with the logical properties of illocutionary operators such as DEC, INT and IMP.

(b) Predicational Logic
which deals with the logical relations between whole predications.

(c) Predicate Logic
which deals with the logical properties of predicates and predicate operators such as Pres, Past, Prog etc.

(d) Term Logic
which deals with the internal logical properties of terms, including term operators such as d, p, 1, m etc.

(e) Lexical Logic
which deals with the logical relevant properties internal to lexical predicates.

We conclude therefore that FG-CNLU is limited to reasoning with, and in, predications. Moreover, its inferencing capacities are limited to deduction, i.e. given "correct" (e.g. true or sincere) premises it deduces a necessarily correct conclusion. In other words, it can only draw valid conclusions (Dik, 1987b: 21).

6.1. Processor 6

In describing the proposals about processor 6 we can be very brief again. Processor 6, according to Dik (1987b: 20), "need only decide which predications are to be retained, and integrate them in the knowledge base."

7. The interpretation model revisited

Now that we have some understanding of the knowledge that is stored in knowledge base D, and of the functions of processors 6 and 7, we will have a short look again at the interpretation model suggested by Dik (1986a). More in particular, we can now present the extended interpretation model (see figure 4, overleaf).

In this extended interpretation model two things have been combined. The interpretation model as presented earlier has been combined with the knowledge typology of D. Together they form a single "integrated model of linguistic interpretation" (Dik, 1986a: 17). This combination follows from the assumption that long-term knowledge is the knowledge from which the Deducer derives its predications, and that short-term knowledge is the analysis which the Analyser comes up with on the basis of the raw input data (Dik, 1986a: 16).

As one can see, there is one new component in this integrated model, namely the Evaluation component. Its function is described as follows:

> This component evaluates the interpretation that the system arrives at with respect to its retention value. If the interpreted information is judged sufficiently important for later use, it is integrated into the systems long-term knowledge. If not, it is discarded and forgotten (Dik, 1986a: 16).

8. Epilogue

In this chapter we have taken a look at the proposals that Dik has made in a variety of papers concerning the building of a process model of the Natural Language User. We have refrained from commenting on these suggestions where possible, trying to represent his ideas as faithfully as we can. Perhaps we should stress once again that FG-CNLU is still under development, and that it may expand and change, given time.
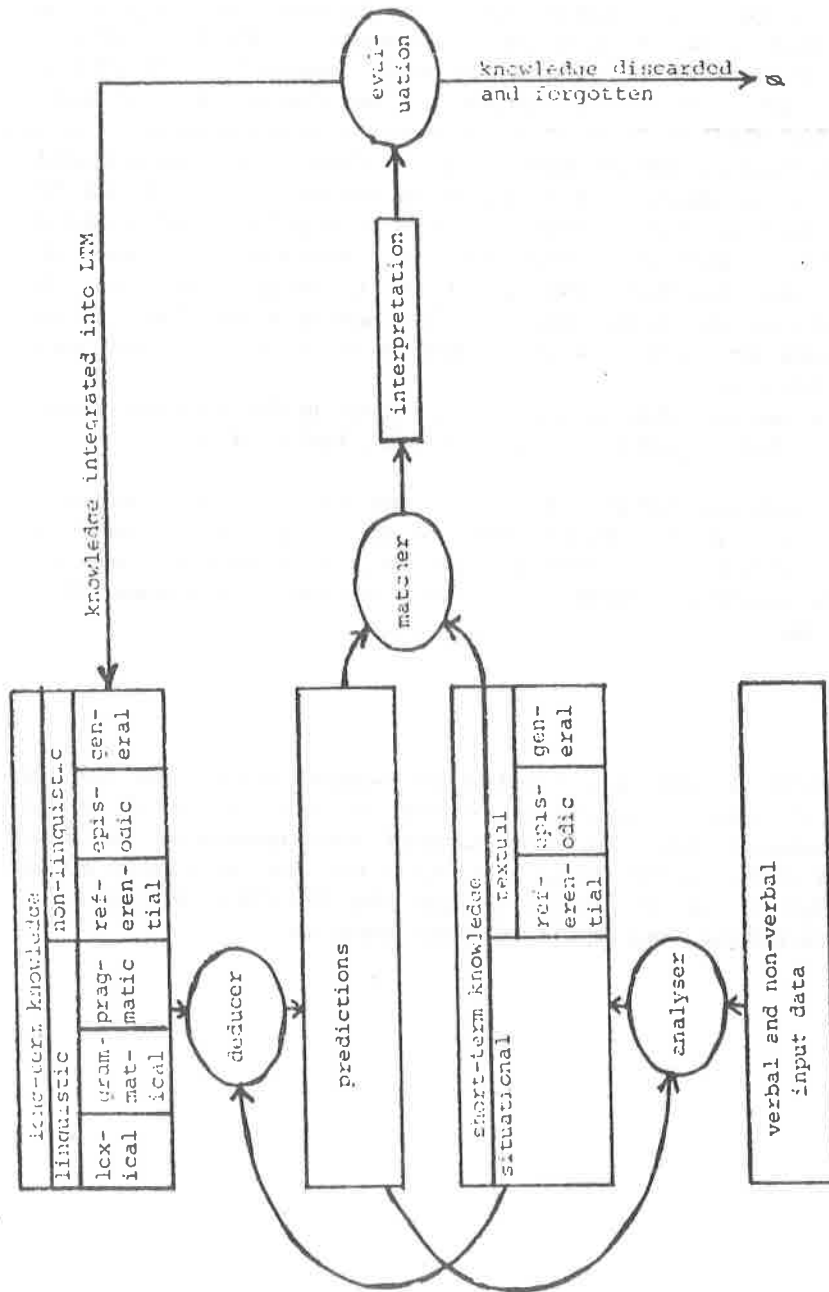
Figure 4.
An integrated model of
linguistic interpretation

Chapter 3

FG-CNLU, technically speaking

1. Introduction

In the previous chapters we have given an overview of the theory of Functional Grammar (FG) and of the proposals that have been made to use its rules and principles as a guideline for the building of a process model of the Natural Language User (NLU). We have also paid attention to the additional assumptions that were necessary to make this possible. As the reader will have noticed, no doubt, we have not gone much further than merely listing the different design objectives, ideas, etc.

In this chapter we will go over the model again. This time, however, we will critically discuss its proposed functioning. Our major interest will be the extent to which the model can be considered coherent. That is to say, whether the proposals that have been made are technically feasible given the architecture of FG-CNLU as presented in figure 1. Since we will have to refer to figure 1 repeatedly, it is reprinted on p. 22 for convenience, as is figure 4.

The second point of interest will be the relation that the model FG-CNLU bears to the underlying theory of how humans process natural language. In general, the same behaviour can be mimicked by many computational models. This, in effect, is the main point of the so-called "Turing test" (Turing, 1950). In this test one has to decide, roughly, whether one is communicating with a computer or with another human being. Any two systems that pass this "test" when or while modelling the same behaviour can be said to be equivalent or, at least, to show an input-output equivalence. So too, Dik's major claim that (all) non-perceptual information or knowledge is processed, stored and retrieved in verbal form could be made operational in different models with different architectures.

The reason why this relation is of special concern to us is the following. If we want to test the psychological reality or adequacy of the different proposals by matching predictions based on them against experimental findings, we have to know what falsifies what. That is to say, some experimental results may falsify (part of) the particular model at hand, but need not necessarily falsify the theory as well. Especially when a model only partly covers the phenomena that the underlying theory makes claims about, the job of falsification is extremely difficult. Any evidence that falsifies certain predictions based on the model runs the risk of being "sidetracked" by the claim that the theory does not commit one to the view taken in the model.

This chapter, then, also serves to make the distinction between theory and model, which have been presented by Dik as if they were one and the same, somewhat more clear.

2. FG-CNLU and perception

In the theory behind FG-CNLU a strict distinction is made between conceptual and perceptual knowledge. The former is thought of as represented in FG predications, and the latter in pictures. As figure 4 shows, no perceptual long-term knowledge is included in the actual model. This is not surprising, since no processor has been postulated that handles
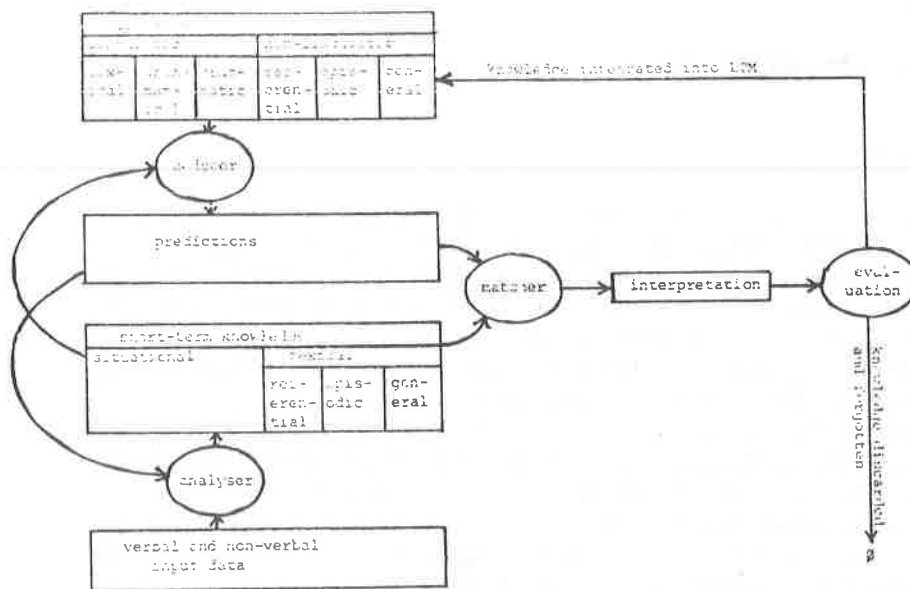
Figure 4.
An integrated model of
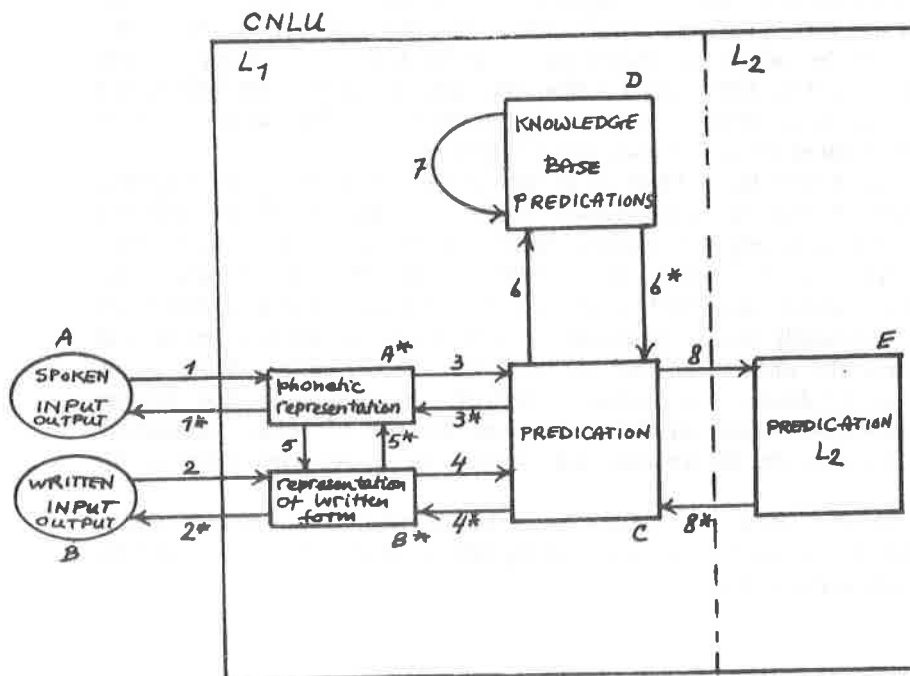linguistic interpretation

Fig.1. Global structure of CNLU.

perceptual input. Nor would processor 7 be able to reason with such knowledge if it were present, given that its current subtheories only cover (parts of) predications. What is surprising, however, is that the system is described as having short-term situational knowledge (knowledge of what is perceivable in the communicative situation) and as having non-verbal input data (see figure 4). Since short-term knowledge (STK) results from the analysis of input, and since no processors to analyse perceptual input have been postulated, this cannot be correct, of course. We will assume, therefore that the model FG-CNLU can only handle verbal material, either as input, as output or as intermediary.

The perception of verbal input, however, also poses problems for the model, both for spoken and for written input. Let us start with spoken input (processor 1).

The knowledge that a phonetic processor would use to attribute higher-level phonemic or diphonic values to its lower-level acoustic input includes knowledge of phonetic rules. This knowledge (presumably) falls under the category "perceptual", and as such is not present in the model. A further type of knowledge that is used in word recognition consists of phonological rules. Phonology, which specifies which sound sequences are (possible) vocabulary items of the language in question, is by definition part of the linguistic knowledge of the speaker/hearer. It is not clear, however, how phonological rules should be represented in UP format. (Nor is it clear whether knowledge of phonological rules is at all present in FG-CNLU, since they are never mentioned.) Apart from these two types of knowledge, processor 1 would also have need of a lexicon in which it could check whether the proposed analysis is a vocabulary item of the system. Since lexical knowledge is stored in D, and since processor 1 is not linked up to D, it is not clear how processor 1 is supposed to function.

The model, however, does not only lack the knowledge needed for the processing of spoken language, it also lacks the "equipment". The two-way interactive interpretation model that has been suggested for the processing of (spoken) input would require the presence of an auditory memory or buffer in which the system could temporarily store its raw acoustic input signals until an interpretation had been settled on. Since no such buffer is postulated, this alone would suffice to make the process of interactive interpretation impossible.

A similar story holds for written input. FG-CNLU has neither the knowledge (i.e. visual or graphemic knowledge) nor the equipment (e.g. graphemic memory) to deal with real written input, despite Dik's (1987b: 18) claim that "the mappings of processors 2/2* are least problematic as long as typed or printed input is used".

The only conclusion that we can draw, then, is that FG-CNLU cannot handle perceptually available input, not even verbal input. Communication with the FG-CNLU model is (in principle) only possible by typing the message directly into the system. But the next paragraphs will give us reason to doubt even that.

3. Parallel and interactive processing

The interpretation model that was suggested for FG-CNLU is claimed not only to function in an interactive way, but also in a parallel fashion as well. Let us first point out that there is a difference between these two notions.

23

Two processes can be said to function interactively if there is feedback from the second process to the first. Two processes can be said to operate in parallel if they work on the same input at the same time. This is illustrated in figure 5, opposite.

The system as depicted in figure 5a operates interactively. During the performance of some task process 1 sends information to process 2. Process 2 then starts working on it, and may send information back to process 1. Because of the serial nature of this system, however, process 1 must first produce a partial analysis before process 2 has something to work on.

The system as in 5b operates in parallel. Both processes are working on the same input at the same time, but are not in contact with one another. The main requirement of such a system is that the processes are indeed capable of handling the same input.

In figure 5c the system functions in a parallel interactive fashion. Both processes receive the same input and provide one another with information concerning (partial) analyses.

Given this distinction between parallel and interactive processing, let us consider FG-CNLU when set to comprehension mode, that is, suppose it is given spoken (verbal) input. Processor 1 is the only processor capable of handling auditory wave forms. It converts them into phonetic representations. The output of processor 1 can only be sent to processor 3 since no other processor is capable of dealing with phonetic representations. Processor 3 converts this phonetic representation into predicates and computes the structural relations between those predicates. The output of processor 3 is then sent to D by processor 6. Once in D, processor 7 can start deducing other facts from it. FG-CNLU is thus a system that operates strictly serially. This conclusion holds not just because the processors are ordered linearly in figure 1, but also because there are no two processors postulated as working with the same type of input at the same time in the interpretation process. Hence, no parallel processing is possible.

As far as interaction between the processors is concerned, it will be clear that figure 4 and figure 1 give totally different pictures. In figure 4, the Analyser and Deducer provide one another with feedback. The architecture of FG-CNLU as presented in figure 1, however, does not show how such interaction is possible. None of the processors depicted there is connected with any of the other processors. Since figure 1 is presented as "specifying its main components and their mutual relationships (Dik, 1987b: 6)", we can only conclude that in this representation of the model no interaction is possible.

4. Knowledge and processes

The knowledge that FG-CNLU utilizes is basically characterized by the following three statements:

1) The system is said to know something if it has such knowledge stored in D or if processor 7 is able to infer it on the basis of the knowledge already stored there.

2) The knowledge in D, furthermore, is divided into linguistic and non-linguistic knowledge on the basis of what it is about.
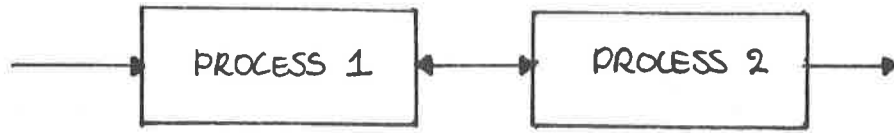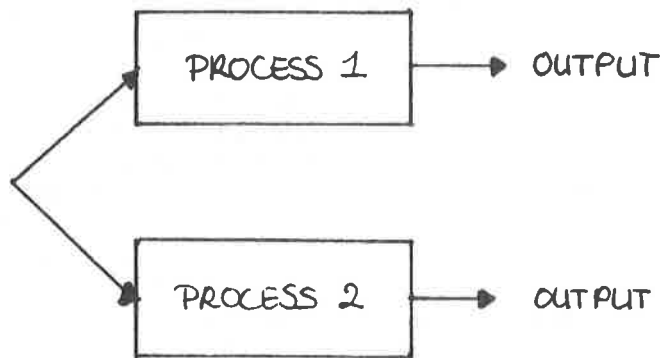
FIGURE 5A. SERIAL INTERACTIVE
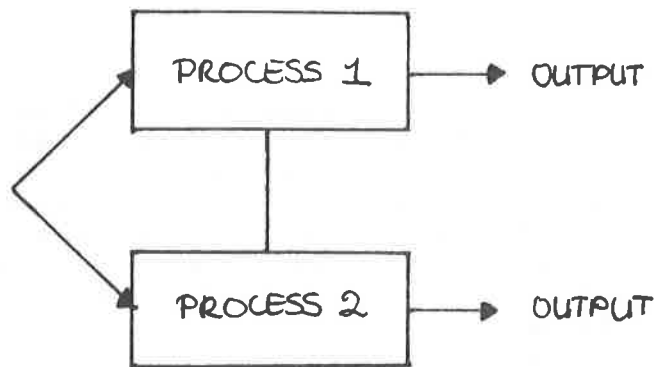PROCESSING

FIGURE 5B. PARALLEL PROCESSING

FIGURE 5c. PARALLEL INTERACTIVE
PROCESSING

FIGURE 5. METHODS OF PROCESSING

3) The knowledge in D, finally, is divided into long-term knowledge (LTK) and short-term knowledge (STK).

This description poses the following problems.

4.1. Inaccessibility of D

It is said that many of the types of knowledge that the system possesses interact in determining the correct interpretation of a linguistic expression. Since interpretation is considered a two-way process in which an analysis and a prediction are matched against each other, it follows that both the analysis and the prediction are made by utilizing the knowledge stored in D. We may, thus, assume that the linguistic knowledge which the system has stored in D would be the kind of knowledge that processors 1 to 4 could use in analysing their (linguistic) input. Notice, however, that none of these processors has access to this knowledge (see figure 1). The knowledge stored in D, therefore, is of little use to the system during comprehension.

The same holds for the non-linguistic knowledge of D that the processors might need to consult. Let us illustrate this point by having a closer look at, for instance, the parser. The parser computes the FSUP underlying the input. Part of this job consists of determining the various functions that the terms have. In order to do so, however, the parser needs extra-sentential information, i.e. information not specified by the input sentence per se, to determine, for instance, which terms have the pragmatic functions of Top(ic) and Foc(us). In a language that does not have specific markers for these functions, the parser must at least have access to knowledge about the context in which the utterance was made. Only then can it determine which information is most relevant in the situation, i.e. focal, and which entity it is that something is predicated about, i.e. topical. Since the parser is not specified as having access to knowledge base D, this will be difficult (if not impossible).

4.2. Processes

Apart from the fact that none of the processors 1 to 4 is linked up with knowledge base D - hence, that input processing will be extremely difficult- the claim that all the knowledge is stored in or deducible from D boils down to the denial of any processing whatsoever. Although this cannot have been intended, of course, the above mentioned claim shows a serious misunderstanding of what, in fact, constitutes a "processor". This confusion may have arisen from the very use of the misleading term "processor" (and of "data store", for that matter).

The term "processor" may have suggested a certain physical entity that exists independently of the rules that govern its actions. This in turn may have suggested that a processor could be described (and represented in a figure) as independent of any knowledge. We have to bear in mind, however, that we are dealing (largely) with computer programs. What in the model is called a "processor" is, in fact, a computer program. What such a program describes (or constitutes) is not called a processor but a process. A process is series of operations that are to be performed on its input. Since processes are devised in such a way as to perform according to certain rules and principles, a process itself represents (or constitutes) a body of knowledge of how to perform certain actions. Any computational system, therefore,

consists of two types of knowledge. On the one hand, there is knowledge of how to perform certain operations on data. On the other hand, there is the knowledge as represented by the data. Without this first type of knowledge there is no process, and thus no processing.

Since this first type of knowledge is completely left out of the presentation of the model, it cannot be determined whether the knowledge of D should only be taken as representing data for the different processes - in which case it should be made accessible to these processes - or whether (some of) that knowledge should, in fact, be thought of as constituting (partial) processes.

4.3. Long-term versus short-term knowledge

Figure 4 and figure 1 give quite different pictures of how short-term knowledge (STK) comes to be long-term knowledge (LTK). In figure 4, the predictions based on LTK are matched against the analysis (which constitutes STK). The final product of this matching is then evaluated by the Evaluator for its retention value. The interpretation settled on is then either integrated into LTK or discarded. Although process 6 in figure 4 has exactly the same functional description as the Evaluator in figure 4, there are some important differences, since process 6 mediates between C and D, whereas the Evaluator operates on data already stored in D (both LTK and STK are part of D).

With respect to process 6, it is also said that knowledge representations do not usually contain illocutionary operators, etc. That is to say, knowledge representation consists of underlying predications (UPs), not of fully specified underlying predications (FSUPs). This claim, however, can only hold for LTK, and not for knowledge representations is general, since STK consists of FSUPs. If D only consisted of UPs, processor 6 should somehow convert its input (i.e. FSUPs) into UPs before it could store them in D. To do this it would need Functional Logic (FL). FL, however, only operates on data already in D. The resulting situation would be truly paradoxical. FL cannot operate on predications unless they are in D, and process 6 cannot store predications in D unless FL has already processed them.

Apart from the differences between process 6 and the Evaluator, there are also problems common to both. The first problem concerns the unspecified conditions that make them decide what to retain in long-term memory (LTM). Suppose that FG-CNLU is offered some input that it has encountered before under exactly the same circumstances. Does the "gatekeeper" (i.e. either process 6 or the Evaluator) then decide to discard the newly input or does it decide to store the new "fact" that it has "heard" or "seen" the same statement twice? If the latter is the case - which is the more realistic of the two options - then the gatekeeper is capable of generating new knowledge. This aspect has been completely overlooked in the presentation of the model. In the model only the analyses of the input processes and the knowledge that FL is able to deduce are considered sources of (new) knowledge.

The second problem is that it is by no means clear whether the new knowledge that FL is able to deduce is stored after deduction - and by which process - or whether the same deduction is made over and over again each time it is needed. The latter seems to be the case, for there is nothing

that indicates that knowledge generated by FL is passed on to process 6 for renewed integration into the existing knowledge structure, processes 6 and 7 not being connected to each other.

The third problem is closely related to this. It concerns the way in which new information is integrated into the existing knowledge structure. Since LTK is stored in a network of thematically organized files that are connected by multiple cross-references, new information can, in principle, affect the entire organization of the network: every piece of knowledge is related to all other pieces, directly or indirectly. The problem, then, is how to limit the amount of restructuring of the network. Especially in light of the fact that new knowledge can come either from without (i.e. through input) or from within (i.e. through processes 6 and 7). This problem is a variant of the well-known "frame problem" (McCarthy & Hayes, 1969). Let us briefly discuss the technical side here.

No mechanism has been postulated that tells FL when to stop or when to start deducing. Nor is there a mechanism that tells FL what to deduce. So, FL either does no processing at all or it starts making random deductions from any new input that appears in LTM. (Notice that the latter possibility is unlikely because process 6 and 7 are not in contact with one another (witness figure 1), hence FL does not know what is new.) Suppose, however, that FL could somehow be coaxed into action. It would then start making every possible deduction from the new input, with the logical properties of that input as its only guidance. Since every other piece of knowledge already stored could in principle be brought to bear on the new knowledge, FL would be deducing new knowledge endlessly. This new knowledge, moreover, would have to be integrated again. This newly-integrated knowledge in its turn would allow new deductions to be made. In short, the system would get trapped in an infinite loop.

5. Figure 1 versus figure 4

Apart from the differences between process 6 and the Evaluator, there are also other differences that do not make it immediately clear whether the interpretation model of figure 4 can be realized on the architecture of FG-CNLU as represented in figure 1. First, as we have seen, interaction between the processes is not possible. Second, although we might perhaps equate the Deducer with FL on account of its being the only process postulated as making deductions from the contents of D, and although we might (perhaps) equate the Analyser with processes 1 to 4 as a unit on account of their being involved in input processing - which do not seem very impressive reasons, all in all - we would still be missing a Matcher.

6. The parser(s)

The parser of spoken language takes phonetic representations as input and delivers FSUPs as output. Before it can compute the structural relations that hold between the constitutive elements, however, it must first translate the phonetic representations into something it can handle. That is to say, it is probably not intended that parsing itself should be performed on phonetic representations - for then there should be different sets of knowledge predications for dealing with written and with spoken input. To make these translations, the parser should have access to a lexicon that pairs phonetic representations of words to predicate frames and grammatical elements,

respectively. (Although no such lexicon has been postulated as being used in the interpretive process as a whole, this, of course, could be the same lexicon that was suggested as being used in the wider system of FG-CNLU for the pronunciation of words.)

When the parser has retrieved the predicate frames and grammatical elements corresponding to the phonetic representations, it can start the actual parsing. It will, however, have need of morpho-phonological knowledge as well. The ending -s, for instance, could signal either a plural noun or a third-person-singular verb. (As we have seen, it is not particularly clear, given the model as in figure 1, whether the parser has access to, or possession of, the knowledge required so far.)

The only suggestion made for the parser concerns the next step in the process. It is suggested that the parser should first identify the main verb, then retrieve its predicate frame, and finally utilize the rich information contained therein to compute the structural relations between the terms. It will now begin to dawn why this proposal will never work. First of all, for the majority of languages the verb is not the first constituent of a sentence or clause (Greenberg, 1966; Ultan, 1969). Before the verb has even been encountered, therefore, the predicate frames of the "words" preceding it will have been retrieved. Second, the parser does not know what the verb is until it has actually made a few parses. Consider what happens when a sentence is parsed, for instance a sentence from an SVO language like English:

(6) The granite rocks fell

(Whether the parser first retrieves all the predicate frames before it starts parsing, or whether it retrieves them one by one as they come in from process 1, is immaterial to our present concerns.) Suppose, however, that at this stage the parser has the following list at its disposal:

(7) /the/ $= d$
/granite/ $= \text{granite}_N \ (x_i)$
/granite/ $= \text{granite}_A \ (x_i)$
/rocks/ $= \text{rock}_N \ (mx_i)$
/rocks/ $= \text{rock}_V \ (x_i: \text{3rd.pers.sing.})_{Ag} \ (x_j)_{Go}$
/fell/ $= \text{PAST fall}_V \ (x_i)_{Proc}$

The parser cannot decide beforehand whether "rocks" or "fell" should be taken as the main predicate, it can only try both. (Whether it tries them both at the same time, or one after the other, will again depend on the actual design of the parser.) One may object, of course, that this type of "garden path" sentence is the exception rather than the rule. However, the same thing would happen, ceteris paribus, if the parser attempted the non-garden path sentence "the rocks fell". Here too, there are two candidates for main-predicate status.

It will be clear that the amount of processing performed before the verb is encountered is substantial. It will also be obvious that encountering the verb is not the same as identifying it, and that identification itself takes a considerable amount of processing. Moreover, the processing load will increase considerably when sentences from SOV languages are parsed, when sentences contain multiple plural nouns ending in -s, or for any number of reasons that it may amuse one to imagine.

Another point is that, if the parser retrieves the (corresponding) predicate frames from the lexicon, work is needlessly duplicated. Since the

phonetic processor should access the lexicon to see whether its (tentative) analysis is part of the system's vocabulary, it could at the same time retrieve the information stored there (i.e. the predicate frames). This, in fact, is the standard assumption (see e.g. Garnham, 1985).

For the parsing of written language largely the same criticisms apply, although in this case a kind of network has been postulated that maps (typed in) morphemes onto predicate frames and grammatical elements (see Dik, 1987b: 19, based on Koskenniemi, 1983 and Karlsson & Koskenniemi, 1985). This, of course, does not explain how the graphemic input of the parser eventually comes to be parsed.

7. Conclusion

We have seen that many parts of the model that are essential for its proposed functioning are, as yet, missing. We have also seen that many parts are missing that have not been proposed, but that are essential all the same. So far, the absences may (perhaps) be attributed to the fact that work on the model has only just begun. What is far more disturbing, however, is that on a number of major points the (different) proposals are incoherent, and even contradictory or downright impossible.

Although the model itself does not live up to the design expectations, i.e. cannot perform as intended, none of the technical criticisms damages the theory in any way. What is seriously damaged, however, is the possibility of making predictions from the model. Too many important stages are missing to determine accurately what the system would do if it worked.

Chapter 4

The psychological adequacy of FG-CNLU: part 1

1. Introduction

In the previous chapters we have seen that the currently available descriptions of the FG-CNLU model only sketch very broad outlines. We have also seen that many of the theoretical proposals are not reflected in the architecture of FG-CNLU as presented in figure 1, partly because many necessary links are missing. Given the lack of (technical) detail and the infeasability of the system under this description, it will be obvious that any effort to draw accurate and specific predictions from the model is seriously handicapped. Moreover, many predictions that one could (seem to) make on the basis of the model would, in fact, be influenced as much by what the model **does** specify as by what it does **not** specify but only hints at.

The difficulty of drawing detailed predictions in its turn makes it difficult to test the psychological adequacy of the model, precisely because we cannot be sure what was intended or could have been intended. Such efforts are even more handicapped when we consider that no account is given of the exact behaviour that the system is supposed to model. What is understood by "human communicative behaviour", which in itself can hardly count as a well-defined area, is not specified.

Despite these difficulties, we will nevertheless attempt to test the adequacy of the model as a psychologically real account of the language comprehension and cognition involved in communication. We will do so by following up on the consequences of what is claimed in/of the model. Claims "in" the model, roughly speaking, are those about which the FG-CNLU literature is to a greater or lesser extent clear. Claims "of" the model are those presented in the texts but not reflected in Figure 1. Both types of claims will be taken in their strongest form, modified only by the actual further considerations made in the model. For instance, when all conceptual knowledge is said to be linguistic in nature this will taken to be so. The consequences, which we will reluctantly call "predictions", will be compared with known experimental findings and/or discussed in the light of more theoretical considerations.

2. Predictions from the linguistic nature of conceptual knowledge

The major claim of the model is, of course, that conceptual knowledge is represented linguistically. A number of more basic predictions follow from this claim, some of which can be further subdivided in turn.

2.1. Conceptual and procedural knowledge

The first prediction that one could base on the model is that people who have never learned a language (including sign-languages), such as the congenitally deaf, should be completely mentally retarded. This prediction follows from the assumptions that all conceptual knowledge is

    a) stored,
    b) reasoned with

c) acquired in linguistic form.

Assumption c) follows from the suggestion that the actual input sentences (minus the syntactic and pragmatic functions of the terms) are stored in memory, hence are the source of knowledge. Assumption c) also follows from the fact that, since languages must be learned, no (conceptual) knowledge can be innately present. Let us consider in some detail why assumptions a) to c) lead to this prediction.

In general, a distinction can be made between declarative and procedural knowledge (e.g. Ryle, 1949). This distinction, according to Winograd (1975), is a revival of the ancient distinction between "knowing that" (i.e. facts) and "knowing how" (i.e. skills). In the computer metaphor of human reasoning, thought is characterized as the manipulation of facts by means of procedures, where one procedure may also take other procedures as facts (i.e. as its arguments). Consequently, if all conceptual knowledge is linguistic in nature, both facts and procedures are such. Therefore, non-linguals should lack both the facts and the procedures to manipulate them, hence do not have conceptual knowledge and cannot think.

It will be clear that this prediction is false. The congenitally deaf are not only not retarded by definition, it can even be shown that they classify, categorize, and form concepts (in other words, think) in much the same way as their speaking/hearing peers (Furth, 1966). Moreover, because of the circularity of the argument, assumptions (a-c) would also lead to the conclusion that people who still have to learn a language, i.e. pre-linguistic children, could never be able to perform this miraculous feat, because they would lack the required cognitive abilities. Since the capacity to learn requires an initial state of being able to learn, there must be some abilities that are not and cannot be learned. Thus, the knowledge required must be innate (e.g. Fodor, 1975; cf. also Chomsky's "Language Acquisition Device", 1972) - hence cannot be linguistic. Since languages must be learned, the ability to learn them cannot be provided by knowledge that is already represented linguistically. Because pre-linguals are able to learn languages, and because non-linguals think, one can only conclude that language is not necessary for thinking (e.g. Lenneberg, 1967; Fodor, 1975; Glucksberg & Danks, 1975; Clark & Clark, 1977).

One of the properties that distinguishes procedural knowledge from knowledge of facts is that procedural knowledge cannot be communicated verbally (e.g. Ryle, 1949). Consider "motor knowledge", for instance, which is a special type of procedural knowledge, and which has been said to call for "enactive representations" (Bruner, Olver & Greenfield, 1966). Motor knowledge controls muscle activity, and is used in walking, riding a bike, in speaking and writing, etc. It will be clear that no one is able to give an exact report of what happens during these activities, and that the processing involved cannot be brought to the conscious. As these activities are difficult to report on, they are also difficult to instruct. Consequently, skills come gradually and with practice - whereas one can learn facts by having been told them once.

It will be clear that much of what is listed in Dik's (1986a; 1987b) typology as being conceptual knowledge cannot be verbalized. Basically, two explanations can be given to account for this. One possibility is that large amounts of unreportable forms of knowledge are procedural in nature. Grammatical linguistic knowledge of the type represented in (constituting) the rules of a parser, for instance, is a likely candidate. Knowledge of how words are combined to form larger units such as clauses and sentences is internalized, as it were, in procedures to parse clauses/sentences - not in the ontogenetic sense of the word, though.

32

Another possibility is that the process of reporting does not have access to some forms of knowledge, irrespective of whether this knowledge is declarative or procedural. This is called "cognitive impenetrability" (e.g. Pylyshyn, 1980; 1984; Fodor, 1983). In this approach it is possible for something to be represented as a fact (i.e. declaratively) without being reportable. (More about this approach will follow in the next chapters.)

If, however, conceptual knowledge were linguistic in nature, and if, as in the model, no principled constraints were posited as to what information is available to what processes, then there would be no way to account for the fact that many forms of knowledge cannot be verbalized. One has to remember, in this respect, that the knowledge of the system is stored in a cross-referenced network where, directly or indirectly, each piece of knowledge can be reached from all other pieces. Moreover, it is explicitly claimed that the tentative analyses of the input form the Short-Term Knowledge (STK) of the system, and that this is stored in Short-Term Memory (STM). Since STM, or "working memory" or "attention" as it is sometimes called, is the conscious part of memory, one should be aware of all the tentative analyses that one makes of the input before an interpretation is settled on. One should be conscious of all the possible semantic and syntactic ambiguities that words, phrases and sentences may have. Introspection shows, however, that this is not the case, and so do experiments (see e.g. Foss & Hakes, 1978; Fodor, 1983). In sentences such as (1), for instance, one will not normally become aware of the alternative meaning of "straw", hence the sentence will only receive a "barnyard" interpretation (Foss & Jenkins, 1973):

(1) The farmer put the straw beside the machine

Also, if conceptual knowledge were linguistic in the sense of the model, and therefore reportable, building a language understanding system would be much easier than it is now. Every Tom, Dik or Harry would be able to specify the rules along which it operates - simply by giving an exact account of his own processing during comprehension.

2.2. Bilinguals

The linguistic nature of conceptual knowledge would also commit one to the view that people who speak several languages should have separate conceptual systems for each of these languages. Moreover, each of these conceptual systems would (presumably) have its own knowledge base. Finally, each of the conceptual systems would have its own interface with the perceptual apparatus.

The first hypothesis has been tested by Kolers (1966), using a list-learning task. In such a task, subjects are asked to memorize lists of unrelated words, which they have to recall later. A well-known fact concerning such tasks is that repetition of a word increases the likelihood of its being remembered. Kolers found out that for English/French bilinguals a translated repetition (e.g. ten - dix) acted just like a verbatim repetition. Since these subjects took no longer to recall the lists of words (than monolinguals on similar monolinguistic tasks) - which one would expect if two separate memories had to be queried - or to memorize them, this provides evidence that the information was stored in a single uniform way in a single memory.

The second hypothesis has been tested by Preston & Lambert (1969),

using a version of the "Stroop" colour word test (Stroop, 1935). In this test subjects are presented colour words that are printed in coloured ink (e.g. the word red in blue ink). They are then asked to name the colour of the ink as quickly as possible. For monolinguals this is quite difficult because the tendency to read the word interferes with naming the colour.

Preston & Lambert reasoned that bilinguals should be able to do better on this test if they could "turn off" one of their linguistic conceptual systems. (The rationale of this condition lies in the fact that bilinguals are not normally hindered by thoughts in language B while, for instance, reading language A. The two languages, thus, are not "switched on" at all times.) For the English/French bilinguals that they tested, Preston & Lambert argued as follows. If one were able to switch off one of the conceptual systems, it should be possible to ignore the word red and answer the question in French with bleu. The subjects, however, performed no better on this task than monolinguals. This was taken as evidence that one is not able to switch off a conceptual system - even if it might be useful. Since on the other hand it is possible to switch off a language, this was interpreted as indicating that in bilinguals the two languages converge upon a common core, and therefore, that there is just one conceptual system for both languages.

One may object, of course, that it is difficult to see how one should differentiate between the linguistic and the conceptual systems in this task. Nevertheless, this experiment does indicate that if bilinguals do have two conceptual systems, these systems are at least not completely separated.

2.3. Conceptual and perceptual knowledge

If conceptual knowledge and perceptual knowledge are stored separately, one in the form of predications and the other in pictures, and cannot be inter-translated (as in the present model), then one could predict that people should always be able to determine whether a particular piece of information was received perceptually or verbally. (This could be done, perhaps, simply by checking the representational format, picture or predication, of the to-be-remembered item.)

This prediction has been tested by Rosenberg & Simon (1974), using a "sentence-picture verification task". In such a task, subjects are usually presented with a set of pictures, and have to decide whether they are congruent with what is said in the corresponding sentences. The time needed to make this judgement is then measured, and taken as an indication of the amount of processing involved, hence of the complexity of the task. In this case, however, subjects were given an additional surprise recall test afterwards. The results were that people confused whether they had heard a sentence or seen a picture that represented event X. Although these results do not exclude the possibility that people sometimes, or perhaps often, are able to remember whether the input was verbal or perceptual, it falsifies the prediction that they always do so.

These and similar results have been taken as indicating that there is just one representational format, and that it is independent of the input mode(s) (e.g. Fodor, 1975; 1983; Anderson, 1976; Pylyshyn, 1980; 1984). Whether or not people are able to remember the input mode of a certain item at a certain moment can then be viewed as a matter of their ability/inability to retrieve the relevant information, rather than as a matter of whether it has or has not been specified in one form or another somewhere.

In all fairness we have to mention that facilities for inter-translating

the two types of knowledge are mentioned as further extensions of the model (Dik, 1987b: 25). The above results could be explained in the model, therefore, by assuming that the possibility of its having been translated makes it difficult to decide for a particular representation whether it was received in verbal or perceptual form (Dik, personal communication). What is troublesome for this assumption is, in turn, that people sometimes do remember the (original) input mode. The difficulty of explaining the latter lies largely in the fact that the model only stores representations of the input and does not keep accounts of the processing involved in the derivation thereof. Therefore, the model is principally incapable of checking such potential sources of information.

There are other experimental findings, however, that are far more easily explained by assuming that information from different sources is used to build up mental representations in a single code than by assuming two codes. Moreover, there are theoretical and practical model-building concerns. First, the number of translation devices that one has to posit in order to account for interaction increases if there are two codes. Second, the amount of translation that these devices will have to perform also increases drastically. One would also have to posit duplicate mechanisms for all other thought processes (e.g. for reasoning with different types of representations, for retrieving them from memory, etc.). An experimental example may serve to illustrate how verbal input may affect already existing knowledge structures of perceptual origin, and to show that the end result is most easily explained by assuming a single format.

Loftus has made a number of studies on eye-witness reports. Loftus & Palmer (1974), for instance, showed subjects a film of a multicar accident, and afterwards had subjects answer a number of questions about the film. One group of subjects was asked, for instance, what their estimate was of the speed of the cars when they "hit" each other. Another group was asked what their estimate was of the speed of the cars when they "crashed" into each other. The speed of the cars was estimated significantly higher by the "crashed group" than by the "hit group". Considerable time later, subjects returned for another round of questions. This time they were asked, for instance, whether they had seen "the glass flying around" in the film (in the film there had not been any). Again, the "crashed group" answered this question significantly more often with "yes" than the "hit group".

To account for this, Loftus & Palmer assumed the following. Subjects first built up a memory representation of the events in the film. When asked later to estimate the speed of the cars, the wording of the question clearly influenced the answers that subjects gave. Apparently, the abstract idea or schema (cf. Bartlett, 1932; Rumelhart & Ortony, 1977) or frame (Minsky, 1975) or mental model (Johnson-Laird, 1983) that people have of "crash(ing)" and which they apply in order to account for their experience involves a "high-speed" component. The answers given at the later session indicate, moreover, that the original memory representations of the second group had been changed to incorporate the idea of crashing, which in turn led them to "remember" flying glass - apparently, another component of the crash schema. The use of the definite article "the" in the later question, suggesting a known and identifiable referent, not only made subjects believe that there probably had been glass flying around - which would explain why some of the "hit group" also answered "yes" to this question - but also fits well with the general account of the film in terms of the crash schema.

Memory representations that incorporate perceptual and verbal aspects (e.g. schemas, frames or mental models) can, therefore, relatively easily account for the influences that one source of information may have on the

other by assuming that the same type of more abstract representation results from them. Additional processing-levels for translating one form of knowledge into another - which complicate matters and are difficult to test empirically - can be left out in this manner. The assumption of a single memory code, therefore, also adheres to the scientific practice of long standing in which the simplest explanation that can account for a body of data is preferred.

3. Lexical definitions

Another problem-area for the model arises from the assumption that words, and therefore concepts, are defined in sets of more basic predicate-frames by the principle of "stepwise lexical decomposition" (Dik, 1978b). One may not only object to definitions in general and to the definitions given, but also to the supposed linguistic nature of these definitions.

3.1. Definitions

As concerns definitions, there is the general issue of whether definitions can adequately capture the totality of the denotative meanings of words. We may question with Wittgenstein (1953), for instance, whether natural concepts can be defined at all. A well-known example of his in this respect is that there is no defining characteristic that all games have in common. We may also wonder whether we should not rather think of concepts in terms of "family resemblances" or "prototypes" (Rosch, 1973a; 1978). A robin, for instance, is a more typical bird than an ostrich and therefore a better example of the concept "bird". Next, we may observe with Fodor et al. (1980) that no satisfactory definitions have ever been proposed. Consequently, it is even more difficult for a definitional approach to provide adequate accounts of the meaning of the structures that are built from these ill-definable parts, namely clauses, sentences and texts.

Secondly, definitional accounts of word-meanings do not and cannot capture their non-denotative aspects. Consider Dik's (1986a: 18) definition of wife, for instance:

$$\text{wife}_N\ (x_1:\ <\text{human, female}>(x_1))\ (x_2)\ =_{def}\ \text{married}_A\ (x_1)\ \text{and}\ (x_2)$$

This definition of the word wife includes neither its associative meaning nor its affective meaning. The associative meaning of a word is "the sum total of all the things a given person thinks of" (Deese, 1970: 109). The affective meaning of a word is measured by the "Semantic Differential" (Osgood, Succi & Tannenbaum's, 1957). It measures whether a word is evaluated as good or bad, active or passive, weak or strong, etc.

Definitional accounts, then, can succeed only partially in covering the denotative meaning aspects of words, and not at all in capturing their other meaning aspects.

3.2. The linguistic nature of the definitions

There are also problems concerning the linguistic nature of these definitions, in particular with respect to the acquisition of words.

Generally speaking, the definiens of a definition must be simpler than the definiendum (e.g. Dik, 1978b: 24). What "simpler" means in this respect is

usually not very clear, but the least requirement one could make would be that the defining concepts must be known already. That is, the definiens must be known before the definiendum.

With respect to children's acquisition of language - for which it is assumed that semantically simpler terms are learned first (e.g. Clark, 1976)- it is immediately obvious, however, that the proposed definitions violate this principle. Consider the following definition (from Dik, 1978b: 30):

$$\text{man}(x_i) =_{def} \text{male } (x_i: \text{person}(x_i))$$

If <u>man</u> is defined as <u>male person,</u> then the latter should be simpler. Consequently, the latter words should be learned first (and, perhaps, used together) by children. Consider the following, however.

There is evidence that children perceive differentiation in the sexes as early as in their second year (Van Tilburg, 1987). This can be taken as indicating the emergence of concepts like male and female. However, the word <u>man</u> clearly appears earlier in the vocabulary of children than <u>male person,</u> both in comprehension and in production. In terms of Rosch & Mervis (1975), the word <u>man</u> names a basic level category and is learned before its hypernym <u>person</u> and/or its hyponyms (e.g. <u>bachelor</u>).

Moreover, the priorities in acquisition as defined by definitions like that of <u>man</u> would ignore the following observations. First, language acquisition develops in stages from one-word utterances to two-word utterances, etc. and not the other way round. Second, the early utterances of children most frequently consist of what linguists would call nouns and not of combinations of adjective plus noun as in <u>male person</u>.

Definitions like these then do not correctly reflect their supposed "simplicity", in terms of what is learned first. Also, it appears that concepts like those that adults name by means of particular words are/can be present before the child has acquired the word to express them. Acquiring a concept and a word/words for it do not go hand in hand, as the model suggests, but the former precedes the latter. This is quite natural, of course, since it is only when something is conceived of as being conceptually distinct from the background that the need arises for a means to refer to it. Moreover, as a means of referring, pointing precedes uttering the appropriate word. This also proves, by the way, that children do "have something in mind" to which they wish to refer even though it is non-linguistic in nature. From this evidence one can only conclude that concepts are not defined linguistically.

3.3. The nature of "the act of defining"

Not all arguments against definitions, however, depend on disputing the particular definitions proposed. Consider the following.

One of the arguments used by Dik to defend lexical definitions is that they come quite close to what native speakers respond with when asked what a word means, or that they come close to dictionary definitions. The implicit suggestion is that the act of defining a word is like reading its definition from the lexicon. If this were true, one would expect the (uttered) definition for a given concept always to be the same.

The giving of a definition, however, can be considered a type of speech act in the sense of Austin (1962) and Searle (1969) in that one does something with words. Much more is involved in this act than the rather passive "reading-out-aloud" account suggests. The kind of definition that one gives of a certain concept is highly dependent on pragmatic factors. It

depends, for instance, on the situation in which the definition is uttered (e.g. to which person, on what occasion, for what purpose). Definitions given in classroom situations, for instance, will differ from the more context-free definitions given in dictionaries. In the former case, the purpose may be to show the teacher that one has done one's homework. In the latter case, the purpose is to inform the general public. The compilers of dictionaries, however, will also have to estimate what they may consider as known or "given" information and what as "new" (cf. Haviland & Clark, 1974), and they may have to provide sentential contexts.

In general, then, defining is quite similar to referring. Olson (1970), for instance, has suggested that we refer to something by giving just enough information to distinguish it from the "background" - no more, no less. So too, one defines something by giving the addressee the amount of information that we estimate as being required in the particular situation.

Given this contextual dependency, there is no reason to assume that lexical definitions have psychological plausibility (cf. Fodor et al, 1975; 1980, for a different argumentation leading to the same conclusion). The fact that people are able to produce definitions that are close to the proposed theoretical constructs does not in and of itself prove that such definitions are actually stored in the lexicon. It only proves that the concept described by the definition is more or less the same for informant and theorist alike.

4. Predictions from the storage of UPs in memory

The FSUP that the language processor derives from the input sentence is stored (in slightly modified form) in LTM - given sufficient retention value. Three types of prediction can be based on this aspect of the model.

First, it may be predicted that some time after hearing/reading a sentence subjects should be able to recognize (if not to recall) the (parts of the) original input sentence that are specified by the stored UP. More in particular, they should recognize

    a) the grammatical categories of the constituents (what was used as a verb, a noun or an adjective);
    b) the exact words that were used to express these notions (both for the verb and for the arguments);
    c) the semantic functions of the arguments;
    d) what was expressed in one sentence and what in others.

Secondly, since UPs are also the input of Functional Logic (FL), one could predict

    e) that it should take longer to reason with information provided in complex sentences than with the same information provided in simple ones.

Thirdly, another body of predictions can be based on the fact that UPs are the only representations stored in memory, hence that all information not specified by them is lost for further use. Since UPs do not indicate, for instance, whether the original input was heard or read, the model predicts

    f) that people should never be able to remember the input modality of sentences.

Also, since UPs do not reflect certain types of processing difficulties that may have occurred - and since no processing account is kept - people should

g) never remember physical attributes of the input (e.g. bad handwriting, whispering voice, etc.).

4.1. What people should remember of input sentences

Prediction a) would be falsified if it could be shown that people do not always remember whether a particular idea was expressed as a verb or as a noun. That this is indeed the case has been demonstrated in an experiment by Johnson-Laird, Robins & Velicogna (1974). In this experiment, subjects were confused as to whether it was (2) or (3) that had originally been presented:

(2) The owner of the magic staff **dispatched** the ship
(3) The **dispatcher** of the ship **owned** the magic staff

Confusion between sentences (2) and (3) indicates, moreover, that the semantic functions of terms are not remembered well either. (Semantic functions are determined by the predicates in which they occur.) The owner is the Agent term of the verbal predicate dispatch in (2), whereas the dispatcher, which refers to the same entity, has zero semantic function in the verbal predicate own.

The matter is even more clearcut in examples (4-5) (from Johnson-Laird & Stevenson, 1970):

(4) John **liked** the painting and he **bought** it from the duchess
(5) The painting **pleased** John and the duchess **sold** it to him

Although (4) and (5) express different States of Affairs (SoAs) in terms of who is in control of the situation (he is the Agent of the predicate buy in (4), whereas the duchess is the Agent of sell in (5)), subjects apparently do not store such information directly in the form of which term has which semantic function in which predicate. Prediction c), therefore, has also been falsified.

Prediction b) would be falsified if one were to show, for instance, that people (may) confuse sentences that contain words not present in the input sentence with the original. Schweller, Brewer and Dahl (1976) demonstrated that people confuse having heard (6) or (7):

(6) The cute little girl **told** her mother she **wanted** a drink
(7) The cute little girl **asked** her mother **for** a drink

Prediction b) would also be falsified if one could show that people mistakenly report having heard a sentence which, in fact, they have not heard at all. The latter has been reported in a number of studies, especially in work on linear arrays (e.g. Barclay, 1973; Potts, 1974). Barclay presented subjects with sentences that described part of the relationships between the elements of a linear array such as that in (8):

(8)   lion bear moose giraffe cow

Thus, subjects might hear sentences such as (9) or (10).

(9)   The bear is to the left of the moose
(10)   The cow is to the right of the giraffe

In a surprise recognition test on the original sentences, it became apparent that those subjects who had been asked to figure out the order of the array responded "yes" or "no" on the basis of whether the test sentence was true or not of the array, irrespective of whether it had been presented before. In other words, they also (mistakenly) reported having seen some sentences before which in fact they had not.

In order to falsify d), one would have to show that information from different sentences may be reported as belonging to the same sentence. Bransford & Franks (1971) did just that. They presented subjects with sentences in which either 1, 2, or 3 of the following types of propositions occurred:

(11a) The rock rolled down the mountain
(11b) The rock crushed the hut
(11c) The hut is at the river
(11d) The hut is tiny

Later, subjects were again presented a set of sentences and were asked to indicate which sentences they had seen before. Moreover, they had to rate their confidence in these judgments. The results were that subjects were most confident in having previously seen sentences in which all four propositions occurred, such as (12), which had not in fact been presented.

(12)   The rock that rolled down the mountain crushed the tiny hut at the river

These results, replicated by Cofer (1973), Flagg, Potts & Reynolds (1975), and Singer (1973), not only prove that information from different sentences is incorporated into larger wholes (e.g. into stories), but also that (normally) sentence representation is in no sense verbatim.

4.2. UPs as the arguments of thought processes

Prediction e) has been tested by Kintsch & Monk (1972) and by King & Greeno (1974). They argued that if memory for sentences is more or less verbatim, then it should take longer to reason with complex sentences than with simple ones. The argument is based on two assumptions:

a)   "reasoning" is the manipulation of representations
b)   complex sentences receive complex representations

To test the hypothesis they gave subjects instructions to perform a particular task. The set of instructions were couched either in (syntactically) simple or in complex sentences. In both studies the results were that, although the second group took longer to understand the instructions (as measured by the time it took subjects to signal that they were ready), performance on the actual task was equally fast and accurate for both groups. This can only be explained, of course, by assuming that the representations with which subjects reasoned were (highly) similar for both

groups, and thus that lexico-syntactic differences had disappeared.

4.3. What people should not remember

Predictions f) and g) are based on what UPs do not specify, and concern matters that are largely non-linguistic in nature.

In the model a sharp, though ill-defined, distinction is made between "verbal" and "perceptual" knowledge/information. We will assume, therefore, that these types of information are processed by different mechanisms. It has been argued, however, that memory for physical attributes of sentences (including input modality) is a natural by-product of the language-understanding processes involved, even though it is not the major concern of the language processor (e.g. Kintsch, 1977). It has, in fact, been established that such information is available on a short-term basis (Brooks, 1968), but it is less clear whether it normally remains so over longer periods of time (Clark & Clark, 1977).

In order to falsify f) and g), however, we need only show that people sometimes are able to remember such aspects - and thus that the language processor does specify the relevant information - irrespective of whether under "normal circumstances" people are always able to retrieve that information later.

That people are able to remember the input modality of verbal material has been shown in a number of experiments (e.g. Hintzman, Block & Inskeep, 1972; Bray & Batchelder, 1972). In the study by Bray & Batchelder, for instance, subjects could remember which words in a list-learning task had been heard and which had been read (both when the test was given immediately and when it was given after a 15 minutes' delay, both when subjects had been warned about this test and when not). They could even remember the modality when they could not recall the item itself. What the latter result indicates, moreover, is that input modality is not stored with the "sentence representation" (e.g. in a label attached to it), but separately and independently (cf. also Kintsch, 1977).

5. Problems for a slightly revised model

It will be eminently clear from the above that the FG-CNLU model does not give a (psychologically) realistic account of what people store in LTM, of what they (should be able to) remember, and of what they reason with. It should be noted in this respect, however, that even Dik himself (personal communication) does not hold the views expressed in the model. That is, he does not believe that the actual UPs that are derived from the input sentences are stored in memory (and serve as input to FL). Instead, he suggests that the FSUPs or UPs may be broken down into more basic predications, and/or converted into perceptual representations, before they are stored in memory - although he largely leaves open how this is done.

To account for the confusion between (2) and (3), for instance, he suggests that both may eventually be stored as two predications of the form "x own ship" and "x dispatch ship".

To account for cases of confusion as between (13-14) (from Barclay, 1973),

(13) John is taller than Peter
(14) Peter is shorter than John

he suggests that, because "taller than" and "shorter than" are each other's converses, subjects may infer one from the other and not remember what the original was.

Confusion between (15-16), as reported by Bransford, Barclay and Franks (1972), is explained by assuming that both are translated into the same perceptual picture.

(15) The woman stood on the chair and the mouse sat beneath **it**

(16) The woman stood on the chair and the mouse sat beneath **her**

(17-18) on the other hand, in which "beneath" was changed to "beside", and which did not confuse subjects when they were asked to indicate which one they had seen originally, would be represented in different pictures.

(17) The woman stood on the chair and the mouse sat beside **it**

(18) The woman stood on the chair and the mouse sat beside **her**

The above suggestions, however, meet with problems too - apart from the problem of not being compatible with the present model.

First of all, it is still maintained that conceptual knowledge is verbal in nature, which as we have seen is not very likely if only because some people (e.g. the congenitally deaf, pre-linguistic children) manage do to without language.

Secondly, on this inferencing account, many of the types of confusions that people actually make would receive rather ad hoc solutions. To explain mistaken reports of identity between (6) and (7), for instance, one could assume (Dik, personal communication) that "ask (for)" is represented in the lexicon with a definition "tell someone you want something". This definition, however, clearly leaves out many of the finer nuances, and certainly cannot be substituted for "ask (for)" in every context. Moreover, even though such a lexical definition would explain how one can be inferred from the other, it does not explain which one is actually stored. One must not forget in this respect that although in all the above mentioned experiments significantly large numbers of people confused test sentences with originals, not everyone did so. It is thus not the case that lexical substitutions with synonyms are made "automatically", or that certain inferences automatically influence recognition. Some of the subjects that did remember which sentences (or part of the sentences) they had seen before may even have made the inferences and still have come up with the correct answers.

Thirdly, psychologically real inferences need not be logically correct ones (cf. e.g. Henle, 1962). It has been shown, for instance, that the inferences that people sometimes confuse for the original sentence need not necessarily be valid ones, but may also be probable ones. Johnson, Bransford and Solomon (1973), for instance, report that subjects confused whether they had seen (19) or (20):

(19)   He slipped on a wet spot and **dropped** the delicate glass pitcher on the floor

(20)   He slipped on a wet spot and **broke** the delicate glass pitcher **when it fell** on the floor

In this case subjects, apparently, reasoned that when a delicate glass pitcher is dropped it usually or often breaks, and therefore that (20) must have been the original sentence. (There is thus a bias towards normality in what people remember.) Such confusion cannot be explained in the model, however, (or in

a model revised along the lines suggested above) because its inferencing apparatus (Functional Logic) is only capable of making "valid" inferences, i.e. deductions. And since there is no way (other than by ad hoc stipulation of the condition that when something breakable is dropped it breaks) in which (20) can be deduced from (19), (19-20) can never be represented by the same set of more simple predications.

6. The parser

Another set of predictions can be based on the proposed functioning of the language processor (LP) itself. Since the parser is the only subprocessor of the LP for which even remotely concrete proposals have been made, we will restrict ourselves to it.

The function of the parser is to retrieve the predicate frames corresponding to the phonetic/graphic representations, and to build (from them) an interpretation that displays the structural and semantic aspects of the input. Moreover, the parser should first identify and retrieve the main predicate. In other words, the parser should work in a top-down fashion. Thus for sentence (21),

(21) John's wife is pregnant

it is suggested that the parser first retrieves the predicate frame of "pregnant", and then uses the information specified therein "to look for a single entity indicating some female entity" (Dik, 1986a: 21). As we have seen (Ch. 3, section 6), this characterization poses two technical difficulties (or better, impossibilities).

First, one cannot explain how the parser knows that "pregnant" is a candidate for being the main predicate unless it already has the predicate frames of the foregoing words available. Only then can it determine that none of these are candidates.

Second, one cannot explain how the parser decides that "pregnant" is the main predicate unless it has actually tried to parse (21) accordingly.

Apart from the technical implausibility, these claims would lead to the psychologically inadequate suggestion that the meanings of individual words do not become available before the main predicate is encountered, identified, etc. It would mean, for instance, that when one hears (21) in the right context one could not start determining who "John's wife" refers to before hearing "pregnant", simply because the predicate frames of these words, and therefore their meanings, would not yet be available. This is not what actually happens, of course.

Besides running counter to intuitive reports of what happens during comprehension - neither in reading or in listening does it seem that understanding only comes at the end of sentences - it would also place serious constraints on the length of sentences that people would be able to understand. Consider the following.

If parsing (and other comprehension processes) does not start until the predicate has been encountered/identified, then the preceding part of the sentence should be held in Short-Term Memory (STM) in unanalysed form for the time being. The maximum length of pre-predicate parts of sentences that people would still be able to understand would then be determined by the capacity of STM for holding unanalysed (and thus unrelated) bits of information. If the number of words that precede the predicate for a particular sentence were to exceed this limit, then the first word(s) of the sentence would have been purged from memory (and thus forgotten) before

actual parsing and comprehension started. As the capacity of STM for holding unrelated units or "chunks" of information has been estimated at 7 plus or minus 2 (Miller, 1956), this would mean that people should not be able to understand sentences in which more than 7 words precede the predicate. Any number of sentences, particularly from SOV languages, will falsify this claim.

However, even if we assume that all predicate frames are retrieved first and that the parser starts building up structures from them as soon as they become available - so as not to exceed STM's capacity of holding maximally 7 units (words or phrases) at a time - it would still be suggested that major syntactic and semantic processing takes place at clause boundaries, because one cannot have all predicate frames available unless the sentence or clause has ended. There are, however, several factors that argue against such a view.

First, there is the existence of single-clause garden path sentences, e.g. (22) (from Milne, 1982):

(22)   The granite rocks during the earthquake

Generally speaking, people report becoming aware of having chosen the wrong interpretation before a garden path sentence has ended. In the case of garden path sentences that consist of one clause, this can only mean that comprehension (including parsing, interpretation, etc) is well under way by that time, and thus that the comprehension process starts as soon as words come in (e.g. Foss & Hakes, 1978; Garnham, 1985).

Second, Marslen-Wilson (1973; 1975; 1976) has demonstrated experimentally that syntactic and semantic analyses are computed before clause boundaries have been reached. In his experiments, subjects had to repeat the sentences they heard over headphones as quickly as possible (a so-called "shadowing task"). Even the fastest shadowers (who lagged only 250 msecs, or about 1 syllable, behind the tape-recorded texts) were able to restore words that had deliberately been corrupted to correct forms that fitted the syntactic and semantic context. Here too, one can only conclude that (partial) analyses must have been made word by word as the clauses were in the process of being processed.

7. Conclusion

The conclusions to be drawn from this chapter are straightforward. Since none of the predictions that the FG-CNLU model warrants are borne out by experimental findings or stand up to theoretical scrutiny, we can only conclude that it is inadequate. From this, in turn, we can only conclude that many of the assumptions on which the model is based are false. Another possibility is, of course, that the aspects taken into consideration in this chapter give only an incomplete account of the model.

Chapter 5

Some aspects of language understanding and cognition

1. Introduction

In Chapter 4 we saw that none of the predictions that can be based on the suggestions that the FG-CLNU model makes is borne out by experimental findings or theoretical considerations. We concluded that there must be serious flaws in the assumptions on which the model is based, some of which we discussed in that chapter. An equal portion of the defects in the model, however, can be attributed to its ignoring many important issues that are critically involved in language understanding and cognition, and which need to be taken into account if one is to build an adequate model. An issue already raised in the previous chapter in this respect was the interfacing of verbal and perceptual knowledge.

Before we can introduce some other issues that the model neglects, however, we must first introduce a general framework in which various aspects of the investigation of human language understanding and cognition can be placed. The introduction of this framework serves two purposes. First, it provides a short survey of work done in the area and, although in no way exhaustive, gives at least a general idea of what phenomena an adequate model will have to account for. Second, the theories placed in this framework offer more plausible explanations for the issues that FG-CNLU makes faulty predictions about, thereby giving us a better view of how human understanding might work, and of what a model might look like.

2. A multi-levelled view of human behaviour

In Cognitive Science - the conglomerate of psychology, Artificial Intelligence (AI), linguistics and philosophy - human behaviour is generally explained in terms of "information processing". Crudely put, man is seen as a robot, and his brain as a computer. What Cognitive Science is basically interested in, however, is not so much the "hardware" as the "software", i.e. the programs that are run. These programs should describe processes that mimic human behaviour in roughly two ways. First, there should be input-output equivalency (Turing, 1950). That is, the models should produce the same output as humans when given the same input. And second, the knowledge structures (abstract representations of information) and the way in which they are manipulated should have psychological reality. In other words they should correspond to human ways of reasoning with respect to their complexity (Pylyshyn, 1980; 1984).

It is customary to distinguish these processes according to the level at which they take place, each level being governed by its own rules and principles. A generally accepted view is that there are three such levels (e.g. Dennet, 1969, 1978; Pylyshyn, 1980, 1981, 1984), although Fodor (1983) recognizes only two. Let us first briefly describe the different levels and then discuss the differences between Fodor and Pylyshyn on this issue.

The basic, or input processing, level mediates between transducer outputs and percepts (Fodor, 1983: 102). This level makes up the "functional architecture" of the system (Pylyshyn, 1984: 120), since it provides the foundations on which the next level is built. The middle level is called the "functional" level because the knowledge types at this level are functional

45

representations of the relations between the input and output of the level below. This functional level is where symbolic representations are manipulated for specific tasks, e.g. problem solving. The top level, called the "semantic" level by Pylyshyn (1984) and the "intentional" level by Dennet (1978), is characterized by the rational connection between "goals", "knowing" or "believing", and "actions", in other words, by "plans" (of action) (Miller, Galanter & Pribram, 1960). It is at this level that "the fixation of belief" takes place (Fodor, 1983: 102), or "the optimal adjustment of one's knowledge to one's experiences" (ibid.: 114).

What Fodor and Pylyshyn agree on is the characterization of the input level. Above that level, however, Fodor recognizes only one further level, i.e. the level of central cognitive processes or "central systems" (1983: 103), where Pylyshyn sees two. The reason for this difference lies in what they view as the domain of Cognitive Science, i.e., in what part of human behaviour they think is sufficiently rule-governed and experimentally observable to be formalizable in computational models.

According to Fodor it is only the input processes that qualify as such, and he therefore has no need for further subdivisions above that level (cf. "Fodor's First Law of the Nonexistence of Cognitive Science", ibid.: 107). In Fodor's view only the functioning of the input processes can be studied under the "black box" model. Although we do not know what exactly is in the box, we can observe what goes into it, and we can get a fair impression of what comes out. Therefore, it should be possible to reconstruct the intervening processes in a limited number of steps. As concerns studying central processes, Fodor is sceptical both about the observability either of what goes in or of what comes out and about the likelihood of a restricted number of formalizable processing steps.

Pylyshyn (1980), on the other hand, has reverse views. He places the domain for scientific research above the input processes, because he considers it too dark to see inside the black box, whereas the higher-level processes are at least open to inspection (see also Fokkema, 1987; Meysing, 1986).

We do not wish to take a stand on this issue here, and therefore will describe some work related to each of the three levels. There is one major problem, however, that arises in describing this multi-levelled view. When we try to allocate the phenomena that the exisitng theories (attempt to) handle to one of the processing levels, we often find that there is overlap, with allocations to more than one level. In part, this is due to the recency of these theories, especially with respect to their characterization of the input processes, for many of the relevant experiments ante-date the (revival of the) "faculty" or "modular" approach. Since the change in perspective or "paradigm" is almost Kuhnian (1970) in magnitude, there are many experimental findings that (still) have to be re-interpreted. To some extent, of course, it may be that the processes defy (rigid) classification.

2.1. Input level processes

Of particular interest to our concern of reviewing the psychological adequacy of the FG-CNLU model is the characterization of the input processes, since it is at this level that language processing (and, partly, comprehension) takes place. We will, therefore, present in some detail the ideas set forth by Fodor in his book **The modularity of mind** (1983). The

input processes, or (perceptual) "modules" as Fodor (ibid.: 47) calls them, will also briefly be contrasted with the more "central cognitive processes". Some additional information in support of Fodor's claims will be given between square brackets [ ].

2.1.1. Fodor's account of perceptual modules

According to Fodor, input processes have the following characteristics:

a) they are domain specific;
b) their processing is mandatory;
c) they are fast;
d) they are (relatively) cognitively impenetrable;
e) they are informationally encapsulated;
f) they have shallow outputs;
g) (tentatively) there seem to be fixed neurological structures associated with them that are not only exhibited in specific breakdown patterns but also in ontogenetic development.

Central cognitive processes, on the other hand, have directly opposite properties on each of these points.

Let us clarify Fodor's use of terminology. "Domain specificity" should not be taken as indicating one perceptual module for each of the five traditional "senses". Rather, what he proposes is that there are (candidate) input processing modules for restricted classes of stimuli. Thus, he proposes modules for the perception of language, for the perception of colour, for the perception of shape, for the perception of three-dimensional spatial relations, and so forth (ibid.: 47). Each of these types of input is thought of as being processed by a special-purpose mechanism, or faculty (ibid.: 51). Central cognitive processes, on the other hand, do not show this "single-mindedness", and are not tuned to specific domains or tasks. Indeed they cannot be so in Fodor's view, since for the fixation of belief they are required to take all relevant information - from the input processes and from memory - into account (ibid.: 101-102).

The processing of input modules is mandatory in the sense that there is no voluntary control over it. To put it more plainly, Fodor (ibid.: 55) points out that, for a language that one understands, one cannot but hear what is said. One cannot, that is, perceive speech simply as "sounds" without identifying the words, without computing the underlying structure, etc. Neither is it possible, for instance, to touch a surface with one's hand without noticing whether it is rough or smooth. Central processes, again, have opposite properties in this respect. One can voluntarily choose the next move in a game of chess, whether to multiply 12345 by 6 or not to bother, etc. In fact, the exertion of control over these central processes is what makes up intellectual sophistication, for this consists largely of thinking along lines that are conducive to the satisfaction of one's goals.

That input processing is fast will not surprise anyone, as it can be observed daily. To give an indication of how fast this can be, Fodor (ibid.: 61-63) gives the following example. In so-called "shadowing" experiments, many subjects are able to repeat continuous speech as they hear it with latencies of only 250 msec. As the syllabic rate of normal speech is about 4 per second, this means that subjects lag only about one syllable behind the presented stimulus sentences. Moreover, since there is evidence that these subjects actually understand what they repeat (Marslen-Wilson, 1973), this quarter-second is sufficient time both to comprehend and to repeat speech.

47

The comprehension process alone, then, must be even faster [estimated at 150-200 msecs by Rohrman & Gough, 1967; Sabol & DeRosa, 1976]. Central processes such as problem solving, on the contrary, can be anything from relatively fast to very slow indeed. The difference between "hard" and "easy" problems at this level, however, may have to be measured in months or even years, rather then in milliseconds.

The claim that input processes are (relatively) cognitively impenetrable amounts to the following. The various "interlevels" of representations that are produced during, say, sentence comprehension (e.g. phonetic, phonological, lexical and syntactic representations) are not generally open to inspection, i.e. cannot be accessed by central processes (Fodor, ibid.: 60). In particular, the information represented at lower levels of language comprehension cannot be brought to consciousness at all, and cannot be reported on (e.g. what happens before a word or even a single phoneme/grapheme has been identified). The information at higher levels of language processing is (under normal discourse circumstances) lost within moments after comprehension, only their gist being retained [cf. Fillenbaum, 1966; Sachs, 1967; Wanner, 1974].

"Informational encapsulation" is more or less the reverse of cognitive impenetrability. It means that input modules do not have access to knowledge represented at central levels, and therefore are not influenced by it. This aspect, according to Fodor, is of vital importance to the survival of the species (ibid.: 70). For an organism to survive it must be able to rely on what it perceives, unhindered by what it would like to see, by what it would hope to see, or by having to check a large knowledge base to decide whether what it perceives might, for instance, be dangerous. ("If one is called by a panther", the advice is "don't anther"). More seriously, there are many instances of perceptual illusions in which knowledge of the fact that it is an illusion does not alter how and what one perceives. The Müller-Lyer illusion is a famous example:
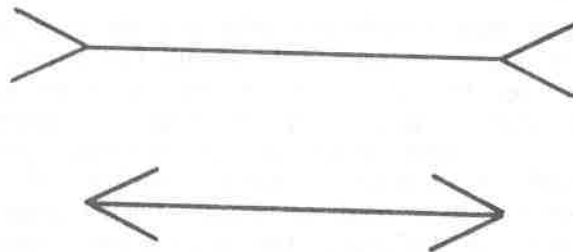
No matter how many times one has been told that these two lines (one ending in forks and the other in arrowheads) are of the same length, no matter how many times one has measured them - in other words, no matter how well one knows that they are actually equally long - one still perceives them as not being of equal length. (Central) knowledge, therefore, has little influence on what we perceive. In Fodor's view, then, perceptual illusions are to be attributed to intrinsic properties of the perceptual apparatus, and not to miscalculations of the cognitive system.

The claim of informational encapsulation, however, only holds for knowledge from outside the particular input module under consideration. Fodor is not as a rule opposed to top-down information flows within

modules. He argues that interaction, or "feedback", between submodules should be postulated with care, since "feedback is only possible to the extent that the information which perception supplies is redundant" (ibid.: 67). As far as language processing is concerned, he points out that knowledge of the context in which some utterance is made will be of little help in parsing its form, since

a)   there are many ways of saying the same thing, and
b)   context is hardly a determinant of syntactic form (ibid.: 78).

He also points out that in cases where it may seem as if context guides interpretation alternative explanations can be found. Contextual facilitation of word-recognition (also called "priming"), for instance, can be explained by associative links between lexical items; hence by information internal to the language module (ibid.: 79-80) [the same
proposal has been made by Forster (1979)]. Central processes are not informationally encapsulated for obvious reasons - all available and relevant information will be used, if possible.

The claim that the outputs of input analysers are "shallow" is meant to cover the following. Input systems only compute those properties that are coded in the input signal, that is, they compute type/token relations (ibid.: 90). In the case of language processing, for instance, only the information carried by words, word-order, stress and intonation is retrieved. Thus, inferences as to the speaker's intentions, and other "possible inferences", are not made at this level when not signalled by the "information carriers" of the input. Similarly, inferencing based on the decomposition of lexical items does not occur during language processing, and neither, therefore, are lexical items replaced by their (successive) definitions (Fodor et al. 1975; 1980) [also, Kintsch, 1974]. For visual input Fodor proposes that the highest level of abstraction reached by the input processes is something like the "basic categorizations" of Rosch et al. (1976). Central processes, in contrast, can have deep outputs, all depending on the task.

The final, tentative, characteristic of input processes, namely their (apparent) fixed neurological structures - or hardware - manifests itself in two ways. First, there exist certain well-defined pathological syndromes (agnosias and aphasias) that seem attributable to the malfunctioning of specialized circuits. Central processes, on the other hand, can take place-roughly speaking - everywhere in the brain (ibid.: 99), and are not as vulnerable to traumas. Secondly, there is evidence to suggest that the ontogenetic development of neural mechanisms subserving input analysis is largely endogenously determined, i.e. innately specified (cf. Chomsky's "Language Acquisition Device", 1972), whereas the more central processes are much more open to external influences (e.g. training and education).

2.2. The functional level

The middle or functional level is the level at which representations of information are manipulated for specific tasks in order to achieve goals that have been established at the top-level. Since all available information may be relevant for carrying out a specific task, the processes at this level have access not only to the outputs of all the input processors but also to the knowledge stored in LTM.

Various aspects of the processes and types of representations at this level have been discussed in the literature, some of which will be reviewed here.

2.2.1. Representation(s)

Since it is at this level that the various input processes interface with one another and with memory, the most important - and controversial - issue is whether there is a mode of representation common to them all. That is to say, is the representational format at this level independent of the input modality or not, and if so to what extent?

Bruner, Olver & Greenfield (1966) argued that there must be (at least) three types of knowledge representations, namely "enactive", "iconic" and "symbolic" ones. The former two were thought of, more or less, as "copies", and the latter as "abstractions" of experience. "Enactive" representations (used for motor knowledge) and "iconic" representations (used for sensory-perceptual knowledge) were thought not to be easily manipulated, as contrasted with symbolic knowledge.

The material with which they worked largely consisted of children's definitions of certain concepts (e.g. "A hole is to dig"). The fact, however, that children defined abstract concepts (cf. the use of the indefinite article "a") in enactive terms (cf. "to dig") indicates that translations between these types of representations are possible - at least partially.

The same point can be argued by the fact that one can give a (partial) description of what one perceives. Because of this intertranslatability it has been argued that there is no reason to assume that different types of knowledge are not represented in the same format. If these knowledge types have enough things in common to be translated into one another, there must be something underlying them all.

Fodor (1975) has argued that this format must have the nature of a symbolic language, because only languages have sufficiently abstract and combinatory possibilities to express every type of knowledge. He compares this "language of thought" to the machine-languages of computers, into which all information is translated before it is further processed. Since machine-languages can express different types of information (e.g. the video-pictures on the computer screen as well as the instructions of higher-level programming languages), he sees no reason why a "language of thought" could not do so as well.

Moreover, as others have also pointed out (e.g. Pylyshyn, 1973; Clark & Clark, 1977), the simple picture-like representations that were initially argued for (e.g. by Paivio, 1971) cannot be used to represent abstract ideas. These images, as they were called, were considered direct copies of some actual original that were re-viewed as it were in the mind's eye. However, if one pictures "a tiger" for instance, the image will not usually be very detailed, e.g. will not specify the number of stripes. Images, then, must be more abstract. Another drawback of pictures is that they are highly ambiguous and therefore not very well suited for the task of representing knowledge. A picture of a man climbing a mountain, for instance, might also represent a man descending a mountain backwards (Fodor, 1975). The representations of perceptual knowledge, according to Fodor (ibid.), must at least be pictures under a certain description.

Various proposals for such Knowledge Representations (KRs) that include conceptual as well as perceptual aspects have been made, notably **frames** (Charniak, 1972; Minsky, 1972; 1975), **schemata** (Rumelhart & Ortony, 1977;

cf. also Bartlett, 1932) and **mental models** (Johnson-Laird, 1983). Kosslyn (1980; 1981; 1987), on the other hand, argues for a more sophisticated form of **images**. He assumes that they are not direct copies of actual experience but rather that they are generated on the basis of analogue information stored in LTM. Kosslyn, thus, opposes the main-stream cognitivist view, as represented by Fodor and Pylyshyn, that the representations at the functional level are (solely) of an abstract, symbolic nature (Fokkema, 1987).

There have been various other suggestions for KRs, but these do not, strictly speaking, include or represent perceptual aspects of the input. In propositional accounts, for instance, only the "meaning" of the input is captured (e.g. Anderson & Bower, 1973; Anderson, 1976; Clark, 1976; Kintsch, 1974; Rumelhart & Norman, 1978). The abstract nature of these propositions is assumed to make them suitable for representing information from different input sources in memory, and for representing internal drives and goals (e.g. Anderson, 1976: 120-121). If we briefly consider the use of "sentence-picture verification" tasks (e.g. Clark & Chase, 1972; Clark, Carpenter & Just, 1973; Trabasso, 1970), however, it will become clear that information from "perceptual" input is represented only in a limited sense.

In experiments by Clark & Chase (1972), for instance, subjects had to verify pictures (e.g. one showing a dot above a cross) against sentences either correctly or incorrectly describing these scenes (e.g. "The dot is above the cross" or "The dot is below the cross"). The time needed to make these judgments was taken to be a measure of the difficulty in judging the congruency between the propositions abstracted from sentence and picture. As for the pictures, however, it was simply assumed that propositions of the type **dot above cross** were abstracted, without explaining how this was done. These propositions, moreover, do not specify information concerning the sizes, colours, relative distance, etc. of the two entities; in other words, they do not represent truly perceptual information.

Another drawback, which holds for the propositional approach in general, is that the sentences in such tasks were supposedly analysed by decomposing the lexical items into more abstract semantic elements - for which there is no psychological evidence, as we shall see.

What everybody agrees on, however, is that the representations at this level are more abstract than then the original input, if only because it is possible to generalize over particular experiences and to induce (general) rules. Schank (1975b) has argued in this respect that the distinction between semantic (abstract, general) knowledge and episodic knowledge (datable facts) as made by Tulving (1972) is misleading. He argues that if lexical knowledge is subtracted from knowledge in general nothing but personal experience remains.

It would seem, however, that there is little reason to make an exception for knowledge of the meanings of words, in this respect. As Labov (1973) has pointed out, people may take two words to be identical as to what they refer to (e.g. "beech" and "larch") in cases where experts can point out differences. (Or they may know that there must be a difference even if they do not know what the difference is.) Lexical knowledge, then, is acquired through personal experience in dealing with and talking about certain topics. An everyday example of this phenomena is formed by the specialized meanings that words may come to have for members of specific subcultures or "in-groups".

A second reason for the need for abstraction is that people are not very good at remembering what the original input was. One of the first to

demonstrate that people do not normally remember the exact wording of sentences was Fillenbaum (1966). In his study, subjects confused whether they had heard sentences like (1) or (2):

(1) The door was closed
(2) The door was not open

and he concluded that people remember the "gist" (e.g. what they inferred from them) rather than the syntax or the semantics of sentences. It may be necessary to point out briefly that what people do remember (best) of sentences has been called different names at different times, which somewhat confuses the picture. Let us clarify the matter.

Although the term "semantics" has sometimes been used to refer to "meaning" in the sense in which (1) and (2) have the same meaning, there is also a more narrow sense. Two sentences differ semantically in the narrow sense if the range of situations that they could describe are distinct. (3) and (4) are such semantically different sentences:

(3) The man with the martini waved to the hostess
(4) The man by the window waved to the hostess.

Garnham (1981) has shown that in cases where sentences have the same referents, as when (3-4) are used in the context of (5),

(5) By the window was a man with a martini

subjects confuse which of the sentences (3 or 4) it was that they originally heard. People, thus, do not normally remember the semantics of sentences in the narrow sense of the word.

(Long-term) memory for semantic aspects in the wider sense has been investigated in a number of experiments. Sachs (1967) and Wanner (1974), for instance, contrasted it with memory for syntactic aspects. Sachs presented sets of sentences to subjects and later gave them surprise recognition tests. The recognition test consisted of making judgments on sentences that were either:

a) the same as the original,
b) syntactically different (e.g. changes in word order, or passivized sentences) or
c) semantically different (e.g. with reversed roles).

When the recognition test was not given immediately after reading the sentences, only the last type of change was detected. In a similar manner, Wanner was able to demonstrate that surface structure is lost within 16 syllables after its being processed. Thus, in the wider sense of the word, "semantic" aspects are remembered better than syntactic aspects. A final example may serve to show this.

In a study by Honeck (1973) it appeared that repetition of sentences with nothing but new words was just as effective for remembering the original stimuli as verbatim repetitions were - which indicates that lexico-syntactic aspects of sentences have little to do with was is remembered.

On the other hand, people may sometimes remember the exact wording of the input (e.g. through memorization). It is therefore not so much a question of what they are in principle able/unable to remember because the relevant information has/has not been specified somewhere, but rather of

what they normally do remember. To summarize, if the different levels at which input material is processed are conceptualized as forming a scale from the more shallow to the deeper processes as in (6), then the products of the deepest processes (e.g. what is inferred from the input) are normally remembered best (cf. Fillenbaum, 1966; Clark & Clark, 1977; Kintsch, 1977; Garnham, 1985):

(6) truly perceptual (auditory and visual) > phonemic/graphemic > lexical > syntactic > (truly) semantic > "gist"

It is nonetheless the case that the more shallow processes do leave memory traces, and, under the right circumstances, can be remembered. A striking example of this is that people are able to memorize large chunks of texts even in languages they do not understand, and which, therefore, can only have been processed at sound-level. Members of a Hausa-speaking tribe from North Africa, for instance, have memorized the entire Koran (Clark & Clark, 1977), whereas they do not speak Arabic.

What can be learned from the variety of memory experiments that have been performed is that what people remember of the input is to a large extent determined by the nature of the task with which they see themselves faced. Different tasks lead people to remember different things. Episodic memory tasks generally lead people to encode words in "acoustic-articulatory formats" (Glucksberg & Danks, 1978: 64). In remembering short lists of unrelated items, for instance, people tend to make errors that sound like the to-be-remembered items (Conrad, 1964). With prose passages, on the other hand, which people do not normally expect to recall verbatim, meaning is remembered far better. Such observations led Glucksberg, Trabasso & Wald (1973) to the conclusion that there is no single way people perceive, remember or comprehend any given type of linguistic material. Attempts to explain differences in encoding strategies have been given in terms of **depth of processing** and in terms of **focus of processing,** and this brings us to processing issues.

2.2.2. Processing

2.2.2.1. Depth and focus of processing

To explain observations concerning the differences in what people remember of input sentences under various tasks, two complementary theories have been proposed. The focus of processing theory (Tieman, 1972) assumes that in different tasks different aspects of the input material are focused on and therefore remembered better. The depth of processing theory (Bobrow & Bower, 1969; Craik & Lockhart, 1972) assumes that the deeper some sentence is processed (i.e., the more attention is paid to its meaning) the better its meaning will be remembered. To assign these theories a place in the framework it is necessary to briefly explicate the underlying view of language processing, and to re-interpret the results obtained in studies designed to test these theories.

Language processing was considered by the propoents of these theories to proceed more or less as follows. A verbatim record of the unanalysed input sentence (e.g. the auditory signal) is available in working memory for a brief period of time to allow tasks to be performed on it. During the successive processing stages more abstract (e.g. phonetic > syntactic >

53

semantic) representations of the input are built up, which gradually leads to comprehension. After some time the verbatim record is purged from memory and only meaning remains. It was assumed that this natural flow could be manipulated by asking subjects to perform only up to a certain level. It was thought that in tasks that require shallow processing no deeper processing occurs and, consequently that no "deep" representation is built up, so as not to waste processing resources. In other words, it was assumed that if people were asked to pay attention to the form of a sentence, they would not extract its meaning.

In studies that were designed to test this assumption (e.g. Craik & Lockhart, 1972; Hyde & Jenkins, 1973; Mistler-Lachman, 1974; Treisman & Tuxworth, 1974), subjects were required to do one of two things with the input; one task was thought to involve shallow processing (i.e. to concern the form of the input) and the other to involve deeper processing (i.e. to concern the meaning of the input). Afterwards, these subjects were given surprise recognition or recall tests. In the study by Treisman & Tuxworth (1974), for instance, half of the subjects were asked to respond when a particular phoneme occurred in the input, and the other half to spot semantic anomalies. When, in delayed tests, recall was better for the second group, it was assumed that only the latter group had built up representations that were sufficient for remembering the input. It was assumed, moreover, that both groups had the verbatim record available during the immediate test, and that they had therefore performed equally well on it.

The fact, however, that the performance of the "semantic anomaly" group was better only on the delayed recall test, and not when it was immediate, can be interpreted differently. Since the groups performed equally well on the immediate recall test, we could assume that both had built up a complete representation of the input, which was available (in STM) for immediate recall. That the "shallow" group performed less well on the delayed recall test could then be explained in terms of the shallowness of the central-level processing involved, and not of that of the input processes. The poor performance of the "shallow" group is then to be attributed to the fact that the representation which they stored in LTM was insufficient, and not to the fact that the output of the language processor was insufficient. Another possible explanation is that the traces left by deep processing, such as spotting semantic anomalies, are better suited in nature for retrieval from LTM (Johnson-Laird, 1977: 212).

These re-interpretations of the test results would bring them into line with the assumption that the functioning of the language processor is mandatory and autonomous (Fodor, 1983). In this view the depth of the output of the language processor is always the same, and cannot be influenced. Accordingly, it is only at the functional level that task-specific requirements come into play. It is only at this central level, then, that there are differences in depth and focus of processing (and in what is finally stored).

Evidence for the autonomy of the language processor, for its insensitivity to outside task-demands, can also be found in a remark by Marslen-Wilson & Tyler (1981: 327). They observe that "even when subjects are asked to focus their attention on the acoustic-phonetic properties of the input, they do not seem able to avoid identifying the words involved". It may well be, therefore, that the so-called "shallow" tasks are performed after language processing rather than during.

The phoneme-monitoring task as invented by Foss (1970), for instance, which requires subjects to respond when a particular phoneme occurs in the input, may be performed on the output of the language processor, rather

then on its input (e.g. Goldiamond & Hawkins, 1958; Foss & Swinney, 1973; Cutler & Norris, 1979; Foss & Blank, 1980; Henderson, 1982). The issue is whether the response is based on an analysis of the acoustic input signal or on an analysis of phonological information read out of the lexicon, i.e. before or after a word has been identified.

Foss & Swinney (1973) argued that, although, phonemes are analysed before words, lower-level responses need not (always) be faster than higher-level ones, as was implicitly assumed in the earlier work (e.g. Savin & Bever, 1970). They argued that the word-level processor is not designed to produce awareness of phonemes or to make responses based on the presence of particular phonemes, and that recognizing individual words, although not its major concern either, is a more natural task for it. Higher-level responses based on the presence of some word, therefore, may be faster than lower-level responses based on the presence of some phoneme.

Evidence for these claims can be found in a experiment by Morton & Long (1976). In this experiment, responses to words beginning with the same phoneme were faster when the words were highly plausible in a given context (as in (7)) than when they were possible but not likely (as in (8)).

(7)   The man sat reading his book until it was time to go home
(8)   The man sat reading his bill until it was time to go home

If the responses had been based purely on a phonemic analysis of the input, then word-level variables such as contextual plausibility should not have had any effect.

2.2.2.2. Data-limited and resource-limited processing

Differences in what people remember of the input have also been explained in terms of limitations on the number of tasks that can be performed at the same time. At the central levels, various different tasks can be executed simultaneously. One may for instance drive a car and have a conversation at the same time. Since processing resources, such as Short-Term Memory space (also called working memory or attention) or access to a General Problem Solver (GPS), are limited, priorities must be assigned as to what (partial) task to perform first, and execution of other activities requiring the same resources may be put on hold temporarily. When a task is not performed as well as it could have been, given more priority and attention, it is said to be resource-limited. When, on the other hand, processing is not optimal because of poor quality of the input, it is said to be data-limited (Bobrow & Norman, 1975).

In general, resource-limited processing is typical of data that is readily accounted for. Such data, including things that one expects, will not require much processing and will not be remembered well (Bobrow & Norman, 1975: 145). Unexpected events, on the other hand, may suddenly require one's immediate attention, will require considerable processing and will be remembered well. In such cases, then, the calm and planned flow of information between the various processes, and the execution of plans, will be interrupted.

2.2.2.3. Selective attention

The possibility of interruption means, however, that there must be a mechanism that can determine the relative weight or importance of the information as delivered by the input processes even before it is further (consciously) processed at this central level. What such a mechanism has to account for, for instance, is the following well-known "cocktail party experience".

When people stand chatting in small groups, one often finds oneself trying to follow the conversation in another group some distance away that seems more interesting. Because of the noise level one will have to strain one's ears to do so. Although it may seem that one could simply filter out the background noise, including the conversation in one's own group, people are nevertheless able to hold up the appearance of being interested in the current conversation by nodding and by making "sympathetic noises" at the appropriate time. Moreover, they are able to respond immediately when their name is called. It is not the case, then, that one "communication channel" is completely unattended or switched off.

A similar argument holds for so-called dichotic listening experiments. In such experiments subjects are presented with two messages, one over each channel of a pair of headphones, and are instructed to attend to only one of them. In the main, subjects are unable to report what was said over the unattended channel. To be sure of this, however, experiments have also been carried out in which the second message was presented subliminally, i.e. below "hearing" level. In experiments by Lackner & Garret (1972) and MacKay (1973), subjects heard an ambiguous sentence in one ear and a subliminally presented disambiguating context in the other. Their task was to paraphrase the attended message. Depending on the disambiguating context that was presented, subjects could be biased towards paraphrasing sentences such as (9) either as (10) or as (11):

(9)   The spy **put out** the torch in the window

(10)   The spy **extinguished** the torch in the window
(11)   The spy **displayed** the torch in the window

This demonstrates not only that "decisions" may be based on information that one never becomes aware of, i.e. cognitively impenetrable, but also that such information is taken into account despite (conscious) efforts at the central level not to. However, the exact nature of the mechanism that weighs processing priorities largely remains a mystery.

What has become clear in any case is that the amount of processing done on a certain task limits the amount of additional processing that can be performed on other tasks (e.g. Brooks, 1968). Various experiments have shown, for instance, that processing load is directly related to the availability of Short-Term Memory space (see Wanner & Maratsos (1978) for an overview).

2.3 The top level

An early criticism by Miller, Galanter & Pribram (1960: 5-10) of researchers working in the cognitivist tradition in Europe, and on Behaviourists in America, was that they too readily assumed that once a system has been

given enough knowledge it will start to act on its own. To illustrate their point, they likened this assumption to the situation in which a laboratory rat, once it knows where to find its food in the maze, will go to its food irrespective of whether it is hungry or not. They argued instead that any cognizant system needs some kind of plan or plans, specifying what goals to achieve and how to achieve them. The mechanism that produces such plans had long before Miller et al.'s remark been known as "the will", of course (e.g. James, 1890).

A more recent version of the same statement can be found in Searle (1980) in that a system (or organism) must have the "intention" of performing certain actions to reach a goal (cf. also Dennet, 1978). Although many of the goals that a system wishes to achieve may remain unconscious, at certain stages the system must contemplate its options, set priorities, plan its actions, and distribute processing resources accordingly. Some of these operations take place at the "top" level.

These issues have received quite some attention from workers in the field of Artificial Intelligence (AI). On the one hand, there has been an extensive amount of theoretical work on so-called higher-order knowledge representations (i.e., representations of knowledge about knowledge), such as "scripts", "plans" and "acts" (e.g. Charniak, 1972; Abelson, 1975; Schank, 1975a; Schank & Abelson, 1977). On the other hand, there has been attention for more practical questions such as how to make computer systems perform multiple actions according to plans, i.e. how to make them perform different, perhaps conflicting, tasks at the same time.

Usually, systems have to be capable of handling different tasks at the same time. Execution of these different tasks in a limited amount of time requires a fair amount of planning. Not only may there be deadlines to meet (Miller, 1983), but some of these tasks may also require the same processing facilities, e.g. a General Problem Solver (GPS) or memory space, and have to compete for resources.

To the extent that top-level processes can be distinguished from processes at the functional level - of which Fodor is sceptical - separate planning mechanisms have to posited. If the system has knowledge about what it knows (just as humans need very little time to decide that either they know or do not know something), different plans or strategies for reaching some goal may be contemplated before one is decided upon and sent to the level below for execution. The top-level planning mechanism may, for instance, compute a cost-benefit analysis to decide what course of action is most effective. The planning mechanism of the functional level, on the other hand, not only has to ensure that the top-level instructions are carried out smoothly but also has to keep an eye on input signals that need immediate attention (and/or cause a change of plans).

3. Conclusion

The above overview of the processes that take place at the various levels should serve to provide an idea of what an adequate model will have to account for. We can now consider FG-CNLU more closely in the light of some issues raised, and test its assumptions against some further experimental results.

Chapter 6

Psychological adequacy of FG-CNLU: part 2

1. Introduction

In the previous chapter we described a framework in which theories and experimental studies of language processing and cognition can be assigned a place in accordance with the phenomena they concern. This framework provides a multi-levelled view of (the entirety) of human behaviour. It will have been obvious that most of the phenomena about which FG-CNLU makes faulty predictions are (or can be) handled more adequately in these accounts. In addition, the framework introduced some issues not accounted for in the FG-CNLU model but which are nevertheless of vital importance to a complete understanding of the processes involved. In this chapter some of these issues will be taken up again, in particular those relating to the relation between language and thought.

2. FG-CNLU and input processes

Even if we disregard the fact that in FG-CNLU there is just one input processor, i.e. a language processor, there are still many differences between FG-CNLU and what we may call the "standard picture". The FG-CNLU model "violates" the characterization of the input processes as given by Fodor (1983) in many respects. The major differences arise from the non-modular approach that the FG-CNLU model takes to the working of the language processor.

2.1 Informational encapsulation and cognitive impenetrability

In the FG-CNLU model, no principled distinction is made between the knowledge available to the input processes and the knowledge available to the central processes. The conceptual knowledge of the system, which comprises both linguistic and non-linguistic knowledge, forms a single body of knowledge which is stored in knowledge base D in its entirety and which is available both to the (Deducer of the) language processor and to Functional Logic.

Now although there are certain forms of linguistic knowledge that are available to central processes and can be reasoned with (e.g. the kind of linguistic knowledge that is formally taught at schools), many forms of linguistic knowledge actually used by the language processor most likely cannot be reasoned with (e.g. because they cannot be brought to the conscious). This "cognitive impenetrability" (cf. Chapter 4, 2.2) concerns both the linguistic knowledge that the language processor uses and, according to Fodor (1983), its interlevels of representation. In Fodor's view, both are unavailable to central processes. Although, in the literature, views differ as to exactly which types of linguistic knowledge are and which are not available for central level processing, it is accepted at least that not all such knowledge is available. The FG-CNLU model not only deviates on this issue but is also less plausible as a result.

Some of the types of knowledge used by the language processor have also been called "tacit knowledge" to reflect the idea that they are and/or remain (largely) unconscious. It may be necessary to point out again, however, that there can be several causes for this.

In the multi-levelled view the distinction between declarative and procedural knowledge does not necessarily coincide with the distinction between central and input processes, or the distinction between penetrable and impenetrable. It need not be the case, for instance, that the linguistic knowledge used by the language processor is "knowledge how to" rather than "knowledge that", as is sometimes (tentatively) suggested (e.g. Nuyts, 1988: 426). In a multi-levelled view, there is no principled reason why knowledge cannot be declarative or procedural at each of the levels.

A well-known example of procedural knowledge at the more central levels is formed by "routine behaviour". The frequent dialling of certain phone-numbers, for instance, may result in the (temporary) inability to (quickly) name the numbers, whereas dialling them occurs more or less automatically, having become a procedure (Anderson, 1976: 119).

On the other hand, there may be declarative knowledge at the input levels. The psychological correlate of knowledge that certain words are verbs or nouns may be represented as facts (e.g. go is a verb), and still not be communicable. (One may object that one can communicate that go is a verb, but that is to express knowledge of a different kind, i.e. central knowledge). It will be clear, for instance, that the lexicons of children must reflect the syntactic categories of words - as judged by their ability to use them correctly - even before they are taught that there are "verbs" and "nouns".

On the assumption that linguistic knowledge at the input levels is (by definition) impenetrable this must remain pure speculation, of course (if we cannot inspect the knowledge itself, we cannot establish its format either). The distinction between declarative and procedural knowledge may, therefore, not be very helpful or enlightening from a theoretical point of view (cf. Nuyts, 1988: 426), and only have practical value.

In discussing his ACT model, Anderson (1976: 78) points out that because of the close analogy between facts and data on the one hand and between procedures and programs on the other, the distinction is not always easy to make (one program may take another program as datum). He suggests, therefore, that we represent declaratively that "knowledge which is subject to multiple, different uses and that knowledge whose eventual use is uncertain", and that we represent procedurally that "knowledge which is used over and over again in the same way" (1976: 118).

In the other direction, the FG-CNLU model assumes that all central knowledge is available to the language processor, e.g. that general knowledge of the world is used in language processing. To see whether this is true we need to consider the nature of language processing and the structure of the language processor first.

2.1.1. Language comprehension and language processing

Although we have reported Fodor's claim that the language processor does not have access to knowledge outside it (hence, is not influenced by it), we have left largely open the question of what knowledge is available to it. To be able to answer this question, however, we need to consider to what extent the language processor is responsible for language comprehension.

Generally speaking, language comprehension (i.e. understanding) is said

to comprise roughly three parts, viz. a word-level stage, a syntactic stage, and a semantic stage - corresponding to the three levels that Chomsky (1965) subsumes under linguistic competence, viz. phonology, syntax, and semantics. The term "semantic(s)", however, can be misleading again in that it may suggest "meaning" in a narrow, linguistic sense.

One of the early and valued contributions of researchers in AI to the field of linguistics was the observation that the meaning of sentences often cannot be determined without taking "knowledge of the world" into account. Winograd (1972), for instance, pointed out that differences in the interpretation of sentences like (1-2):

(1) The city officials banned the women's demonstration, because they **feared** violence
(2) The city officials banned the women's demonstration, because they **advocated** violence

in which the usual interpretation is that "they" refers to "officials" in (1) and to "women" in (2) - can only be explained by reference to general knowledge of the kind that

a) city officials do not want disturbances of the peace;
b) demonstrations/demonstrators are sometimes/usually violent.

Thus, there is more to the "meaning" of sentences than linguistic semantics: the knowledge needed to interpret the meaning of linguistic structures comprises more than lexical knowledge, for instance. Besides general knowledge, one also needs knowledge of the con- and cotext to determine the referents of noun phrases or terms, to resolve anaphors (i.e. to find antecedents), to determine the referents of deictic expressions, etc. In other words, to establish the full meaning (in the sense of import or "significance" (Johnson-Laird, 1977)) of a sentence in a given context one may need information from more than one input source and from memory.

In Fodor's (1983) terms, then, a great amount of "central" knowledge is needed to assign "meaning" to input. This should not be taken, however, as running counter to his claim that the language processor does not have access to such knowledge since the purpose of input modules such as the language processor is (primarily) to establish what the input is. The language processor, for instance, need only specify which words are there and what structures (e.g. what clause sentence) they form. The language processor, thus, accounts for only part of the "meaning". The language processor outputs "what is said" rather than "what is meant", to use a classical distinction (e.g. Clark & Lucy, 1975), or the "literal meaning" of the input sentences.

Evidence for the "shallow" output of the language processor can be found in a number of experiments that show that the (input) words of a sentence are not replaced by their (lexical) definitions (e.g. Kintsch, 1974; Fodor, Fodor & Garret, 1975; Fodor, Garret, Walker & Parkes, 1980). In other words, the output of the language processor contains the actual (content) words of the input sentences. Despite some claims to the contrary (e.g. Lakoff, 1971; McCawley, 1971; Katz, 1972; Clark, 1976), there is indeed very little evidence that decomposition of lexical items occurs during language processing.

In experiments by Kintsch (1974) it has been demonstrated that the speed with which sentences are comprehended is hardly influenced by the definitional complexity of the lexical items as measured by the number of

embeddings. The definitional approach would require, for instance, that (3) ought to be more complex than (4) and therefore should take longer to comprehend.

(3) Cats chase mice
(4) Cats catch mice

The semantic analysis of "chase" involves decomposition into some structure that explicitly refers to an intention to catch (Katz, 1972). No such delays in reaction time (RT) have been found, however.

Fodor, Fodor & Garret (1975) also provided evidence against a definitional account of sentence comprehension. They conducted experiments in which subjects had to judge the truth values of arguments either containing

a) explicitly negative free morphemes (e.g. not),
b) explicitly negative bound morphemes or morphological negatives (e.g. in-, un-, im-, etc.),
c) implicitly negative morphemes (e.g. doubt, fail, deny),
d) or PDNs, i.e. Pure Definitional Negatives (lexical items that supposedly have "not" in their definitions, as in (5)).

(5) kill $=_{df}$ cause to become **not** alive

Subjects had to judge the validity of arguments like those in (6). The speed with which they did so was taken as a measure for the complexity of the argument.

(6) If practically all of the men in the room are **not married** / **bachelors**, then few of the men in this room have wives.

What they found was that, contrary to what the definitional account would predict, arguments containing PDNs were significantly easier (i.e., were responded to more quickly) than the paired arguments containing explicit negatives. Moreover, the difference in reaction times (RTs) between these two was significantly greater than the difference in RTs between explicit negatives and either implicit or morphological negatives.

A different approach to the same problem was taken by Fodor, Garret, Walker & Parkes (1980). Instead of measuring RTs to sentences supposedly containing definitionally complex lexical items they used a modified version of Levelt's (1970) paradigm, in which subjects have to rate the structural relationships between constituents. Subjects were presented, for instance, with sentences containing causative verbs, such as in (7):

(7) John broke the glass

On the definitional account such a sentence would be analysed as (8):

(8) ((John) ((caused) ((the glass) (break$_{intrans.}$))))
   "John did something which caused the glass to break"

Note that whereas John and the glass are subject and object respectively of the verb break in (7), there is no verb at the semantic level of which both John and the glass are arguments. Rather, in (8) John is the subject of cause, and the glass is subject of break$_{intrans.}$ Consequently, subjects" intuitions about the structural relationships should differ considerably for

sentences in which there is a shift in the grammatical relations as expressed by the surface structure and between that of the deep structure constituents. More concretely, in sentences like (9a-9b) subjects should rate the relatedness of John and Mary differently:

(9a) John killed Mary   (complex)
(9b) John bit Mary   (simplex)

The results clearly indicated, however, that no such shifts occurred, hence that the relationship that the verb expresses between the arguments does not differ for "complex" or "simplex" verbs. Fodor et al. took this as strong evidence against claims that the internal representation of lexical items consists of definitions by which they are replaced during comprehension, and as proving that there is no semantic interpretation stage in the linguistic sense of "semantic".

In order to reflect observations about the difficult distinction between linguistic meaning and meaning in general, and about the psychological unreality of the decompositional approach, the term "semantic" has largely been dropped as naming a processing stage in language comprehension. Modern models of language comprehension, therefore, comprise the following three processors (e.g. Forster, 1979):

(i)   a word-level processor, which establishes which words are there;
(ii)   a syntactic processor, which establishes the structure of the clause/sentence;
(iii)   a message-level processor, which computes what the input "means", given the context.

In such models, the message-level processor is the cognitive system (e.g. Garnham, 1985). In language comprehension as a whole, then, there are only two processors that handle strictly linguistic input, viz. the word-level processor and the syntactic processor (parser). Accordingly, only these are included in what Fodor (1983) calls the language module/processor - whereas others include all three processors under that name (e.g. Garnham, 1985: 183). Fodor's claim that the language processor does not have access to central knowledge, therefore, is not incompatible with the claim that such knowledge is used in language comprehension. Rather, he restricts the use of such knowledge to message-level processing.

An alternative view of the structure of the language processor in the wider sense has been proposed by Schank and his colleagues (e.g. Schank, 1972; 1975; Schank & Colby, 1973; Small & Rieger, 1982). In strong forms of this view, it is assumed that a (complete) structural description of the input (i.e. a syntactic parse) can be forgone. In other words, that one can derive sentence meaning directly from word meaning. (According to Schank "the rules of French grammar are not crucial to understanding French", 1975a: 12).

    The output of the word-level processor is assumed to be fed directly into a semantic parser that analyses the input by means of "case-frames" (Fillmore, 1968). (In weaker forms, word-level output is also passed on to a syntactic parser that provides back-up, as it were, in case of structural ambiguities.) In this approach the meaning of a sentence such as (10) would be given by (11):

63

(10)   The boy watered the flower

(11)   ACTION: water
               TENSE: past
               AGENT: boy_1 (some particular boy)
               PATIENT: $flower_1$ (some particular flower)

It may seem that one has effectively bypassed syntactic analysis by using the selection restrictions on water, which requires an animate agent, to decide that the boy must be the agent, and thus that the flower must be the patient. Nevertheless, there has been some syntactic analysis (Garnham, 1985: 70), since on two occasions the definite article the has implicitly been grouped with the following noun to produce an NP. Moreover, there are many sentences, even simple ones, that cannot be interpreted by this kind of case-frame analysis because they do not contain semantic cues to the roles of the NPs in them (ibid.). This approach cannot distinguish, for instance, between simple active and passive sentences of the type in (12), because both NPs fit the selection restrictions on the arguments of chase:

   (12a) The dog chased the cat
   (12b) The dog was chased by the cat

The most obvious way to correctly interpret such sentences, then, is to use structural information. It has been argued in this respect that the function of syntax is to allow less plausible meanings, as in (12b), or implausible meanings to be conveyed (Forster, 1979).

   Other problems for this approach arise from its confusion about Chomsky's (1965) formal autonomy principle, which states that syntactic concepts should be defined without reference to semantic notions. In "semantic grammars" with syntactic categories such as AGENT, PATIENT and TIME, the structural descriptions display the meaning of sentences transparently, and are assumed to have been derived without recourse to syntactic analysis of the input. Consequently, (11) - which is a structural description - is easily mistaken for a semantically interpreted one. However, labels such as AGENT (rather than NP or VP) are empty but for the meaning that someone attaches to them. That is, these labels only receive meaning when someone interprets them.

   One may conclude that, in general, it is not possible to forgo syntactic analysis, also because every other strategy that has been proposed fails on some simple sentences (Garnham, 1985: 69-70). A syntactic parser, therefore, must be included in the language processor.

2.1.2. Interaction

The question of whether central knowledge is used in language processing in Fodor's sense thus boils down to whether such knowledge is used in word-level processing and in parsing. How does this relate to the FG-CNLU model?

   In the model, general knowledge of the world aids the analysis of linguistic input from the earliest stages. In Dik's (1987b: 17-18) example that we considered in Chapter 2, 4.2.1, for instance, general knowledge is used to infer that men are only likely to have "scars" on their cheeks, and not "stars" or "cars", and thus that the input must have been scar. For this to be possible, the entire sentence The man had a long ... on his cheek must

have been phonetically, syntactically and "semantically" processed before such knowledge could be used to decide the identity of the blanks (how else could the system know that <u>cheek</u> was part of the input and that it had relevance to the identity of the blanks?).

It is difficult to determine, however, at which stage the language processor is influenced by general knowledge in the model. On the one hand, it is said that only a model that interprets in a parallel, top-down and bottom-up processing fashion is sufficiently powerful to explain a number of well-known linguistic phenomena (Dik, 1986a: 24), which we take as stating that the language processor and the cognitive system should interact in the FG-CNLU model. But on the other hand, if, as suggested, the identity of the missing word is established only after the entire sentence has gone through the various processing stages, then the actual decision is not made interactively but serially and independently. The model, thus, is not very clear about its supposed interaction - also because interaction is not reflected in the outline of the model as presented in Fig. 1 of chapter 2. Moreover, Dik's claim shows a serious misunderstanding of the issues involved in that it suggests that interactive models are to be preferred over serial models of human language understanding (cf. also ibid.: 12).

One of the problems that one encounters in establishing whether the subprocessors used in language comprehension work serially and independently, i.e. autonomously (e.g. Forster, 1979; Marcus, 1980; Garret, 1982), or interactively (e.g. Marslen-Wilson, 1975, 1976; Marslen-Wilson & Tyler, 1980), is that autonomy models and interactive models are empirically not as distinct as they are conceptually (Norris, 1982). A model operates serially if there is a strict order in which the sub-processors receive information about a given part of the input. Serial operation between a word-level processor and a syntactic processor does not have to mean, however, that all word-level processing takes place before all syntactic processing. While the current word is being identified previous word(s) may undergo syntactic analysis. Also, various subprocessors may produce multiple outputs simultaneously, instead of one at a time. Many forms of serially operating models are therefore possible.

There are also many forms of interaction - interaction can be defined as the exchange of information between two processors during the performance of a task (Norris, 1982). Crain & Steedman (1985), for instance, distinguish between strong and weak forms of interaction. In strong forms of interaction, the higher-level processor directs the lower-level processor as to what analysis it should try. In weakly interactive models, on the other hand, the higher-level processor can only signal to the lower-level processor that its current (low-level) analysis does not fit into the (high-level) structure being built. In the latter manner, "autonomy" for the individual subprocessors is maintained. Some weakly interactive models therefore do not differ much from some forms of serial models.

Serially operating models, thus, can be made far more powerful than their opponents (e.g. Marslen-Wilson, 1975: 227 and also Dik, 1986a: 12) give them credit for. By using multiple outputs per process, by supplying the (sub) processes with Problem Solvers (e.g. Forster, 1979) etc., defenders of processing autonomy have been able to explain away all the evidence that seemed at first sight to favour an interactive view of the structure of the language processor (Garnham, 1985: 202). According to Garnham, theorists are in an impasse on this issue because no single test has been devised (by the interactionists) that would have settled matters once and for all for each and every autonomy model proposed, in favour of an interactive view.

Although no experimental results have been found that cannot be explained in some (possible) autonomy model, evidence has been obtained against certain strong forms of interaction. There is evidence that word-level processes are not influenced by (some forms of) semantic/general knowledge. It has been shown, for instance, that even in strongly biasing contexts all the meanings of an ambiguous word are initially retrieved from the lexicon (e.g. Swinney, 1979; Tannenhaus, Leiman & Seidenberg, 1979), regardless of whether they make sense. It is thus not the case that, due to knowledge of the context, one picks out the relevant lexical entry for some word only. Swinney (1979), for instance, presented subjects with fragments of texts, like (13), over headphones:

(13) Because he feared electronic surveillance, the spy searched his room. He was not surprised to find several **bugs** [1] in the corner [2] of his room

Simultaneously, he had them make "lexical decisions" on words presented on a screen - i.e., subjects had to indicate when they had identified the word. What he found was that at point [1] subjects responded equally fast when a contextually likely reading of "bugs" was presented visually - which would have produced (14),

(14) He was not surprised to find several **microphones**

as when a contextually unlikely reading of "bugs" was presented - which would have produced (15):

(15) He was not surprised to find several **ants**

Moreover, by varying the position in the sentence at which the word was presented visually, he was able to establish that the ambiguity was resolved quickly. At point [2], responses to microphones were faster. These results, thus, provide evidence against the use of general knowledge on (some aspects of) word-recognition.

As concerns parsing, it has often been suggested that semantic or general knowledge may help the parser decide between possible syntactic structures. In Winograd's (1972) SHRDLU program, for instance, choices between alternative parses (16a-b) of an ambiguous sentences such as (16):

(16) Put the block in the box on the table

(16a) Put (the block in the box) (on the table)
(16b) Put (the block) (in the box on the table)

are based on knowledge of the world (e.g. of whether the block is in the box to begin with or not). It is not clear, however, whether this has psychological relevance, i.e., it has never been demonstrated that this is what humans do.

Rayner, Carlson & Frazier (1983), for instance, studied the eye movements of subjects reading structurally ambiguous sentences. They found that the patterns of fixations and regressions for a given syntactic structure were always the same, regardless of the plausibility of its possible interpretations in the context. In sentences (17) and (18) alike, regressive eye movements were made from was pleased to florist/performer sent,

(17)   The florist sent the flowers was very pleased
(18)   The performer sent the flowers was very pleased

Apparently, both sentences were parsed the same way. However, sentence (18) has two plausible readings (19-20), whereas (17) has only one (21), since it is somewhat unlikely that someone would send flowers to a florist:

(19)   The performer **who was** sent the flowers was very pleased
(20)   The performer **who** sent the flowers was very pleased ·

(21)   The florist **who** sent the flowers was very pleased.

Parsing, then, seems little disturbed by such general knowledge as to what is likely or not (at least for such heavily decontextualized sentences as are often used in experiments).

These experiments, thus, provide evidence against the availability of (some forms of) semantic/general knowledge to the word-level and syntactic processors of the language module. They show that not all forms of central knowledge that the system has available are, in fact, used to determine what the input is. So, contrary to what Nuyts (1988: 610) claims, there are arguments against certain forms of interactive bottom-up and top-down processing. It may be, therefore, that to the extent that top-down processing occurs, if it occurs at all, it serves to restrict the number of candidate analyses rather then to determine the exact analysis (Marslen-Wilson, 1980). The FG-CNLU model, in which all knowledge of the system is available, and used, for language processing, is in any case much too powerful to be psychologically adequate.

2.1.3. Local explanations

If the language processor does not have access to some forms of higher knowledge, then it may be sensible to adopt the working hypothesis that it does not have access to such knowledge at all (e.g. Forster, 1979; Fodor, 1983). As a consequence, alternative explanations must be found for the influence or influences that context does have on language processing. These explanations could assume that the relevant information can be found within the language processor (i.e. they must be relatively "local"). Another possibility left open is that the information available to the language processor is influenced only **indirectly** by some forms of central knowledge, i.e. that there is no direct exchange of information between the two systems but that there is an intermediary that has access to both. In Forster's (1979) model, for instance, all processors (including the message-level processor or cognitive system) are in two-way contact with the lexicon and with a General Problem Solver (GPS). Finally, there is the possibility that some difficulties (e.g. ambiguities) cannot be resolved at the input level, in which case the various analyses could be sent on to the cognitive system for a final decision. In none of the above cases would interaction be required.

2.1.3.1 Local explanations for context effects on word-recognition

To see what explanations are possible in a system in which language processor and cognitive system do not interact we will have to detour briefly.

67

There are some phenomena in which higher-level knowledge is used to guide lower-level processing, but according to Fodor (1983) this is done only on a limited scale. One such phenomenon is the restoration of incomplete input - e.g. the "phoneme restoration effect" (Warren, 1970). When people listen to a tape-recorded sentence in which one of the speech sounds of a word has been spliced out and replaced by a cough or by silence, subjects nevertheless report having heard the word - and only very few will volunteer that they heard something "in the background". In such cases, the lexicon is probably queried by as much phonetic information as has been securely identified. The best match from the lexicon is then used to fill in the blank (or to replace the cough). Here, then, higher-level knowledge has been used to determine what the input must have been. The explanation, however, has been kept relatively "local" in that it assumes that the (higher-level) information is obtained closest to where it is needed - at any rate well within the limits of the language module. Consequently, this explanation adheres as closely as possible to the ideal encapsulation principle.

Similarly, a "local" solution can be found to explain why a system in which the language processor and cognitive system do not interact would opt for the "scar" interpretation in Dik's example. Although it is highly unlikely, of course, that the word-level processor would not itself be able to distinguish between this combination of "candidates" (scar, star and car) in the first place - if only because one of them contains fewer phones than the other two - consider the following.

It is a well-known fact that associations between words speed their recognition (e.g. Meyer & Schvaneveldt, 1971), and it has also been established that such word association patterns are quite homogeneous and predictable among adults (Entwhistle, Forsyth & Muuss, 1964). The word butter, for instance, is recognized faster after its high frequency associate bread than after some low frequency associate - e.g. nurse. The semantic and syntactic properties of high frequency associates, therefore, are also retrieved from the lexicon more quickly, and the parser can start working with them sooner. Now suppose that the lower-level processes within the word-level processor were not able to decide between star and scar as correctly representing the input. Suppose also that when a word is accessed from the lexicon, it primes or activates words related to it (e.g. Morton, 1969, 1970; Collins & Loftus, 1975). The more closely words are related, the more strongly they are primed and the more easily they are recognized. The strength of associations between words could then help decide what the most likely interpretation is. Since, presumably, the association between man and scar is relatively stronger then the association between man and star, the choice would be obvious. In case this is not very convincing, e.g. because the associations between words in a sentence need not always be very strong, absolutely speaking (as in the example), sentential context effects on word-recognition can also be explained by invoking a special type of mental model.

Foss (1982) found that associations between gills, fins and fish did not speed recognition of the latter in (unrelated) lists of words as in (22) - as compared to recognition of fish after the much weaker associates spots and stripes;

(22) Group the examined the entire gills the and fins the of fish caught in the river Nile

whereas recognition of fish was speeded in sentence (23) - again as compared to a sentence with the weaker associates:

(23)   The entire group examined the gills and fins of the fish caught in
the river Nile

He assumed that (23) allows one to build up a (linguistic) mental model in
which the actual lexical items of the sentence appear. The model may prime
related words, and so speed their recognition. In the example, therefore, the
decision that scar most probably represented the input need not solely be
based on a possibly stronger association between man and scar than between
man and star. Rather, the decision may be based on the better fit of scar
in the association chain or model man - have - ... .

Another possibility is that of "backwards priming". Suppose that the
word-level processor as a whole could not decide between star and scar. It
might pass both of them onto the parser, which could then build up in-
dependent structures for both of them. Just prior to the system's processing
of the final input word there would be two structures (24-25):

(24)   (The man) had (a long scar) (on his ?)
(25)   (The man) had (a long star) (on his ?)

Since the final input word cheek is a closer associate of scar than it is of
star, it will be recognized faster in the case of (24). The word-level
processor might then decide that it must have been scar after all. Yet
another possibility is that, since cheek is recognized faster in the case of
(24), parse (24) is completed sooner, and therefore preferred.

The imaginary data, thus, can be accounted for in many ways without
taking recourse to general knowledge, either in the final serial decision or
being used interactively. What we need to assume is that many forms of
knowledge are reflected in the lexicon, as it were, by means of associations
between lexical items. Many associative models of "semantic memory" have
been proposed, of course (e.g. Quillian, 1968; Collins & Quillian, 1969;
Anderson & Bower, 1973; Collins and Loftus, 1975; Eco, 1976; etc.). According
to Fodor, however, these associations are more like "reflexes" than
"knowledge".

2.1.3.2. Frequency effects

Another possibility is that many decisions in the language processor are not
so much based on explicit knowledge but on frequency. In many cases it may
appear as if some (top-down) predicted analysis was made when in fact the
most common analysis may simply have been tried first.

In word-recognition, such a frequency effect has often been observed
(e.g. Rubenstein, Garfield & Millikan, 1970; Forster & Chambers, 1973). Words
that occur frequently (e.g. door are recognized faster than infrequent ones
(e.g. cask) even when they are matched for length, as in the examples.
Instead of postulating some form of rather explicit knowledge as to what are
frequent and what are infrequent words, one may consider more implicit
methods. Frequency effects may (come to) be built into the architecture of
the system. For instance, in Morton's (1969; 1979) logogen system there is a
set of feature counters, called logogens. There is one logogen for each word
that a person knows. The feature count of a word is incremented as soon as
perceptual analysis reveals that one of its features (e.g. curves or straight
lines) occurs in the input. For each logogen there is a "threshold" level at
which it "fires". A threshold is reached when enough of a word's features
have been recognized, i.e., when it is almost certain that the input is the

word corresponding to the logogen. The logogen is then said to fire, and the corresponding word becomes available. To account for the frequency effect one could assume that threshold levels of logogens are (slightly) reduced when they fire often (Morton, 1970: 207).

Another approach is taken by Forster (1976; 1979). He assumes that a complete perceptual representation of the input word is constructed first. This perceptual representation is then compared to access files of the lexicon, i.e. files with orthographic and phonological representations of words plus pointers to the location in LTM where the actual lexical entries are kept in master files. Access files are subdivided into bins, e.g. one bin for each initial letter or sound of words. The words in bins are arranged in order of frequency, and matching the perceptual form against words takes place in that order. Frequent words are, thus, matched faster because they are on the top of the list. The semantic and syntactic information in the master files also becomes available faster as a result.

Similarly, it makes good sense that in parsing, the most common analysis is or should be tried first (e.g. Garnham, 1985: 202), since it will apply in the majority of cases. For structurally ambiguous sentences with two possible parses in which there is no syntactic information about which parse is correct, the less frequent one could be terminated (Norris, 1982). This could explain why (26), from Frazier & Fodor (1978), produces a garden path:

(26) Tom said that Bill will take the cleaning out yesterday

If (26) produces a garden path, we may assume that the wrong analysis was chosen. Since there is no syntactic information to indicate that yesterday cannot be a modifier of Bill will take the cleaning out in (26), the parser apparently prefers (27) to the less frequent (28) - e.g. because of the length of the intervening clause.

(27) Tom said that (............................)
(28) Tom said that (...........................) yesterday

That is to say, the parser does not "expect" structures like (28).
The parser is probably so constructed that frequent parses are made faster then infrequent ones, just as the word-level processor is faster in recognizing frequent words.

2.1.3.3 Conclusion

The extent to which local explanations can be found for context effects on language processing has yet to be determined in its full variety. What the above accounts buy us, however, is that the encapsulation of the language processor (i.e. word-level processor + syntactic processor) is not endangered. In this way, we are able to avoid, or rather to exclude, the possibility of "wishful seeing" (Fodor, 1983) or hallucinating, since (higher-level) knowledge from outside the language processor, including wishes and expectations, does not influence its functioning. Dik's characterization of the top-down processes, on the other hand, leaves this possibility wide open:

> The top-down component should be powerful enough to reconstruct certain pieces of information even in the total absence of any spoken input. (1987b: 18)

2.1.4 Speed

What the above account also buys us is "speed" - one of the other charac-
teristics of input processes. Since in FG-CNLU all the system's knowledge is
available for language processing, there is no principled reason why it should
not be consulted either. This, however, will take (considerable) time.
Moreover, since it is possible to make an almost endless number of
inferences on the basis of the input plus the resident knowledge, even more
time will be needed. It is doubtful, therefore, that the present system could
ever simulate real-time language comprehension, which has been estimated as
taking somewhere in the range of 150-200 msecs (Rohrman & Gough, 1967;
Sabol & DeRosa, 1976), unless it were provided with large quantities of ad
hoc, and theoretically unfounded, heuristics.

2.1.5. Shallow output

Another characteristic of input processes is their shallow output. In the case
of the language processor this means, among other things, that its output
contains the actual (content) words of the input sentence. We have seen that
this approach is also taken by Dik (1987a; b). The FG-CNLU model,
therefore, is correct in this aspect. Another matter, however, is FG-CNLU's
pragmatic functions.

In FG-CNLU, the parser determines the functions that terms have in a
particular predication. Among these functions are the pragmatic functions of
Theme, Tail, Topic and Focus. Notice, however, that whereas Theme and Tail
have specific linguistic structures associated with them (Dik, 1978b: 130) (e.g.
left- and right-dislocation, respectively; the use of special purpose phrases
such as as for ...) Topic and Focus do not. It is not the case, for instance,
that Topic always coincides with the first constituent in the clause,[1] or that
Focus aways coincides with the constituent receiving main stress (ibid.: 131).
It is doubtful, therefore, whether the parser can determine which terms have
Topic or Focus function purely on linguistic grounds. Determination of
Topic function (i.e., which constituent it is that presents the entity about
which the predication predicates something - Dik, 1978b: 130) and of Focus
function (i.e., which constituent it is that presents the most important or
salient information - ibid.) seem to be based primarily on central knowledge.
Therefore, unless it can be shown that Topic and Focus functions are also
signalled by linguistic cues, as in cases of "emphatic focus" constructions
(Hannay, 1985 - e.g. It is ... that ...), the output of such function
specifications must be deemed to be too deep for the language processor (i.e.
in those cases where they are not formally signalled or marked).

Especially with respect to Focus function, it may be that pragmatic
functions, when not formally marked, should be thought of as being associ-
ated with linguistic expressions rather than being part of them. Such
pragmatic functions would then be in the same league as "presuppositions",
which Dik (1978a: 130) holds not to be part of the linguistic expression
proper. (Cf. Wilson (1975) and Gazdar (1979), who assume that presup-
positional phenomena are all **pragmatic** in nature.) According to Kintsch
(1977: 369) presuppositions are inferences, i.e., inferences as to what (the
hearer believes that) the speaker takes for granted. And inferences, as we
have seen, are not made by the language processor.

---

[1]This has, however, been recently proposed by Hannay (forthc.).

3. Central level processes

Two notes on FG-CNLU's central level processes may be relevant to Dik's argumentation for the supposed linguistic nature of conceptual knowledge. One argument involves FG-CLNU's (implicit) conception of communication, and the other its conception of language-independence.

3.1. Communication

Dik observes that even those theorists who firmly believe in a language-independent code of thought always express their examples in the words of some particular language. He concludes that "N(atural) L(anguage) expressions are quite efficient means for representing knowledge" (1986a: 3). However, the only thing that is proved by the fact that people express their ideas in words is that NL expressions are used for communicating; it does not prove that language is used to represent knowledge. It should be made absolutely clear from the beginning that this misconception has serious consequences for the FG-CNLU model. What is basically at stake here is what communication is - which, after all, is what the system is supposed to model.

To reason along the lines that Dik does is to fall into the trap of what Reddy (1979) has called the "conduit metaphor", and into a very strong version of it at that. What this metaphor suggests is that human communication is the direct transfer of knowledge from one person to another by means of a pipe, i.e. language. Roughly speaking, the speaker has knowledge in his head, he wraps it up in words, he sends it down the pipe to the addressee, and all the addressee has to do is unwrap the parcel and take out the knowledge. In Dik's version, moreover, the wrapping-up stage is kept to a minimum by the assumption that the knowledge in one's head is already linguistic in nature. All one has to do to as receiver in an act of communication is to reverse the mappings of expression rules, and there the message (the FSUP) is.

Moreover, what makes the model's view of human communicative behaviour even more crude is that the messages/packages that one receives convey explicit instructions, orders rather, about what to do with them. When the illocutionary function is DECL(arative), the contents should be stored; when it is INT(errogative), one should reply (Dik, 1987a).

Together, these two (implicit) suggestions provide a rather passive, almost slavish view of communication. Although in FG language is appropriately seen as a means of communication, as an instrument (Dik, 1978a: 1), in FG-CNLU it is inappropriately suggested that communication is an end unto itself. If we take into consideration, however, that people communicate in order to achieve something, in other words, that communication is itself a tool, then it will become clear that something is missing in the model.

The failure to understand that communication is not an end but a tool is reflected in the fact that the model is completely input-driven, i.e. only reacts to external stimuli. The model does not have any intentions of its own, is not trying to perform a particular task or tasks for which it might need pieces of information that it does not possess itself, in fact, the model does not have any reason or purpose to communicate at all. If all humans were as the model suggests they are, this view of communication would lead to everlasting silence since there would be no one to initiate conversation. Although one may object that is impossible to give a computer

(program) intentions, this can be mimicked at least by giving it tasks to perform. To perform these tasks as well as external/internal conditions allow can then be seen as its intentions.

An adequate model of the Human Natural Language User - which we take to be a complete human being, will therefore have to take into account (if not to account for) the intentions of the system. Since different intentions, or goals to be achieved, may conflict or compete for processing resources, there will also have to be a control mechanism (which is also completely missing in the model) to smooth operations, to set priorities, etc. In other words, a whole new level of processes will have to be added to the present system.

3.2. Language independence

From the observation that words of one language often cannot be translated into another without adding or subtracting elements of content, Dik concludes that a language-independent code of thought, which should make direct translation possible, is unlikely to exist (1986a: 4-5). One of the examples that he gives is the English word <u>supper</u>, for which the closest Dutch translation would be something like the equivalent of (29):

(29) "(potential) evening meal, distinct from dinner" (ibid.: 5)

This argument is fallacious for least two reasons.

3.2.1. Language and cognition

Firstly, it is highly questionable what the absence of one-to-one correspondences between words of one language and that of another proves anything with respect to cognition. Does the absence of a single word <u>supper</u> in Dutch prove that Dutchmen cannot understand what <u>supper</u> means? Does it prove that they cannot form a concept(ion) of it, as opposed to the English? In Dik's approach this is indeed the case, for the Dutch would only succeed in approximating to the concept, either adding or subtracting elements of content not present in the original. Consider, however, that the English themselves do not do much better. They, too, succeed only in giving a partial description of the meaning of <u>supper</u> by contrasting it with something else, cf. (30):

(30) "evening meal, less formal and substantial than dinner" (<u>Concise Oxford Dictionary</u>, 1977)

It will be clear therefore that no such conclusion can be drawn.

Secondly, the fact that the Dutch are able to give an exact translation of the English definition/description also throws doubt on such a conclusion. Moreover, it suggests that Dutchmen can imagine what it would be like to have supper and thus to understand what it means. One should be careful, therefore, in drawing inferences from the non-occurrence of lexical items to conceptual systems and cognitive abilities (cf. Black, 1962).

3.2.1.1. Linguistic relativity

A case in point is offered by a number of psycholinguistic experiments that were originally designed to test just such claims.

The theory of Linguistic Relativity holds, in weaker and stronger forms, that languages influences thought. A strong version, sometimes called the Sapir-Whorf hypothesis after two of its advocates (Sapir, 1941; Whorf, 1956), holds that languages are free to cut up our experiences of the world into categories as they see fit, and that the words we learn to name these categories determine they way we think and perceive.

Most experimental studies designed to test this hypothesis have focused on the naming and perception of colours. Berlin & Kay (1969), for instance, were interested to find out whether the different segmentations of the colour spectrum by different cultures actually entail that people perceive colours differently. They presented speakers of 20 diverse languages with 329 coloured chips, which the subjects then had to name. Berlin and Kay then distilled the basic colour terms for each language by establishing from what "atoms" compound descriptions were built up. After this, they asked their subjects to indicate on the spectrum which basic term covered what area, and to mark the best examplar of each term. The best examplars were called "focal colours".

When they compared the results of the different languages, they found two things. First, the number of basic colour terms that languages have is highly restricted, ranging from 2 to 11. Second, and more importantly, the divisions into focal colours are respected across languages. That is, what is considered a focal colour by speakers of one language is never subdivided by others. E.g., if one language has 4 basic colour terms and another language 6, then the four focal colours as chosen by speakers of the first language will correspond to four of the six focals chosen by speakers of the latter. These results thus provide evidence against the view that there are no intrinsic (universal) properties of, or constraints on, perception. Also, it suggests that language reflects rather than causes the distinctions that we make. As Minsky puts it, "words should be are servants, not our masters" (1985/86: 79).

Further evidence for this view was provided by Heider & Olivier (1972), Heider (1972) and Rosch (formerly Heider, 1973b) in a series of studies on the Dani. The Dani are a Stone Age tribe living in New Guinea, who have only two colour words (corresponding to bright/warm and dark/cool). In the Heider & Olivier study, English and Dani subjects performed equally well in a recognition memory experiment, in which they had to indicate which one of the colours on a chip had been shown before. The Sapir-Whorf hypothesis, in contrast, predicts that English subjects should perform much better because their language offers greater opportunity for distinctly naming the colours and thus for keeping them apart in memory. Moreover, both groups tended to confuse and to keep apart the same colours, indicating that the memory representations must have been highly similar for both groups.

Second, Heider showed that both the Dani and the English are better at recognizing (i.e. correctly identifying) previously presented colours that are focal to the English than at colours that are non-focal (e.g. a focal red is recognized better than a non-focal red), even though the Dani did not have distinctive names for the items in the pair, and even though they did not have a word that names the pair as a whole.

Third, Rosch showed that the Dani, when required to learn names for English colours, did better when the test material consisted of focal (English) colours then when it consisted of non-focal or intermediate colours.

From all these studies, one may conclude that there is something universal about the perception of focal colours, hence about man's perceptual apparatus. Data obtained from speakers of one language (i.e. English), for instance, could be used to predict the behaviour of others (e.g. the Dani) in memory and learning experiments. All in all, then, this can hardly be a case of linguistic determinism (Foss & Hakes, 1978).

3.2.2. Language-independent is universal?

Another dubious assumption is that language-independent (necessarily) means "common for all mankind" (Dik, 1986b: 4). There is no reason, however, why concepts should not be independent of the language in which they are expressed, and, at the same time, not be universal. The concept of "wedding" that (modern) Britons have, for example, can be independent of the word they use to express it and still be different from, say, the Turkish concept "wedding" - e.g. in that it no longer includes a "dowry" component.

An alternative interpretation of language-independence, in this respect, is that conceptual codes are independent of actual languages but heavily dependent on the common experiences, social mores, customs, etc. that speakers of a particular language (tend to) share. In other words, conceptual codes are language-independent but culture-specific. This approach has been taken by Fodor (1975), who suggests different conceptual codes, different "languages of thought", for speakers of different languages qua cultural beings.

One suggestion that quite naturally flows from this view is that speakers of different languages understand each other only to the extent that their cultures have things in common, to the extent that their concepts overlap. Although there seems to be some truth in this observation - judging by the amount of misunderstanding and misperception involved in hostilities between cultures, the question to be answered (namely, is the nature of the conceptual apparatus such that it can only handle predefined units of meaning?) is rather one of what people are able to understand than of what they normally do understand.

3.3. Concepts

Despite philosophers' warnings (e.g. Wittgenstein, 1953), a persistent view of human reasoning is that it somehow operates on and in discrete units. Even if one adopts this view as a working hypothesis, however, one may want to avoid the consequence that one cannot understand that for which one has no concepts. An implicit assumption most theories seem to agree on, irrespective of their stand on the linguistic or non-linguistic nature of concepts is that one is compelled to use the concepts that one has, that one cannot help thinking in them. A quite different approach to concepts can be taken, however. Suppose, for instance, that one views concepts not as compulsions but as **habits.** To see what this means let us first clarify "concept".

The concept "concept" is itself highly complex. For our present purposes, however, it will suffice to give an instrumental definition only. A concept can be defined as a set of stimuli that produce the same response (Bolton, 1977), and that therefore can be seen as forming a whole. It will be obvious that there are different processing levels at which something can serve as a stimulus to produce a response. Spotting an actual red rose, for instance, produces a visual pattern (with specifications of e.g. size, shape and colour),

which in turn can be variously classified and named (e.g. <u>red rose</u>, <u>rose</u>, <u>flower</u>) depending on how many of its characteristics are included in the classification/naming process, which in turn may produce associative responses such as <u>love</u> or <u>thorny</u>. Decisions as to what is the same, or sufficiently the same to receive the same response, can also be made at different levels, of course, since for each level of stimuli there is also a level of responses. Perceptual stimuli can also combine with conceptual stimuli to produce responses. In principle, therefore, an unlimited number of concepts can be formed by grouping stimuli together recursively, i.e. by assuming that a set of common responses can in turn function as a set of stimuli.

Even though theoretically there are unlimited ways of grouping stimuli together to form concepts, there are constraints. We will not go into this matter deeply here, but let us mention some of the principles involved so as to provide at least the general idea. One such set of constraints falls under the heading of what we may call the "functionality" criterion, i.e. there must be some purpose for which it is useful to group certain things together and to keep others apart. Another set of constraints follows from the principle that there is a "basic level" of categorizations (Rosch, 1973a; Rosch & Mervis, 1975) at which the distinctions that are made are maximally informative, i.e. a level at which category members are maximally similar to each other and maximally distinct from non-members. Closely connected to this is a third set of constraints that is provided by the perceptual apparatus itself.

At the lowest input-processing level there is the perception of the stimulus itself. As we have seen, there is reason to assume that the (basic) perceptual apparatus is the same for all mankind, in other words, that it is universally constructed to detect certain input characteristics. Let us assume therefore that the outputs of these basic processes - in the example, the visual pattern plus a number of specifications - are also the same across cultures. In as far as perceptual stimuli are the building blocks of concept formation, therefore, we can expect universal patterns (e.g. focal colours).

In principle, then, responses to stimuli can differ across cultures (or, alternatively, across ways of thinking) at each level except at the strictly perceptual level, since it is at this level that the stimuli are the same for all humans. Above that level groupings may differ to produce different responses that in turn lead to different responses again, and consensus disappears.

To say in this respect that concepts are the habits rather then the compulsions of the mind is to say that people are in principle able to make responses to sets of stimuli that differ from their usual responses, that people are in principle able (to come) to see different sets as forming wholes (e.g. the Dutch may not have a ready-made concept "supper" but they can easily form one by regrouping a set of stimuli). To say that concepts are habits is also to acknowledge that there may be concepts that are difficult not to use, just as there are habits that are difficult to quit. There also may be ways of thinking, of grouping things together differently, that are so utterly strange to a particular individual or group of individuals that they are not understood.

In the habit approach, however, this non-understanding is not simply and solely a matter of the incapability in principle to understand incompatible concepts but also of the unwillingness to give up known and trusted ways of thinking. This idea is not altogether new, of course. In the schema-approach, for instance, it is reflected in the claim that schemas (tend to) resist

restructuring (Rumelhart & Norman, 1978).

Another well-known characteristic about habits is that they may conflict with "better judgment". Knowing that smoking is bad for one's health, for instance, does not necessarily make quitting any easier. So too with concepts, one may "know" the one thing but "think" the other (cf. prejudices).

Also, habits and concepts are both sensitive to fashions, especially with respect to emotive and normative aspects (e.g. Eco, 1976: 288). Any number of other similarities offer themselves.

The "habit" metaphor, perhaps, raises more questions than it answers. Its only point, however, was to make clear that many problems concerning (natural) concepts and the conceptual apparatus have still been left untouched, and that no simple and direct view does justice to the complexity of the matter. If the above metaphor (or any other one) is successful as a replacement of the "compulsion" metaphor, it may appear that altogether the wrong questions used to be asked.

4. Conclusion

In comparing the FG-CNLU model to some relevant issues from the literature that together outline a multi-levelled view of human behaviour, it has again become clear that much of what it assumes is difficult to reconcile either with experimental results or with theoretical and philosophical considerations.

Chapter 7

General conclusion

1. Overview

After the outline of Functional Grammar (FG) that we gave in chapter 1, we presented in chapter 2 Dik's recent proposals for the building of a process model of the Natural Language User (NLU) based on the principles of the Grammar. In chapter 3 we saw that the proposals made in various papers do not add up to a single, coherent model. In fact, if the suggestions are taken literally, no potentially working model is presented at all. In chapter 4 we nevertheless followed up on some of the model's implications. The result was that there is no experimental evidence for (the majority of) the model's predictions about human language processing and cognition. Theoretical considerations led us to conclude, moreover, that too many important issues have not been taken into account in the model. To remedy these gaps we introduced a general framework in which various aspects of human cognition treated in the literature were discussed (chapter 5). In chapter 6, finally, we reviewed some of FG-CNLU's underlying assumptions in the light of this framework, again with an unfavourable outcome for the model.

2. Conclusion

As a general impression of the FG-CNLU model and its underlying theory, one cannot help feeling that too many bold statements are made that are not backed up by (substantial) evidence in any way, either by its author or by reality as we know it from current psycholinguistic research. Those issues that are addressed explicitly in the model and theory are mostly handled in a grossly simplified and all or none fashion (e.g. the strict dichotomy between conceptual and perceptual knowledge), and those issues about which one would like to see some explicitness are either not or only indirectly addressed (a good example of the latter is what behaviour the system is supposed to model). If anything can be learned from our discussion of the complex matters involved, then it is the kind of detailed argumentation that is necessary to account for even the most basic facts.

All in all, we cannot but conclude that the model does not achieve psychological adequacy. In fact, one may seriously doubt whether FG-CNLU (certainly in its present form) could ever meet the requirement of "providing insight into how NLU's work" (Dik, 1986a:1). To use Weizenbaum's comment on Turing tests (reported in Dennet, 1983/4:104), if one had to decide whether one was "talking" to FG-CNLU or to an autistic child and typed in one's questions, one would not receive any answers, and not notice the difference.

3. Areas for improvement

If FG/FG-CNLU is intended not to be incompatible with what is known about language processing and cognition then it may be wise to bring it into line with that knowledge. A first step would be to adopt a multi-levelled view to human behaviour (e.g. of the type discussed in previous chapters). In broad outline, a model based on that view would consists of (at least) two

processing levels.

As concerns the input level, a modular approach to the language processor such as advocated by Fodor (1983) could be taken. The main design objective should be to limit the amount of knowledge that its subprocessors require, either as data or as procedures, since endless supplies of knowledge are no more realistic than "brute force" methods. Initially one could adopt the principle of "local explanations". Information about "context", for instance, can be provided by making the subprocessors sensitive to previous processing (e.g. through priming and through frequency effects). The depth of its output should be limited.

As concerns central processing levels, the model should be provided with tasks to perform, and with intentions (e.g. to perform them as well as circumstances allow). What the model infers, and what it stores in LTM, should be made to depend on these tasks. The (resulting) knowledge representations should at least be such that information from different input processors can combine or, at least, interface. This, of course, would require something more abstract than either FSUPs or UPs, also because the model should be able to integrate different sentences into texts. The inference engine, finally, should be capable of abduction.

4. Epilogue

It will be clear that incorporating the above improvements into the model would involve major renovation - if not rebuilding the system - since we would need not only an entirely different language processor, but also an entirely different cognitive system.

Although such a new language processor would work in a different fashion, its output, which would serve as one of the inputs of the cognitive system, could well be something like FG-type underlying predications. The grammar incorporated in this language processor could possibly be called "Cees grammar" after its proponent, but would involve only minor changes to the present Functional Grammar. Care should be taken, however, that the characterization that the language processor (and thus the grammar) gives of a certain linguistic expression should contain only those aspects that are coded in the signal (e.g. by such linguistic cues as case-markers, agreement, sentence position and specific constructions, etc.). Such an undertaking must be deemed possible.

(Re)modelling the cognitive system, on the other hand, would involve many more changes, since the first task would be to develop a cognitive code. By repudiating FSUPs/UPs as forming the (cognitive) code in which knowledge is represented and manipulated, and thereby giving up on unification, we are back to square one. It is here, probably, that the principle of "functional architecture" is bursting at its seams. Although we may learn something from studying and modelling the brain by only considering what it does (what functions it performs) and not by how it does it, i.e. without in-depth knowledge of its anatomy and physiology, this learning is of a Popperian nature. We may falsify some theories, because they do not work even in theory, and in that sense we approximate to the "truth" (Popper, 1959; 1963), but whether we would actually learn from such an experience how the human brain might work is very dubious. It is not by chance that many workers in the field of artificial intelligence see their work more as "knowledge engineering" than as having to do with psychology (e.g. Charniak & McDermott, 1983). The endeavour to develop a code for representing knowledge may therefore be more of practical than theoretical value.

80

References

Abelson, R. (1975). Concepts for representing mundane reality in plans. In: Bobrow, D.G. and A. Collins (eds), **Representations and understanding: studies in cognitive science**. New York: Academic Press, pp. 273-312.

Anderson, J.R. (1976). **Language, memory and thought**. Hillsdale (NJ): Erlbaum.

Anderson, J.R. & Bower, G.H. (1973). **Human associative memory**. Washington: Winston.

Austin, J.L. (1962). **How to do things with words**. Oxford: OUP.

Barclay, J.R. (1973). The role of comprehension in remembering sentences. **Cognitive Psychology 4**, 229-254.

Bartlett, F.C. (1932). **Remembering: a study in experimental and social psychology**. Cambridge: CUP.

Berlin, B.R. & Kay, P. (1969). **Basic color terms: their universality and evolution**. Berkeley: University of California Press.

Black, M. (1962). Linguistic relativity: the views of Benjamin Lee Whorf. In: Black, M., **Models and metaphors**. Ithaca (NY): Cornell University Press.

Bobrow, D.G. & Bower, G.H. (1969). Comprehension and recall of sentences. **J. of Exp. Psych. 80**, 455-461.

Bobrow, D.G. & Norman, D.A. (1975). Some principes of memory schemata. In: Bobrow, D.G. & A.M. Collins (eds), **Representation and understanding: studies in cognitive science**. New York: Academic Press, pp. 131-150.

Bolton, N. (1977). **Concept formation**. Oxford: Pergamon.

Bransford, J.D., Barclay, J.R. & Franks, J.J. (1972). Sentence memory; a constructive versus interpretive approach. **Cognitive Psychology 3**, 193-209.

Bransford, J.D. & Franks, J.J. (1971). The abstraction of linguistic ideas. **Cognitive Psychology 3**, 331-350.

Bray, N.W. & Batchelder, W.H. (1972). Effects of instruction and retention interval on memory for presentation mode. **JVLVB 11**, 367-374.

Brooks, L. (1968). Spatial and verbal compounds of the act of recall. **Canadian Journal of Psychology 22**, 349-368.

Bruner, J.S., Olver, R. & Greenfield, P. (1966). **Studies in cognitive growth**. New York: Wiley.

Charniak, E. (1972). Context and the reference problem. In: Rustin, R. (ed.), **Natural language processing**. New York: Algorithmics Press, pp. 311-331.

Charniak, E. & McDermott, D. (1983). **Introduction to Artificial Intelligence**. Reading (MA): Addison-Wesley.

Chomsky, N. (1965). **Aspects of the theory of syntax**. Cambridge (MA): MIT Press.

Chomsky, N. (1972). **Language and mind** (enlarged edn). New York: Harcourt Brace Jovanovich.

Clark, H.H. (1976). **Semantics in comprehension**. The Hague: Mouton.

Clark, H.H., Carpenter, P.A. & Just, M.A. (1973). On the meeting of semantics and perception. In: Chase, W.G. (ed.), **Visual information processing**. New York: Academic Press.

Clark, H.H. & Chase, W.G. (1972). On the process of comparing sentences against pictures. **Cognitive Psychology 3**, 193-209

Clark, H.H. & Clark, E.V. (1977). **Psychology and language**. New York: Harcourt Brace Jovanovich.

Clark, H.H. & Lucy, P. (1975). Understanding what is meant from what is said: a study in conversationally replied requests. **JVLVB 14**, 56-72.

Cofer, C.N. (1973). Constructive processes in memory. **American Scientist 61**, 537-543.

Collins, A.M. & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. **Psychological Review 82,** 407-428.

Collins, A.M. & Quillian, M.R. (1969). Retrieval time from semantic memory. **JVLVB 8,** 240-247.

Conrad, R. (1964). Acoustic confusion in immediate memory. **British Journal of Psychology 55,** 75-84.

Craik, F.I.M. & Lockhart, R.S. (1972). Levels of processing: a framework for memory research. **JVLVB 11,** 671-684.

Crain, S. & Steedman, M.J. (1985). On not being led up the garden path: the use of context by the psychological parser. In: Dowty, D., L. Kartunen & A. Zwicky (eds), **Natural language parsing.** Cambridge: CUP.

Cutler, A. & Norris, D.G. (1979). Monitoring sentence comprehension. In: Cooper, W.E. & E.C.T. Walker (eds), **Sentence processing: psycho-linguistic studies presented to Merril Garret.** Hillsdale (NJ): Lawrence Erlbaum.

Deese, J. (1970). **Psycholinguistics.** Boston: Allyn & Bacon.

Dennet, D.C. (1969). **Content and consciousness.** London: Routledge & Kegan Paul.

Dennet, D.C. (1978). **Brainstorms, philosophical essays on mind and psychology.** Brighton: Harvester.

Dennet, D.C. (1983/84). Kunnen machines denken? **Wijsgerig perspectief 4,** 98-108.

Dik, S.C. (1968). **Coordination: its implications for the theory of general linguistics.** Amsterdam: North-Holland.

Dik, S.C. (1978a). **Functional Grammar.** Amsterdam: North-Holland.

Dik, S.C. (1978b). **Stepwise lexical decomposition.** Lisse: Peter de Ridder.

Dik, S.C. (1986a). Linguistically motivated knowledge representations. **Working papers in Functional Grammar 9.** Amsterdam: University of Amsterdam; also in M. Nagao (1987) (ed.), Language and artificial intelligence, Amsterdam: North-Holland, 145-170.

Dik, S.C. (1986b). Functional Grammar and its potential computer applications. In: Dik, S.C. (1986d); also in W. Meijs (1987) (ed.), Corpus linguistics and beyond, Amsterdam: Rodopi, 253-268.

Dik, S.C. (1986c). Generating answers from a linguistically coded knowledge base. In: Dik, S.C. (1986d); also in G. Kempen (1987) (ed.), Natural language generation: new results in artificial intelligence, Dordrecht: Martinus Nijhoff, 301-314.

Dik, S.C. (1986d). Two papers on the computational application of Functional Grammar. **Working papers in Functional Grammar 18.** Amsterdam: University of Amsterdam.

Dik, S.C. (1987a). Concerning the logical component of a natural language generator. Paper, University of Amsterdam; also in M. Zock and G. Sabah (1988) (eds), Advances in natural language generation: an inter-disciplinary perspective, London: Pinter, 73-91.

Dik, S.C. (1987b). Functional Grammar Computational Natural Language User, a research program. Paper, University of Amsterdam. Also as "FG*C*M*NLU: Functional Grammar Computational Model of the Natural Language User" in J.H. Connolly and S.C. Dik (1989) (eds), Functional Grammar and the computer, Dordrecht: Foris, 1-28.

Eco, U. (1976). **A theory of semiotics.** Bloomington: Indiana University Press.

Entwhistle, D.R., Forsyth, D.F. & Muuss, R. (1964). The syntactic-paradigmatic shift in children's word associations. **JVLVB 3,** 19-29.

Fillenbaum, S. (1966). Memory for gist: some relevant variables. **Language & Speech 9,** 217-229.

Fillmore, C.A. (1968). The case for case. In Bach, E. & R.T. Harms (eds), **Universals in linguistic theory**. New York: Holt, Rinehart & Winston.

Flagg, P.W., Potts, G.R. & Reynolds, A.G. (1975). Instructions and response strategies in recognition memory for sentences. **Human Learning and Memory 1**, 592-598.

Fodor, J.A. (1975). **The language of thought**. New York: Thomas Y. Crowell.

Fodor, J.A. (1983). **The modularity of mind**. Cambridge (MA): MIT Press.

Fodor, J.A., Garret, M.F., Walker, E.C.T. & Parkes, C.H. (1980). Against definitions. **Cognition 8**, 263-367.

Fodor, J.D., Fodor, J.A. & Garret, M.F. (1975). The psychological unreality of semantic representations. **Linguistic Inquiry 6/4**, 515-531.

Fokkema, S.D. (1987). **Niveaus, modules en modellen in de cognitive psychologie**. Valedictory lecture. Amsterdam: Free University Press.

Forster, K.I. (1976). Accessing the mental lexicon. In: R.J. Wales & E. Walker (eds), **New approaches to language mechanisms**. Amsterdam: North-Holland.

Forster, K.I. (1979). Levels of processing and the structure of the language processor. In: Cooper, W.E. & E.C.T. Walker (eds), **Sentence Processing: psycholinguistic studies presented to Merril Garrett**. Hillsdale (NJ): Lawrence Erlbaum. 27-85.

Forster. K.I. & Chambers, S.M. (1973). Lexical access and naming time. **JVLVB 12**, 627-635.

Foss, D.J. (1970). Some effects of ambiguity upon sentence comprehension. **JVLVB 9**, 699-607.

Foss, D.J. (1982). A discourse on semantic priming. **Cognitive Psychology 14**, 590-607.

Foss, D.J. & Blank, M.A. (1980). Identifying the speech code. **Cognitive Psychology 12**, 1-31.

Foss, D.J. & Hakes, D.T. (1978). **Psycholinguistics**. Englewood Cliffs (NJ): Prentice Hall.

Foss, D.J. & Jenkins, C.J. (1973). Some effects of context on the comprehension of ambiguous sentences. **JVLVB 12**, 577-589.

Foss, D.J. & Swinney, D.A. (1973). On the psychological reality of the phoneme: perception, identification and consciousness. **JVLVB 12**, 246-257.

Frazier, L. & Fodor, J.D. (1978). The sausage machine: a new two-step parsing model. **Cognition 6**, 291-325.

Furth, H. (1966). **Thinking without language**. New York: Free Press.

Garrett, M.F. (1982). Remarks on the relation between language production and language comprehension systems. In: Arbib, M.A. & D. Caplan & J.C. Marschall (eds), **Neural models of language processes**. New York: Academic Press, pp. 209-224.

Garnham, A. (1981). **Mental models as representations of discourse and text**. Unpublished D. Phil. thesis, University of Sussex.

Garnham, A. (1985). **Psycholinguistics: central topics**. London: Methuen.

Gazdar, G.J.M. (1979). **Pragmatics: implicature, presupposition and logical form**. New York: Academic Press.

Glucksberg, S. & Danks, J.H. (1978). **Experimental psycholinguistics: an introduction**. Hillsdale (NJ): Lawrence Erlbaum.

Glucksberg, S., Trabasso, T. & Wald, J. (1973). Linguistic structures and mental operations. **Cognitive Psychology 5**, 338-370.

Goldiamond, I. & Hawkins, W.F. (1958). Vexierversuch: the logical relationship between word-frequency and recognition obtained in the absence of stimulus words. **J. of Exp. Psych. 56**, 457-463.

Greenberg, J.H. (1966). **Language universals**. Mouton: The Hague.

Hannay, M. (1985). **English existentials in functional grammar**. Dordrecht: Foris.

Hannay, M. (fc). Speaker strategies and placement rules.

Haviland, S.E. & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. **JVLVB 13**, 512-521.

Heider, E.R. (1972). Universals in color naming and memory. **J. of Exp. Psych. 92**, 337-354.

Heider, E.R. & Olivier, D.C. (1972). The structure of the color space in naming and memory for two languages. **Cognitive Psychology 3**, 337-354.

Henderson, L. (1982). **Orthography and word recognition in reading.** London: Academic Press.

Henle, G. (1962). On the relation between logic and thinking. **Psychological Review 69**, 366-378.

Hintzman, D.L., Block, R.A. & Inskeep, N.R. (1972). Memory for mode of input. **JVLVB 11**, 741-749.

Honeck, R.P. (1973). Interpretive vs structural effects on semantic memory. **JVLVB 12**, 448-455.

Hyde, T.S. & Jenkins, J.J. (1973). Recall of words as a function of semantic, graphic and syntactic orienting tasks. **JVLVB 12**, 471-480.

James, W. (1890). **The principles of psychology.** New York: Henry Holt.

Johnson, M.K., Bransford, J.D. & Solomon, S. (1973). Memory for tacit implications of sentences. **J. of Experimental Psychology 98**, 203-205.

Johnson-Laird, P.N. (1977). Procedural semantics. **Cognition 5**, 189-214.

Johnson-Laird, P.N. (1983). **Mental models.** Cambridge: CUP.

Johnson-Laird, P.N., Robins, C. & Velicogna, L. (1974). Memory for words. **Nature 251**, 704-705.

Johnson-Laird, P.N. & Stevenson, R. (1970). Memory for syntax. **Nature 227**, 412.

Karlsson, F. & K. Koskenniemi (1985). A process model of morphology and the lexicon. **Folia Linguistica 19**, 207-231.

Katz, J.J. (1972). **Semantic theory.** New York, etc.: Harper & Row.

King, D.R.W. & Greeno, J.G. (1974). Invariance of inference times when information was presented in different linguistic forms. **Memory & Cognition 2**, 233-235.

Kintsch, W. (1974). **The representation of meaning in memory.** Hillsdale (NJ): Lawrence Erlbaum.

Kintsch, W. (1977). **Memory and cognition** (2nd edn). New York: Wiley.

Kintsch, W. & Monk, D. (1972). Storage of complex information in memory: some implications of the speed with which inferences can be made. **J. of Exp. Psych. 94**, 25-32.

Kolers, P.A. (1966). Interlingual facilitation of short-term memory. **JVLVB 5**, 314-319.

Koskenniemi, K. (1983). **Two-level morphology: a general computational model for word-form recognition and production.** Helsinki: Department of General Linguistics.

Kosslyn, S.M. (1980). **Image and mind.** Cambridge (MA): Harvard University Press.

Kosslyn, S.M. (1981). The medium and the message in mental imagery: a theory. **Psychological Review 88/1**, 46-66.

Kosslyn, S.M. (1987). Seeing and imagining in the cerebral hemispheres: a computational approach. **Psychological Review 94/2**, 148-175.

Kuhn, T.S. (1970). **The structure of scientific revolutions** (2nd edn). Chicago: Chicago University Press.

Labov, W. (1973). The boundaries of words and their meaning. In: Bailey, C.N. & W.R. Shuy (eds), **New ways of analyzing variation in English.** Washington: Georgetown University Press, 340-373.

Lackner, J.R. & Garret, M.F. (1972). Resolving ambiguities: effects of biasing contexts in the unattended ear. **Cognition 1,** 359-372.

Lakoff, G. (1971). On generative syntax. In: Steinberg, D.D. & Jakobovits, L.A. (eds), **Semantics: an interdisciplinary reader in philosophy, linguistics and psychology.** Cambridge: CUP, pp. 232-296.

Lenneberg, E.H. (1967). **Biological foundations of language.** New York: Wiley.

Levelt, W.J. (1970). A scaling approach to the study of syntactic relations. In: Flores d'Arcais, G.B. & W.J. Levelt (eds), **Advances in psycholinguistics.** Amsterdam: North Holland, pp. 109-121.

Loftus, E. & Palmer, S. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. **JVLVB 13,** 585-589.

McCarthy, J. & Hayes, P. (1969). Some philosophical problems from the standpoint of artificial intelligence. In: Meltzer, B. & D. Mitchie (eds), **Machine intelligence 4.** New York: American Elsevier.

McCawley, J.D. (1971). Prelexical syntax. **Georgetown University monograph series on languages and linguistics 24,** 19-33.

MacKay, D.G. (1973). Aspects of the theory of comprehension, memory and attention. **Quarterly Journal of Experimental Psychology 25,** 22-40.

Marcus, M.P. (1980). **A theory of syntactic recognition for natural language.** Cambridge (MA): MIT Press.

Marslen-Wilson, W.D. (1973). Linguistic structure and speech shadowing at short latencies. **Nature 244,** 522-523.

Marslen-Wilson, W.D. (1975). Sentence perception as an interactive parallel process. **Science 189,** 226-228.

Marslen-Wilson, W.D. (1976). Linguistic descriptions and psychological assumptions in the study of sentence perception. In: Wales, R.J. & E.C.T Walker (eds), **New approaches to language mechanisms.** Amsterdam: North-Holland.

Marslen-Wilson, W.D. & Tyler, L.K. (1980). The temporal structure of spoken language understanding. **Cognition 8,** 1-71.

Marslen-Wilson, W.D. & Tyler, L.K. (1981). Central processes in speech understanding. In: Longuet-Higgens, H.C., J. Lyons & D.E. Broadbent (eds), **The psychological mechanisms of language.** London: The Royal Society and The British Academy, pp. 103-118.

Meyer, D.E. & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. **J. of Exp. Psych. 90,** 227-235.

Meysing, M. (1986). Geest, hersenen, computer. In: Hagoort, P. & R. Maessen (eds), **Geest, computer, kunst.** Utrecht: Stichting Grafiet, pp. 64-88.

Miller, D. (1983). Scheduling heuristics for problem solvers. Report 264, Yale University Department of computer science.

Miller, G.A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. **Psychological Review 63/2,** 81-96.

Miller, G.A., Galanter, E. & Pribram, K.H. (1960). **Plans and the structure of behaviour.** New York: Holt, Rhinehart & Winston.

Milne, R.W. (1982). Predicting garden path sentences. **Cognitive Science 6,** 349-373.

Minsky, M. (1975). A framework for representing knowledge. In: Winston, P. (ed.), **The psychology of computer vision**. New York: McGraw Hill.

Minsky, M. (1985/6). **The society of mind**. New York: Simon & Schuster.

Mistler-Lachman, J.L. (1974). Depth of comprehension and sentence memory. **JVLVB 13**, 98-106.

Morton, J. (1969). Interaction of information in word recognition. **Psychological Review 76**, 165-178.

Morton, J. (1970). A functional model for memory. In: Norman, D.A. (ed.), **Models of human memory**. New York: Academic Press.

Morton, J. (1979). Word recognition. In: Morton, J. & J.C. Marschall (eds), **Psycholinguistic Series 2: structures and processes**. London: Paul Elek.

Morton, J. & Long, J. (1976). Effect of transistion probability on phoneme identification. **JVLVB 15**, 43-51.

Norris, D.G. (1982). Autonomous processes in comprhension: a reply to Marslen-Wilson and Tyler. **Cognition 11**, 97-101.

Nuyts, J. (1988). **Aspekten van een kognitief-pragmatische taaltheorie: een poging tot een bijdrage aan de theorievorming over de kognitie van natuurlijke taal**. Dissertation, University of Antwerp.

Olson, D.R. (1970). Language and thought: aspects of a cognitive theory of semantics. **Psychological Review 77**, 257-273.

Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. (1957). **The measurement of meaning**. Urbana (IL): University of Illinois Press.

Paivio, A. (1971). **Imagery and verbal processes**. New York/Toronto: Holt, Rinehart & Winston.

Popper, K.R. (1959). **The logic of scientific discovery**. London: Hutchinson.

Popper, K.R. (1963). **Conjectures and refutations**. London: Hutchinson.

Potts, G.R. (1974). Storing and retrieving information about ordered relationships. **J. of Exp. Psych. 103**, 431-439.

Preston, M.S. & Lambert, W.E. (1969). Interlingual interference in a bilingual version of the Stroop color word task. **JVLVB 8**, 295-301.

Pylyshyn, Z.W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. **Psychological Bulletin 80**, 1-24.

Pylyshyn, Z.W. (1980). Computation and Cognition: issues in the foundations of cognitive science. **Behavioural and Brain Sciences 3**, 111-169.

Pylyshyn, Z.W. (1981). The imagery debate: analogue media versus tacit knowledge. **Psychological Review 88**, 16-45.

Pylyshyn, Z.W. (1984). **Computation and cognition: toward a foundation for cognitive science**. Cambridge (MA): MIT Press.

Quillian, M.R. (1968). Semantic memory. In: Minsky, M. (ed.), **Semantic information processing**. Cambridge (MA): MIT Press.

Raynor, K., Carson, M. & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: eye movements in the analysis of semantically biased sentences. **JVLVB 22**, 403-417.

Reddy, M. (1979). The conduit metaphor. In: Ortony, A. (ed.), **Metaphor and thought**. Cambridge: CUP, pp. 284-324.

Rohrman, N.L. & Gough, P.B. (1967). Forewaring, meaning and semantic decision latency. **Psychonomic Science 9**, 217-218.

Rosch, E.R. (1973a). On the internal structure of perceptual and semantic categories. In: Moore, T. (ed.), **Cognitive development and the acquisition of language**. New York: Academic Press.

Rosch, E.R. (1973b). Natural categories. **Cognitive Psychology 4**, 328-350.

Rosch, E.R. (1978). Principles of categorization. In: Rosch, E.R. & B. Lloyd (eds), **Categorization and cognition**. Hillsdale (NJ): Lawrence Erlbaum.

Rosch, E.R. & Mervis, C.B. (1975). Family resemblances: studies in the internal structure of categories. **Cognitive psychology 7**, 573-605.

Rosch, E.R., Mervis, C.B., Gray, W., Johnson, D. & Boyes-Braem, P. (1976). Basic objects in natural categories. **Cognitive Psychology 8**, 382-439.

Rosenberg, S. & Simon, H.A. (1974). Modelling semantic memory: effects of presenting semantic information in different modalities. **CIP working paper 270.** Carnegie Mellon University.

Rubenstein, H., Lewiss, S.S. & Millikan, M.A. (1970). Homographic entries in the internal lexicon. **JVLVB 9**, 487-492.

Rumelhart, D.E. & Norman, D.A. (1978). Accretion, tuning and restructuring: three modes of learning. In: Cotton, J.W. & R.L. Klatzky (eds) , **Semantic factors in cognition.** Hillsdale (NJ): Lawrence Erlbaum.

Rumelhart, D.E. & Ortony, A. (1977). The representation of knowledge in memory. In: Anderson, R.C., R.C. Spiro & W.E. Montague (eds), **Schooling and the acquisition of knowledge.** Hillsdale (NJ): Lawrence Erlbaum.

Ryle, G. (1949). **The concept of mind.** London: Hutchinson.

Sabol, M.A. & deRosa, D.V. (1976). Semantic encoding of isolated words. **Journal of Experimental Psychology: human learning and memory 2**, 58-68.

Sachs, J.S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. **Perception and Psychophysics 2**, 437-442.

Sapir, E. (1921). **Language: an introduction to the study of speech.** New York: Harcourt, Brace & World.

Savin, H.B. & Bever, T.G. (1970). The nonperceptual reality of the phoneme. **JVLVB 9**, 295-302.

Schank, R.C. (1972). Conceptual dependency: a theory of natural language understanding. **Cognitive Psychology 3**, 552-631.

Schank, R.C. (1975a). **Conceptual Information Processing.** Amsterdam: North-Holland.

Schank, R.C. (1975b). The structure of episodes in memory. In: Bobrow, D.G. & A.M. Collins (eds), **Representation and understanding: studies in cognitive science.** New York: Academic Press, 237-272.

Schank, R.C. & Abelson, R.P. (1977). **Scripts, goals, plans and understanding.** Hillsdale (NJ): Lawrence Erlbaum.

Schank, R.C. & Colby, K.M. (eds). (1973). **Computer models of thought and language.** San Francisco: Freeman.

Schweller, K.G., Brewer, W.F. & Dahl, D.A. (1976). Memory for illocutionary forces and perlocutonary effects of utterances. **JVLVB 15**, 325-337.

Searle, J.R. (1969). **Speech acts: an essay in the philosophy of language.** Cambridge: CUP.

Searle, J.R. (1980). The intentionality of intentions and actions. **Cognitive Science 4**, 47-70.

Singer, M.A. (1973). A replication of Bransford and Franks (1971) "The abstraction of linguistic ideas". **Bulletin of Psychonomic Society 1**, 416-418.

Small, S. & Rieger, C. (1982). Parsing and comprehending with word experts (a theory and its realization). In: Lehnert, W.G. & M.H. Ringle (eds), **Strategies for natural language processing.** Hillsdale (NJ): Lawrence Erlbaum.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. **J. of Exp. Psych. 18**, 643-662.

Swinney, D.A. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. **JVLVB 18**, 545-569.

Tannenhaus, M.K., Leiman, J.M. & Seidenberg, M.S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. **JVLVB 18**, 427-440.

Tieman, S.A. (1972). Recognition memory for comparative sentences. Unpublished doctoral dissertation. Stanford University.

Tilburg, W. van (1987). Syllabus geestelijke gezondheidszorg: speciële psychiatrie. Unpublished manuscript. Amsterdam: Free University.

Trabasso, T. (1970). Reasoning and the processing of negative information. Invited Address, Division 3, 78th Annual Convention of the American Psychological Association. Miami Beach, September 1970.

Treisman, A.M. & Tuxworth, J. (1974). Immediate and delayed recall of sentences after perceptual processing at different levels. **JVLVB 13,** 38-44.

Tulving, T. (1972). Episodic and semantic memory. In: Tulving, E. & W. Donaldson (eds), **Organization of memory.** New York: Academic Press.

Turing, A.M. (1950). Computing machinery and intelligence. **Mind 59,** 433-460.

Ultan, R. (1969). Some general characteristics of interrogative systems. **Working papers on language universals 1,** 41-63.

Wanner, E. (1974). **On remembering, forgetting and understanding sentences: a study of the deep structure hypothesis.** The Hague: Mouton.

Wanner, E. & Maratsos, M.P. (1978). An ATN approach to comprehension. In: Halle, M., J.W. Bresnan & G.A. Miller (eds), **Linguistic theory and psychological reality.** Cambridge (MA): MIT Press.

Warren, R. (1970). Perceptual restoration of missing speech sounds. **Science 16,** 392-393.

Whorf, B.L. (1956). **Language, thought and reality** (ed. J.B. Carroll). Cambridge (MA): MIT Press.

Wilson, D. (1975). **Presupposition and non-truth conditional semantics.** New York: Academic Press.

Winograd, T. (1972). Understanding natural language. **Cognitive Psychology 3,** 1-191.

Winograd, T. (1975). Frame representations and the declarative-procedural controversy. In: Bobrow, D.G. & A.M. Collins (eds), **Representation and understanding: studies in cognitive science.** New York: Academic Press, pp. 185-210.

Wittgenstein, L. (1953). **Philosophical investigations** (transl. G.E.M. Anscombe). Oxford: Blackwell.