CHAPTER II

REVIEW OF LITERATURE

What is Style?

ENKVIST has listed the various approaches to style under the following six heads which would serve as a convenient starting point for our discussion of style :[3]

1. Style as a shell surrounding a preexisting core of thought or expression
2. Style as a choice between alternate expressions
3. Style as deviation from a norm
4. Style as a set of collective characteristics
5. Style as a set of individual characteristics
6. Style as those relations among linguistic entities that are statable or may be statable in terms of wider spans of text than the sentence.

Style as a shell surrounding a preexisting core of thought or expression

Bally sees the origin of style in the addition of a contenu affectif to expression. Any expression is composed

---

[3] ENKVIST (53).

of two layers, one of which, the effective, consists of the core of the thought necessary for comprehension and the other, the affective, which consists of elements due to the ego of the person, etc. Bally has given an illustration to demonstrate the difference between 'language usuel et banal' and the affective language.[4] Selecting a short passage from DAS RECHT DER FRAU by L.Fulda, Bally gives two renderings in French which convey the same message but differ in their affective content. An emotional state is portrayed in the second version by means of structural and lexical devices : the dependence on semi-sentences, choice of lexis from different registers etc. According to Bally a text sans aucune nuance affective lacks style. Sol Saporta on the other hand says, "it seems desirable to distinguish between the so-called expressive features from the linguistic features that will be relevant to discussion of style," adding further that performance or delivery does not belong to the message and is irrelevant for style (SAPORTA (153)). He cites as example :

There is a big bear in the woods

There is a Bi::G bear in the woods.

In our opinion it is a moot point whether such sharp distinctions can be made between suprasegmental features which

---

[4] cited in KAYSER (98), pages 298 ff.

belong to Langue and those which belong to Parole. As paraphrase of the second expression one may take :

There is a very big bear in the woods.

Here the syllabic length and stress have been replaced by the modifier very, which would have to be taken as a necessary part of the message.

In order to separate those linguistic features in an utterance which go to make up the affective element, we have to make first of all the assumption that an utterance U is composed of two elements C and A, where C represents the core of the thought and A the affective part. The second assumption is that the addition of C and A does not bring in a third element into play. In other words the addition of A does not have an influence on C and vice versa. Within semantics itself this is often not the case. Granting these assumptions however A has to be arrived at either by a process of reduction (using segmentation and deletion) or by comparing with a paraphrase (which may be theoretically looked at as derived through substitutions, though the psychological reality might be so, that one total expression has been replaced by another total expression). By the first we can arrive at features like stress, pitch, which are affective and also at syntactic features like repetition of a word or phrase which show

affective function. In the second method, in comparing paraphrases, we assume the semantic invariance of deep structures of a set of expressions (the assumption being supported by intuition or by testing inter-subjectively on informants) and investigate what, where and how affective features have come into play. In practice both methods will have to be used so that ultimately it is a case of comparing alternate expressions. The existence of A cannot be demonstrated without comparing C+A with (C+A)-A. This leads us to the concept of style as a choice.

Style as a choice between alternate expressions

Where the choice of an item is made on extra-linguistic grounds : X loves Mary, X = John or X = Peter, or on grammatical grounds the choice is non-stylistic. Stylistic choice is said to occur between two expressions which are almost equal, having a common referent and appearing in the same frame :

He is a X man, X = fine or X = nice.

In order to circumvend the near-equality-in-meaning criterion ENKVIST brings in the notion of context as being more amenable to objective study. The element of choice is emphasised differentially by different scholars. WELLS (184) states that style is understood to be optional like vocabulary as contrasted with grammer. Operationally this

would mean the setting up of classes of equivalent expressions and comparing them. If the premise is, "style is a choice," then it is an act of verbal behaviour and ceases to be the concern of linguistics. If however the statement is amended to read that in such a choice different elements of style are manifested in different expressions, the question is : exactly what elements are to be characterised as stylistic. For instance, in the statements

Joe is a fine man

Joe is a nice man

there is apriori no grounds to say one is more stylistic than the other. It becomes therefore necessary to check the macrocontext, establish a norm and then make the comparison.

Style as deviation from norm

Thus both the perspectives of style described above envisage the setting up of a norm as a yardstick of measurement. ENKVIST (53, page 28) has stated :

> The style of a text is a function of the aggregate of the relations between the frequencies of its phonological, grammatical items and the frequencies of the corresponding items in a contextually related norm.

BERNARD BLOCH attempts to avoid the subjective selection of a norm by wording his definition as follows :

> The style of a discourse is the message carried by the frequency distributions and transitional probabilities of its linguistic features especially as they differ from those of the same features in language as a whole.

Even by means of a computer it is possible to cover 'language as a whole' only by taking random samples and applying statistical methods. WERNER MUELLER (255) has also argued that even a "choice" will become style only when taken in comparison with a norm. If every text shows the same choice, reflected in the same frequency and transition probability of the unit then it will not belong to 'style' but to 'language'.[5] MUELLER denies the very

---

[5] Here one might argue that agreement in statistical measures between all authors belonging to the same era belong to 'the period', between authors of the same literary genre to genre, etc., leading finally to language common characteristics. FUCKS (209) has demonstrated that the length of words in terms of syllable (syllable per word) is a language specific characteristics.

possibility of discovering even one stylistic index on the basis of one text alone. A norm has to be chosen which has to be so determined that it can be taken to be representative of a totality called "texts". Style then exhibits itself in the difference between the chosen form and the mean values of the norm.

The statistical definition of a norm assumes that the factor of choice involved in the production of any literary work is nullified when numerous works are being considered. Where the variables are formal parameters like sentence length, syllable length, a quantitative approach to norm is easy, but when one is working with psychological categories like, for instance, expressions revealing fear extensive preliminary searches will have to be made before a reasonable norm with acceptable means are available, in other words before the norm can be said to represent the population and the statistical measures of the norm like mean, standard deviation can be taken for the corresponding values of the population. Even when the reliability of these ratios have been tested it is questionable whether they do full justice to all aspects of language. We may quote FUCKS and LAUTER (209) :

> Ob eine selbst erschoepfende Erfassung der Formalstruktur Authentizitaetsprobleme grundsaetzlich

> zur Entscheidung bringen kann oder nicht, kann nicht ohne weiteres gesagt werden. ...
> Obwohl die Angaben ihrer Natur nach exakte Zahlenwerte sind, ist immer zu bedenken, dass es Wahrscheinlichkeitsaussagen sind.[6]

Further, with a quantitative norm, it is for the researcher to interpret the message carried by the differences in the various probabilities. A qualitative norm is equally well conceivable and is presumably intuitively applied by every literary critic. In this context ENKVIST introduces the interesting notion of style markers. The style markers are the linguistic items that only appear or are most or least frequent in one group of texts. Stylistic choice involves the choice of style markers whereas non-stylistic choice involves selection from stylistically neutral items. The style markers that appear in the same text form a stylistic set for that text. An overlap of stylistic sets within a

---

[6] Whether even a very exhaustive covering of formal structure can or cannot thoroughly decide problems of authenticity, cannot be said straightaway. ...
Although the data are by their very nature exact numerical values one has still to bear in mind that they are all statements of probability. (My translation)

given passage or text could mean that the criteria for defining the sets will have to be revalued or there might be a mixture of styles or a shift in styles in the text itself. Enkvist wishes to separate literary effect from style, the two belonging to two different categories, literary criticism and stylistics.

The term 'norm' as used here need not carry with it a prescriptive value attached to it. RENE WELLEK and AUSTIN WARREN (182, page 151) designate as norm the small part of every individual experience of a work of art so that the real work of art is conceived of as a structure of norms. But any norm implies automatically deviation from norm, where again, deviation need not be associated with any approving or disapproving judgement.

Deviations can be classified binarily into poetic deviations and non-poetic deviations. Poetic deviations have an aesthetic function. The function of non-poetic deviations is to provide the contrastive background to the poetic deviations. (Deviation is here understood merely as difference from norm.) In any literary work there is always a constant interaction between the two so that poetic language may be said to operate as a chain reaction. DOLEZEL (202, page 278) speaks of <u>automatisation</u> - conventional patterning evident in communicative language - and

actualisation - individual variants which transgress the conventional patterning as happens in poetic language. There is a constant attempt to standardise actualisation so that there is always tension between the need to automatise and the need to actualise. NARASIMHAN (257, page 5) distinguishes between expected deviations and unexpected deviations.

The fundamental question posed by such distinctions in regard to style is : Should style concern itself only with unexpected poetic deviations or should it also include deviations in general? BAUMGAERTNER (15, page 69) has pointed out that there is no statistical explanation for ambiguities, a remark which would apply also to poetic deviations which very often result from the usage of individual lexical items. Although theoretically it should be possible to deal with 'expected' and 'unexpected' deviations with a probabilistic model, in practice the near-zero probabilities of any word or construction makes it hardly a feasible venture. With transitional probabilities also it is not much better. Hence by restricting style to poetic deviations alone not only is the purpose of stylistics restrictively defined, but also the methods available to it are restricted.

The more generalized view of stylistics permits the

use of mathematical methods and bears also relation to the fourth and fifth concepts of style enumerated on page 9 : style as a set of individual characteristics and style as a set of collective characteristics. CARROLL (29) argues that in a deliberately coded message the author by reason of the skills at his command might choose particular styles or style which may not reflect his personal "enduring characteristics." Carroll is of the opinion that personal non-deliberate conversation might exhibit the enduring invariant characteristics of style. Related to a literary work this would imply that poetic function is to be distilled out and that what we aim at is the personality of the author. But most literary critics are however agreed that a work of art is autonomous. KAYSER (98, page 289) writes : "Dichtung kann und muss zunaechst als ein Gebilde betrachtet werden, das voellig selbstaendig ist und das sich restlos von seinem Schoepfer geloest hat und autonom ist."[7] Carroll's conception of style is in essence psychological and has nothing to do with

---

[7] "Poetry can and must be at the first instance regarded as a thing that is completely independent, that has separated itself entirely from its creator and is autonomous." (My translation)

linguistics. Under collective characteristics JENKINS understands a style concerned with art form, to be distinguished from a second style which comprises of what the individual does to vary the characteristics of the art form.[8] The implications of this will be discussed after dealing with the final concept of style which is rather isolated and beare little relation to the previous views.

Style as relations statable above the sentence level

HILL (83, page 406 ff.) calls all those relations among linguistic entities which are statable or may be statable in terms of wider spans of text than the sentence as constituting style. The definition excludes structural or lexical relationships within the framework of the sentence itself. Linguistics deals with phonology, morphology and syntax as manifested within the sentence, but when literary analysis is undertaken, when the structural relationships between the sentences of a work of art are being studied the investigation is metalinguistic.[9]

---

[8] Jenkins, comments to part 8 in STYLE IN LANGUAGE, page 332

[9] Hill (84, page 405).

The grammatical similarities between the sentences (on the levels of phonology, morphology, and syntax) form 'the first level of microliterary structure'. From this level analysis moves on to the lexical level and by investigating the interrelationships between the units used at these two levels the total pattern of the poem is arrived at. Hill cautions that 'any further statements of meaning are in the metaliterary sphere of correlation between the literary structure and known facts of patterned cultural behavior and values.'[10] In this context it is interesting to compare WELLEK & WARREN (182, page 179) :

> We can write the grammer of a literary work of art or any group of works beginning with phonology and accidence, going on to vocabulary (barbarisms, provincialisms, archaisms, neologisms), and rising to syntax (e.g., inversion, antithesism and parallelisms).

Hill conscientiously avoids any appeal to meaning and searches for formal patterns, and the similarities in the patterns between sentences. Obviously the grammer of stylistics will depend upon the basic assumptions made regarding the units of description, the definition given

---

[10] Hill (84, page 406).

to style and the methods employed. In all methods there is comparison, either work-intern (HILL) or work-extern norm (ENKVIST). The definition covers only the affective (BALLY) or the aesthetic elements (WELLEK) or all characteristics (JENKINS). The current work takes the most general view which is amenable to quantitative description since all other 'styles' can be derived from it by a process of abstraction. (vide Chapter I)

Transformational Grammer - an exposition and a critique

Introduction

The merits and demerits of transformational generative grammer (TG) must be judged in the light of the objectives set forth by the transformational grammarian. However if any theory is to be accepted as a part of general linguistic theory or even as "the" theory then it must prove its worth independently both of the claims made by the theoretician as well as the deficiencies admitted by him.[11] In the following paragraphs a historical background to TG is presented followed by a description of TG, an example of TG in practice to serve as sample as well as for purposes of

---

[11] This refers in particular to the absence of descriptive or taxonomic procedures in TG.

comparison with other methods, and major criticisms levelled against it.

Coseriu points out that G. von Gabelentz[12] suggested a form of grammar similar to that of TG as early as 1901 in his book "Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse" (Leipzig).[13] Chomsky himself does not fight shy of tracing the origin of some of his ideas to medieval ages.[14] However two modern branches which appear to have sparked off the advent of TG are Automata Theory and Mathematical Logic. Hilbert's axiomatic approach in mathematics is mentioned by Katz.[15]

---

[12] Coseriu's lectures during the Summer term (1968) at Tuebingen contain a good account of TG.

[13] Gabelentz's grammar consisted of two parts, a synthetic one to show the formation of sentences and analytic one, to identify the functions in a given sentence. For instance the following constituted some propositions in it :

1. The subject stands in front of the predicate.

   Eg.: $A^{Subj}{}_{B}Vact.{}_{C}^{obj.} = \emptyset$ (Sentence)

   $BC = P$ (Predicate)

2. The active verb stands in front of its object.

3. If an otherwise active V stands at the end of the

An axiomatic system is characterized by 4 things : a vocabulary (list of symbols), a set of axioms, a set of formation rules and a set of inference rules.

13.. 3. sentence, then it is passive.

$C^{Subj.}{}_{B}{}^{Vpass.} = \emptyset$ (sentence)

$C^{Subj.}{}_{B}{}^{Vpass.}$ iūe $A^{Agent} = \emptyset$

Sentence = ABC

Biūe = P

4. Active can be changed into passive - logical object comes before verb etc.

5. Every predicate can be transformed into an adnominal attribute if it occurs in front of the logical subject and the relative particle $\overset{Y}{c}$i occurs between the relative particles.

$P^{Attr.}$ ci A (or C)$^{Substantive} = \underset{n}{\emptyset}{}^{Substantive}$

(subtantivised part of the sentence)

6. Relative sentence so formed is replaced by a substantive using the relative pronoun $\overset{Y}{c}$e :

paò min ce one who protects the people

paò iūe wang $\overset{Y}{c}$é who/what is protected by the king.

It may be noted that 1-3 are constituent structure, 4 is a transformational rule and 5-6 are replacement rules.

The inspiration from mathematics becomes more evident when one considers the fundamental theorems and basic ideas in recursive function theory. Thus the set of integers is infinite, but given any number, however big, we have an algorithm for deciding in a finite number of steps whether the number is a prime. Further the said number is formed by concatenation of a small finite set of elements, namely the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Analagously one wants to set up formal algorithms for deciding grammaticalness or wellformedness or acceptability.[16] Chomsky's own contributions in the field of formal languages are well-known.[17]

The theory of universals[18] and the creative aspect of language[19] are both quite old. The contribution of TG to these ideas lies in its attempts to formalize them.

---

[14] vide Chomsky (32).

[15] vide Katz (96).

[16] These concepts are dealt with in greater detail later.

[17] Some context-free languages bear the name "Chomsky languages". Vide Nelson (258), chapter 8 on Generators for a lucid exposition of Chomsky's contributions to formal languages.

[18] Chomsky (32).

The assumption is that "each natural language is organized on the same formal pattern."[20] The theory of language shall concern itself with a "formulation of the universals of language."[21] TG differentiates two kinds

---

[19] Chomsky acknowledges his debt to Wilhelm von Humboldt for this aspect. Humboldt subscribed to numerous ideas on language, with not all of which the transformationalists would agree. In his chapter on "Verschiedenheiten des menschlichen Sprachbaues" (Humboldt, 91, page 153) he writes : Die Vorstellung, dass die verschiedenen Sprachen nur dieselbe Masse der unabhaengig von ihnen vorhandenen Gegenstaende und Begriffe mit anderen Woertern bezeichnen und diese nach anderen Gesetzen, die aber, ausser ihrem Einfluss auf das Verstaendniss, keine weitere Wichtigkeit besitzen, an einander reihen, ist, ehe er tiefer ueber die Sprache nachdenkt, dem Menschen zu natuerlich, als dass er sich leicht davon losmachen koennte." ... "Die wahre Wichtigkeit des Sprachstudiums liegt in dem Antheil der Sprache an der Bildung der Vorstellungen." (The idea that different languages designate with other words the same mass of things and concepts, existing independently of language, and that they arrange these words according to other laws

of universals : substantial ones actually manifest in language like consonant, vowel, animate, inanimate and formal ones, which relate to rules appearing in grammar. Distinctive features would constitute a substantial set of universals in phonology.[22]

19.

which however have no further importance except for their influence on understanding, is so natural to man, before he begins to more deeply on language, that he cannot get away from it so easily." ... "The real importance of language study lies in the participation of language on the formation of concepts.") Brown (27, page 111) describes at least 3 pairs of partly antithetical ideas in Humboldt's thinking which he had inherited and the contradictions and conflicts of which he never finally resolved : i) a belief in the validity of deductive theoretical thought as against a belief in the importance of inductive, empirically based thought, ii) a belief that there are important universals characteristic of all nations and all ages, as against a belief that each nation and age shows important individual peculiarities, iii) a belief in the power of the individual to shape the collectivity through his own actions as against a belief in the

In explaining the creative aspect of language Chomsky specifically rejects both Skinner's ideas on verbal behaviour as well as the structuralistic conception of innovation through analogy.[23] The generative grammar as a system of rules is or shall be capable of reflecting the creative aspect of language. Chomsky admits that structuralistic grammars can be reformulated as rules of phrase structure grammar, but they have not paid sufficient attention to the "production and interpretation of new, previously unheard sentences."[24]

19.. power of the collectivity over the individual. Humboldt's initial interest in language universals shifted later to emphasis on fundamental differences (in keeping with developments in the empirical science of comparative anthropology).

20 Katz (96, page 11). Katz does not claim that languages exhibit isomorphic formal structures. Only in some undetermined sense there is organizational similarity.

21 Katz (96, page 107)

22 Chomsky (31, page 65). Chomsky is careful not to present any set of substantive phonetic features as the proper universal set; the only claim made is that there is justification for assuming the existence of such a set.

Theory of Grammar

Investigations of a child learning language have led to the concept of competence. A child is able to sort out the rules of the language and generate sentences beyond what it has heard concretely. Apparently intuition is at work. It is not mere learning of patterns.[25] According to Chomsky the child possesses the scheme of a general grammar, the totality of formal universals even before language learning.[26] Chomsky distinguishes between the speaker-hearer's knowledge of his language i.e. his competence

---

[23] In MIT Quarterly Progress Report No.88 pages 283-285, Halle and Chomsky setting forth their research objectives write : "... creative aspect, that is, its unboundedness and freedom from stimulus control ... Nor are these utterances "generalizations" from past experience ... Nor can language use be described in terms of "habits" or "repertoires" of responses." See also notes 21, 22, Chomsky (32), pages 81-82.

[24] Chomsky (31), page 67 and note 30, page 205.

[25] Chomsky, Formal Nature of Language, page 400, in Lenneberg (110).

[26] Chomsky (31), page 25. The child possesses innately even before language learning "a linguistic theory

and performance (the actual use of language in concrete situations).[27] This distinction is parallel to the distinction between langue and parole, but whereas Saussure regards langue as an inventory of items Chomsky regards competence as a system of generative process.[28] A grammar of a language describes the "ideal speaker-hearer's intrinsic competence."[29] The term 'ideal' implies that there is an element of abstraction involved, that grammar "neither synthesizes particular sentences as does the speaker nor does it recognize the structure of presented sentences as does the hearer."[30]

26.. that specifies the form of the grammar of a possible human language, and, second, a strategy for selecting a grammar of the appropriate form that is compatible with the primary linguistic data."

27 Chomsky (31), page 4. "The problem for the linguist, as well as for the child learning language, is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance."

28 Chomsky (31), page 4.

29 Chomsky ibid, page 4.

30 Chomsky (34), page 120.

Limitations of Phrase-Structure-Grammar (PSG)

Constituent structure grammar has no provision for recursion or for transformational rules, both these factors leading either to combersome repetition of rules or the exclusion of apparent regularities from grammatical description.[31] As an extreme case Chomsky gives the example :

the man was old, tired, tall ... but friendly.

There is, Chomsky maintains, no internal structure in the co-ordinated items, but a constituent structure grammar would assign a rule for each co-ordination.[32] Binary splitting (characteristic of and central to constituent structure analysis imposes often superfluous structure.[33] Descriptive linguistics has held the view that syntactic structure is determined exclusively by operations of segmentations and classification. Chomsky lists sequential treatment, restriction to set of base strings, as against full set of actual sentences, introduction of complex symbols, and the separate treatment of the lexicon as

---

[31] Postal (141).

[32] Chomsky (34), page 128.

[33] Chomsky (31), page 88.

departures from the 'taxonomic' view.[34] Theoretically the greatest drawback in Immediate Constituent analysis is its assumption that deep and surface structures are actually same. IC-analysis hence does not provide an adequate account of deep structure.[35]

In part Chomsky's views on Grammar of "human" languages are influenced by his works on formal languages where different systems of rules generate different languages. Thus he speaks of grammars of a "particular" language (as opposed to universal) and the need for comparing them on the basis of adequacy, simplicity and explicative power.[36] Nevertheless he expects of grammar as such to fulfil certain basic requirements. A grammatical theory about a particular language must specify or predict "all and only the sentences" of the language.[37] As opposed to the taxonomic approach which views langue as a mere inventory of rules regarding regularity and exception, the

---

[34] Chomsky ibid, page 16.

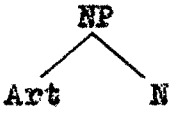
[35] vide page 80 regarding the claims of IC-Analysts.

[36] Chomsky (31), page 60. A descriptive theory must contain i) a universal phonetic theory that defines the notion 'possible sentence, ii) a definition of structural description, iii) a definition of generative

generative approach views langue as a system of production rules.[38] TG does not work with an actual corpus since a corpus is never complete and scientifically unsatisfactory in many ways. The corpus of TG is the ideal set of infinite sets of all the sentences of the language. Grammar shall be a system to produce or generate these infinite sets of sentences. If grammar merely generates sentences it is called weak generation and if it generates also a structural description to each of the sentences indicating how this sentence is understood by the ideal speaker/hearer,[39] it is called strong generation.

36..
grammar, iv) a method for determining the structural description of a sentence, given a grammar. For explanatory adequacy an additional condition is imposed, v) a way of evaluating alternative proposed grammars. Bach (6), page 106 : "when two grammars generate the same strings, the simpler one is that which has fewer context restricted rules."
The preoccupation which TG has with notations is explained by Chomsky in (31, p.42) .. "The obvious numerical measure to be applied to a grammar is length in terms of number of symbols. ... Thus it is the notational conventions used in presenting a grammar

The major components of a generative grammar are a syntactic, a phonological and a semantic one. The syntactic component specifies for each sentence two structures : a deep structure and a surface structure.[40] Structure can be described in the form of labeled phrase markers or equivalently as rules : a Noun phrase will bear the label NP and will be branched into Article and Noun, i.e.



or NP ——— Art + N

36.. that define "significant generalization" if the evaluation measure is taken as length".

37 Bach (6), page 5; Chomsky (34), page 119 defines Grammar as a device of some sort (that is set of rules) ...

38 Production rules of simple Markov type would rewrite a non-terminal symbol as either a terminal symbol or a terminal symbol + a non-terminal symbol.
A ——— Bc, A ——— c (where A, B are non-terminal and c is terminal). (Postal (141), page 149)

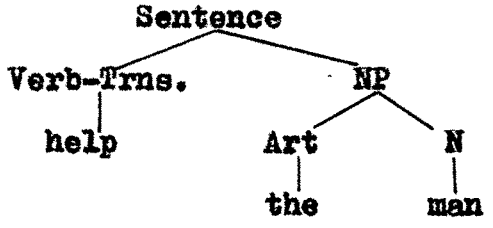
39 Chomsky (31), page 60.

40 vide Chomsky (32), pages 31-51, chapter on "Deep and Surface Structure". A criticism follows in our later paragraphs.

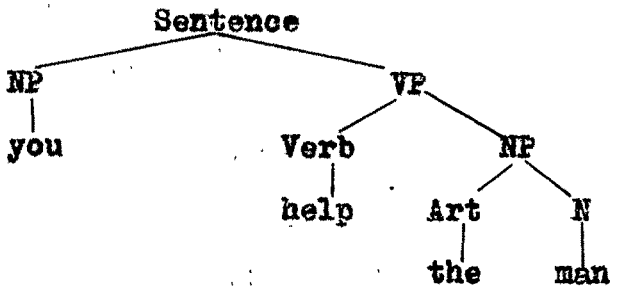
Thus the sentence

help the man[41]

will have the surface structure represented by the set of rules or markers :



But from the fact that imperatives can contain only, if at all, the reflexive "yourself (we have "help yourself", but not "help myself, themselves etc.) Katz posits the underlying structure of "help yourself" as "you help you". Correspondingly the underlying structure of "help the man" is given as



The Subject "you" will be deleted by a transformational rule, when the surface structure is derived from the deep structure. The purpose of such an analysis is not to establish that the subject of an imperative sentence is

---

[41] Katz (96), page 130 ff.

"really there" or "understood" in the mind of the speaker.[42] The analysis is presented merely on grounds of simplicity.

The Transformational grammarians claim that their demonstration of the necessity for 2 phrase markers or 2 structures in a syntactic description is a major achievement of TG.[43] The function of the surface structure is to serve as input for the phonological components, the output being the phonetic interpretation.[44] Similarly the semantic component will assign a semantic interpretation to the deep structure.[45] The full (syntactic) conceptual machinery consists of a base component that generates deep structures through a system of rules and a transformation part which maps these deep structures into surface structures.[46] The phonological and the semantic components interpret the underlying phrase markers and derived phrase markers respectively.

---

[42] Bach (6), page 98.

[43] P.Postal (142).

[44] The concept of phonemes has no place in TG. In fact Halle has shown that the phoneme theory is empirically unwarranted. The phonological components in TG consist on the other hand of a bundle of features.

[45] Chomsky (31), page 16.

[46] Originally the base component was to be strictly limited to a system of phrase structure rules. Chomsky (31,

Coseriu distinguishes four phases in the development of transformational generative grammar. In the first phrase sentences were conceived of either as kernel sentences or those that could be derived transformationally from Kernel sentences.[47] The transformational rules were obligatory for the phonetics and morphology of sentences, but optional for active/passive etc.[48] In the next phase ideas of deep structure thrust 'kernel sentences' to the background. The formation rules refer to deep structure while the transformational rules refer to the transition from deep structure to surface structure. TG in this form was a transitional phase. In the third phase semantics is introduced in the form of a lexicon. Active/passive transformation is taken over to deep structure, as they have different but similar structures.[49] Previous optional rules become now obligatory[50] and the notion of kernel sentences is given up. In the fourth phase proposed by Lackoff and Lees, but not accepted by Chomsky, deep structure is simply equated to meaning.

46.. page 99) later viewed its main role as one of determining grammatical relations.

47 Bach (6), page 69 : "Sentences which are derived from PS terminal strings by the application of obligatory transformation only and phonological rules are called

Transforms

The concept of transform itself has several variations though within the TG it has got a clear meaning. Lenneberg calls transformations the abstraction from two seemingly different structures and noting their inner sameness.[51] Akhmanova and Mikael'an write "in our case 'transformation' would imply giving an English form to a certain content either as purely 'conceptual' to begin with or already (previously) materialized in the vernacular sentence.[52] HARWOOD has pointed out that works on syntax usually give some information regarding the equivalences between some sequences and others as for instance

John discovered the path
= the path was discovered by John.[53]

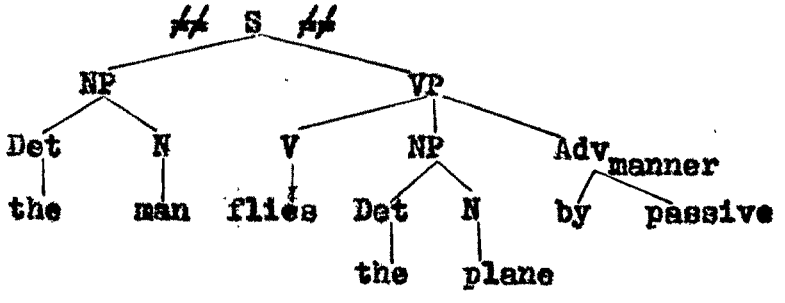
47.. Kernel sentences. ... We choose those strings as belonging to the kernel which lead to the simplest over-all grammar."

48 Chomsky (30), page 61.

49 As Verbs in English which take passive are those which take Adverbials of manner, passive sentences are derived from sentences with this label. A rewriting rule replaces ^{Adv}manner through "by + passive" and a

HJORTH[54] defines a statement as a transformational rule, following Harwood, if and only if it makes certain assertion about two expressions A and B of a language. Hjorth discusses then the adequacy of the definitions based on assertions. Any assertion that A and B are, in one of their meanings each, synonym or equivalent, will be called a transformational rule. If one wants to construct the other expression B on the basis of A by means of the transformational rule then this definition is considered inadequate. Other definitions suggested and discussed are in respect of conditions i.e., A, B must be cognitively synonymous or logically derivable.

49..
passive transformation takes care of the rest.



50 Chomsky (31), page 132.

51 Lenneberg (110), pp.299-301.

52 Akhmanova (2), p.105, note 27. The reference is to teaching English as foreign language.

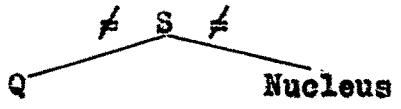
Although TG exhibits a bias towards logical structural relations and a concern for "thought schemes" (Denkschemata as the basis of language universals) paraphrase relations i.e. equivalence of propositions have not yet come under the purview of TG. Transformations are thus defined not on the basis of logical definitions but serve purely to mediate between deep structures and surface structures. Transformations introduced at various appropriate derivational stages of a sentence help to give abstract deep structures overt surface form. Originally transformations were of two types : optional transformations and obligatory transformations on the one hand and singulary transformations and generalized transformations on the other. An optional transformation converted an active sentence into a passive one. Subsequent work on TG showed that grammar could be simplified if the passive is not conceived of as a surface structure transformation, but both passive and active are already manifested in the deep structure. (See footnote 49, page 38) A dummy marker in the deep structure under the category Adverbial of manner enables the generation of passive sentences. The choice of the dummy marker would

---

53 P.W.Harwood, "Axiomatic Syntax", page 409, cited in Hjorth (86), p.91.

54 Hjorth (86), pages 91-94.

make the passive transformation then obligatory. In the same manner questions are generated by means of markers in the deep structure :



In the earlier work

did Bill see John and Bill saw John

as well as will Bill see John and Bill will see John

had the same underlying P-markers or deep structure. The transformational rule brought along with it a change from assertive to question as well as a change of meaning. Arguing that questions can be paraphrased somewhat like imperatives,

will Bill see John

being same as : I request that you answer 'X Bill will see John'

where X is "one of special class of sentence adverbials" including yes, no, of course, etc, the full paraphrase would be

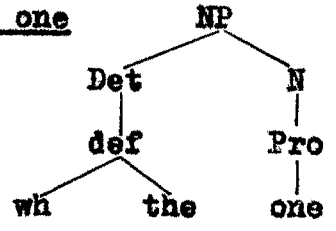
I request that you answer 'yes, Bill will see John or no Bill will not see John.'

The Q-marker has then the function of giving the appropriate reading, distinguishing yes - or - no questions from their corresponding declaratives. In order to differentiate the various types of "wh" - questions which can be asked apart

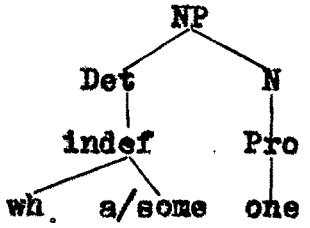
from yes-no questions, the Q marker is intended as global marker to signify a question i.e. introducing a reading "I request that you answer" and subcategorization is effected by means of the presence of "wh" marker and "either-or".[55]

An interesting feature of Katz and Postal's proposal is the derivation of the "wh" marker as a subconstituent of the definite or idefinite article and not of the noun in NP. The argument is that there are two types of wh questions in English, one where the wh-word represents the definite article + noun, and the other where it represents the indefinite article + noun or pronoun.

<u>which one</u>



<u>who</u>



Whether a transformational rule is optional or obligatory thus depends upon the deep structure markers posited.

A transformational rule consists of three parts, a <u>structural description</u>, the <u>structural change</u> effected and the <u>conditions of application</u> of the transformation.

---

[55] Katz & Postal (97).

Such a rule can be of the form

$\not$, X, Noun Phrase, Y —— 1, 3, 2, 4 (optional except where 2 does not contain Q)
1 2 3 4

where 3 dominates wh.[56]

This will bring the NP dominating the question marker to the front. In the earlier stages of TG complex sentences were derived by means of generalized transformations. Out of the matrix sentence (1) and the constituent sentence (ii)

i) du kannst ihn + C_1 + fuehlen

ii) er dreht sich

C_1 implying the empty slot where the constituent structure is added on, the generalized sentenced :

du kannst ihn sich drehen fuehlen.[57]

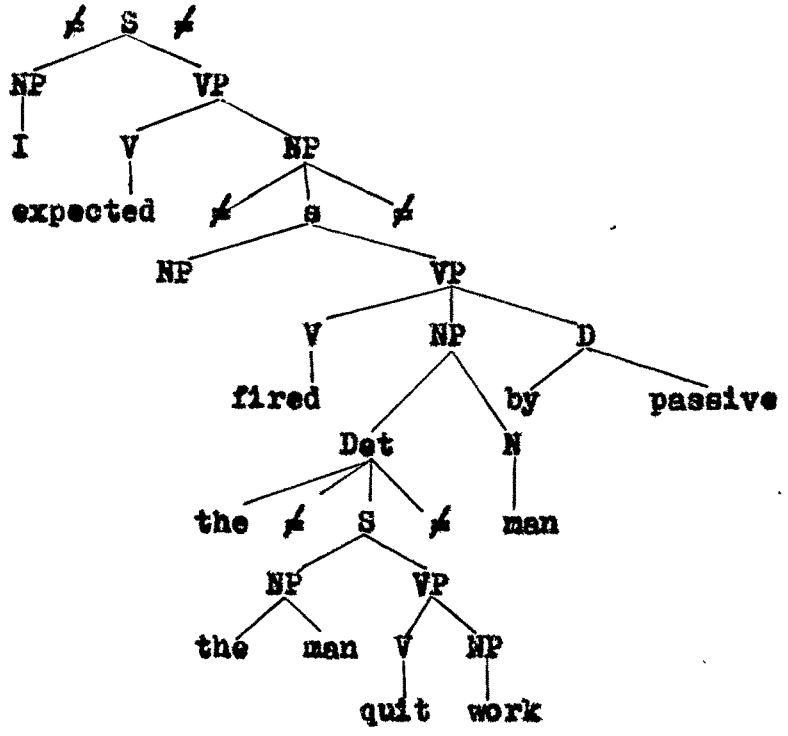
is obtained. Following Katz and Postal however, the need for generalized transformations has been eliminated for two reasons, namely that TRANSFORMATIONS SHALL NOT INTRODUCE MEANING BEARING ELEMENTS and that deep structure can generate straightaway complex sentences by recursively introducing $\not$ S $\not$ at intermediate nodes also (i.e. not only as initial symbol).[58] By specifying a dummy symbol in the

---

[56] Katz & Postal, ibid.

[57] Bierwisch (23) p.124, points out that the reflexive depends upon the subject of the constituent sentence. Another way of looking at it is that it is dependent

underlying structure then the need for generalized transformations is altogether eliminated. Hence a generalized phrase marker can be derived as follows :



Summing up one might say, transformations, employed to delete, permute or substitute elements transform abstract underlying structures into concrete surface structures.

57..

upon the object of the matrix sentence. When the object is not there, the constituent element will also not be there :

du kannst etwas fuehlen — du kannst ihn sich drehen fuehlen
du kannst sein Drehen fuehlen — du kannst fuehlen, dass er er sich dreht

Categorization in TG

Mention has already been made of the fact that TG sees no reason for setting up an intermediate level of representation between the morphemic and the phonetic levels. (page 34 , note 44)[59] As the rewriting rules of the form A —— X are arranged according to the generative needs of the grammar, there appear to be no fixed set of levels such as word level, clause level, phrase level, sentence level, in the phrase structure rules.[60]

---

[58] Chomsky (39), pages 39/40. Weinreich (181) also accepts the stipulation that meaning must be restricted to deep structure. vide also (31) p.132.

[59] Chomsky (39), pp.49/50.

[60] Bach (6), page 59. As units, however, lexis and S (sentence) play a fundamental role in TG. Both clauses and phrases are derived from S as sentences. (the old man is here = the man is here, the man is old). Fries (56, page 18) says, words as parts of speech cannot be determined in terms of their syntactic function. This problem is nevertheless circumwended by TG since the members of the various major categories and minor ones are defined by enumeration. Subcategorization is effected on the basis of syntactic functions, like, for

The categorial component consists of grammatical morphemes and the symbol Delta. In the preterminal strings generated by this component lexical items will be inserted from the lexicon (the other part of the base component). A lexical item will be substituted for Delta only if the item belongs to the same category as the one which immediately dominates Delta.

Apart from the category feature a lexical item may have other features. A Verb can be + or - Transitive, a Noun can be + or - Human. Features like Transitive are called contextual features and features like Human are called non contextual features. Features which stipulate the frames in which the lexical item can occur are called strict subcategorization features. Thus Transitive feature means the frame —— NP. A Verb like persuade has the frame —— NP PP (the feature will be denoted as (—— NP PP). Eg. I persuaded John of the pointlessness of his actions.

60..

instance, environment NP PP in the case of verb "persuade". But "John" and "of the pointlessness..." are surface structures. If the labels NP and PP refer not to superficial surface characteristics, but to some deeper underlying structures TG has never shown

The other types of contextual features are called selectional features. "Whereas the strict subcategorization features specify categorial frames in which an item appears, the selectional features of a lexical item X specify lexical features of the items with which X enters into grammatical relations.[61] Thus one meaning of "eat" will indicate that its subject must be specified as (+Animate).

Perhaps the most interesting aspects of TG are the studies in subcategorization and the derivation of corresponding grammatical morphemes from higher order grammatical

---

60..
what these structures are. Weinreich treats the preposition "like" as zero morph in the sentence :

He seemed like a man,

a morph needed purely for surface structure. The crucial question here is, what is the semantic role assigned to prepositions and how far are they treated as automatic concomitant variations of the verb. Chomsky and other transformational grammarians assign semantic role or deep structure status to prepositional phrases, but the categorization has not gone beyond what traditional grammarians have been saying since the Middle Ages (categories like rate, manner, degree, frequency etc. are being used just as in traditional grammar).

morphemes. Example of deriving the question word from the definite or indefinite article in the Noun Phrase has already been given. As a detailed example of the type of work done by TG we may cite the treatment of the Comparative in English.[62] The comparative has been derived four ways : as complex adjectival, adverbial, nominal and as complex verbal.

As complex adjectival[63]

Modifiers of nominals exhibit certain regularities. For instance, if the noun phrase is indefinite the modifiers generally follow the head.

Last night I saw something very strange
I chased the wretched thief
\+ I chased the thief wretched

Hence these modifiers are derived from relative clauses by rules like :

1) Last night I saw something
Something was very strange —— Last night I saw something which was very strange

---

61 Chomsky (39), page 45.

62 Hale (68).

63 Hale (68), page 8. The study is by Carlota S. Smith : 'A class of complex modifiers in English Language (July-Sept. 1961).

Similarly I chased the thief who was wretched

ii) Reduction of relative clause - Post nominal position

Last night I saw something very strange

I chased the thief wretched

iii) A rule for prenominal permuation - obligatory where the noun is definite. Non-applicable elsewhere.

I chased the thief wretched ——— I chased the wretched thief

Secondly Adjective + non-verbal complements are always postnominal

I bought a book which is yellow with age

=== I bought a book yellow with age.

\+ I bought a yellow book with age

\+ I bought a yellow with age book

With definite article even postnominal position is not possible : + I bought the book yellow with age

In the case of Adjective + verbal complements, relative clause reduction is not possible with definite article :

She heard the sound which was too frightening to ignore

\+ She heard the sound too frightening to ignore

\+ She heard the too frightening sound to ignore

However with indefinite Noun phrase reduction is possible and permutation is optional :

She heard a sound too frightening to ignore

She heard too frightening a sound to ignore

Hence when these conditions of relative clause reduction and permutation are properly stated, it is possible to derive complex adjectivals starting with the formation of the relative clause.

To introduce the comparatives, the three rules (relative clause formation, reduction, and permutation) are preceded by the following rules :

1) Comparative conjunction :

a man is tall
Bill is tall —— -er than —— a man is taller than Bill is tall

2) Obligatory deletion :

A man is taller than Bill is tall —— A man is taller than Bill is

3) Optional deletion :

A man is taller than Bill is —— A man is taller than Bill

Such rules generate a wide range of comparative sentences, but they also generate strings like :

\+ John is a taller man than Mary.

Hale suggests the possibility of regarding <u>-er than N</u> as a kind of coherent structure which modifies the adjective (here 'tall').[64]

---

[64] Hale ibid, page 14.

Comparative as Complex Adverbial[65]

Lees has as a starting point the similarity in constructions involving *that* and *more than* :

John is that intelligent

John is more intelligent than Bill.

The comparative is directly embedded to a dominating mode, ADV_a (attributive Adverb). The rules take the form :

1) Adverbial embedding :

Jack built a (Adv_a) large house
Joe built a large house
——— Jack built a -er than Joe built a large house large house

2) *Than*-Complement Permutation :

——— Jack built a (-er) large house (than Joe built a large house)

3) Deletions and Morphophonemics :

——— Jack built a larger house than Joe did.

This analysis has the advantage that it can take care of graded comparisons like :

Bill is taller than Mary by more than Sam is taller than Pete.

The English Comparative as a Complex Nominal[66]

This is based on the derivation of the attributive

---

65 Lees (109).

adverbial from prepositional phrases

John is very tall

= John is tall <u>to a great extent</u> (degree)

Thus John is taller than Bill

= John is tall to an extent to which Bill is not tall.

The negative is then removed optionally and replaced by surface comparative elements, (optional because the sentence is acceptable as it is also). The motivation lies in the fact that negative preverbs and negative verbs as well as <u>not</u> are excluded from <u>than</u> clauses.

\+ Bill is taller than John isn't.

\+ Bill is taller than John hardly is.

<u>The English Comparative as a complex Verbal</u>[67]

In this view comparison is posited simply as a two-place predicate with a class of verbs such as, <u>exceed</u>, <u>is greater than</u>, <u>is less than</u>, <u>is as great as</u> and the like. The paraphrases are of the following type :

More people came than were invited	The number of people who came was greater than the number of people who were invited
John was taller than Bill	John's height was greater than Bill's height
John ran faster than Bill	The rapidity of John's running was greater than the rapidity of Bill's running

---

[66] Hale ibid pages 32 ff

[67] Hale ibid pages 45 ff

Critique of TG

Much of the criticism of TG is rather ill-founded as it rests on misinterpretations of the goals and claims of TG. Although TG attempts to explicate the intuition of the speaker/hearer, Chomsky categorically says that it is not a model for a speaker/hearer. He further disclaims that the interpretative or generative processes taking place in the brain correspond to the branchings and rules suggested in his grammar.[68] If the pronoun is taken as understood in the imperative it is done so for a simpler exposition of the grammar and not due to psychological reality. With these preliminary remarks we proceed to a critical analysis of the various aspects of TG.

TG and creativity

Coseriu points out that innatism, taken as the basis of TG although not proven, leads to a primitive banalisation of the creative character of language. The real character of the creative activity has been dealt with before by Aristotle, Kant, Schelling and others. In our opinion it is difficult to make proper assessment since the usage of this term is ambivalent. Unboundedness and freedom from stimulus control of human language is called its creative aspect

---

[68] Chomsky (31), page 140.

by Chomsky.[69] But creativeness is ascribed also to the mediating role played by the syntactic component in the semantic interpretation of a phonetic representation.[70] Chomsky assumes that the set of grammatical sentences is somehow given in advance.[71] Properly speaking it is an infinite set of sets of sentences and formal apparatus to generate these sentences has perforce to employ recursion in some manner or other. The recursiveness provided for the symbol ≠ S ≠ in the base component is too powerful a creative device to describe by itself natural language. The transformational rule acts at times like a filtering device to prevent ungrammatical sentences being generated.

Putnam suggests that the grammatical sentences can be considered a recursive lot provided 3 arguments are accepted :

i) the self-containdedness of language. A person who considers 'many good home' deviant would always consider it so irrespective of the context

---

69 see note 23.

70 Chomsky (31), pages 135/136.

71 Chomsky (30), page 103.

ii) meaningless sentences can be classified whether grammatical or non-grammatical

iii) grammatical intuition of a speaker, namely that any person knows in his language what sentences are grammatical and what are not.[72]

By self-containedness of language is meant apparently the non-appeal to extra-linguistic factors like situation. Appeal to extra-linguistic factors may be necessary to decide which of the possible interpretations of an utterance are meant, but the 'resolution' of ambiguity, ascribing possible but semantic or syntactic structures to the same utterance (surface structure) does not require appeal to situation. Computational linguistics has clearly demonstrated how explicit rules expose ambiguities which may not even be noticed by a hearer or reader.[73] Secondly consistency in considering certain utterances as always deviant presupposes decidability about grammaticalness. This condition is in our opinion rather too strong. When formal procedures to decide grammaticalness are not available one has to have recourse only to grammatical intuition and it is not clear if the constancy of grammatical intuition can be accepted.

---

[72] Putnam (144).

[73] Kuno & Oettinger (238) point out, for instance, how <u>They are flying planes</u> can receive a third interpretation (though absurd), analagous to : <u>The facts are smoking kills</u>.

Psychological tests have shown that it is impossible to define grammaticality empirically.

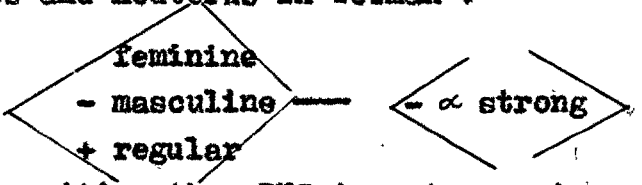
TG and Syntax

Syntax occupies a central position in TG. It consists of a set of structures which associate the deep structure of a sentence with its surface structure. Criticism of this aspect of TG refers mostly to the neglect of the paradigmatic axis in favour of the syntagmatic.[74] Although paradigms have not been treated explicitly in TG they are implicit in the rules of surface structure (phonological component).[75] TG's disregard of the usual levels of phrase, word, clause, and of paradigms is motivated partly as a reaction to American structural linguistics. The concept of

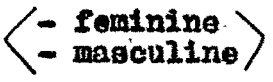
---

[74] Akhmanova (2), pages 38/39 regarding criticism by Saumjan.

[75] Zwicky (189) for instance gives the following rule for incorporating the feature plural er for irregular masculines and neuterus in German :



If ∝ is positive then RHS is -strong, i.e. weak.
Left hand side = regular feminines.
If ∝ is negative then RHS is + strong, i.e. L.H.S. = regular neuters are strong, neuter being defined as



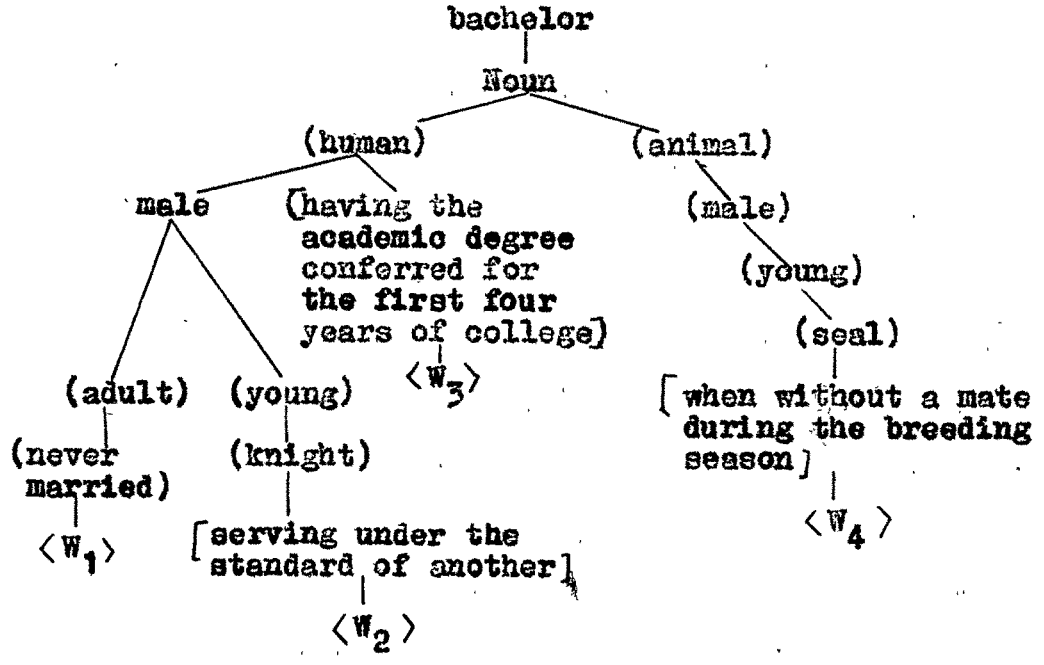
function has also been dealt with differently by Chomsky. Notions like subject, object were considered purely as relations existing between different parts of a sentence and were hence defined on the basis of the relationship existing between the different nodes in the (generalized) phrase marker.[76] But it has become increasingly clear that formal subcategorization is motivated considerably by semantic functions and one cannot avoid using functions as labels for the nodes, if the deep structure is to reflect fully the changes in meaning. TG which defines word classes like Adjectives extensionally[77] tries to overcome this difficulty in formal description by making the lexicon a powerful device, wherein lexical items are associated with semantic, syntactic and phonological features with all the paraphernalia necessary to introduce selectional restrictions in Grammar for guaranteeing the generation of only grammatical sentences. Computational linguistics has shown the way, but the failure of mechanical translation so far raises misgivings whether TG will be fully successful in its goals.[78]

---

[76] Chomsky (31), p.71. The definitions proposed are for instance : Subject-of (NP,S); Predicate of (VP,S); Direct object of (NP,VP)

TG and Semantics

Weinreich has criticised that TG has been unable to draw a distinct line of demarcation between the domains of syntax and semantics. However it has not been shown that such a strict demarcation is possible at all. The failure of automatic translation seems to imply in fact the very opposite. Another criticism leveled is that the content is not determined in opposition to other contents but a signifiant is taken as the starting point and from that one comes to content. This is merely identifying the signs. How far is this objection valid? Let us take the classical example of Katz and Postal[79] :



77 Bach (6), page 28 : Any item will "be" an adjective

Here the word bachelor is not defined with respect to other semantic oppositions like spinster, married man, mate etc., but with respect to features which characterize the different senses in which bachelor is used, these features being given in the form of binary branchings. The branchings are however not oppositions, since "Noun - human - female - adult - never married" giving 'spinster' is not mentioned. The object of specifying such features is not to differentiate between different significants but to assign different semantic markers to the same word-form which might be used in different contexts.

77.. in a given derivation if it is derived from the symbol Adjective by the application of a chain of rules beginning with a rule of the form : Adjective —— X.

78 Kuno and Oettinger (238).

Table 1. Sample from an English Dictionary (page 307)

English word	Class code	(comments not stored in machine)
THEY	PRN	personal pronoun in the nominative case
ARE	BE1	finite complete intransitive verb, as in "they are in the sky". (A prepositional phrase according to the present grammar is considered to be adverbial, and cannot fulfil the role of a complement or object of a verb.

78..

Table 1 contd..

English word	class code	(comments not stored in machine)
ARE	BE2	finite copula, as in "They are students." and "They are good."
	BE3	finite auxiliary verb for the progressive form, passive voice, and be-to form, as in "They are coming", "They are seen", and "They are to come here".
FLYING	RI1	present participle of complete intransitive verb, as in "They are flying to Boston" and "It is a flying plane".
	RT1	present participle of single-object transitive verb, as in "He is flying a plane".
	GI1	gerund of complete intransitive verb, as in "Flying is pleasant".
	GT1	gerund of single-object transitive verb as in "Flying a plane is a pleasant".
PLANES	NOU	noun, as in "They are planes"
	VI1	finite complete intransitive verb as in "The glider planes",
	VT1	finite single-object transitiver verb as in "He planes the surface of the board".
	PRD	period as end of sentence punctuation

In the above table the various interpretations of "ARE" syntactically defined on the basis of its functions, but semantic homonymy like for instance in the case of PLANES can be resolved only like :

Structural semantics partitions a global semantic category into subsets which are mutually disjunctive, but contiguous.[80] The disjunction based on semantic oppositions implies that there is some feature by which the word-form differs from other related word-forms. TG on the other hand is interested purely in the componential structure of a single word-form only, the purpose being to explicate how the meaning of a sentence is understood.[81]

78..

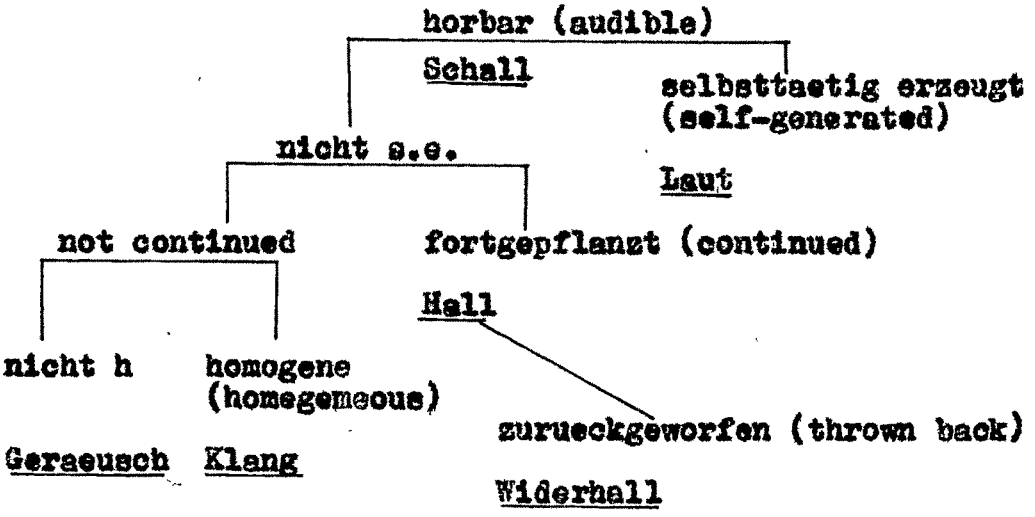
NOU1	transport machines of specific type
NOU2	concept in geometry

Similarly a word like "run" will have different interpretations like run_1, run_2, run_3 etc.

79 Katz & Postal (97), p.14. The <u>semantic markers</u> (Human), (Male), (Adult) etc. express formally general semantic properties. <u>Optional semantic distinguishers</u> "are the formal elements employed to represent what is idiosyncratic about the meaning of a lexical item." (page 14, ibid.) Thus $\langle w_1 \rangle$, $\langle w_2 \rangle$, $\langle w_3 \rangle$, $\langle w_4 \rangle$ are semantic distinguishers.

80 Coseriu (44), describes an analysis of the lexical field "sound" by Heyse (posthumously published in 1856): In Greimas scheme it will look like :

80..

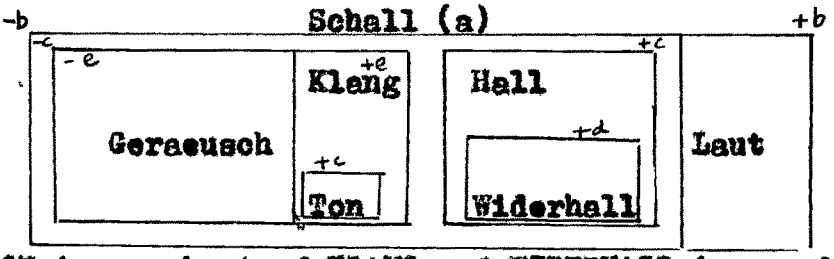


In Pottier's matrix form it will look like :

feature / lexis	a	b	c	d	e	f
Schall	+	o	o	o	o	o
Laut	+	+	o	o	o	o
Hall	+	-	+	o	o	o
Widerhall	+	-	+	+	o	o
Klang	+	-	-	o	+	o
Geraeusch	+	-	-	o	-	o
Ton	+	-	-	o	+	+

a = audible,
b = self generated,
c = continued,
d = echoed
e = homogeneous
f = qualified
o = feature immaterial

In Coseriu's diagrammatic representation :



Here TON is a subset of KLANG and WIDERHALL is a subset of HALL.

Another criticism levelled against the semantic theory of TG is that paraphrase relations have been touched only in passing.[82] Although paraphrases have not been dealt with in theory, TG has in practice accorded paraphrase relations a vital role as all transformations are concerned with synonymy.[83] In fact it has been suggested by Irene Bellert (18) that all paraphrases which a speaker recognizes as equivalent must have the same deep structure.[84] WHAT we say is same, but HOW we say it is different. Following this line of argument she suggests that the two utterances :

John has sold a house to Jim

and Jim has bought a house from John

should have the same abstract deep structure. Similarly Lakoff gives the examples :

Galileo observed Jupiter with a telescope

Galileo used a telescope to observe Jupiter

Galileo observed Jupiter, he used for it a telescope etc.

all of which have the same deep structure. Coseriu remarks

---

81 Katz & Postal (97), pp.12/13.

82 Weinreich (181). See also Thuemmel (175) for a review of Weinreich's article.

83 Quine (145), page 22.

that from the assumption 'deep structure determines meaning' TG infers 'deep structure is the meaning' and through a long and complicated way TG comes to the conclusion that all languages can express the same facts. (Since meaning is equal to the stated facts, deep structure is itself the stated fact.)

According to Coseriu the theoretical mistake made here lies in not differentiating between what is linguistic and what is extra-linguistic. This must be stipulated before the

---

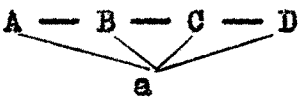
[84] Bellert (18), p.169 ff.
In a footnote to this article Irene Bellert mentions that 41,124 paraphrases of a 17-word sentence taken at random from a physics book contained only a small proportion as 'grammatical' paraphrases. Hence sentences which a speaker easily recognizes as equivalent must all have the same deep structure. Another argument put forth by her is that whenever inadequacies have been detected in the formulation of TG it has been precisely because of the incompatibility of the deep structure representation with the notion of semantic invariants, which should be taken as the basis for grammar construction." (page 170 ibid.)
Typical examples cited are changes proposed for interrogatives, imperatives, Klima's work on negation.

formulation of the concept of deep structure. The other theoretical mistake is that transformations and deep structures are defined only operationally. Deep structures can have something to do with one another as :

1) paraphrase (as in the dictionary)
2) identity of reference (active / passive)
3) differentiation of primary and secondary levels (word/their derivatives e.g. Haustier)

Here one must admit that TG is not unaware of the theory of meaning or the theory of reference.[85] The notion of subject is considered a grammatical universal and hence the subject is considered to be, for instance, in active/passive same. As a grammatical universal it has to be also semantically invariant and it is this 'assumed' condition which necessitates regarding passive as similar to active, but not identical and introducing the passive transformation as an adverbial element belonging to the verb phrase.

Another pertinent question posed by Coseriu is : is the deep structure horizontal or vertical? In other words are active/passive relations transformations or do they have common features which lie deeper? Schematically the two ways of looking at it might be represented as :



---

[85] Katz (96), p.47.

A similar question is : how deep is the deep structure? If we take the example of positing "you" as understood in the imperative "help him", then it is merely one underlying structure. If however this underlying structure is further analyzed into constituent units, then the structure so derived can be thought of simply as a phrase structure representation. Depth in the real sense of the term will be there only if in the generalized phrase marker subcategorization is introduced not only for syntactic reasons, but also for semantic considerations, considerations which are now taken care of by the lexicon.

Grammar, grammaticalness, wellformedness and deviations

TG maintains that its metatheory proposes a hypothetical model which need not be motivated and which derives its importance in that it explains more facts more simply. As a model it can only be falsified, but not confirmed with respect to the truth or reality of the objects. As such transformations are not established in the facts themselves (eg. active/passive), their acceptance being only a hypothesis. The confirmation for the hypothesis lies in deep structure. Coseriu points out that linguistic facts are also facts and while extensions of transformation can be made to non-existant linguistic facts as method, the origin is certainly based on linguistic facts. Hence TG it cannot be

claimed, has nothing to do with facts. In Coseriu's opinion the epistemological theory attempts to abridge the clift between nature and mental sciences, but leads to an old error, to a paradoxial form of new positivism. In our opinion Chomsky does maintain a rather "sophistic" attitude in formulating the foundations of his theory. Although he claims that his method seeks to take the best out of both branches, modern structural linguistics and classical humanistic approaches, a rigid anti-empirical posture is evident in his writings; utter disregard for descriptive procedures and rejection of corpus, are two glaring reatures of his theory, which give rise occasionaly to inconsistencies.

TG has presented well-formedness or grammaticality as a serious scientific issue for the first time.[86] It is interesting to note that according to Saumjan TG is not really explanatory, being based not on a calculus, but on a list or inventary of "well-formed" sentences. If this is accepted, then the question of wellformedness as a problem is pre-scientific because the wellformed sentences are first given and then the grammar is developed.

As far as Chomsky is concerned the only assertion made by him is that grammaticalness needs a scale and that there are degrees of grammaticalness.[87] Katz has shown that the

---

[86] Akhmanova (2), pages 41, 102.

levels proposed by Chomsky are not satisfactory as there are sentences which are less grammatical but comprehensible and there are sentences which are more grammatical but incomprehensible. Katz suggests associating a comprehension set with each deviationary utterance and transfer rules which effect this association. Thus man bit dog will have its comprehension set the sentences :

The man bit a dog

a man bit a dog

The man bit some dog etc.

Understanding a 'semi-sentence' will be possible only if the utterance has sufficient structure permitting association of the utterance with a set of grammatical sentences. Empirical tests have been conducted based on these ideas,[88] but it is doubtful if even extensive psycholinguistic tests can establish any scales of grammaticalness since too many variable factors are involved : individual background and idiosyncracies, changes of mood etc. A related aspect of linguistic performance which has not received sufficient attention is that of pragmatics or usage, as a pragmatic norm for linguistic performance. It has been pointed out

---

[87] Chomsky (31), pages 78/79.

[88] vide Dingwall (45, 47).

that sentences like "the cello is played by Casals" (exhibiting syntactic incompatibility) are more worthy of investigation that "casals is played by cello" since "casals plays the cello" is perfectly acceptable.[89]

Katz has rejected the suggestion of Ziff namely the notion "simplest relation to the set of sentence" or "the simplest route from the grammar to the semisentence" as being unworkable because the rules will be infinite in number and secondly they will not partition the set of ungrammatical sentences into semisentences which are comprehensible and nonsense sentences which are incomprehensible. Two assertions implicit in the above statement are i) that the infinite set of deviationary rules can be generated by a finite set of rules just as the set of grammatical sentences can be generated by a finite set of rules, ii) that it is possible to partition the set of ungrammatical sentences into semisentences and nonsense stringe, for which first of all it would be necessary to give a formal procedure by which one can decide whether a given utterance is grammatical, and if ungrammatical whether it is a semisentence or a nonsense string. The first assumption would be justified only if one can show that the infinite set of grammatical sentences and the infinite set of ungrammatical

---

[89] Akhmanova (2), p.104.

sentences are equipolent. Naive formalism in the case of (ii) is unlikely to stand the test of empirical evidence judging from the state of the research as on today in this field.

Intuition, competence and performance

Theoretically Coseriu sees the most fundamental contradiction of TC in that they want to explain intuition on the one side and on the other have a typically positivistic conception of the science. How can one show that the interpretation of the scientist corresponds to the intuition of the speaker? I know it as native speaker, is not enough because the native speaker does not know what the linguist knows. The intuitive internalization of grammar has been called competence, but Coseriu wonders whether competence can be reduced to the generation of sentences. Competence seems to be an intuition of oppositional difference, an intuition about the purpose and function of sentence structure. The fact that a hearer is able to understand or assign an interpretation even to a deviationary sentence implies that competence must have this built-in ability. Further problem is how does one recognize competence? If through performance, which is a linguistic and material experience, it need not go back to innatism. TG disregards corpus and taxonomy totally, but the latter

atleast has a part to play since linguistic oppositions are important for the "understanding" of sentences. If taxonomy is disregarded, if sentences and situations are separated then the question arises: how at all is competence to be discovered?

STRUCTURAL LINGUISTICS

Introduction

Although generative grammar might be considered as a development in the history of Linguistics which has made structuralism obsolete,[90] the impact of the latter on modern linguistics has been sufficiently great to merit a careful analysis of its significance. The unquestionable systemic nature of language, the fact that linguistic units combine to form larger units and similar realisations about the nature of language were sufficient to induce linguists to accept structuralism as basic idea.[91] If synchronic

---

90 „Der Strukturalismus ist tot" (Structuralism is dead) - Posner (140).

91 Glinz (64), page 38 : "Diese Faehigkeit, nicht nur jedes Zeichen fuer sich und unveraendert zu gebrauchen, sonderndurch Zeichenverbindung beliebige hoehere Zeichen zu schaffen, ist die Grundlage der unerhoerten Dehnbarkeit der Sprache." (This capacity, not only to

linguistics was a reaction to diachronic linguistics, so was also the shift to whole structure a reaction to the preoccupation with historical development of individual sounds or words.[92] Like other ideas which were powerful enough to generate movements in fields wide apart, structuralism has become rather a global cover term than a label for a particular method. A 'structure' may be defined as an entity composed of certain elements which bear some relationship to one another.[93] A restriction might be imposed on the elements by stipulating or insisting that the elements attain significance only by virtue of their position in these relationships.[94]

91.. use each sign without any change by itself, but also to create any desired higher signs through combination of signs is the basis for the unheard of elasticity of language).

92 Glinz (62), page 34: "Nicht vom Einzelzeichen ausgehen und dieses in seiner Geschichte deuten, sondern von der ganzen Struktur, vom ganzen Systemzusammenhang, in den das Einzelzeichen eingebettet ist und in dem es gilt und als Einzelzeichen ueberhaupt nur funktioniert wie seit je selbstverstaendlich gesehen im Recht und und in der Mathematik." (... not proceed from individual sign

Hans-Heinrich Baumann cites four points as characteristic of structural linguistics : a) it is unconcerned with ideology, being technology oriented, b) it does not compare different entities inter-phenomenally, but determines relations ('rapports') among these entities inner-phenomenally, c) it does not explain, but makes available possibilities in description, d) it does not lay any claims to exclusiveness but remains a procedure or method among several others.[95] How far these remarks hold good of American and European structural linguistics will be made clear in the following paragraphs.

92..
and interpret it historically, but from the whole structure, from the whole systemic connection, in which the individual sign is embedded and in which it attains its validity, in fact in which only it functions as individual sign at all, as has been self-understood from time immemorial in jurisprudence and mathematics).

93 Lutz (119) writes: Im allgemeinen meint man, wenn man von der Struktur eines Gegenstandes oder eines Systems spricht, die Elemente aus denen es aufgebaut ist, und "die Art und Weise wie sie (die Elemente) mit einander zusammenhaengen" (Russel, 1148) (In general, when one speaks of the structure of an object or a system one means the elements of which it is constructed and

Methods in Structural Linguistics

The theory and methods of American structural linguistics are a direct consequence of the innumerable studies conducted in unknown American dialects. The fact that grammars were being written for languages which were unknown implied that there was no room for 'Sprachgefuehl'. Language-specific linguistic intuition could neither aid nor hinder. Semantic notion became quite irrelevant to the problem of describing linguistic structure. Field methods had to be evolved to overcome the limitations under which the research worker worked and also to ensure that the linguistic

---

93. "the manner in which they (the elements) are connected to one another".)

94 Eimermacher (49), page 141: "Als äusserlich gleichbleibende formale Elemente nehmen derartige Beziehungen auf Grund ihrer Stellung im jeweiligen Bezugssystem, Bedeutungen' an, die sie an sich nicht besitzen." (As formal elements externally remaining same, such relations achieve 'meanings', which they do not possess by themselves, by reason of their position in the respective relationship-system.
This provise has far-reaching theoretical implications.

95 Baumann (14), page 158.

description was free of premature false inferences. The method was thus inductive, and the analysis proceeded from larger units to smaller ones.[96]

---

[96] vide Akhmanova and Mikael'an (2), pages 36-38. The most fundamental difference between descriptive linguistics and transformational generative grammar is that the former is inductive, whereas the latter is based on a 'mathematical' model! See Dieter Wunderlich (186), pp. 64-66, for a short but clear exposition of the background to Chomsky's ideas. He points out that the algebraic description of linguistic structures does not mean that language is itself mathematical. A Model starts with some initial abstract elements, defines certain formal operations on the sets of initial objects and derives final structures. Of Linguistic interest is Revzin (149). It seems almost as if empirical studies and generative grammars approach natural language from two different poles. Natural languages inspite of all their moments of arbitrariness, display regularities, both overt and hidden which seem therefore suited for empirical analysis. On the other hand natural languages have 'potentialities' of various kinds, i.e. some linguistic structures which could easily have been present in a particular language have not found actual

The first aim was to set up by means of a quasi-mechanical procedure the phonemes of the language under investigation by segmenting utterances and then write out the grammar by analysing its morphophonemics and syntax. Differences in procedures and perspectives among the leading American linguists were large enough to encourage the formation of schools around them until the advent of SYNTACTIC STRUCTURES exposed their weaknesses.

FRIES' description of the structure of English[97] rested on four main form-classes established by means of substitutions in frames : I) units like food, family, concert, II) units like is, was, become, remember, III) units like there, always, suddenly. Further there were 15 groups

---

96..
manifestion at any particular time for reasons of historical accidence. Studies of phonological grammars by Sol Saporta, Ramasubramanian (TIFR Report No.39, Bombay, 1968), Karl D.Buenting's study of the morphological structures of German words (197) prove the capacity of generative grammars to describe such phenomena. It is of course possible for structural linguistics to enlarge its powers of description by some kind of extrapolation, say, incorporating a theory of deviations, but then 'corpus' will have to be given up and along with it the theoretical foundations on which descriptive linguistics has been based.

of function words (prepositions, conjunctions etc.). The function words were crucial for understanding the structural meaning signalled by the formal arrangements of sentences. To get at abstract structures therefore one had to classify formal arrangements which in turn meant recognition of patterns.[98]

Patterns or rather perception of patterns enabled one to segment larger strings into groups of smaller strings i.e. a constitute into its constituent parts. The result of every segmentation was a division into immediate constituents. BLOOMFIELD was content to depend upon native speaker's ability for verifying the correctness of the IC-cut, but PIKE and in particular WELLS have striven to

---

97 Fries (56).

98 The realisation of the enormous role played by patterns in language acquisition and language teaching has at times led to confusion between structures and patterns. See for instance PATTERNS OF ENGLISH by Paul Roberts, 1956. The utility of pattern-based definitions like :

> A noun is a word like apple, beauty or desk (page 13)
>
> A verb is a word that patterns like sing, beautify or arrive (page 13)

is questionable either for theory or for practice.

provide objective criteria for establishing immediate constituents.[99] STREET (172) mentions the following criteria for effecting the cuts :

a) internal cohesion

b) internal diversity

c) independence

d) juncture

e) simplicity

a) internal cohesion : This is the degree to which a group of morphemes function as a unit. Substitution test will reveal whether the criterion has been satisfied or not.

b) internal divereity : This refers to sequences which are substitutable for a constitute. For instance 'King of England' can be cut two-ways :

King / of England (giving King/ who was angry with his prime minister
... / accompanied by his ministers
... / himself etc)

and King of / England (giving King of / ..., where ... is substitutable by proper names of kingdoms or by the common noun 'a country, followed by various possible attributive phrases or clauses). The first cut gives larger number of internally diverse sequences.

---

[99] Pike (139), Wells (183), Street (172).

c) independence, i.e. the capacity to occur in many different environments and constructions. This is actually a corollary of criterion 'a'.

d) "By juncture is meant the degree to which the members of a group combine in a given series." 'au garçon' can be cut 'au/ garçon' because 'au' displays high degree of juncture. If it is cut however 'a/u garçon' where 'u' represents the morphophonemic variation of 'le' then the segmentation is on the basis of internal cohesion between the article and the noun. Perhaps fusion would be a better term than juncture. Cohesion should get priority over juncture.

e) The criterion of simplicity generally calls attention to itself when one is setting up the IC-system of a language as a whole.

IC-Analysis has been modified to permit multiple cuts, and even to accommodate discontinuous elements. HOCKETT has introduced the notion markers to denote morphemes which may be left out when a constitute is segmented into two constituents. For instance in men and women, one can consider men and women as the two constitutes with and serving as a marker and not belonging to either constituent.

Agreeing that phrase structure grammar alone is inadequate for describing the grammar of language, STREET

takes up issue with POSTAL for claiming that IC-Analysis in its conception is essentially equivalent to PSG. The main reasons advocated by STREET for the superiority of IC-Analysis over PSG are that Phrase structure grammar does not permit discontinuous constituents and leaves also cases of overlapping ICs to the transformational part, while IC-Analysis has been provided with descriptive devices to meet these situations.

The Tagmemic analysis is an improvement on IC-Analysis in that it provides an explicit representation for much of the ideas or results implicit in IC-Analysis. "The Tagmeme as a grammatical unit is the correlation of grammatical function or slot with a class of mutually substitutable items occurring in that slot."[100] The form-class is not the grammatical unit, but both form and function are involved. [2]The wagon pulled the pioneer's family across the[3] prairie[1] would be represented in the clause level as :

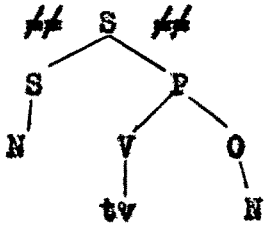
tcl = + S:N +P:tv + O:N ± L:RA

Here the slots are the Subject-slot, Predicate-slot, Object-slot and Locational-slot. These functions are filled in by noun phrases, transitive verb and relater-axis. It has been suggested that because Tagmemic representation carries labels, with slight modification it can be reshaped as a

---

[100] Pickett (138), page 57.

generative grammar.[101] The above clause would then look like :



Such a view disregards the transformational and recursive features incorporated in TG and it also doesn't do much justice to the descriptive capacities achieved by tagmemic grammars.

GARVIN has given an excellent exposition of the methods used in descriptive linguistics stoutly defending the inductive approach.[102] He posits three sets of levels as basic to the language system :

> two levels of structuring, the phonemic and morphemic respectively
>
> two levels of organization, namely selection and arrangement
>
> and several levels of integration, along which the scale of units of increasing complexity is arranged.

Functionally equivalent elements will be considered linguistically equivalent or same. They need not be substantially identical. Relevant is only that which affects functional

---

101 Cook (42), p.43 ff.

102 Garvin (58).

equivalence. Phonemes are meaning differentiators, morphemes are meaning carriers. Consequently in morphemes, the association of form and meaning is in the nature of covariance. By keeping one of the variables of this covariance constant, one elicits for instance the paradigms.

Two commonly used procedures are : dropping and substitution. In dropping, portions of the sequences are omitted from the whole, for testing dependency relations.

Substitutability in controlled frames will help in drawing up inventories of classes based on same external functioning. "Substitution can be defined as right if the result of the substitution is not merely a viable utterance, but an utterance containing the same structural relations of a defined kind as the original utterance. These relations will be of agreement, government and/or dependencies as the case may be for the given language.[103] Garvin defines three types of dependences found useful in analysis :

A presupposes B, but not vice versa

A presupposes both B and C, but neither B nor C presuppose A

Both A and B are allowed to occur without either presupposing the other (mutual tolerance or negative dependence)

---

[103] Garvin ibid, p.61.

Parallel to the developments in American linguistics there appeared in the continent also scholarly attempts to define grammatical concepts objectively and to develop rigorous methods in linguistic analysis. Of the various schools the one relevant to our study is the structural theory and practice of HANS GLINZ, the Swiss-born German grammarian. GLINZ has pointed out that suprasegmental features, especially intonational pattern can serve to delimit the largest unit, namely the sentence, but words cannot be separated without recourse to content.[104] The sentence is primarily a unit of expression, Hervorbringungseinheit, but it also exhibits arrangement of lower order structures, Gliederungseinheit, and is not devoid of content.[105]

---

104 Glinz (64), page 44 : "Dieses kleine Experiment lehrt uns de Saussure verstehen, welcher sagt, dass sich die Einheiten, die Woerter, erst aus dem kontinuerlichen Stimmstrom herausgrenzen, wenn dieser Stimmstrom mit Inhalten verbunden wird." (This little experiment teaches us to understand de Saussure, who says, that the units, the words separate from the continual stream of speech only when this speech stream is connected to contents).

105 Glinz ibid, pp.416-421.

Glinz buildsup his categories step by step starting with the verb, das Leitglied, which is identified by its fixed position and replaceability by single words. The tests used in the entire description are :

i) substitution tests in the first instance to give merely correct sentences (of same pattern)
ii) displacement test (to check position)
iii) transformation test (Umsetzprobe) to determine cohesion between the units
iv) dropping tests to test rank among units
v) subtitution tests of second order keeping the general meaning constant, for determining classes of Nomosphere.
vi) intonation tests (Klangproben) 106

In his categorizations Glinz has freely taken into consideration the inflections of word classes (like declension in the case of nouns). Instead of attempting to show false rigour, he strives to put grammar in the proper perspective by showing how form, function and 'Inhalte' interact differently. The main divisions into parts of speech are :

Vorgangeglieder (the verbals)
Groessen (nominals)
Angaben (modifers and adverbials)
Fuegteile (connectives)

The main word-classes are :

Vorgangswoerter (verbs)

Groessenwoerter (nouns)

Artwoerter (adjectives)

Lagewoerter--i) Stellwoerter (pure adverbs-placing in Time, Space etc.)

ii) Fuegwoerter (conjunctions) 107

The morphosyntactic categories used by Glinz in his analyses and found insightful are :

nicht satzgliedbildendes Stueck (Konjunktion, Interjektion, auch Anrede)	ng
Subjekt	Sub
Objektsakkusativ	OAkk
Gleichsetzungsnominativ und -akkusativ (zusammengenommen)	GN(A)
Objektsdativ	ODat
Objektsgenitiv	OGen
Adverbialakkusativ und -genitiv	AdvK
Praepositionalkasus speziell (= akkusativ und Dativ bei der gleichen Praeposition moeglich, mit Sinn-Unterschied)	Pr.spez.
Praepositionalkasus allgemein (= alle uebrigen Praep.-	

Akkusative, Praep.-Dative, Praep.-Genitive - natuerlich nur soweit sie eigene Satzglieder sind, nicht in attributiver Funktion) Pr.allg.

Qualitativ (fallfremdes Satzglied, nach dem durch "wie" gefragt werden kann; Artangabe i.e.S.) Qual.

Situativ (fallfremdes Satzglied, nach dem nicht durch "wie" gefragt werden kann; Lageangabe) Sit

reine Negation ("nicht"-"nicht mehr" - "noch nicht" - gar nicht") (aus praktischen Gruenden von den Situativen zu unterscheiden, zu denen ich es bisher rechnete) Neg

Verbzusatz (wenn allein stehend oder mit der Personal-form des Verbs zusammengeschrieben mit Infinitiv oder Partizip zusammengeschriebene Verbzusaetze werden als Bestandteile des Infinitivs (resp. Partizips betrachtet) VZs

Infinitform (Infinitiv und Partizip II in verbalem Gebrauch) vInf

Personalform (finiter Teil des Verbs) vpef 108

In the Nomosphere i.e. the level of contents, Glinz has not attempted any categorization; however he has set up classes (transformation-bound-semantemes or TbS-classes) on the basis of substitution tests of second order (i.e. substitution tests in frames keeping the general meaning constant).[109]

In a penetrating study of Glinz's methods Dieter Wunderlich has criticised it on the following grounds :

i) Glinz leaves soon the theoretical framework and goes over to surface-structure-bound tests of selected texts for performance-analysis;

ii) The structural tests cannot establish the fundamental concepts of Theory, at best they can only motivate the latter;

iii) Glinz always restricts himself to texts, although the concrete linguistic utterances consist of spoken talk;

---

[106] Glinz (59), Note 1, Page 120. In qualifiers (Artwoerter) the application of transformation test before asking the question "how" (wie?) brings too many constructions in, which do not exhibit the same functional capacity. Hence order of applying tests is important.

iv) Glinz avoids formalization and subscribes to the romantic theory about the artistic basic nature of language;

v) One does not know how the TbS-classes hang together; according to their structure they are all linear strings of Taxolemes with decreasing cohesion to verb. The concept TbS-class does not even permit the step-wise construction of sentences with increasing complexity, each containing wholly the earlier sentence;

vi) In short it is a specific German 'ordinary language philosophy'.

Glinz's primary aim is to describe the systemic nature of language, i.e. to extract the 'Langue' values in so far as it is possible at all.[111] The basic assumption underlying his work is that syntactic connecting possibilities throw light on the internal structure of language, an assumption which is implicit in a different form also in TG.

---

[107] Glinz (64), pages 456, 461.

[108] Glinz (213).

[109] Glinz (60).

[110] Wunderlich (186).

Parole or Performance is the only mode of ingress to Langue. It is also unfair to use the term 'romantic'. Glinz is only being realistic in cautioning against pseudo-rigour.

While most objections to structural linguistics stem from a difference in basic attitudes towards the nature of language and the notion of scientific method, the criticism of TbS-Classes is well-founded. Though conceptually they are meaningful they display in practice circularity which is reminiscent of dictionaries :

haben (to have)	Vb des Besitzens, fuer Eigentum (Verb of possessing, for property)
haben (to have)	Vb des Habens, fuer Charakteristisches, Person-Eigenes (Verb of having, for characteristic things, personal things) 112

Substitution tests have also been criticised as circular : to find the constituent relationships in a sentence, one must do substitutions, to do substitutions one must know the constituent relationships.[113] This circularity is the bane of all hermeneutic procedures and Glinz is

---

[111] Glinz (64), page 393 writes referring to the boundaries between lexicology and grammar or transition from Langue to Parole, 'diese Grenzen sind freilich fliessend' (these boundaries are to be sure flowing

not unaware of it. Even when transformational grammarians 'resolve' ambiguity they do not 'disambiguate' an ambiguous expression, but show in an explicitly formal way how the ambiguity arises. The ambiguity itself is known beforehand.[114]

Linguistics and literary analysis

Glinz's concern with literary text arises from his pedagogic background and his desire to apply linguistic methods in literary analysis. Glinz's technique of literary analysis has as its goal the objectivisation of comprehension of texts. By working with a group of informants and working out the consensus of the participants regarding the various possibilities in interpreting each passage an inter-subjective appreciation of a text is achieved.[115] Glinz argues that it is impossible to prove or disprove agreement between how one understands a text and what the author had really intended. His concept of 'das Gemeinte' is therefore "what a text means, what an utterance (spoken or written) calls forth in the hearer's (reader's) mind (thoughts, feelings, emotions, etc.).[116] It is text-immanent. Three columns are drawn up to help maintain both the direction of abstraction as well as its depth.[117]

111.. into one another). Glinz sees precise demarcation between various elements of language as impossible due to the Behelfsnatur (make-shift nature) of language.

---

[112] Glinz (60), pages 123 and 150.

[113] Seiler (162), pages 14-15.

[114] Glinz points out that the valuable results of comparative historical linguistics were derived not purely by acoustic perception but such observation of already demarcated linguistic units whose content and value were known. "Sie setzen ein gewisses grammatisches und lexikalisches Verstaendnis der Sprache schon voraus." (64), page 45 : they presuppose a certain grammatical and lexical understanding of the language.

[115] Glinz (66), see also I.A.RICHARDS (150)

[116] Glinz, 'The relation between inner and outer form' (62).

[117] In the Winter-Semester session, 1968/69, the columns were as follows :

A	B	C
Dargestelltes Geschehen, moeglichst gleichmaessig zusammengefasst (nicht enfach Paraphrase) ...	Besonderer Ton, auffaellige Bilder, besondere Stilmerkmale ueberhaupt ...	Erzaehltechnisches Verhaeltnis zu einem chronologisch - kausalen "Normalablauf" Erfuellen von Erwartungen oder Abweichen von Erwartungen ... (*)
(Events pictures, summarised as far as possible with unform abstraction - not just Paraphrase) ...	(special tone/ accent, striking images, special stylistic features...)	

(*) re.narrative technique, relationship to a chronologically casual "Normal run" Fulfillment of expectations or deviations from expectations ...

An interesting exercise in the application of structuralistic principles to literary analysis has been provided by Roman Jakobson and Claude Levi-Strauss on the one hand and Michael Riffaterre on the other in their two interpretations of 'Les Chats' by Baudelaire.[118]

Jakobson and Levi-Strauss attempt to show that their study of the poem from different levels demonstrates the unitary nature of the poem. Compositionally they however discover three ways of looking at the poem for each of which they find ample grounds: 1) 1st quartette, 2nd quartette and third, the two triplets together, 2) 1st and 2nd quartettes together as a group against the two triplets as a group, 3) the 1st quartette and the last triplet against the 2nd quartette and 1st triplet. The elaborate analysis however borders at times on the speculative side, falling into the same errors as literary critics who allow their judgement to be swayed by their enthusiasm.[119] The fundamental notion they utilise is "equivalence relation", which may be defined as commonness of some feature considered relevant for possibility of occurrence in a particular environment.[120] By assigning formal elements poetic function

---

[118] Jakobson and Levi-Strauss (93), Michael Riffaterre (151).

they discover the poetic structural unity of the poem. Riffaterre concentrates on the 'reception' of the message, the expectations it evokes and the corrections it receives. He sees in "Les Chats" three symbolic structures the mysterious, and two kinds of contemplation, realised in one code, namely "the cats". Lovers and scholars are two types in opposition, two types which are mutually exclusive, but the cats have similarity to both, the ambivalence of the cats being symbolised by the opposition "puissants" and "doux". The cats exhibit the insatiable striving for the absolute, characteristic of Don Juan and Faust. Two paths lead to the absolute : le voyage and the other, inner mediation. In the poem contemplation is preferred to adventure.

---

[119] Roland Posner (140), pp.45/46, points out as pseudo-objective formulations the places; „En songeant", les chats parviennent á s'identifier aux „grands sphinx" ...„ce sont ces parcelles incandescentes qu'une nouvelle identification, la dernière du sonnet associe avec le ..sable fin" at transforme étoiles. ('contemplating; they achieve identification with the big sphinxes... It is the incandescent particles, which in a new identification, the last of the sonnet, are associated with 'fine sand' and transformed into stars).

Both the interpretations however draw largely on their knowledge of poetic tradition and maintain so to say one leg on the traditional ground of literary criticism. An instance of a purely linguistic approach which has become a classic is HALLIDAY's interpretation of Leda and The Swan (70).

---

120 Jakobson's famous theorem reads: The poetic function projects the principle of equivalence from the axis of selection into the axis of combination". For eg. 'horrible' is combined with 'Harry', out of other possible choices, dreadful, terrible, frightful, disgusting, due to phonological equivalence between the initial segments of the two words. K$laus Baumgaertner (15), page 71, points out the inadequacy of the theorem: „Die strukturelle Erklaerung eines Textes laesst sich nicht mit primitiven oder ad hoc gebildeten Paradigmen und nicht in der Form zeilenweiser Symbol-Zuordnung vornehmen, erst recht nicht in der Form bloesser umgangssprachlicher Umschreibung."
(The structural explication of a text cannot be undertaken with primitive or ad-hoc formed Paradigms or in the form of line-wise Symbol-arrangement, and certainly not in the form of mere colloquial transcription.)

The importance of linguistics for literary analysis has been realised by all literary critics today. But to make substantial contributions it has to come to grips with the basic concepts of literary criticism and redefine them on a sound linguistic basis. Only then can a grammar of a literary work be written.

Stylistics and the Computer

The formidable problems posed by machine translation has induced a considerable number of research workers in this field to turn to problems of basic linguistic interest.[121] In spite of the tremendous outlay in man power and time initially required before the computer can start producing outputs, concordance projects are still being carried out optimistically.[122] Computational linguists working on other problems have come up with results and theories which linguists cannot fail to ignore.[123]

---

121 Simmons (274), p.22.

122 The power of the computer as a tool in linguistic research can be seen in the wide variety of problems tackled. As early as in May 1967, 120 projects were reported : 53 dealing with concordance, word lists, indices etc., 7 bibliographical projects, 6 editing, collating, formatting, 6 various aspects of context

Even if the computer can be regarded as a mechanically operating moron, even mere countings raise problems of theoretical interest, as the first question which has to be answered is : is the data formalizable in respect of the categories and units which are being investigated? Formalizable need not necessarily imply formal since even semantic content can be made accessible to the machine in the form of indices.[124]

122..
and semantic analysis, 5 attribution studies, 5 studies of meter and rhyme, 4 in Machine Translation, 2 in information processing, 25 various kinds of linguistic studies, and 7 strictly literary studies. (vide LOUIS T.MILIC in (252). SCHANZE (271) p.316 writes : In der Erkenntnis, dass literarische Phaenomene immer an den "Woertern" anhaengen, koennen "Wortindices" sinnvoll vom Literaturwissenschaftler gebraucht werden".
(In the knowledge that literary phenomena are always attached to "words", 'word indices' can be meaningly employed by the literary scientist.)

123 NARASIMHAN (257) has pointed out that computational languages "can cope with expected deviations but not with unexpected deviations in the input" the compiler receives. (page 5). The distinction of expected from unexpected deviations is extremely important for

In any computer application the steps are generally as follows :

(1) Once the problem is clear, the units to be analyzed and the relations to be studied are fixed.

(2) A check has to be made to find out if the units and the relations are formalizable, i.e. whether the criteria which determine the units and relations are formalizable. Some criteria might have to be given up or modified. An algorithmic procedure must exist to decide whether an element belongs to the unit or the relation.

(3) Preliminary studies with data on a small scale have to be made to determine feasibility of the approach.

123.. esthetic evaluation and opens out possibilities of surer interpretations in literature. SILVIO CECCATO's Italian Operational School (191) have given a philosophical basis to their grammar for mechanical Translation, vide IVANOV (225).

124 For instance the term "thermometer" can be coded in abstract machine language as an intersection of the concepts "measurement" and "instrument", where in turn measurement and instrument will be characterised by one or more semantic factors. This has proved a very economical device in scientific translation. (see IVANOV, ibid, page 1080)

(4) Once the initial results are promising, a coding, suited to the problem, has to be evolved.

(5) The program is written and tested with sample data whose results have already been worked out manually. This will expose concealed logical errors in the program if any.

(6) Punch the data and verify.

(7) Give final run.

(8) Interpret the results.

ANNELY ROTHKEGEL (269) has analyzed functional verb complexes as matter for mechanical linguistic analysis. Her analysis shows clearly the crucial role played by formal criteria in linguistic analysis with the computer. The assumption made is that expressions like "to come to a decision" and "to decide" are not mere stylistic variants of the same verb, but that there is considerable semantic difference between such pairs of expressions.[125] The problem is to find out if on the formal level also a difference can be established between a verb-complex like "to come to standstill" and the superficially similar construction "to come to the station".

---

[125] along with structural difference, whether there is semantic difference also has to be tested in each case individually. The difference might be one of 'register'. Some semantic differences are 'potential'. They need not always be actualized.

Both the functional verb complex (FVC) and the 'simple verb complex' (SVC) have the same form :

Verb + Preposition (opt.) + Article (opt.) + Noun

A possible differentiation is that

i) "FVC" and "SVC" differ in their syntactic behaviour in a sentence

ii) The four word-classes might be represented by one set of lexical items in the first case and a different set of items in the case of simple verb complexes. The membership of each class however depends upon the members of the other classes so that a system of combinatorics has to be worked out.

By further analysis of the linking possibilities the following tentative criteria are set up :

(1) With FVC the possibility of linking a genitive attribute or a relative clause to the nominal part of FVC is highly restricted.

+ He took his exception to my remark

+ He gave expression, which was very forceful, to his resentment

(2) Likewise there is severe restriction in modification of the Nominal in FVC by means of an adjective; mostly an adverbial stands in its place.

(3) The nominal part in FVC is not substitutable (it cannot be 'questioned').

(4) The nominal part is a nomen actionis.

(5) Only the enclitic article is possible. 126

No.5 relating to the article can be readily tested without any restriction. The program will check if the word is an article and if so, whether definite or indefinite and if the next word is noun in the singular. (Chance juxtapositioning of words which also combine to form verb complexes must be shifted in a prior program, if possible at all.)

No.4 is useless since what a nomen actionis is, cannot be determined independently of each individual case.

Re.3 substitutability needs apparently the intuition of the speaker. The machine can substitute, but cannot decide if the grammaticality changes thereby.

For the same reason 2 is also not useful.

Formally also there are five types possible :

type 1 - Verb + (Preposition +) Noun in Singular
2 - Verb + (Preposition +) def.Article + Noun
3 - Verb + (Preposition +) noun in plural
4 - Verb + (Preposition +) (indef,Art.+) Noun
5 - Verb + Preposition + enclitic art. + noun

---

126 Rothkegel, ibid, page 9.

For these we can say that if type 3 or 4 has a genitive attribute or if type 5 has a relative sentence then the verb complex in these cases is a simple one. But those expressions which do not occur with such attributes, will escape this rule since the machine cannot decide whether there might be a genitive attribute or relative sentence as modifier of the noun. In the corpus being analyzed there are a number of such instances. In order to tackle these a new method is chosen. Since the computer centre works with a dictionary, all nouns which have occurred atleast once as FVC are marked in the dictionary with an index. Similarly verbs and prepositions which have taken part in a functional verb complex are marked. In other words further work is carried on with subclasses like FVC-nouns, FVC-verbs, FVC-prepositions. Their number is approximately : verbs 30, prepositions 18, nouns 150. Out of the potential 85000 combinations possible, only 500 are actually manifest, as these elements in combinations are subject to restrictions, partly grammatical, partly specific to FVC. Given any complex the program checks against the inventory list successively to determine if the verb, preposition, and noun in a given complex belong to the sub-classes. In relevant types further check will be made if there is genitive attribute or relative sentence. Thus by a process of inventory check the resolution of a given complex into FVC or EVC is made. This rather detailed description KLEIN and ROTHKEGEL's problem

illustrates the pivotal position occupied by the dictionary in a syntactic analysis by means of the computer. At the same time it must be mentioned that solutions are still far in the case of semantic polysemy :

The prices reached a peak

The climbers reached a peak.

Coding is another aspect of the work which is as important as criteria. The major aim here is to represent the desired information in the most economical manner possible. A punch card consists of 80 columns of which only, say 60 columns might be free for this purpose. Grammatical data about the particular word might be entered in the 37th column, say, as follows :

0 - indicates Article	5 - Verb
1 - Noun	6 - Adverb
2 - Adjective	7 - Preposition
3 - Pron.	8 - Conjunction
4 - Numeral	9 - Interjection

In combination with this information on 37th column, the 38th might be used as follows : If 37 shows 1, i.e. noun, then 38 will indicate 0 - singular

1 - plural

2 - with def. article

3 - without def. article

If 37 shows 3, i.e. pronoun, then 38 will differentiate

further as follows :

0 - personal	1 - reflexive	2 - reciprocal
3 - possessive	4 - demonstrative	5 - relative
6 - interrogative	7 - indefinite	8 - singular
9 - plural		127

While syntactic analyses of languages have to be carried out as large scale projects individual studies on a smaller scale are also possible and have been carried out. BUENTING (197) has investigated word derivation in German by taking 2759 roots and seven types of derivations on the basis of the following affixes :

/ ...-CHEN/, / ...-UNG/, /BE-...-EN/,
/BE-...--IG-EN/, /...-BAR/, /...-LICH/, /UN-...-LICH/

These were compared with relevant material from Mackensen's dictionary. For each affixation 2759 derivations are possible, but the percentage of derivations which are actually found in the language amount to about 5'3% in the average. Some derivations produced by the computer, but not found, were acceptable to informants thereby showing how the artifical production does justice to the openness of the system.

The ease with which strings of signs can be searched for and compared by a computer has lead to the rise of a

---

127 Kukenheim (237).

new branch, namely computer aided literary research. Mention has later been made of KNAUER's filter method.[128] RIEGER (268) has studied the 'so called trivial poems' of students for their formal, thematic and material aspects using the computer as a searching tool. KLEIN and ZIMMERMANN (231) are analysing Trakl's poems thoroughly starting with phonemic counts. Probably the most famous of all such computer-based studies is that SALLY Y. SEDELOW (272, 273).

Sedelow first makes a list of primary words in the text, a primary word being defined by a frequency threshold, which would vary from text to text. For Hamlet all words occurring more than 5 times are considered primary words, whereas for the article on Soviet Military Strategy the cut-off point is 50 words. Synonym dictionaries, thesauri are then manually consulted to prepare a list of words associated with the primary words. The computer then makes a complicated search for possible associated words, first for each word associated with the primary word, then words linked to such of those associated words as having primary status in their own right. The frequencies of such words in various parts of the text show the shifts in concepts and conceptual associations. This typically information retrieval approach is aided by a MAPTEXT program, which blanks out words showing only those words in which the

---

[128] vide page 111 . Vide also Knauer (232).

analyst is interested. A single line output of this program looks like :

115 ---------S-P----W

where S,P, and W stand for specific words.[129] Sedelow investigates not only synonyms but also the occurrence of antonyms.[130]

In the context of her work Sedelow has defined style as patterns formed in the linguistic encoding of information; stylistic analysis, then, is the perception of pattern in language.[131] She writes further : "The use of two translations of one word for analysis probably provides the strongest evidence obtainable that content word choice is an aspect of style. This point certainly should no longer be a matter for the dispute implicitly contained in all discussions of style that entail distinctions between style and content, or form and content. So long as such distinctions are made, the problem of the analysis of style will be obscured." (Stylistic Analysis, Third Year report, page 89).

---

129 compare Knauer (232) vide page ... , this report.

130 Similar studies of cluster of concepts have been done also for psychological categories. Vide Stone (278). Sometimes the question is posed : "what can such analyses bring out which an informed sensitive reader -

Difficulties in the establishment of rigorous methods for the objective evaluation of style

If it is not possible to define unambiguously and clearly what an image, what a motif, a metapher is,[132] it equally difficult to define what style is. All discussions on stylistic analyses end up with the question : what is style? In order to achieve clarity it is therefore essential to decide first of all what is understood by this term for the purpose on hand. Assuming that we define style as deviation from norm, we have to define at least for working purposes a norm. Assuming that we have fixed a norm, then question arises whether stylistic deviation will now be measured on a scale or whether it will consist of descriptive statements like : Thomas Mann frequently uses a number

---

130..

in the extreme case the super reader - fails to note in his careful reading of the work? Sedelow mentions that her search revealed familial and non-familial relationships (father, mother, daughter - King, man) in Hamlet which had not been commented upon previously by critics. The importance of the computer lies in its capacity to produce evidence in favour of hypotheses. SCHANZE (270) has demonstrated how frequency counts taken in conjunction with careful study of each individual occurrence throw light on the diction of the author.

of adjectives in order to make an expression complete in all precision thereby transcending the stylistic ideal of precision just in the effort to achieve infinite precision.[133] Measurement and scaling of deviation are also conceivable, but in all these methods a major drawback is that the yardstick of comparison is determined by the investigator and different yardsticks will give different results.[134]

---

[131] Sedelow (272), p.7.

[132] Schanze (271), p.317.

[133] Baumgart (16), p.24.

[134] Simmons (274) suggests the possibility of testing whether two expressions A and B are paraphrases by generating strings from A till B is reached. If there is no path then B is not a paraphrase of A. By means of a dictionary (synonyms, antonyms with association paths) it might be possible to measure deviation by the number of steps (rules) taken to generate the deviationary expression. Here again the measurements would vary with the dictionary.
WERNER MUELLER (255) argues that even a "choice" will become style only when taken in comparison with a norm. If every text shows the same choice, reflected in the same frequency and transition probability of the unit, then it will not belong to 'style', but to 'language'.

134..

Mueller denies the very possibility of discovering even one stylistic index on the basis of one text alone. A norm has to be chosen which has to be so determined that it can be taken to be representative of a totality called "texte". Style then exhibits itself in the difference between the chosen form and the mean values of the norm. Where the variables are formal parameters like sentence length, syllable length, a quantitative approach to norm is easy, but when one is working with psychological categories like for instance expressions revealing <u>fear</u> extensive preliminary searches will have to be made before a reasonable norm with acceptable <u>means</u> (average values) are available.

The Development of Mathematical Linguistics

Explanation of the usage of the term "mathematics" in linguistics

Eight years ago the question arose whether mathematical linguistics is a trend in fact? More in conformity with the methods used, Spang Hanssen, himself a mathematician, pointed out that only mathematical terms were being used and the application of mathematics ended at the symbolization stage.[135] Today considerable progress has been made but "the field is not yet an established one with interrelated problems and methods".[136] On the whole there are 3 directions or approaches discernible based on the methods used, to which the cover term "mathematical" would be applicable :

1) the statistical approach to determine various distributions in language which are subject to statistical laws of chance
2) the automata-theoretic approach to evolve formal models to simulate language behaviour
3) the set theoretic approach to investigate the relational features manifest in language.

The last two have been influenced in their origin and development to some extent by the developments in computer

---

[135] Spang-Hanssen (168).

[136] Preface to Z.Harris (216).

technology. The ideals of automatic translation and information retrieval brought forth problems of analysis which traditional grammarians or structural linguists did not have to tackle.

Statistical methods have been used sporadically even from early times. In Appendix 1 a list of such works is given. Word counts far from being an academic hobby, were in those days in fact of some professional interest. Scribes who copied manuscripts and were paid by the number of letters or words, were interested to keep track of the progress of their work.

The chronology of platonic texts engaged the attention of Prattizi in 1593. Cases of disputed authorship have intrigued scholars sufficiently strongly to cause more than one investigation of the same case. KRALLMANN has traced the historical development of statistical studies of language.[137] Further references may be found in WILLIAMS, LORD, MENDENHALL, PARKER-RHODES, PLATH, and HERDAN.[138]

---

137 Krallmann (235).

138 Williams (288), Lord (241), Mendenhall (250), Parker-Rhodes (263), Herdan (220).

The different approaches in statistical linguistics

a. frequency counts of phonemes, syllables etc :

Mere frequency counts have found among linguists little favour, owing to "the unfathomable naivete, from the linguistic point of view, of the units being counted - for example, the "word"."[139] Nevertheless frequency counts of the sounds of language have been put to practical use.[140] KNAUER has used frequency counts of consonants and vowels to measure the sound-texture of Baudelaire's sonnets. By giving code numbers to the vowel and consonantal segments, like for instance :

p = 45; b = 27;...

the text, in this instance Baudelaire's sonnets were transcribed into a series of numbers, which were then written on punched cards. The computer is then programmed to filter out all unwanted sounds as blanks, the particular code number/s appearing for the sound/s which are the object of study at the moment. Thus the output looks like :

a = 10

-10----------
-10--------1010
---------101010
---------10----

------10-------
---10--10--1010
--10---10--101010
---10---10------

10----10-------
--------10------
--------10--10--

-----10-----1010

A refinement introduced by KNAUER in the study of "fine structures" for their tonal value is the concept of drift.[142] Suppose the vowel sound a is under investigation. Then the occurrence of a is denoted by F (for feature) and non-occurrence by F'. The totality of spaces where F or F' can occur is divided into a right half and left half. For each line it is noted how often F occurs in the right half, how often F' occurs, similarly for the left half. Summing up we get four values, an F and an F' for the right half and the left half respectively. These values are written down in the form of a contingency table and chi-square test is applied to find whether the drift is significant or not. The direction of the drift whether left or right (rhyme-drift) is seen straightaway from the table. This gives an accurate measure for comparing not only individual vowel or consonantal sounds, but it can be used to find out how a bundle of features occur in a sonnet or group of sonnets.

---

[139] Wexler (266), p.122.

[140] Kaeding (227). vide also Meier (249). Kaeding counted 20 million sylables with a view to giving a scientific basis to the German stenographic system.

[141] Knauer (232).

[142] Knauer (233).

Phonetic counts can thus be put to good use. Another field wherein they offer fruitful results is the study of dialects.[143] In a similar way syllable counts have also been usefully employed for linguistic studies, although the theoretical status of the syllable is an undecidable controversy at the present stage of research.[144]

KRALLMANN took vowel monograms and vowel-digrams as forming one nucleus of a syllable, groups of 3 or 4 vowels, if they occurred in sequence, being counted as two syllable-nucleii. The consonanantal environment was irrelevant for the purpose of counting the syllables. KRALLMANN has shown that the margin of error in his case was only 0.812%.[145] Both word length as well as sentence length has been counted in terms of syllables. FUCKS has made extensive counts of distribution of sentence length and word length on the basis of number of syllables. The method of partial summation series is used to show that generally stability is achieved when a particular sample size is taken.[146] The same is shown also by WEISS.[147] Suppose in random counts of 250 sentences each, the mean sentence lengths in terms of syllables (or words) from a particular sample is as follows :

250	250	250	250	250	250	250	....	250
x_1	x_2	x_3	x_4	x_5	x_6	x_7	....	x_n

---

143 Vide Meier (249).

144 Haugen (217).

Then the running means for the partial sums are calculated as follows :

$$\frac{x_1 + x_2}{2} , \frac{x_1 + x_2 + x_3}{3} , \frac{x_1 + x_2 + x_3 + x_4}{4} , \ldots .$$

These values plotted on the y-axis against units 250, 500, 750, 1000 ... on the x-axis tend to stabilize around a particular value although initially they show great fluctuations. Theoretically this must always occur, for suppose at a particular stage of our calculations we have

$$X_r = \frac{X'}{N} , \text{ where } N = 250 \times r$$

Then the next mean is given by $\frac{NX_r + x_r + 1}{N + 1}$

The difference between this mean and the previous mean is

$$\frac{NX_r + x_r + 1}{N + 1} - X_r = \frac{x_r + 1}{N + 1} - \frac{X_r}{N+1}$$

For sufficiently large N therefore the difference $\frac{x_{r+1}-x_r}{N + 1}$

will be a small fraction. Hence the partial summation method has the effect of flattening out the differences, provided N is taken sufficiently large. The significance of this is therefore not that stability is achieved after a particular

---

145 Krallmann (235), p.144.

146 Fucks (210), fig.11, p.19.

147 Weise (286).

level, but when this stability occurs, i.e. for what value of N. In samples of WEISS even with random samples of 750 to 1000 sentences the partial means lie between the levels of confidence limits thus pointing to the level of accuracy of the results, and the homogeneity of the population.

JOSEF LAUTER (240) has investigated three ratios :

S_1/W, B/W, and S_1/S.

(syllable per word, letters per word and syllables per sentence).

For the first two there was sufficient material for comparison and the third characteristic was chosen for the following reasons :

i) it is mechanically easily computable

ii) even large tests can be covered by this characteristic

iii) FUCKS has shown that the number of syllables per word and number of words per sentence are frequently correlated, that is, if the sentence length of an author measured in terms of words is large, then he has a tendency to use also longer words. The ratio

$$S_1/S = S_1/W \cdot W/S$$

consequently produced reinforced effects. It is shown that it is a very effective means of describing Kant's style.

Lauter investigates the fit of poisson distribution as given by FUCKS for S_1/W :

$$p_i = \frac{e^{-(\bar{I}-1)} \cdot (\bar{I}-1)^{i-1}}{(i-1)!}$$

$$\text{or} \quad p_i = \frac{e^{-t} \cdot t^{i-1}}{(i-1)!} \quad \text{where } t = \bar{I}-1$$

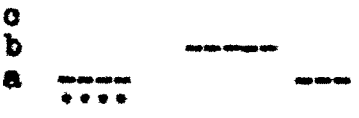
The assumption is that the formation of words out of syllables is describable uniparametrically. LAUTER comes to the conclusion that the sequence of words with respect to their length in terms of syllable follows randomly. But the sequence of sentence length is not random, generally long sentences stand near long sentences and short sentences near short ones. The correlation test for 17 paragraphs showed that for 12 out of 17 the r_1 values were larger than the standard deviation and all r_1 values were positive. The mean of syllable per word lies higher than the mean for German literature in general. Comparisons with a theoretical distribution and a distribution of Goethe show that Kant preferred long words. Further LAUTER investigated the frequency of word classes and their environment. By means of entropy Lauter showed that a statistically encompassable connection between the classes exists only upto an interval of 5 words. Kant's text showed preference for many membered predicates and for beginning a sentence with a preposition.

With the help of frequency polygons and the mean and standard deviation FUCKS shows how far philosophers agree or disagree in their style, how far politicians use similar styles and how it is with modern prosaists.[148] On the x-axis mean number of words per sentence is given and on the y-axis the mean number of syllables per word. Any particular text by an author will be represented by a point in the graph space according to the two mean values it possesses.

In order to go deeper into the analysis FUCKS brings in the concepts of sections and ranks. The total sentence may consist of one or more main clauses and one or more subordinate clauses. The main sentence is of rank a or 1, the dependent clauses of rank b, c ... or 2,3 etc. A section is a continuous main clause or subordinate clause string, until it is broken by a segment of another class. For instance,

the man who visited me yesterday works in a firm consists of 3 sections.

the man/ who visited me yesterday/ works in a firm. Graphically it is represented as follows :



148 Fucks (210).

The length of the lines depend upon the number of syllables contained in each section. The distribution is plotted with co-ordinates as follows :

x-axis : rank number in sentences (1,2,3 ...)

y-axis : frequency of the maximum rank in each sentence

These values taken in conjunction with the mean and dispersion give an idea how complex the authors have written their texts.[149] As a further characteristic of the complexity of the sentence FUCKS takes the sum of all the ranks of the different sections of the sentence and maps their frequency polygon. For the sentence given above the sum will be 1+2+1 = 4. For each sample there will be a distribution of rank sums. The mean values of the rank-sum frequency for each sample is plotted on the x-axis against the respective standard deviation on the y-axis. Five samples each from Lucas Evangelium and the Apostle story give completely separate polygons, giving thereby evidence that the former is less complexly structured than the latter.[150]

The syllable count is used by FUCKS also to determine if there is any force of attraction between the sentences.[151]

---

149 Fucks (210), p.50.

150 Fucks ibid, page 51.

151 Fucks ibid, pages 61 ff.

In order to answer this question the sentence length in terms of syllables was plotted with x-axis giving the order of the sentences, 1,2,3 etc. and the y-axis the length of the sentence measured in terms of syllables. A line parallel to the x-axis was drawn in such a way that half the number of sentences lay above the parallel and half below. If this line was taken as the dividing line for short sentences as against long ones it would give exactly the same number of short sentences k and long sentences l. The text can then be described as a sequence of k's and l's in some text-specific manner. To determine the attractive force between the sentences the following values were calculated :

frequency of the sequence lk = 7, say

frequency of the sequence ll = 13, say.

Since the sentences have been equidivided with reference to length, the matrix on left hand side equals the r.h.s. :

$$\begin{vmatrix} ll & lk \\ kl & kk \end{vmatrix} = \begin{vmatrix} 13 & 7 \\ 7 & 13 \end{vmatrix}$$

If the total number of sentences is known then it is necessary to count only one value, the rest can be calculated.

The proportion 13/7 or 1'86 shows that there are 86% more similar pairs than unsimilar pairs. For a large sample of 2000 sentences the value (excess) was 12%. In a perfectly ideal random case the two values would be equal, i.e. there would be the same number of short sentences as the number of

long sentences. Their proportion would be equal to 1. But in practice this ratio differs from 1 and as a quantitative measure for the follow up of short and long sentences the coefficient of correlation c_1 given by the equation

$$c_1 = \frac{(ll + kk) - (lk + kl)}{ll + kk + lk + kl}$$

is taken. The values range from 2 to 31%, but they are always positive.

How far does this binding force between sentences reach? To test this the coefficient of correlation was calculated not for two consecutive sentences but for each sentence and the third one from it. This coefficient c_2 was for instance 5% for sample from Bismark's works. The intervals were increased and similar coefficients were calculated, c_3, c_4, c_5 etc. These coefficients were calculated upto c_{10}, i.e. for neighbours removed by 10 sentence intervals.[152] One can see from the index how different authors take different positions in the correlograms.

A similar procedure is followed by FUCKS with respect to poetry. The stressed syllable is marked 'b' and the unstressed 'u' (betont-unbetont). The verse is represented as a sequence of b's and u's, the last syllable is joined to the first and counts are made how many times a stressed

---

[152] Vide Fucks (210), fig.19, page 62.

syllable follows an unstressed syllable etc. As in the previous case apart from immediate neighbours, syllables removed by a fixed interval (i.e. number of syllables) can be taken. The indices calculated serve as in the earlier case to characterise numerically the affinity (or otherwise) between the sound segments in the verse. Different authors are seen to have different values.[153]

WEISS has shown that there exists a correlation between the mean word length in syllables and mean sentence length measured in terms of words. In other words shorter sentences tend to have shorter words and longer sentences tend to have longer words also.[154] He investigated 6 works of Goethe for sentence length and found that between LEHRJAHRE and WAHLVERWANDTSCHAFTEN on one side and between FARBENLEHRE and WANDERJAHRE on the other exists a close relationship. Samples from both parts of Werther show that they vary strongly with respect to sentence length. Plotting the year of publication on the x-axis and and mean syllable length of sentence on the y-axis shows that a gradual change in Goethe's style took place. Investigations of similar values for KANT's works showed that the distribution of sentence length of the works written before 1780 (the year of publication of THE CRITIQUE of PURE REASON) and the

---

[153] Fucks (210), pages 71 ff.

[154] Weiss (286), page 54.

works later to that show marked differences. By drawing parallels to the x-axis at mean length = 70 syllables, and 90 syllables it is seen that the precritical writings do not exceed 70 syllable mean, while works after 1780 do not go below 90 syllable mean length. The region from 70 to 90 is a transition range where both pre- as well as post-critical writings are found with large difference in the year of publication. Histograms of 4 precritical works (mean sentence length against number of sentences) and 4 post-critical works show that in Kant's works we find two styles : one characterized as worldly elegant of the precritical period and the other the sober matter of fact style characterized by the occurrence of extremely long sentences which is characteristic of the time of critical undertakings. The latter has a broader and flatter distribution.

The distribution of letters into syllables or words within a language is fixed and hence an investigation of the syllable or words on the basis of letters would not lead to differentiation of the style of an author from that of another or even within the same text. It is of greater interest only from the information theoretic point of view. The number of syllables per word is also characteristic primarily for the language as shown by FUCKS.[155]

---

[155] Fucks (208), page 22.

b. frequency distribution of words

Linguists' interest in word counts have also been not very high since usually the statistical count cuts across homonyms, inflections nonchalantly. Thus in KAEDING the morphological variations of "nehmen" (to take) are counted separately as follows :

genommen	2322	nahmen	593	nehm	65	nehmenden	14
genommene	39	nahmest	7	nehme	661	nehmendes	1
genommenen	48	nahm'e	2	nehmen	3917	nehmet	47
genommener	5	nahmst	19	nehmend	85	nehmt	157
nahm	2265	nahmt	4	nehmende	16	nehmt'e	5
		nimm	357				
		nimmst	10				
		nimmet	72				
		mimmt	1774				
		nimmts	5				

156

On the other hand no distinction is made between $Bauer_1$ and $Bauer_2$ (der Bauer, das Bauer) or between "weise" as verb in ich weise and "weise" as adjective in das Papier ist weise. Nevertheless word counts have fascinated scholars with statistical inclinations especially since the notion "richness of vocabulary" is intimately connected with it.

GUIRAUD takes the class of form words for the calculation of a hypothetical unit, namely the potential vocabulary of an author. In the actual composition of the text only vocabulary V is used by the author. From V and the total length n Guiraud calculated potential vocabulary

---

156 Meier (249), page 9.

of an author as

$$L = 2n(2V_1/V)^{alpha}$$

where V_1 is the number of words used only once. Alpha the interpretation of which is difficult depends upon the structure of the vocabulary. Some words will be relatively more frequent in one text, but in another text relatively less frequent. Guiraud divides each text of length n into smaller texts of length n' with corresponding vocabulary V'. The factor alpha is then given by

$$\text{alpha} = \frac{\log n' - \log n}{(\log V' - \log V) - (\log V_1' - V_1)},$$

the subscript 1 denoted the number of words occuring once in the corresponding texts.

The Type-Token ratio has been suggested as a measure of verbal diversification. In the equation

$$fr^x = C$$

where x is related to the slope of the standard curve, x is an index of verbal diversification.[157] If an author has a wide range of different words x lies between 0 and 1. If words are repeated over and over x falls between 1 and 2. Carrol (1938) suggests another method. The number of words that separate successive occurrences of the most frequent word is counted and their mean is taken. With diversified

---

157, 158
Miller (128).

vocabulary there are more words in between. The average value does not change systematically with the length of the sample.[158]

In recent years HERDAN has intensely occupied himself with word counts, modifying the work of YULE to a great extent.[159] The main themes of HERDAN are the characteristic v_m, the Type/Token ratio and the random partitioning function. The characteristic v_m is given by the equation

$$v_m = \frac{v_x}{N} = \frac{\sigma_x \sqrt{N}}{M_x} \qquad 160$$

The number of words constituting the vocabulary of a writer, though large, is finite. Hence as occurrence increases particular words are used often than other words. The average number of occurrences per word, M_x, shows an upward trend as the number of occurrences increases. Similarly the standard deviation also increases but not at the same rate as the frequencies of words. A characteristic depending upon mean or standard deviation alone would thus be dependent upon the vocabulary N, while v_m on the other hand is free of this disadvantage.

This characteristic which is an index of the type token relation, has been the object of study of several

---

[159] vide Yule (290), chapter 3 and 4 on characteristic, chapter 7 for random partitioning.

statisticians. YULE regarded it as a measure of concentration of vocabulary.[161] WILLIAMS applied it to biological distribution of animal species and called its reciprocal a measure of diversity.[162] GOOD conceived of the characteristic constant as the repeat rate of species, words, etc. in a great number of occurrences of such events.[163] Though the characteristic K as given by YULE or v_m as given by HERDAN is not amenable to an 'interpretation' it can serve as an indicator in questions of disputed authorship, taken in conjunction with other factors.

The results on word counts boil down to : The repeat rate of vocabulary items in a given universe of discourse is sensibly the same for the whole universe as for any part thereof, and this is quite regardless of the sample size. The distributive law for the frequency of words is best approximated by dividing the distribution into three parts, and treating them separately :

high frequency words like grammar words
rare words
those in the middle region [164]

A good approximation for high frequency words is given by the binomial distribution $nC_r\ p^r\ q^{n-r}$, where p represents probability of occurrence and q is probability of non-occurrence, i.e. 1-p.

---

160 Herdan (220).

161, 162, 163 Herdan (221), p.85.

In the case of rare events it goes over to Posson's law. For p=0'001 HERDAN finds the approximation quite close.

The gap length between successive occurrences of grammar forms as seen in a sample from Russian follow the law of chance. If L = occurrence of a particular word divided by total occurrences, i.e. L is its relative frequency (in a big sample) then the probability that the number of words separating successive K's (K being the preposition in Russian) less than or equal to x is given by

$$F(x) = 1 - e^{-x}.$$

Except for small values of x where the discrete nature of the gaps play a disturbing role the fit between theory and observation is very satisfactory.[165]

Herdan cites the results of CH.MULLER as revealing that the stylistic differences are best brought out by nouns, less so by adjectives and least of all by verb forms and grammar forms.[166] The main theorem of HERDAN is :

> The proportion of linguistic forms belonging to one particular level of understanding or to one stage of linguistic coding - phonological, grammatical, metrical - remain sensibly constant for a given language, at a given time of its development and for a sufficiently great and unbiassed number of observations.[167]

Examples of these are the relative frequency distribution of phonemes, of letters, of word length in terms of number of letters and syllables, of grammatical forms.

The random partitioning function : This has been used previously by YULE and Ch.MULLER. Its object is to test the vocabulary connectivity in two or more texts. The texts may be partitions from the same text, say A,B,C,D. By using capital letters to denote the set of words present and small letters to denote that no word comes from that particular text, we can calculate all the words common to all the texts, i.e. (ABCD); (ABcd) will denote the number of words which are common to texts A and B, but which do not appear either in C or in D. (AbCd) will denote the number words common to texts A and C, but not appearing either in B or D. Against these observed values one might check estimated values obtained as follows. If the 4 texts are represented by four quadrants of a rotating circular tray and counters represent each occurrence of a type, then the probability of a counter falling in one quadrant is 0'25 and its probability of non-occurring is 0'75. If the

---

164 Herdan (222).

165 Herdan (223), page 129.

166 Herdan (223), page 170.

167 Herdan ibid, page 15.

same word occurs, say 4 times, then the probabilities are increased four times, i.e. $(0'25)^4$ and $(0'75)^4$ respectively, i.e. 0'0625 and 0'5625.

In general the probability of appearing in one of k quadrants is k/r, the negative probability being (r -k)/r, where r is the total number of samples and k is the number of samples in which the word occurs. If there are f_x such words (the sum of all f_x being the total number of words in all samples or texts put together), the expected number of them not appearing in any of the k sections is therefore

$$f_x\left(\frac{r-k}{r}\right)^x$$

The entire number of words not falling into any one of the k sections is the sum of such expressions for all values of x :

$$\sum f_x\left(\frac{r-k}{r}\right)^x$$

With r = 4, $\quad (a) = \sum f_x(\tfrac{3}{4})^x = (b) = (c) = (d)$

$$(ab) = \sum f_x(\tfrac{1}{2})^x$$

$$(abc) = \sum f_x(\tfrac{1}{4})^x$$

which gives 14 frequencies since (a)=(b)=(c)=(d) and so forth. Given N, the total number of nouns in the frequency distribution and also (abcd)=o, we can calculate the complete set of fourth order frequencies etc. There is good fit between the observed values and estimated values which shows that there is a hidden regularity in the use of words.[168]

c. Distribution in terms of transitional probabilities have been studied although the necessary calculation work reaches unmanageable proportions even with the deployment of computers. Working with five classes the possible transitions of first order are 5x5 =25; for second order it is 125 and if grammatical words are somehow included in the study, the number of possibilities will have to be multiplied by the number of grammatical words. In order to be able to make meaningful statistical statements about these various possibilities a proportionately larger sample will have to be taken. In spite of such difficulties interest has been evinced to find out the effect of context on the statistical behaviour of linguistic units.

WICKMANN (287) has tested the authorship of NACHTWACHEN by using transition probabilities between word classes. The disputed text was compared with works of WETZEL, JEAN PAUL, BRENTANO, and HOFFMANN. The 19 classes considered were :

1. Noun
2. finite verb form
3. infinite verbform
4. auxiliary verb
5. flected adjective
6. article
7. pronoun - nonattribut.
8. attribut. used pronoun
9. numerals
10. adverb
11 Verbal complementary particle
12 preposition
13 conjunction
14 proper noun
15 comma
16 period
17 uninflected adjective
18 nominalized verb
19 nominalized adjective

---

168 Herdan (223), pp.219 ff.

The comma was blanked out later for various reasons. Preliminary studies showed that out of the 19 x 19 transition possibilities many were weakly represented. Statistical considerations of sampling led to the elimination of such transitions so that 70 transitions were finally chosen for study. The problem posed was: on the basis of the distribution of the transitions selected, what is the probability that two samples drawn at random belonged to the same population. Since the degrees of freedom was small the generalized t-test was not used, but chi-square distribution with 2 degrees of freedom was used to sum up the different transitional frequencies. The results showed that Wetzel, Jean Paul and Brentano have to be rejected as authors, at 5% significance level.

Among the most famous statistical studies in literary analysis one may count "Inference and Disputed Authorship : The FEDERALIST" by Frederick Mosteller und David L. Wallace (254). Over 50 assistants participated in the calculations, notwithstanding the use of electronic computers. Their results ascribe Madison's authorship to the disputed papers as more likely. The statistical basis on which the study rests is Bayes' theorem, although in the four main studies and several ancillary ones the inferential methods and interpretation have been to some extent modified.

The Bayesian inference operates on two factors, one an initial probability of authorship, say Madison wrote the 52nd paper, the other set of sampling distributions if Madison had written the paper and alternately if Hamilton had written the paper. The values investigated are the frequencies of selected words. With these two factors Bayesian theorem is used to calculate the final odds that the one or the other wrote the disputed paper.

The papers comprise short essays, 77 in the first instance and then 8. Of the initial 77, John Jay wrote 5, Madison 14, Hamilton 43. 3 were jointly written by Madison and Hamilton. The later 8 were by Madison. 12 papers are however disputed. All the essays were written during 1787-88.

The average sentence length and standard deviation does not help much because they are nearly equal for both authors :

Mean : Hamilton 34.55 (words per sentence)
Madison 34.59
s.d. : Hamilton 19'2
Madison 20'3

The large standard deviation means that some sentences were very long. The authors argue that in such investigations a single variable however carefully selected does not give so much satisfaction as a large pool of variables. The rate of

usage for each word can be regarded as a variable and a large collection of words would then provide overwhelming inference one way or the other for deciding the question. Previous investigators had noticed for instance that whereas Hamilton uses "while" Madison uses "whilst" in a corresponding situation. The problem was then to select a set of such "discriminatory" words. Contextual words were omitted from the list since their variation with respect to change of topic was to be expected and it was also not possible to evaluate their value for individual authorship. Some meaningful words like commonly, innovation seemed relatively free from context.

Frequencies distribution for upon : (page 19, table 2.1-4)

rate per 1000 words	Hamilton	Madison
0 (exactly)	-	41
0-1	1	7
1-2	10	2
2-3	11	
3-4	11	
4-5	10	
5-6	3	
6-7	1	
7-8	1	
	48	50

Thus low rates of upon favour Madison, high rates Hamilton.

The spread of the distributions seems also appropriate for both authors : Hamilton's with 3/1000 and Madison's with 0'18/1000.

The authors studied the effect of time on the word usage by taking any, by, from, her, in, must, one, some, there, where, would and dividing the texts into 5 groups in chronological order and calculating the rate of occurrence per 1000 words. The only word suggesting a trend with time was in and possible from.

The words any, her, there and would vary considerably from one group to another, while by, from, must, one, some and where appear fairly stable. The general conclusion of the authors is that pronouns and auxiliary verbforms are potentially contextual and offer risky discriminations. Elaborate procedures were followed for selecting the test words. The Miller Newman Friedman list of 363 function words produced 70 "unselected" high frequency words and 20 random low frequency words. Screening study based on all different words (about 3000) in 11 federalist papers produced 28 selected words. Index with frequencies based on some Hamilton federalist papers outside the screening study and some Madison non-Federalist writings produced 103 selected words. In the initial screening words were eliminated in hundreds for their potential dependence on context. The papers were studied in lots of 10 papers nearly equally

divided between the authors. The counts were made according to the number of Madison and Hamilton papers in which they occurred, (3,2) meant that the word occurred in 3 Hamilton and 2 Madison papers. The cumulative scores were made and after two pools the index

$$z = \frac{(x-y)^2}{(x+y)}$$

was introduced to measure the discrimination power of the word. Only words whose z value was greater than 3'6 were retained. The authors show that the discriminating power can be taken to be real by showing that the frequency distributions in the third pool for 22 Madison markers and 24 Hamilton markers do not distribute about zero, which should be the case if they had had equal chance.

The authors found that for most words the binomial distribution gives a fair fit, but also that it is not entirely adequate. The word frequencies were examined in blocks of texts, say 1000 words. Table 2.3-3 Mosteller (254), p.29 gives the occurrence of 51 words in the 247 blocks. Generally if the number of repetitions of a word is increased, then there are less blocks which will have this many repetitions of the particular word. Exceptions are <u>her</u> and <u>his</u>:

	0	1	2	3	4	5	6	7	8	9	14
<u>her</u>	241	3	1	1						1	
<u>his</u>	192	18	17	7	3	2	4	1	2		1

The occurrences for 10 words which were typical in the above set were compared for Poisson and the negative binomials distributions. The words chosen were : an, from, any, may, upon, can, every, his, do, my. The poisson distribution is given by

$$P(x) = \frac{\lambda^x e^{-x}}{x!} \qquad (x=0,1,2,\ldots)$$

where e = 2'71828... and λ is a given constant

where x is the average number of occurrences, and the probability is for exactly x occurrences for a given block. The Poisson and the observed distribution fit well for some words, but for may and his there is marked difference. To describe words like may and his one needs a family that offers a fatter tail than the Poisson. The negative binomial is such a distribution. For large values of x it gives substantially larger probabilities than the Poisson distribution.

Mosteller and WALLACE's analysis is a highly sophisticated statistical approach calling for real mastery of mathematical statistics.

MORGENTHALER's study of the statistics of the vocabulary of the New Testament does not employ advanced statistical methods, but is of interest for the wide spectrum of variables considered. Statistics of word classes like prepositions, verbs with prepositions, prefixes, etc.

are carefully analyzed giving due weight to thematic influence.[169] Next words which occur commonly in the works of the new testament are analyzed taking two texts at a time. The Object of study is not merely the frequency of such words but also which items of vocabulary are common, and which words may be considered "preferred words" of one or the other author. Grammatical and lexical preferences are evaluated, compared and contrasted keeping in view all the time the possibility that such items might be due to the specific nature of the subject matter and not due to style.

d. Rank frequency correlation

The relationship between rank and frequency has been the subject of intensive study starting with Zipf and undergoing modifications through JOOS, MANDELBROT and others.[170] The most satisfactory way of regarding it seems

---

169 Morgenthaler (253, p.13) points out : "Sachliche und stilistische Momente gehen dauernd durcheinander." (topical and stylistic moments continually go through one another). For instance the fact that ἐγώ is a preferred word in Johannes evangelium and not in the Johannes letters is a matter of theme. But in comparison with the 'environment' (i.e. of the other writings) it is considered Johannes-type.

to be the tripartite division suggested by Wagner (285) (vide fig.1, page 139)

the regions where H (frequency)=1		(Hapax Legomena)
rank-frequency upto 31		(JOOS region)
-"-	above 31	(Mandelbrot region)

Wagner sets the limit r_u at u=31 on empirical grounds. However as CARROL pointed out (1938) the results depend to a large extent on sample size. With newspaper language optimum size seems to be 120000 tokens.

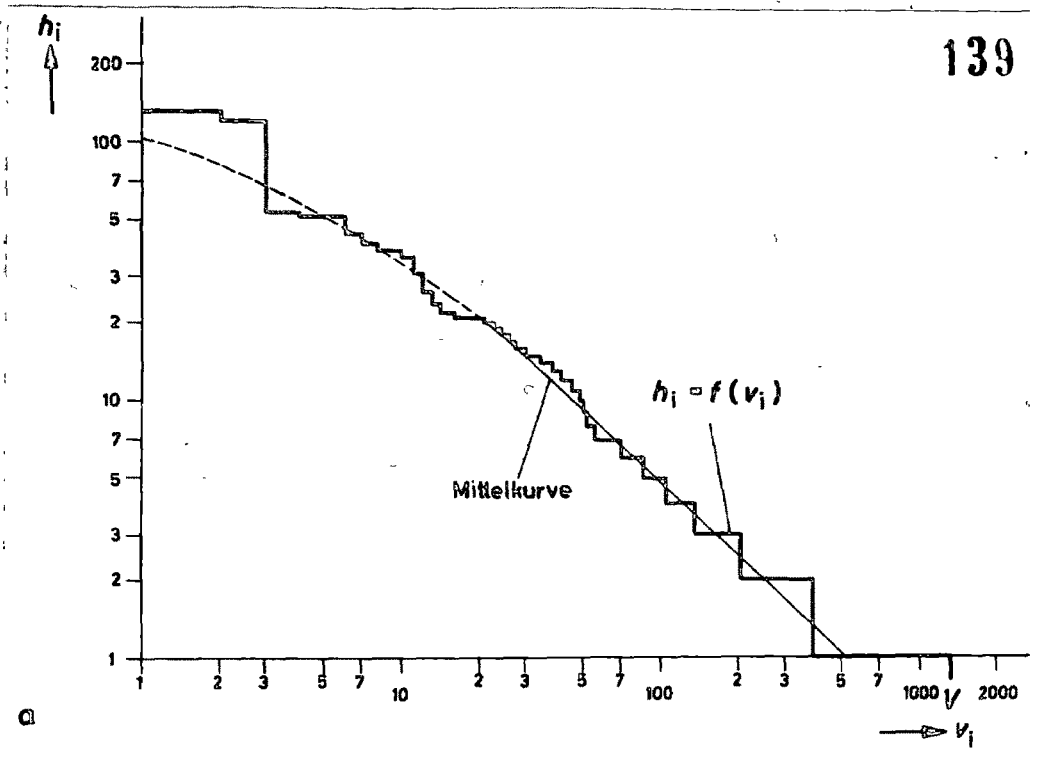
HERDAN has severely criticised repeatedly all the rank-frequency approximations, nevertheless such rank estimates serve a practical purpose, for instance, in information retrieval systems. WAGNER has for instance modified the formulae and uses the following criteria to establish the significance of words :

a) word rank r_e
b) word frequency h_e
c) word length L
d) no. of word pairs.

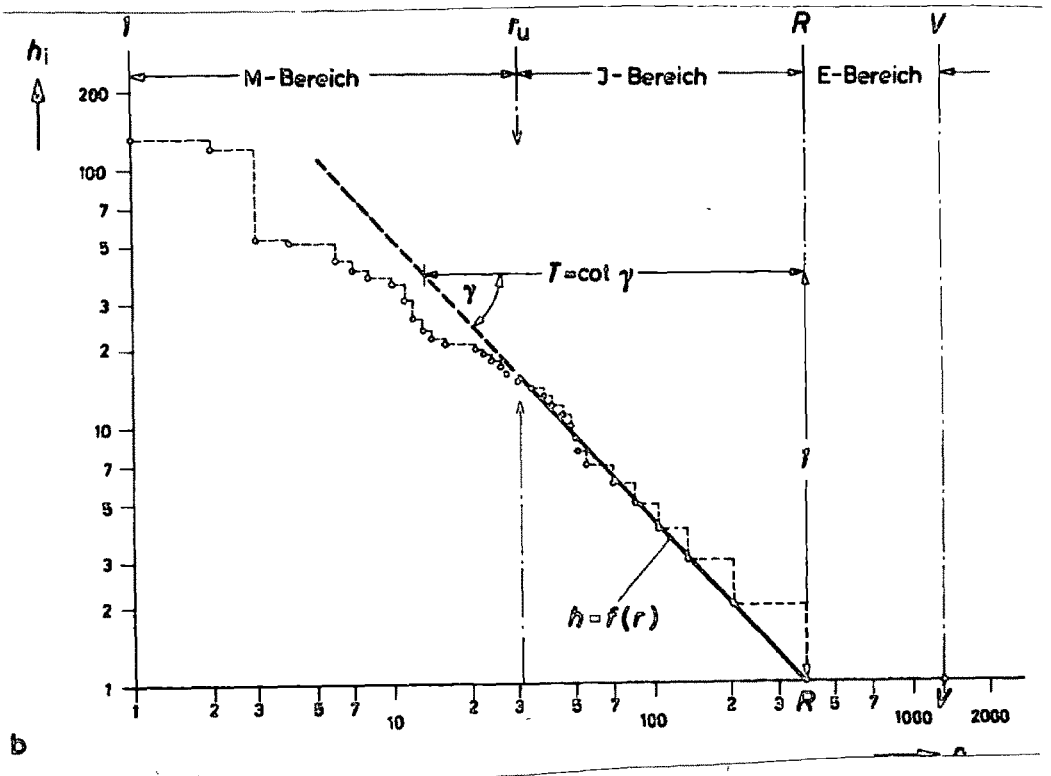
H.P.LUHN (243) makes the inclusion of a suspect clue word in the final list dependent upon the fact whether it appears in the same sentence with another clue word.

---

170 Zipf (291), Joos (226), Mandelbrot (246), Simon (275).



a



b

Fig. 1

Frequency-Rank distribution of a
Text with N = 3216 Words

a) conventional representation

b) special interpretation (division into 3 regions with introduction of parameters R and T)

(S.W. Wagner)

Summary

Summing up one might say that both descriptive statistics as well as inferential statistics have found important application in language analysis. Descriptive statistics can be used in practically any field of the language study, the interpretation of the statistical data will of course be purely the responsibility of the linguist. The data can be summarised by means of an index which will range from one particular value to another - vide the indices of FUCKS, but again the interpretation of the index is the sole responsibility of the analyst.[171] The range might be from 0 to 1 (as in probabilities), from -1 to +1 (as in correlation) or from 0 to 100% or in fact any other range found suitable for the particular problem on hand.

---

[171] For example, in Johannes I substantive pronouns occur 14'7% while in Johannes 2 the value is 13'7. Fucks sets up a difference-index U as follows:

$$U_1 = (14'7 - 13'7)/ (14'7 + 13'7) = 0'0352$$

The difference is taken always positive and the U-values are added for the various word classes. This sum gives the U-index which will summarise the results with reference to the distribution of word classes in two works and can serve for purposes of comparing various works.

A common application of statistics has been to draw inferences regarding cases of disputed authorship. In such problems the methods of statistics have to be carefully applied due to the very nature of language. Basic questions which have to be settled first are : does the author change his style very much? If yes, in statistical terms this would mean that the work or works considered are heterogeneous. By taking different samples from the same text one can establish the field under which the work lies.

As a test variable can be used for instance the mean number of words per sentence. The other works of the same author are tested and their fields are established. If now these fields lie close together one can conclude that the author is not susceptible to much variation. If now the disputed work shows values which lie beyond the fields established for a particular author then one can infer that in all probability the concerned author did not write the disputed work. As Fucks puts it, in finger-prints the overwhelming factors recommending their use are that they remain unchanged for each person, that they differ for each person and thirdly, that they can be established easily. In authorship studies one has to check if the author is addicted to changing his style in respect of the parameters investigated; one must establish that the parameter gives different results for different authors so that a relevant

comparison is possible. But as in determining the paternity of a child according to the blood group the negative test is only possible. Agreement of the statistical results can mean little for establishing the question of authorship, but disagreement with the results can be of great weight for deciding the question negatively.[172]

Statistical predictions in general take the following shape. An assumption about the distribution of the parameter under consideration is tested for its validity, say, that it conforms to normal distribution. Samples drawn are usually by using random sampling methods and if population statistics are known then one can estimate whether any sample came from the population. Given two samples one can estimate if they came from the same population, even when population statistics are unknown. These and a variety of similar questions can be answered by means of various tests with varying degrees of accuracy. If say the level of accuracy is 95% it means that if the same experiment or sampling is done 100 times, then 95 times the prediction will come out true. In 5% of the cases we might be making a wrong judgement.

A particularly useful fact in all these applications is that various common parameters like mean, standard

---

[172] Fucks (210).

deviation are themselves approximately normally distributed, i.e. if 10 different samplings are done and the standard deviations calculated for each of the 10 samples, then although the population from which the samples had been drawn might not be a normal one, the standard deviations themselves behave normally, so that a number of standard tests become applicable.

In language statistics one has to remember that there might be correlation among the units considered so that special methods will have to be adopted. Secondly the population might be 'fictitious' as in dealing with the vocabulary of an author. No one knows what the entire potential vocabulary of an author is, not even the author himself! Thirdly as in dealing with noun statistics, theme might have enormous influence. Fourthly what exactly is the variable under 'statistical' study might not be so obvious in language statistics.

Consider for instance the following distribution of nouns from a random selection of 21 pages from BUDDENBROKS :

x_i	f_i	$x_i f_i$
1	x 994 =	994
2	163 =	326
3	53 =	159
4	29 =	116
5	14 =	70
6	9 =	54
7	6 =	42
8	2 =	16
10	3 =	30
11	5 =	55
13	2 =	26
14	2 =	28
9 ⎫	1 =	130
12 ⎪		
15 ⎪		
16 ⎪		
17 ⎪		
18 ⎪		
21 ⎪		
22 ⎭		
	2046	

f_i represents the number of words which are repeated x_i times, i.e. 994 words are occurring only once, 163 nouns occur twice and so on.
Here the variable is not the noun nor is it the number of nouns which occur once or twice etc., but the number of times nouns can occur.
The mean of this distribution i.e. the mean repeat rate of a word is 1'586 and standard deviation is 2'423. This does not tell the linguist much, especially since it is not clear if the distribution can be described uniparametrically.

The linguist would like to pose rather questions of a different sort : given a sample of 2000 nouns from a known

author, what is the probable number of words which occur only once, how many words are likely to occur twice etc. To answer such questions a proper fit has to be established first of all with a known distribution. Secondly estimates of the population parameters must be available.

Notwithstanding the limitations imposed by the nature of language and by the methods of statistics, the application of statistics has thrown considerable light on linguistic problems. But a unified approach is yet to develop.