P-model Alternative to the T-model

Mark D. Roberts,

Department of Mathematics and Applied Mathematics,
University of Cape Town,
Rondebosch 7701, South Africa
roberts@cosmology.mth.uct.ac.za
http://cosmology.mth.uct.ac.za/$\sim$ roberts

&

California Institute for Physics and Astrophysics,
366 Cambridge Ave., Palo Alto, CA 94306.

August 12, 2000

Eprint: http://arXiv.org/abs/cs-CL/9811018

Comments: 28 pages 70596 bytes, six diagrams, 50 references.

Mathematical Reviews Subject Classification: http://www.ams.org/msc/

03D55, 92K20, 68S05.

ACM Classification: http://www.acm.org/class/1998/overview.html

I.2.7,J.4,I.2.6

Keywords: T-model,Movement, Frege Representation.

Abstract

Standard linguistic analysis of syntax uses the T-model. This model requires the ordering: D-structure $>$ S-structure $>$ LF, where D-structure is the deep structure, S-structure is the surface structure, and LF is logical form. Between each of these representations there is movement which alters the order of the constituent words; movement is achieved using the principles and parameters of syntactic theory. Psychological analysis of sentence production is usually either serial or connectionist. Psychological serial models do not accommodate the T-model immediately so that here a new model called the P-model is introduced. The P-model is different from previous linguistic and psychological models. Here it is argued that the LF representation should be replaced by a variant of Frege's three qualities (sense, reference, and force), called the Frege representation or F-representation. In the F-representation the order of elements is not necessarily the same as that in LF and it is suggested that the correct ordering is: F-representation $>$ D-structure $>$ S-structure. This ordering appears to lead to a more natural view of sentence production and processing. Within this framework movement originates as the outcome of emphasis applied to the sentence. The requirement that the F-representation precedes the D-structure needs a picture of the particular principles and parameters which pertain to movement of words between representations. In general this would imply that there is a preferred or optimal ordering of the symbolic string in the F-representation. The standard ordering is retained because the general way of producing such an optimal ordering is unclear. In this case it is possible to produce an analysis of movement between LF and D-structure similar to the usual analysis of movement between S-structure and LF. It is suggested that a maximal amount of information about a language's grammar and lexicon is stored, because of the necessity of analyzing corrupted data.

Contents

1 Introduction.

1.1 Forward.

For the purposes of theory language can be split up into segments: paragraphs, sentences and so forth. It is not always clear what the segments are in natural language, especially spontaneous speech, see for example Rischel (1992) [40]. Linguists and psycholinguists give priority to the analysis of sentences and approach this in different ways. There are a variety of ways of approaching sentence word order. Linguists usually approach word order by invoking the T-model in which various principles are used to change word order from a primitive form (D-structure) to the audible or written form (S-structure); linguists do not seem to realize that the T-model is a psycholinguistic processing model. Other ways of accounting for word order include the Markov cascade models of Brants (1999) [3].

Linguistics is split up into several subdisciplines which include: phonology, morphology, social, historical, semantics, and syntax. According to James McCloskey (1988) [33]:

> the study of syntax has always been a more acrimonious business..... than the pursuit of [its] sister disciplines in linguistics.

Modern syntax grew from the need to record the grammars of North American Indians whose languages where rapidly becoming extinct, much of this work was done by Bloomfield in the 1920's. In the 1950's Chomsky started applying an analogy with pure mathematical category theory to syntax - then syntax began to take its modern form. The way that syntactical investigations usually take place is by analyzing contrived sentences, rather than naturally occurring sentences. Similarly psychology is split up into several subdisciplines which include: social, developmental, psychometric, neuropsychology, and cognitive.

Linguistics and psychology are not as sociologically closely related as one might expect. For example, psycholinguists rarely produce phrase trees of their test sentences: this would be the starting point of any syntactic analysis; an exception to ignoring the work of linguists being Hall (1995) [20] who discusses potential contributions of psycholinguistic techniques to Universal Grammar. Similarly syntacians rarely refer to the measurements of psycholinguists. Some well known linguistic textbooks on syntax, for example Chomsky (1986) [5], invoke and usually start with a processing model called the T-model. Here I point out that this is indeed a psycholinguistic model, and can be subject to the methodology of that discipline. This model requires the ordering: $D(eep) - Structure > S(urface) - Structure > LF(logicalform)$. Between each of these representations there is movement, described by various principles, which alters the order between constituent words.

Frege analyzed the meaning of a sentence to depend on three qualities: sense, reference, and force. In §2.2 it is argued that Frege's *three* qualites which describe a sentence: sense, reference, and force, should be replaced by the *five* qualities: external referents, lexical referents, formal declarants, formal string, and force. These qualities form a representation here called the F(rege)-representation. In the F-representation the order of elements is not necessarily the same as in the LF. It is then suggested that the correct ordering is: $F - representation > D - structure > S - structure$. This ordering leads to a more natural view of sentence production (or processing) called the P-model. Within this framework movement originates as the outcome of the emphasis applied to the sentence; rather than as it occurs in linguistic models where movement is unmotivated and *ad hoc*. The requirement that the F-representation precedes the D-structure needs a picture of the principles and parameters which pertain to movement of words between representations. In general this would imply that there is a preferred or optimal ordering of the symbol string in the F-representation. The general way of producing such an optimal ordering is unclear; but might be found by invoking an extremal principle as discussed in §3.1. In §6 a new model called the P-model is presented, this model uses the standard ordering, as the optimal ordering is still unknown. For the P-model it is possible to produce an analysis of movement between LF and D-structure similar to the usual analysis of the movement between S-structure and LF.

In §3.3 it is suggested that a maximal amount of information about a language's grammar and lexicon (vocabulary) are stored. At first sight this might seem inefficient, but could occur because it allows for a quick analysis of speech which is often only partial heard.

1.2 Language Acquisition.

McDonald (1997) [35] reviews how language learners master the formal structure of their language. *Three* possible routes to the acquisition of linguistic structure are investigated: *firstly* the use of prosodic and phonological information, *secondly* the use of function words to syntactically classify co-occurring words and phrases, and *thirdly* the use of morphology internal to the lexical items to determine language structure, and the productive recombination of these subunits in new items. Evidence supporting these three routes comes from normal language acquirer's and from several special populations, including learners given improvised output, learners with Downs syndrome, and late learners of first and second languages. Further evidence for the three routes comes from artificial language acquisition experiments and computer simulations, see also Ferro *et al* [11]. Language acquisition has also been reviewed by Gleitman and Bloom (1998) [18]. There is also the problem of how the segments of language, as discussed in §5.1, occur in language acquisition. Josephson and Blair (1998) [26] view language acquisition primarily as an attempt to create processes that connect together fruitfully linguistic input and other activity. From a philosophers point of view this is a COHERENCE theory of language, see the discussion in Roberts (1998) [41]. Hymans (1985) [22] discusses how parameters are set in language acquisition. Valian (1990) [50] argues that in a child's acquisition of whether to have a null subject or not, an initial dual switch allows both options; the previously accepted model being either null subjects or not at one value being predetermined at one choice. There is also the question of how language evolved in the first place, the main differences being over whether most of it evolved recently *or* whether it evolved gradually, see the review of MacWinney (1998) [30].

1.3 The T-Model.

The T-model is the basic framework within which linguistic syntax is currently understood. There are various levels with different word order, the word order being altered by movement. The object of linguistic syntax is to find the rules which describe movement. The T-model is usually pictured by diagram one of an upside down Y, (not an upside down T from which its name derives) see for example: Chomsky (1986) [5] p.68, Cook (1988) [7] p.31, Haegeman (1994) [19] p.493, Hornstein (1995) [21] p.2. Brody (1998) [4] presents a system of principles relating the LF representation to lexical items that are compatible to his assumption of no externally forced imperfections in syntax; his assumption is a generalization of the linguistic projection principle.

The T-model illustrated in diagram one consists of several levels. The D-structure (deep structure) level is supposed to hold a given sentence in a primitive form. The D-structure level cannot be the same as the message level in serial psychological models as individual words are supposed to be

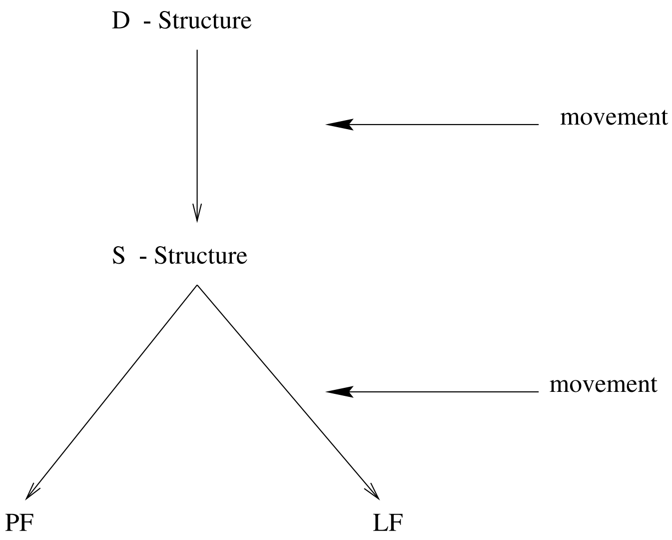

Figure 1: Diagram One - The T-model.

already delineated. In the Garret model (1980) [14] (see also Garman (1990) [13] p.394) the D-structure level could be interpreted as being about midway in the sentence level. This level is then subject to various rules pertaining to how the order of the words can be moved. These rules make up the bulk of the principles and parameters approach to linguistic syntactical analysis. The communicative purpose of such movement is to alter the emphasis of D-structure. The consequence of movement on the D-structure level is to produce another level called the S-structure (surface structure) level. The S-structure sentences are physically realized by the PF (phonetic form) and this corresponds to the positional level representation in serial models. To produce a *post-hoc* analysis of S-structure for its formal content it is postulated, by analogy with movement from S to D-structure, that movement again occurs to bring the sentence into its LF (logical form). The "T" diagram is sometimes made more complex by the addition of other factors, e.g. Cook(1988) [7] p.33. Here movement from S-structure to LF is illustrated with examples of quantifier-raising and Wh-raising taken from Haegeman (1994) [19] Ch.9.

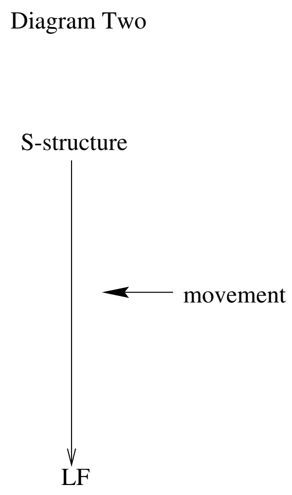

Figure 2: Diagram Two - Quantifier Raising

1.4 Quantifier-Raising.

Consider the surface structure sentence (compare Haegeman (1994) [19] p.489 eq.3):

$$Jones\ saw\ everyone. \tag{1}$$

To this surface structure word order movement is applied to yield the LF word order given by the string 4. This can be represented by diagram two which can be expressed in terms of symbols

$$\forall x, (x \varepsilon H \rightarrow J \varepsilon S x), \tag{2}$$

where H,J and S denote "human", "Jones", and "saw" respectively, and $\forall$ denotes 'for all'. Using traces this can be expressed in terms of words

$$For\ all\ x\ it\ is\ the\ case\ that\ if\ x\ is\ human\ then\ Jones\ saw\ x. \tag{3}$$

or in terms of the quantifier $everyone_i$

$$[Everyone_i[Jones\ saw\ x_i]]. \tag{4}$$

The idea here is that the universal quantifier $everone_i$ can be put first in the LF representation. The universal quantifier could also be put last, but by convention it is taken to come first. To achieve this there is movement from $everyone_i$ being last to being first in the S-structure this leaves a trace x_i in place of $everyone_i$

1.5 Wh-Raising.

Consider the sample surface tructure question (compare Haegeman (1994) [19] p.494 eq.9):

$$Who\ did\ Jones\ see? \tag{5}$$

In this case the word order of the surface structure and the LF remain the same. In term of symbols

$$Wh(x) \quad , \quad x\varepsilon H \quad , J\varepsilon Sx \tag{6}$$

where $Wh(x)$ denotes "Who ...". Using traces this can be expressed in terms of words

$$For\ which\ x, \quad x\ is\ human, \quad is\ it\ the\ case\ that\ Jones\ saw\ x? \tag{7}$$

This has S-structure representation

$$[_{CP}Who_i\ did[_{IP}Jones\ see\ t_i]]? \tag{8}$$

and LF representation

$$[_{CP}Who_i\ did[_{IP}Jones\ see\ x_i]]? \tag{9}$$

The S-structure and the LF have the same form, but with different traces. The S-structure ordering of Wh-phrases is not the same in all languages; in some Wh-words do not appear on the left, so that in these cases both the S-structure and the LF would not have the same form: having LF depending on specific languages is contrary to its name. There can be ambiguity in the scope (or domain of applicability) of the quantifiers; this is illustrated by the sentence (compare Haegeman (1994) [19] p.490 eqs.4 and 5)

$$Everyone\ saw\ someone. \tag{10}$$

which has the two interpretations:

$$For\ every\ x\ there\ is\ some\ y\ such\ that\ it\ is\ the\ case\ that\ x\ saw\ y. \tag{11}$$

and

$$There\ is\ some\ y, \quad such\ that\ for\ every\ x, \quad it\ is\ the\ case\ that\ x\ saw\ y. \tag{12}$$

2 F-level Generalization of Logical Form.

2.1 Drawbacks with the Standard Approach to Logical Form.

The notion of logical form is described in May (1985) [32] and Hornstein (1995) [21], Stanley (1998) [45] discusses the origin of logical form. Logical form as currently understood has at least two drawbacks. The *first* is that it is restricted - in the sense that it is based on the simple calculi of logic such as the predicate and propositional calculi. For example sentences such as

$$The\ probability\ of\ snow\ is\ 80\%. \tag{13}$$

require an understanding of probability and hence of real numbers, $\mathcal{R}$, which have continuous properties as opposed to the discrete properties of simple calculi and standard LF. This implies that sentence 13 requires a larger formal structure, encompassing real numbers $\mathcal{R}$, than is usually included in LF. Sentence 13 can be represented

$$Prob, \quad p\varepsilon[0,1], \qquad (snow = 0.8) \tag{14}$$

with similar representations for other formal mathematical statements such as occur in fuzzy logic. The *second* drawback is that a string such as

$$(x\varepsilon H \rightarrow J\varepsilon Sx)\forall x, \tag{15}$$

unambiguously means the same as equation 2; but the symbol order is different, $\forall x$ (for all x) coming first in equation 2 and last in equation 15. In the principles and parameters approach the order of the symbols is essential, otherwise it is hard to know where to insert traces.

2.2 A New Approach to Logical Form.

To overcome these drawbacks consider Frege's approach to the meaning of a sentence. To quote Dummet (1973) [10] p.83

> Frege drew, within the intuitive notion of meaning, a distinction between three ingredients: sense, tone and force.

Here the variation of these that is used can be represented by diagram three: Reference and force are essentially unchanged, but sense is decomposed into three. Lexical referents are words and the concomitant symbol used in the

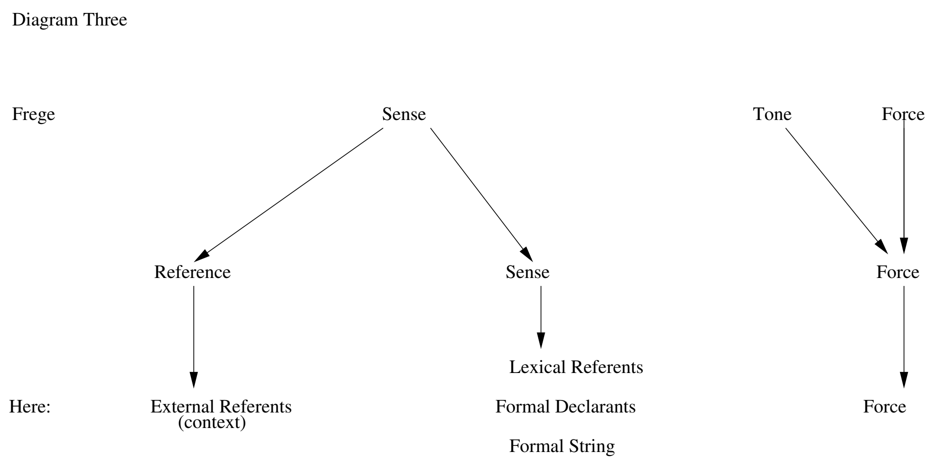

Figure 3: Diagram Three - modification of sentence content from sense, tone & reference.

formal language. Formal declarants are similar to the beginning of computer programmes where the following three things are specified the language used, the parameters used, and the scope (local or global, or other specified ordering to give precedence in case of ambiguity) are specified. The formal string is an ordered set of symbols which are well formed (well defined) in the language. For example the sense for equation 1 would be decomposed as follows

$$\begin{array}{lll} LexicalReferents & : & g = Jones,\ S = saw,\ h = human, \\ FormalDeclarants & : & Predicate\ Calculus,\ x\varepsilon H, \\ FormalString & : & \forall x(x\varepsilon H \rightarrow J\varepsilon Sx). \end{array} \tag{16}$$

Another advantage of the new approach is that sentence 10 is ambiguous leading to either 11 or 12. This ambiguity can be removed by limiting the scope of the variables in the formal declarants.

3 Maximal Verses Minimal Encoding of Information.

3.1 Extremal Principles.

In physics there are *least action principles* which, when the action is minimal, give differential equations which describe the dynamics of systems. The analogy has been carried through to other areas of science, see for example Roberts (1998) [42] and references therein. There is a minimalist program in theoretical linguistics, Lasnik (1998) [27] and Culicover (1998) [8], which invokes an *economy principle* where the steps, symbols and representations in the principle and parameters approach are minimal. Gibson (1998) [16] proposes a theory which invokes *economy of processing*; this theory relates sentence processing to available computational resources. The computational resources have two components, *firstly* an INTEGRATION COST COMPONENT and *secondly* a MEMORY COST COMPONENT which are quantified in the number of syntactic categories that are necessary to complete the current input string as a grammatical sentence. These cost components are influenced by *locality* which entails both 1) the longer a predicted category must be kept in memory before the prediction is satisfied, the greater the cost for maintaining that prediction and 2) the greater the distance between an incoming word and the most local head or dependent to which it attaches, the greater the integration cost. Gibson claims his theory explains a wide range of processing complex phenomena not previously accounted for by a single theory. Lee and Wilks (1999) [28] suggest that it is implausible that there is a highly nested belief structure computing the nature of speech acts, rather there is a minimal set of beliefs. The garden path model §5.4 uses a minimum principle.

3.2 Minimal Principles and Word Order.

If a principle and parameters approach is going to be used in order to legislate movement between the F-level and the sentence level, then some order must be given to both the elements of the formal declarants and the formal string. It is difficult to justify such an order *a priori*: an optimistic hope is that it could be explained by a minimum encoding of information and thus to economy of processing. Sentences, both formal and informal, can contain redundant information. For example in the predicate calculus strings which are always true can be added to a given string without effecting the resulting truth value. It is hard to see how this could be compatible with a minimum encoding of information. Also sentences can be expressed in several ways. For example in the predicate calculus using the four connectives { and, or, not, implies}, or by using the one connective { Sheffer stroke } or { Pierce symbol}, see for example Prior (1962) [39] p.31. For the purposes of the P-model in §6 it is assumed that an ordering of the familiar (or standard) type as used in §1 can be used. For other approaches to word order see Downing and Noonan (1995) [9], and Bozsahin (1998) [2].

3.3 The Maximal Encoding of Information.

It is suggested that a maximal amount of information about a language's grammar and lexicon (vocabulary) are stored. This is also an extremal principle however it is the precise opposite of more common minimal principles. From the point of view of language acquisition, see §1.2, what happens is that when a *lexical item* (a particular sample word) is first heard understanding of it is limited, so that it is only partially learnt, resulting in it being only used in limited contexts. As exposure to the lexical item increases more about its semantic and grammatical properties are learnt so that it can be used in wider contexts. This fits in with aptitude tests for word meaning, where understands word nuance is more important than understanding esoteric words. What is happening here is that a *maximum* *encoding* of information about the word is taking place. Generalizing, this mechanism happens not only to lexical items, but to many other aspects of language, such as understanding of intricate grammatical structures, this is how linguistic *performance* is learnt, compare Garman (1990) [13] §3.1.2. At first sight requiring maximal information about a lexical item might seem inefficient and contrary to economy principles, but could occur because it allows for a quick analysis of corrupted data, such as speech, which is often only partial heard. What is a minimum and what is a maximum has to be kept track of: although the information stored about a lexical item might be maximal, the method of obtaining this information could be minimal.

4 Psycholinguistics Models of Word Production.

4.1 Serial verses Connectionist models of word production.

Psychological models of a segment (or part) of a language come in two basic types: serial and connectionist. Serial models work like a serial computer programme with each operation being performed sequentially. Connectionist models have objects which interact to alter one another's connection weights. These ideas can be applied to whole sentences or individual words. For the T-model or something similar to work there must be serial processing at a late stage in sentence production, because of the discrete nature of its lexical items and representations, see diagram one. This does not preclude connectionist processing before the D-structure representation, or even before the understanding of an individual word. There are PROLOG models of restricted sentence production, Johnson and Klein (1986) [24, 25].

4.2 The Stemberger Interactive model of individual word production.

There are psycholinguistic models of individual word production. There is evidence that the reception of speech is interactive, for example there is evidence that seeing the speaker speak influences the word heard, McGurk and McDonald (1976) [34]; also Tanenhaus *et al* (1995) [48] examine how visual context influences spoken word recognition and mediated syntactic processing, even during the earliest moments of language processing. Furthermore there is evidence that verbal speech production interacts with gesture, McNeil (1985) [37], and various other physiological activities, Jacobson (1932) [23] p.692, and these observations suggests that the verbal

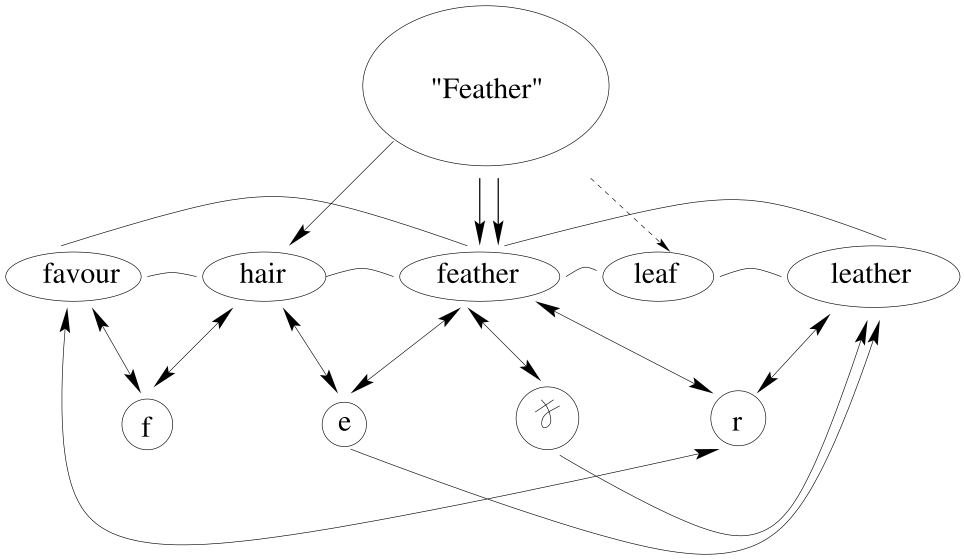

Figure 4: Diagram Four - The Stemberger Diagram.

part of speech production is interactive. A model which allows both phonological and semantic influences to interact is the psycholinguistic interaction speech production model, c.f. Stemberger (1985) [46]. In this model, when the word "feather" is activated a lot of other words are also activated with varying weights according to how closely they resemble "feather". This can be pictured by diagram four, to quote Stemberger's p.148 text: to quote the text:

> Semantic and phonological effects on lexical access. ... an arrow denotes an activating link, while a filled circle is an inhibitory link. A double line represents a large amount of activation, a single solid line somewhat less, and a broken line even less. Some of the inhibitory links have been left out ... for clarity. The exact nature of semantic representation is irrelevant here, beyond the assumption that it is composed of features; ... a word in quotation marks represents its meaning.

There is suppression (also called inhibition) across a level, and activation up or down to the next level. This model accounts for syntax by giving different weights to the different words so that words on the left come first. Speech errors come from the noise in the system. There are *three* kinds of noise. The *first* is that the resting level of a unit node is subject to random fluctuations; with the result that it is not the case that the unit nodes degree of activation remains at the base line level. A fluctuation could produce a random production of a part of a word. The *second* is that words that are used with a high frequency have a higher resting level, and therefore reach activation threshold, or "pop out", quicker. This implies that there should be less error for these high frequency words; furthermore it implies that when real words occur as an error, higher frequency words should occur as errors more often, and this does not happen. The *third* is the so-called systematic spread of activation; this means that the weights in the interaction allow an inappropriate activation of word. There are connectionist programmes of word recognition, e.g. Seidelberg and McClelland (1989) [44].

4.3 Atomist Semantic Feature Models.

There are atomist semantic feature models (Garman (1990) [13] p.388) to which serial models of individual word meaning could be built, however the connectionist interaction picture seems to have less drawbacks. From now on it is assumed that words in some concrete form can be assumed and the question becomes how are they related to make longer structure such as sentences.

4.4 The Marslen-Wilson model of word recognition.

Marslen-Wilson (1987) [31] claims that the process of spoken word recognition breaks down into three basic functions: access, selection and integration. *Access* is concerned with the mapping of the speech input onto the representations of lexical form. At the earliest stages of the lexical access process, all lexical memory elements whose corresponding words begin with a particular acoustic sequence will be fully activated, these activated words are called a COHORT, see also Garman (1990) [13] p.286. *Selection* is concerned with the discrimination of the best fitting match to this input. *Integration* is concerned with the mapping of syntactic and semantic information at the lexical level onto higher levels of processing. Context comes into play after there is information about the syntactic and semantic properties of members of the cohort. Such models embody the concepts of multiple access and multiple assessment, allowing for a maximally efficient recognition process, based on the principle of the contingency of perceptual choice.

5 Psycholinguistic Models of Sentence Production.

5.1 Sentence production in larger structure.

Research in psycholinguistics can be split up according to the size of the structure under study, for example text processing, sentence processing and word meaning. Here structure larger than sentences, for example discourse (see Graesser *et al* [17]), are not looked at. Sentence processing is concerned with how the syntactic structures of sentences are computed, and text processing is concerned with how the meanings of larger units of text are understood. Research in both domains has begun to use the information that can be obtained from a large corpora of naturally occurring texts. In text processing, recent research has focused on what information the words and ideas of a text evoke from long term memory quickly, passively, and at low processing cost; text processing is not looked at here. According to McKoon and Ratcliff (1998) [36] in recent sentencing research, a new and controversial theme is that syntactic computations might rely heavily on statistical information about the relative frequencies with which different syntactic structures occur in language. Gerdemann and van Noord (1999) [15] discuss various rewrite rules used in several areas of natural language processing, such rewrite rules might change word order. Hall (1995) [20] discusses the representations of various linguistic competences. Sentences can mean different things in different contexts, see Akman and Surav (1998) [1] and MacDonald *et al* (1994) [29]. Ferro *et al* (1999) [11] produce learning transformation rules that find grammatical relations and find that grammatical relations between core syntax groups bypasses much of the parsing phrase. Tabor *et al* (1997) [47] describe a model which works by analogy with dynamical systems. Attractors are taken simultaneously to have properties of syntactic categories, with some encoding of context dependent lexical information. Various experiments were contrived which examined the interactions of simple lexical frequencies, and their results favoured their dynamical model over traditional approaches. Truswell *et al* (1994) [49] devise two eye-movement experiments which show that animate nouns where harder to disambiguate when parsing.

5.2 The Clarke & Clarke serial model of sentence production.

There are numerous serial models of sentence production and textbooks on the subject for example Rosenberg (1977) [43]. Here **three**, the Clarke and

Clarke serial model, the Garret serial model, and the garden path model are very briefly presented before going on to the P-model. Clarke and Clarke [6] p.278 mention the formulation of an articulatory [sic] program [sic] which has five steps:
(1)*Meaning Selection*: The first step is to decide on the meaning the present constituent is to have.
(2)*Selection of a Syntactic Outline*: The next step is to build a syntactic outline of the constituents. It specifies a succession of word slots and indicates which slots are to get primary, secondary, and zero stress.
(3)*Content word selection*: The third step is to select nouns, verbs, adjectives, and adverbs to fit the appropriate slots.
(4)*Affix and function word formation*: With the content words decided on, the next step is to spell out the phonological shape of the function words (like articles, conjunctions, and prepositions), prefixes, and suffixes.
(5)*Specification of phonetic segments*: The final step is to build up fully specified phonetic segments syllable by syllable.
By *step (5)*, the articulatory program [sic] is complete and can be executed. Typically, however, people monitor what they actually say to make certain it agrees with what they intended it to mean. Whenever they detect an error, they stop, correct themselves, and then go on. It seems likely that the more attention is required elsewhere - in planning of various sorts - the less likely they are to detect an error. Indeed, many tongue-slips go unnoticed by both speakers and listeners.

5.3 The Garret serial model of sentence production.

The Garret (1980) [14] (see also Garman (1990) [13] p.394) has three basic levels: the message level, the sentence level, and the articulatory level. At the message level a mental model or image of what is about to be expressed is formulated. Loose ideas of the form of the individual words and the overall structure in which they are expressed are formulated to give the sentence level. Here the actual words to be used and the structure in which they occur are crystallized to give the positional level representation which is articulated.

5.4 The Garden Path Model.

More recent garden path models are reviewed in Frazier (1987) [12], who says on pages 561-562 (slightly adjusted):

"In the garden path model, perceivers incorporate each word of an input into a constitute structure representation of the sentence, roughly as each item is encountered. At each step in this process, the perceiver postulates the minimal number of nodes required by the grammar of the language under analysis, given the structure assigned to preceding items. This leads to the *two* principles of the garden path model:
(1)*Minimal Attachment*: Do not postulate unnecessary nodes.
(2)*Late Closure*: If grammatically permissible, attach new items into the clause or phrase currently being processed (i.e. the phrase or clause postulated most recently)."
In other words minimal attachment entails that a perceiver, given so much of the beginning of a sentence chooses the minimal completion of it which makes sense both grammatically and semantically. The minimal attachment principle is just the requirement that the ultrametric height of a sentence should be a minimum, see Roberts (1998) §3, which was written before I had heard of the garden path model. Minimal attachment is also in accord with the minimal principles discussed in §3.1. McRae *et al* (1998) [38] use time measuring experiments to see how event specific knowledge resolves structural ambiguity. Their results suggest that the structure of sentences is best described by a semantic constraint model, then a garden path model with a very short delay, and finally by a one region delay garden path model. Their models and experiments show that event specific knowledge is used immediately in sentence comprehension, and this agrees with maximal encoding of information in §3.3 above.

6 The P-model.

6.1 Description of the P-model.

The P-model can be pictured by the diagram five. Semantic intent produces a F-level representation of the sentence and its context and force. Thought (semantic intent), occurs prior to words (lexical referents). The external referents and the lexical referents combine to produce binding referents which constrain movement between the D-structure and S-structure representation. The D-structure is constructed from the lexical referents (words) and the formal declarants and string. The reason that the D-structure sentence changes to the S-structure sentence is because of the emphasis (or Fregean Force) the speaker wishes to convey in the sentence. It is known, see the beginning of §4.2, that in some cases gesture and other behaviour co-occur

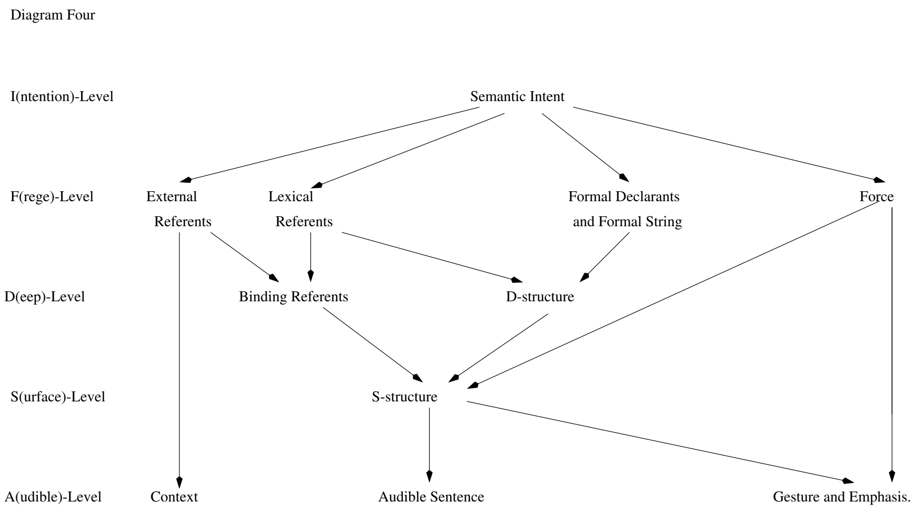

Figure 5: Diagram Five - The P-model.

with spoken sentences: hence the Fregean force and the S-structured sentence interact to produce the gesture and emphasis concomitant with the audible production of the sentence. It is possible to produce variants of the above analysis.

6.2 The P-model from a principle & parameters viewpoint.

From a principles and parameters view point: for the T-model movement between S-structure and LF has to be explained, however for the P-model movement between the F-level and D-structure has to be explained. Disregarding the caveats of §2 concerning LF, the P-model requires an explanation of movement between LF and D-structure, as depicted by diagram six. This can be illustrated by choosing similar examples to §1.4,1.5, but now suggesting quantifier lowering and Wh-lowering.

6.3 Quantifier Lowering.

This can be represented by diagram six or in terms of symbols

$$\forall x(H\varepsilon x \rightarrow J\varepsilon Sx) \tag{17}$$

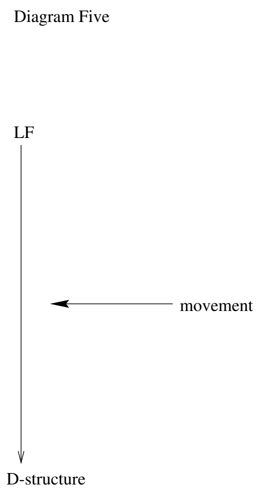

Figure 6: Diagram Six - Quantifier Lowering

where H, J and S denote "human", "Jones", and "saw" as in §1.4. This can be expressed in words

$$y_i \text{ } Jones\ saw\ everyone_i \tag{18}$$

or in terms of the quantifier $everyone_i$

$$Everyone\ Jones\ saw. \tag{19}$$

where y_i the trace left by the movement of the word $everyone_i$.

6.4 Wh-lowering.

This can again be represented by diagram six, and corresponds to §1.5. In term of symbols

$$Wh\ (x), x \varepsilon H, J \varepsilon Sx. \tag{20}$$

which can be expressed in words

$$y^i\ did\ Jones\ see\ who^i? \tag{21}$$

or filling in the trace

$$Who\ did\ Jones\ see? \tag{22}$$

7 Conclusion.

The T-model is explicitly a serial sentence production model. It is the basic framework in which modern linguistics is currently understood, linguists being mainly occupied with describing movement between D-structure and S-structure. In psychology there are also serial sentence production models which bear little resemblance to the T-model. Rather than try to directly accommodate the T-model into a psycholinguistic serial model here an entirely new model has been constructed based upon altering the T-model from what the author perceives to be its defects. One of the defects is that the T-model requires that a S-structure sentence can after it has been uttered be analyzed into a formal component called LF. This analysis requires movement analogous to that between D-structure and S-structure. There is no apparent sentence production purpose for doing such a *post-hoc* analysis, the reason seems to be that historically given sentences were analyzed for their logical content. Here it was argued that the correct place for LF is before D-structure. This fits in naturally into a processing model of sentence production, where the LF can be thought of as part of the message level. Despite expressing reservations about string order in LF (at least as it is commonly understood), it was shown in some simple cases how movement can be described between LF and D-structure. In this case description of movement has a purpose as it elucidates part of a sentence production model. From the point of view of a linguist the most important requirement of any model is whether there is movement between a D-structure and S-structure. This is retained in the P-model.

8 Acknowledgements.

This work has been supported in part by the South African Foundation for Research and Development (FRD).

References

[1] Akman, Varol & Surav, Mehmet §5.1
Steps Towards Formalizing Context.
`http://cogprints.soton.ac.uk/abs/comp/199806020`
AI Magazine **17**(1998)55-72.

[2] Bozsahin,C. §3.2
Deriving the Predicate-Argument Structure
for a Free Word Order Language.
`http://arXiv.org/abs/cmp-lg/9808008` 20 Aug.1998

[3] Brants, Thorsten §1.1
Cascaded Markov Models.
In: Comutational Natural Language Learning (CoNLL-99). A Workshop at the 9th Conference of the European Chapter of the Assoc. for Computational Linguistics (EACL-99), Bergen, Norway, June 1999.
`http://arXiv.org/abs/cs.Cl/9906009`

[4] Brody, Michael §1.3
Projection & Phrase Structure.
Ling.Inq. **29**(1998)367-398.

[5] Chomsky,N. §1.1, 1.3
Knowledge of Language: Its Nature, Origin, and Use.
Praeger(1986), New York.

[6] Clarke,H.H. & Clarke,E.V. §5.2
The Psychology of Language,
Harcourt Brace Jonanovich, (1977) New York.

[7] Cook,V.J. §1.3
Chomsky's Universal Grammar: An Introduction.
Basil Blackwell, (1988)Oxford.

[8] Culicover, Peter W. §3.1
The Minimalist Impulse.
Syntax & Semantics **29**(1998)47-77.

[9] Downing,P. and Noonan, M.(eds) §3.2
Word Order in Discourse, Typological Studies in Language (TSL)
John Benjamin Publishing Company, (1995)Amsterdam.

[10] Dummet,M. §2.2
FREGE Philosophy of Language,
Duckworth London,(1973) Second Edition (1981)p.83.

[11] Ferro, Lisa; Vilain, Marc & Yeh, Alexander §1.2, 5.1
Learning Transformation Rules to Find Grammatical Relations.

In: Comutational Natural Language Learning (CoNLL-99)pp 43-55.
A Workshop at the 9th Conference of the European Chapter of the Assoc. for Computational Linguistics (EACL-99),
Bergen, Norway, June 1999.
`http://arXiv.org/abs/cs.Cl/9906015`

[12] Frazier, Lyn §5.4
Sentence Processing: A Tutorial Reviw.
Attention & Performance **XII**, The Psychology of Reading.
edited Max Coltheart(1987)559-586.

[13] Garman,Michael §1.3, 3.3, 4.3, 4.4, 5.3
Psycholinguistics, (1990)
Cambridge University Press.

[14] Garret,M.F. §1.3, 5.3
Levels of Processing in Sentence Production,
Ch.**8** in B.Butterworth (ed.), Language Production,Vol.**1**,
Academic Press(1980), London.

[15] Gerdemann Dale & van Noord Gertjan §5.1
Transducers from Rewrite Rules with Backreferences.
`http://arXiv.org/abs/cs.CL/9904008`

[16] Gibson, Edward §3.1
Linguistic Complexity: Locality of Syntactic Dependencies.
Cognition **68**(1998)1-76.

[17] Graesser, Author C.; Millis Keith K. & Zwaan Rolf A. §5.1
Discourse Comprehesion.
*Annu.Rev.Psychol.***48**(1997)163-89.

[18] Gleitman, Lila & Bloom, Paul §1.2
Language Acquisition.(1998)
`http://mitpress.mit.edu/MITECS`

[19] Haegeman,Liliane §1.3, 1.4, 1.5
Introduction to Government and Binding Theory.
Second Edition, Basil Blackwell, (1994) Oxford.

[20] Hall, Christopher J. §1.1, 5.1
Formal Linguistic & Mental Representations: Psycholinguistic Contributions to the Indentification and Explanation of Morphological &

Syntatic Competence.
Lang,& Cog.Processes **10**(1995)169-187.

[21] Hornstein,N. §1.3, 2.1
Logical Form: from GB to minimalism.
Blackwell, (1995) Massachusetts.

[22] Hymans,N. §1.2
The Theory of Parameters & Syntatic Development.(1985)
p.1 in Parameter Setting, eds. T.Roeper & E.Williams.
D.Reidel Publishing Company, Dordrecht.

[23] Jacobson,E. §4.2
Electrophysiology of Mental Activities.
Am.J.Psy., **44**(1932)677-94,especially page 692.

[24] Johnson,M. & Klein,E. §4.1
Discourse, Anaphora and Pharsing,
CSLI Report(1986)86-63.

[25] Johnson,M. & Klein,E. §4.1
A Declarative Formulation of Discourse Representation Theory.
J.Sym.Logic **51**(1986)846.

[26] Josephson Brain D. & Blair David G. §1.2
A Holistic Approach to Language.
`http://cogprints.soton.ac.uk/abs/ling/199809001`

[27] Lasnik,H. §3.1
Minimalism.(1998)
`http://mitpress.mit.edu/MITECS`

[28] Lee, Mark & Wilks, Yorick §3.1
An Ascription-based Approach to Speech Acts.
In: Proceeedings of Coling -96, Copenhagen.
`http://arXiv.org/abs/cs.CL/9904009`

[29] MacDonald,Maryellen C.; Pearlmutter, Neal J.
& Seidedberg, Mark S. §5.1
Lexical Nature of Syntatic Ambiguity Resolution.
*Psy.Rev.***101**(1994)676-703.

[30] MacWinney, Brian §1.2
Models of the Emergence of Language.
*Annu.Rev.Psychol.*__49__(1998)199-227.

[31] Marslen-Wilson,William D. §4.4
Functional Parallelism in Spoken Word-recogition.
Coginition **25**(1987)71-102.

[32] May,R.C. §2.1
Logical Form: its structure and derivation.
MIT Press (1985),Massachusetts.
& Logical Form in Linguistics.
`http://mitpress.mit.edu/MITECS`

[33] McCloskey, James §1.1
in Linguistics (1988): the Cambridge Survey.
Cambridge University Press.

[34] McGurk,H. and McDonald,J. §4.2
Hearing Lips and Seeing Voices.
Nature **264**(1976)746-8.

[35] McDonald, Janet L. §1.2
Language Acquisition: The Acquisition of Linguistic Structure in Normal and Special Populations.
*Ann.Rev.Psychol.*__48__(1997)215-241.

[36] McKoon,Gail & Ratcliff,Rober §5.1
Memory-based Language Processing:
Psycholinguinguistic Research in the 1990's.
*Annu.Rev.Psychol.*__49__(1998)25-42.

[37] McNeil,D §4.2
So You Think That Gestures are Non Verbal?
Psy.Rev. **92**(1985)350-371.

[38] McRea, Ken & Spivey-Knowlton, Michael J. §5.4
Modeling the Influence of Thematic Fit (and other Constraints) in On-line Sentence Comprehension.
*J.Mem.&Lang.*__38__(1998)283-312.

[39] Prior,A.N. §3.2
Formal Logic.
Second Edition (1962), Claredon Press, Oxford.

[40] Rischel, Jørgen §1.1
Formal Linguistics & Real Speech.
Speech Communication **11**(1992)379-392.

[41] Roberts,Mark D. §1.2
Radical interpretation Described using Terms
from Biological Evolution.
`http://arXiv.org/abs/cs.CL/9811004`

[42] Roberts,Mark D. §3.1
Ultrametric Distance in Syntax.
`http://arXiv.org/abs/cs.CL/9810012`

[43] Rosenberg,S. §5.2
Sentence Production: Developments in Research and Theory.
Lawrence Erlbaum Associates, (1977)New Jersey.

[44] Seidenberg,M.S. and McClelland,J.L. §4.2
A Distributed Development of Word Recognition and Naming.
Psy.Rev. **96**(1989)523-68.

[45] Stanley,J. §2.1
Logical Form, Origins of.(1998)
`http://mitpress.mit.edu/MITECS`

[46] Stemberger,J. §4.2
An Interactive Activation Model of Language Production.
Ch.**5** in A.W.Ellis(ed.)Vol.**1** (1985) Psychology of Language.
London,Erlbaum.

[47] Tabor, Witney; Juiano, Cornell & Tanenhaus, Michael K. §5.1
Parsing in a Dynamical System: An Attractor-based Account of the Interaction of Lexical and Structural Constraints in Sentence Processing.
Lang.& Cog.Processes **12**(1997)211-271.

[48] Tanenhaus, Michael; Spivey-Knowlton, Michael J.; Eberhard, Kathlenn & Sedivy, Julie C. §4.2
Integration of Visual and Linguistic Information in Spoken Language Comprehesion. *Science* **268**(1995)1632-1634.

[49] Trueswell, John,C. ;Tanenhaus Michael K.& Garnsey, Susan M. §5.1
Semantic Influences on Parsing: Use of Thematic Role Information in Syntatic Ambiguity Resolution.
J.Mem.& Lang. **33**(1994)285-318.

[50] Valian, Virginia §1.2
Null Subjects: A problem for Parameter-setting Models of Linguistic Acquisition. *Cognition* **35**(1990)105-122.