TRAVAUX
DU
CERCLE LINGUISTIQUE
DE COPENHAGUE

VOL. X_1

LOUIS HJELMSLEV & H. J. ULDALL

Outline of Glossematics

PART I: GENERAL THEORY

by

H. J. ULDALL

Second Edition

COPENHAGUE

Nordisk Sprog- og Kulturforlag
1967

OUTLINE

OF GLOSSEMATICS

I

TRAVAUX
DU
CERCLE LINGUISTIQUE
DE COPENHAGUE

VOL. X_1

LOUIS HJELMSLEV & H. J. ULDALL

Outline of Glossematics

A Study in the Methodology of the Humanities
with Special Reference to Linguistics

Second Edition

COPENHAGUE

Nordisk Sprog- og Kulturforlag

OUTLINE OF GLOSSEMATICS

A STUDY IN THE METHODOLOGY
OF THE HUMANITIES WITH SPECIAL
REFERENCE TO LINGUISTICS

PART I: GENERAL THEORY

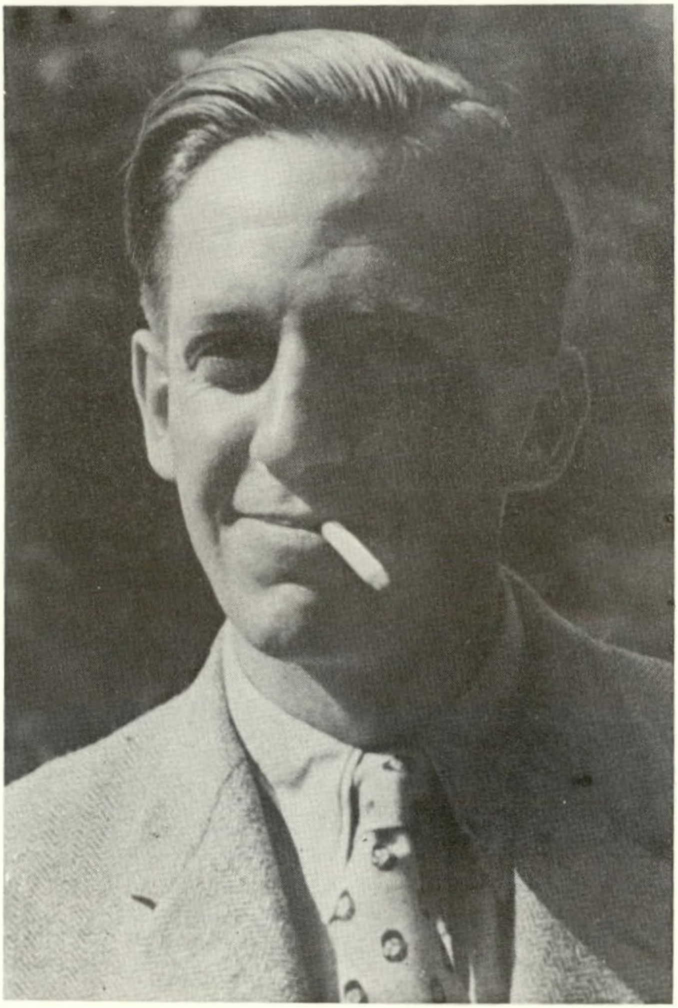
by

H. J. ULDALL

Second Edition

COPENHAGEN

Nordisk Sprog- og Kulturforlag



Hans Jørgen Uldall about 1938

Reprint in offset by H. Veenman & Zonen N.V. - Wageningen - Holland

PREFACE

The manuscript of this first part of "Outline of Glossematics" was delivered to the publishers in July 1952, and by the devoted efforts of publishers and printers a page-proof was produced in time for the International Congress of Linguists held in London in the following month. Circumstances beyond the control of publishers, printers, and author have prevented any further progress until now.

The original plan was to issue "Outline of Glossematics" in two parts, but it has since been decided that a more flexible schedule would be more convenient, and it cannot, therefore, now be said when or in how many instalments the rest of the book will appear. The sequence will appear as parts of vol. x of the *Travaux du Cercle linguistique de Copenhague.*

We wish to express our gratitude to the University of Copenhagen, the Rask-Ørsted Foundation, and the publishers for making it possible for Uldall to write this book without pressure of other duties.

The publication of "Outline of Glossematics" has been subsidized with generous grants from the Carlsberg Foundation and the Rask-Ørsted Foundation.

Copenhagen, July 1956.

LOUIS HJELMSLEV H. J. ULDALL

PREFACE TO THE SECOND EDITION

The present second edition of *Outline of Glossematics* I is a photostatic copy of the first edition, which appeared in 1957, and which has been sold out for some years. A small correction has been made on p. 79 ("specifications" changed to "selections"); an introduction by Eli Fischer-Jørgensen and a bibliography of the author's works have been added.

Hans Chr. Sørensen
Linguistic Circle of Copenhagen

INTRODUCTION[1]

A. Hans Jørgen Uldall (May 25, 1907–October 29, 1957) and his collaboration with Louis Hjelmslev.

The *Outline of Glossematics* was Hans Jørgen Uldall's main work and the last that he published, two months before his premature death in October 1957 at the age of fifty. It was the result of twenty years of struggling with the most fundamental problems of linguistic description under ever-changing, sometimes very difficult, material conditions.

Hans Jørgen Uldall was born in Silkeborg in Denmark on May 25, 1907, and would thus have completed his sixtieth year this spring. His father was a doctor, later chief physician of a hospital in Jutland. His mother died when he was a boy, and he was sent to the boarding-schools of Stenhus and (later) Herlufsholm. In 1924 he matriculated at the University of Copenhagen, where he started as a medical student, but soon changed over to English, which he studied until the spring of 1927 under Otto Jespersen and C. A. Bodelsen. Already at school he had been interested in phonetics, and in 1927 he went to England in order to study under Professor Daniel Jones in the Phonetics Department of University College, where he obtained certificates in English Phonetics and Spoken English, and impressed everyone by his unusual gifts as a practical phonetician. He soon acquired a complete command of the English language without any trace of foreign accent. In 1928 he published his first paper (on the Danish *r*)[2], and in the following years he published various short articles in *Le Maître Phonétique*. In May 1929 he married Inge Ottesen, the sister of an old Herlufsholm friend.

From May to December 1929 he taught at the University of Cape Town, acting as deputy for Professor D. M. Beach, who was on leave;

---

[1] Mrs. Elizabeth Uldall has asked me to write this introduction. I am grateful to her and to Mrs. Vibeke Hjelmslev for giving me access to the correspondence between Uldall and Hjelmslev. Mrs. Uldall has also gone through the manuscript and made various suggestions.

[2] No. 1 in the bibliography p. 91

he took over not only his teaching, but also, in spite of his mere twenty-two years, Professor Beach's various administrative responsibilities. In later years he talked with a certain sardonic humour about this very promising start of his academic career.

In 1930 he taught at the School of African Studies in London as an assistant to Professor A. Lloyd James. At the end of 1930 Uldall and his wife left for the United States to do field work on American Indian languages. Franz Boas had asked William Thalbitzer to find a young Scandinavian phonetician for this job, and Otto Jespersen had recommended Uldall[1]. In 1931 and again in the summer of 1932 Uldall worked on Southern Maidu (now called Nisenan) in California. He learnt to speak the language and made an extensive collection of texts. This work gave him great pleasure, he met many congenial linguists and anthropologists, and he found the problems of Indian languages extremely stimulating. He became warmly attached to his old Maidu informant, Bill Joe, and would relate with great gusto the tales of the old man's life and exploits. The time "when I was an Indian" and sat under the pepper tree workiŋg with Bill Joe, was something he felt a nostalgia for for the rest of his life.

During his stay in America he also studied anthropology under Franz Boas at Columbia University, receiving the degree of M. A. in anthropology there in 1933, though he never sent in his thesis, which was to have been on Maidu. He also lectured on phonetics at Columbia in 1932–33. Besides Maidu he also worked on Achumawi and Pomo, and in 1933 he wrote "A Sketch of Achumawi Phonetics"[2], a clear and systematic description based on Daniel Jones's phonemic theory, and "Pomo"[3] with Jaime de Angulo. But he hesitated to publish his Maidu material, except for very short sketches[4]. The slightly longer article on Maidu phonetics[5], published in 1954, was based on a manuscript from 1932. He found it difficult to analyse and describe his Maidu material in traditional linguistic terms. Later, in 1937–38, he worked out part of the grammar (the verbal system) according to glossematic principles, but he never published either the texts or this grammatical analysis. The texts have, however, recently been published by William Shipley[6]. They com-

[1] According to correspondence with Boas and Thalbitzer from 1929.
[2] Bibl. No. 12.
[3] Bibl. No. 10.
[4] Bibl. No. 9 and 16.
[5] Bibl. No. 30.
[6] Bibl. No. 34.

prise (with translation) 175 pages, and are, obviously, of the greatest interest, not only from a linguistic, but also from an anthropological point of view. The grammatical material is being prepared for publication by Niels Ege.

In 1933 Uldall returned to Denmark and stayed here until 1939. The thirties were a difficult time for scientific workers (as for other workers too); there was much unemployment, and there were few research scholarships. Moreover, the fact that Uldall did not have a Danish university degree created some difficulties. He did not succeed in getting a permanent post, but had various temporary jobs: he had a few classes in English as an assistant at the University, was co-editor of the great Danish-English dictionary, assisted Otto Jespersen with various publications, and, from 1935 to 1939, had a grant from the Carlsberg Foundation. In 1935–36 he lived in Udby, a village in Jutland, working on a description of the local dialect, but most of the time he was in Copenhagen. He sometimes thought of going abroad again, but he could not leave his wife, who had become seriously ill with tuberculosis. In the autumn of 1937 she died after two years' illness. This was a very hard blow for him.

One more reason kept him in Denmark: his collaboration with Louis Hjelmslev. Shortly after his arrival in Copenhagen in 1933, Uldall had become a member of the Linguistic Circle and of its phonological committee. This committee had been set up in 1931 with the purpose of working out a phonological description of Danish according to the theories of the Prague School. This was never completed, but the discussions in the committee led to a close collaboration between Hjelmslev, Uldall, and Paul Lier, which was decisive for Uldall's whole future work. Together they worked out a new theory which they called phonematics, and which was presented by Hjelmslev and Uldall at the Congress of Phonetic Sciences in London in 1935, where Uldall gave a paper on "The Phonematics of Danish"[1]. The main idea was that phonemes should not be defined on the basis of their relevant phonetic features, but on the basis of their functions, viz. combination, alternation, and implication. The choice of functions points to the influence of Sapir and Bloomfield.

In 1934 Hjelmslev had been appointed reader in comparative linguistics at the University of Aarhus, where he stayed until 1937, and this was

[1] Bibl. No. 17.

one of the main reasons why Uldall spent the year 1935–36 in Jutland. For the same reason Lier, who had a post in Copenhagen, could only occasionally take part in these discussions.

In 1935 it was planned that Hjelmslev should write a treatise on phonematics and Uldall a phonematics of Danish. At the end of 1935 they introduced a distinction between phonematics and cenematics, the latter being purely formal, and consequently the title was changed to "Outline of Cenology". This book was planned as a joint work, to be written by Hjelmslev and Uldall, and in March 1936 they asked for a subvention for publication from the Rask-Ørsted Foundation. About Christmas, however, they discovered that cenematics and formal grammar, as described by Hjelmslev in his earlier work *Principes de grammaire générale*, could be combined into a single discipline, in which content and expression were to be analysed according to the same principles. At the suggestion of Uldall they called this new discipline "glossematics". The name is found for the first time in their correspondence in a letter from Hjelmslev, January 5, 1936. In April they decided to let their planned joint work deal with the whole of glossematics, and the title was accordingly changed to "An Outline of Glossematics". In their youthful optimism they hoped to finish and publish this book before the International Congress of Linguists in September 1936. Consequently they worked very hard during the first half of 1936. Udby, where Uldall lived, was not far from Aarhus, and he often went there to see Hjelmslev. During some months in the winter he had a room in Aarhus. Very often they worked together the whole night, "and in the early morning I would bicycle home through the snow", as Uldall writes in a provisional preface from 1942[1]; in between they exchanged long letters.

This collaboration is, I think, something unique in the history of linguistics. They worked together so intimately and harmoniously that afterwards neither of them could tell who had contributed what (they have both used this expression several times). Hjelmslev was eight years older than Uldall, and he had a much broader linguistic background and had already published several works. It also appears from their correspondence that the main ideas stem from Hjelmslev, but Uldall's sharp intelligence was not only receptive and critical, it was also constructive, and he certainly contributed much to the whole development of the theory, and he was a very stimulating co-worker. In one of his letters from

[1] Printed in the introduction to the edition of Uldall's Maidu texts 1966, Bibl. No. 34.

1935 Hjelmslev writes: "My brain is always buzzing with ideas when we have talked together"[1]. From time to time Uldall was depressed and had serious doubts about his own capacity for scientific work and in particular for theoretical work. "I am increasingly aware of the fact that my faculties – such as they are – go in the direction of concrete research, and I shall hardly reach pure theory before I am an old man" (!) (1934). But he was always encouraged by Hjelmslev, and on the whole one has the impression of an almost passionate enthusiasm, which was, however, always balanced by their sense of humour and self-irony. "Dear Hjelmslev, after having sent off the letter, I am seized by an intoxicating idea: the stød is not at all phonematically relevant in Danish!!! (take a deep breath)"[2]. They certainly enjoyed themselves, and they also enjoyed shocking traditional linguists.

Of course they did not finish the book before the Congress, they only succeeded in publishing a small pamphlet of a few sample pages, with the title "An Outline of Glossematics", bearing the note "to be published in the autumn". By accident, but a curious accident, they failed to indicate the year. However, they had not yet lost their optimism. Hjelmslev planned to publish two big books in 1937 besides the books they planned together, and Uldall hoped to publish his Maidu material and to qualify for a chair of phonetics in Copenhagen. One must not forget that they were still young men, Uldall not yet 30. But the work was delayed, partly because they both had many other things to do, partly because they kept changing the theory. It was not until June 1937 that the three fundamental glossematic functions were formulated, and at this time the first procedure was still inductive.

From the autumn of 1937 they were again both of them in Copenhagen.

Early in 1939 Uldall accepted a post in Greece under the British Council, but he hoped to be able to spend almost half the year in Denmark. In the summer of 1939 their work progressed steadily, and they were approaching a final version of the whole theory (it seems to be at this time that the first and decisive part of the procedure was made analytic ("deductive")). But at the outbreak of the war Uldall had to leave for Greece with his second wife, the phonetician Elizabeth Uldall, whom he had married in the summer of 1939. Uldall's five-year-old daughter had to remain in Den-

---

[1] 9.10.1935.
[2] Postscript to a letter 4.7.1935.

mark. During the first year of the war Hjelmslev and Uldall were still able to continue their correspondence, and both worked constantly on the theory. In July 1940 Hjelmslev wrote that in a week he hoped to start making a fair copy of the whole procedure (which had at this time been supplemented by a final "paradigmatic procedure"), and would send Uldall the whole manuscript. But just at this moment all further correspondence became impossible, and they were cut off from each other for the rest of the war.

For Hjelmslev the theory seems to have reached an almost definitive form in 1941. In this year he wrote a summary of all the definitions and rules. This manuscript was revised in 1943, and made ready for printing, but he did not publish it because he hoped that the war would soon be over, so that he could show it to Uldall first. Instead he wrote the *Prolegomena*, 1943, which was meant as an introduction to the theory.

In the meantime Uldall and his wife worked for the British Council and were sent from one place to another: they spent 1939–40 in Athens, 1940–42 in Cairo, 1942–43 in Baghdad, 1943–45 in Alexandria, a good deal of the time living in hotels and pensions, and during this Odyssey various boxes with important papers and books disappeared. But Uldall kept on working on glossematics and seemed to find a sort of refuge in this work. In 1942 he planned to send out an English version on his own, but gave it up again.

In the summer of 1945 Uldall came to Denmark, but only for a short visit, as he had to continue his work for the British Council. In the autumn and winter of 1945 he was in London, and during these months he and Hjelmslev had an intensive correspondence, particularly on Hjelmslev's *Prolegomena*, which Uldall criticized at various points.

In February 1946 Uldall went to Buenos Aires to take up a post as linguistic adviser to the British Council. This was a very demanding and time-consuming job, which did not leave him much time for scientific work. Only in June 1946 did he manage to do some work on glossematics. In January 1948 he accepted a post as professor of linguistics and phonetics at the University of Tucumán in Argentina. Although the conditions there were not ideal (the climate was very warm, and he had to wait for half a year for his furniture, books, and papers) he nevertheless found time to take up glossematics again. In the short periods during and after the war when he had been able to work in this field, he had concentrated most of his efforts on setting up the system of possible duplex and triplex paradigmatic and syntagmatic categories and on working out an algebraic notation for them. During the summer and autumn of 1948 he continued this work intensively, and he also managed to write a paper "Ciencias culturales" (only found in manuscript) which is more or less identical with the first 35 pages of *Outline* I. – Moreover he found time to write an interesting paper on the semantic system of English tenses[1], and during this period he sent Hjelmslev a long series of letters, sometimes with only a few days between them, mostly containing algebraic formulae. His paper "On Equivalent Relations" published in 1949 in the volume *Recherches structurales*[2] gives a brief account of some of his results in this field. But Hjelmslev never reacted to all this, except for a few short letters where he promises to take up the problems later. No doubt Hjelmslev was extremely busy during these years; he had taken up much administrative work, and he travelled a great deal. But a further reason was that he was not very enthusiastic about this elaborate algebraic system, which he found too complicated to be practical in linguistic analysis. Moreover, as the details of the system kept changing from one letter to the next, he preferred to wait for the final version before really trying to understand it and taking up the discussion. As a consequence this whole system was worked out by Uldall alone without any contribution from Hjelmslev.

In 1949 Elizabeth Uldall accepted a post as lecturer in phonetics in Edinburgh, and six months later Uldall was appointed lecturer in linguistics in Edinburgh, where he was to work on the dialect survey of Scotland. However, he was not really suited to this task; what he wanted was to finish his glossematic work, and he did not find it possible to do both. In 1951 he gave up the lectureship and accepted an offer from Copenhagen to work there for a year; both he and Hjelmslev were now determined to finish their treatise on glossematics together.

During this year Uldall also lectured at the University on glossematics and on field methods with great success. He was an excellent teacher: clear, concise, friendly, and tolerant, and in addition his strong sense of humour and his unusual personal charm made his lectures extremely stimulating.

It was planned that Uldall and Hjelmslev should divide the work between them so that Uldall should write the introduction and describe the algebra and Hjelmslev should treat some other aspects of the theory

---

[2] Bibl. No. 28.
[1] Bibl. No. 29.

and give an account of the glossematic procedure containing all the rules and definitions. During the winter of 1951–52 they discussed the general principles and tried to come to an agreement about the definitions, and it appears from their papers that Hjelmslev accepted several of Uldall's new terms and formulations[1]. But they do not seem to have entered into any serious discussion about the algebra. Uldall was very fond of it, and Hjelmslev avoided talking about it, hoping that Uldall would change it himself. Thus, when Hjelmslev left for the United States in May 1952, Uldall finished Part I including the algebra, thinking that they agreed on all essential points.

As Hjelmslev had not succeeded in writing his part, and probably would not have time to do so within a reasonably short time, Uldall proposed that Part I should be published immediately, so that it could be presented at the International Congress of Linguists in London in September 1952. In August he sent the manuscript to Hjelmslev and asked for his consent by telegram. Hjelmslev accepted this proposal, and a second proof of the book was presented at the Congress. It was agreed that Hjelmslev should write a preface to be published in Part I, and continue his work on Part II. However, he found it very difficult to write the second part on the basis of Uldall's algebra and therefore, after a while, suggested that they should try to reach an agreement on the algebra and postpone the publication of Part I until the whole work was finished. Uldall, on the other hand, wanted to get his book out, and Hjelmslev, of course, realized that it was important for his career. Finally, after many years' hesitation, he signed the preface, and the book appeared in 1957. After a while Hjelmslev decided to write his part on the basis of the original algebra, but his working capacity was, already then, very much reduced, and Part II was never written. It is easy to see now that it would have been better for both of them if they had realized that their ways had parted, so that each of them must publish separately in his own name. But they hesitated to give up a collaboration which had started and continued for many years in an almost ideal spirit of mutual understanding. And they did not allow the tragic consequences of this hesitation to destroy their friendship.

---

[1] An introduction with translation of the terminology and explanations of the differences in theory was also planned (cf. letter from Uldall 9.2.1953 and from Hjelmslev 14.3.1953), but never written. – In the following notes *OG* I refers to the present book, *OSG* to Hjelmslev's *Prolegomena* with page references to the original Danish edition, *Summary* to the summary of glossematic definitions and rules, written by Hjelmslev in 1941, and revised 1943 (it will probably be published in 1969), *Lectures* to Hjelmslev's lectures on "Linguistic Theory" 1942–43, and U. or Hj. plus a date to the correspondence between Uldall and Hjelmslev.

After leaving Copenhagen Uldall tried, in vain, first to obtain a research fellowship in Edinburgh, then a post as reader in Danish in America; finally, in 1954, he accepted a post as senior lecturer in English and phonetics at the University of Ibadan in Nigeria. Here his unusual gifts as a teacher and as an organizer, and his talent for understanding people of all types and dealing with them in a completely natural way, were fully exploited. He not only taught at the University, he also arranged week-end and holiday courses for school-teachers, and took a very active part in planning and starting an extensive survey of West African languages.

In the summer of 1957 he was in Europe and took part in the International Congress of Linguists in Oslo. He was obviously happy to see his book out, and glad to meet old friends, witty and charming as always. Two months later, on October 29, 1957, he died from a heart attack the day after a minor sinus operation.

B. Uldall's Glossematic Theory Compared to Hjelmslev's.

As appears from the preceding historical account, glossematics is not simply one coherent theory. It is necessary to distinguish between Hjelmslevian and Uldallian glossematics, although of course, seen from without, and compared to other linguistic theories, they have very much, indeed almost all the fundamental ideas, in common. In this situation it may be useful to attempt to give a short survey of the features which distinguish Uldall's glossematic theory, as it is set forth in this book.

I. "General Principles".

I.1. The first part of the book ("General Principles" pp. 1–35) was meant as an introduction to their joint work, and Uldall therefore did not find it necessary to refer to Hjelmslev's *Prolegomena*, although he has, deliberately, used approximately the same wording in many places, e.g. in the formulation of the empirical principle and of the principles of economy, reduction and generalization. In this introduction one finds, on the whole, the same ideas as in the *Prolegomena*, and in a letter to Uldall Hjelmslev says: "I can subscribe fully to everything in this chapter"[1].

I.2. Common to both are: the endeavour to make linguistics an exact science, the emphasis on functions (not things), the concept of language as consisting of two planes (content and expression, each comprising form and substance), the distinction between syntagmatics and paradigmatics, and the conviction that linguistic analysis must start by a syntag-

[1] Hj. 23.7.1955.

matic analysis of texts according to strict procedural rules; finally, the general principles (the empirical principle etc.) mentioned above.

I.3. At some points there are deviations from the *Prolegomena* but agreement with Hjelmslev's paper "La stratification du langage"[1], which was written two years after Uldall's *Outline*, but published earlier. These points represent innovations upon which they had agreed during their discussions in 1951–52, viz. (1) the description of content substance as, primarily, a "body of opinion" (elements of collective evaluation), (2) the conception of content, expression, form, and substance as the four "strata" of language, and (3) the distinction between intrinsic units, defined by relations within one stratum, and extrinsic (or projected) units, defined by interstratic relations. The latter distinction is not completely new in glossematics. In Hjelmslev's *Summary* a distinction is made between extra-defined and intra-defined units[2] (relating to each of the two planes, content and expression), and as early as 1936 Uldall emphasized the importance of this distinction[3]. In 1945 Uldall uses the terminology intrinsic and extrinsic[4], in this case, it seems, in connexion with the distinction form – substance, but it is not until *OG* I that the distinction is used with reference to all four strata. – The rule that the strata should not be separated in the analysis, as long as there is conformity, was also adopted by Hjelmslev in "Stratification"[5].

I.4. The most obvious difference between Uldall's first chapter and Hjelmslev's *Prolegomena* is a difference of style. Whereas Hjelmslev writes a very homogeneous, academic prose, Uldall's style is extremely varied; in some passages it is an objective and concise matter-of-fact style demanding very much of the reader; in other passages, predominantly in this first chapter, the style approaches relaxed everyday speech, characterized by anecdotal digressions and striking comparisons with facts outside language, often elaborated with much humour.

I.5. The differences are, however, not only stylistic. First of all there is a certain difference in their conception of the scope of glossematics. For Uldall glossematics is a formal theory, which is not defined by any specific material, but designed explicitly to be used for all human activity[6].

[1] *Word* 10, 1954, pp. 163–88, reprinted in *Essais linguistiques* 1959, pp. 36–68.
[2] *Summary*, Definition 177 and 370.
[3] U. 12.1.1936.
[4] U. 26.1.1945.
[5] U. 24.7.1952 and *OG* I, p. 28; Hjelmslev "Stratification", p. 168 (*Essais*, pp. 42–43).
[6] *OG* I, p. 96, and U. 3.3.1941, 2.11.1945 and 30.7.1952.

Glossematics is only defined as a non-quantitative science[1]. Hjelmslev does not enter upon the distinction between quantitative and non-quantitative sciences or methods, and he is more modest in his pretensions. For him, glossematics is a linguistic theory, which, implicitly, may serve also as a model for other humanistic disciplines[2].

I.6. In some cases a difference in views can be deduced from the fact that some aspects of the theory, treated at some length in the *Prolegomena*, are left out or touched upon very briefly in *OG* I.

I.6.1. According to Hjelmslev the process presupposes the system[3]. Uldall tended to the opposite view[4], and based the definition of system on the definition of sequence.

I.6.2. More important is the difference in their conception of the relation between form and substance. For Hjelmslev it was a basic idea of glossematics that substance presupposes form[5], and he was shocked by Uldall's heretical views on this point. Uldall did not believe that this unilateral dependence could be maintained[6], and he did not consider it an essential point in glossematics. He even wanted to get rid of the terms "form" and "substance". For him the names of the strata were purely conventional, and the glossematic description in principle the same for all four strata[7]. After a while, this probably helped Hjelmslev to realize that the idea that substance enters into the commutation test[8] could be accepted without giving up the very basis of glossematics.

I.6.3. It is remarkable that neither the commutation test nor the difference between variants and invariants is mentioned in Uldall's book. It is true that it was agreed that variants and invariants should be treated in the second part, but these concepts are so fundamental in Hjelmslevian glossematics that one would expect them to be mentioned in a general introduction. As a matter of fact Uldall found the distinction between variant and invariant superfluous and even harmful[9], because, in his view, it over-emphasized the importance of the relation between content and expression.

[1] *OG* I, p. 18.
[2] *OSG*, p. 72 and Hj. 21. 7. 1951 and discussion of *OG* I, Linguistic Circle 18.2.1958.
[3] *OSG*, p. 39.
[4] U. 26.10.1945 and 7.11.1945.
[5] cf. Eli Fischer-Jørgensen, "Form and Substance in Glossematics" (*Acta Linguistica Hafniensia* X, 1966, pp. 1–33).
[6] U. 2.11.1945 and 24.7.1952.
[7] *OG* I. p. 28, and U. 7.1.1953.
[8] "Stratification" (cf. footnote 17), p. 171 ff.
[9] U. 2–3.11.1945.

II. "Glossematic Algebra".

It is in the second chapter of the book, "Glossematic Algebra", that the difference between Hjelmslev and Uldall becomes really apparent. This chapter contains not only the algebra, but also a number of definitions of terms used in the glossematic description.

II.1. Some of these have the same (or approximately the same) definitions as in Hjelmslev's theory, e.g. "function", "functive", "chain", "paradigm", "derivate". In some other cases the names are different, but the meaning the same, so that a direct translation is possible: U. "connexion" = Hj. "relation"[1], U. "equivalence" = Hj. "correlation" (for Uldall's use of "relation" and "correlation" see below), U. "sequence" = Hj. "process", U. "unit" = Hj. "part". (In Hjelmslev's terminology "unit" is a syntagmatic class.)

II.2. In still other cases the same designations are used, but with partly different meanings. This is, for instance, true of "analysis", "synthesis", "induction", and "deduction".

II.2.1. By "analysis" Hjelmslev understands "a description of an object by the uniform dependences of other objects on it and on each other", i.e. a division of an object into components (the object subjected to analysis is called a "class"). By "synthesis" Hjelmslev understands "a description of an object as a component of a class"[2], i.e. the opposite. Uldall, on the other hand, defines "analysis" as the registration of a connexion field and "synthesis" as the registration of a paradigm[3]. This means that analysis and synthesis, in Uldall's theory, belong to syntagmatics and paradigmatics respectively.

Both define, in different wordings, "deduction" as a continued analysis[4], going from larger to smaller components, and "induction" as a continued synthesis, going from smaller to larger classes; but as "analysis" and "synthesis" mean different things in the two theories, this agreement is only apparent. Uldall's "deduction" is, in Hjelmslev's terminology, "syntagmatic deduction" (in contradistinction to Uldall Hjelmslev can also talk of a "paradigmatic deduction"[5]), and Uldall's "induction" is, in Hjelmslev's terminology, "paradigmatic induction". This more res-

---

[1] In *OSG* Hjelmslev uses "connexion" as a common term for determination and interdependence, in the English version it is replaced by "cohesion".

[2] *OSG*, pp. 27 and 29.

[3] *OG* I, pp. 45 and 57.

[4] *OSG*, p. 29 and *OG* I, pp. 45 and 57.

[5] *OSG*, p. 89, *Summary*, rule 151.

tricted use of the words is quite understandable, since Hjelmslev has, in most cases, used "deduction" in the sense of "syntagmatic deduction", and moreover often used "synthesis" for "paradigmatic deduction"[1]. The whole terminology is further complicated by the fact that both Hjelmslev and Uldall use "deduction" also in its more normal sense[2]. Uldall's use of "deduction" in the specific sense of a series of analyses is probably a concession to Hjelmslev. The term is not found in his proposals for glossematic definitions from 1951 nor in earlier drafts, and in his first version from 1948[3] he gives the following definition: "By deduction I understand the method of constructing a hypothesis for the purpose of explaining a material".

II.2.2. Besides the difference in terminology there is also a difference in procedure. Hjelmslev starts with a syntagmatic deduction, followed by a paradigmatic deduction, which may, in its turn, be followed by a synthesis of units (syntagmatic classes)[4]. Uldall states explicitly that each operation comprises both an analysis and a synthesis. "An analysis is a registration of a function and of its terminals, and the synthesis consists in classing all those components together which can be terminals of the same functions"[5]. The classes are set up inductively[6]. – Hjelmslev was against this inductive procedure. His own "paradigmatic deduction" seems to have been planned as a distribution of the smallest elements (the glossemes) found in the syntagmatic deduction on a paradigmatic hierarchy of categories[7].

A further difference of procedure is that Hjelmslev starts from a definite basis of analysis: First the whole text is analysed with solidarity as a basis of analysis, then with selection as a basis of analysis[8]. Uldall, on the other hand, does not start from a definite basis of analysis, but determines in each case the orientation of the function[9]. Hjelmslev considered this an important difference between their procedures[10].

These differences between their procedures should be seen in con-

---

[1] Hj. 12.7.1940, *Lectures* p. 70 (April 1942).

[2] OG I, p. 34.

[3] "Ciencias Culturales" in English, destined for a South-American periodical, but not published.

[4] *OSG*, p. 89.

[5] OG I, p. 25.

[6] U. 3.11.1945.

[7] *OSG*, pp. 89–90, *Summary* p. 117 ff. and p. 138.

[8] *OSG*, p. 77. *Lectures* p. 162 and pp. 183–86 (Oct. and Nov. 1943).

[9] U. 7.11.1945.

[10] Discussion of OG I, Linguistic Circle 18.2.1958.

nexion with the fact that for Hjelmslev the glossematic procedure was, above all, an instrument of final control of the analysis of a language already known to the investigator, whereas Uldall seems also to have had the field-work situation in mind. When talking about the inductive establishment of classes, he writes: "This is the procedure you would use in practice. When sitting with your Indian under a tree, you will, of course, do your best to register as many differences as possible"...[1]

II.3. The most obvious common feature of Uldall's and Hjelmslev's analysis is the use of the three glossematic functions, which in syntagmatics are called "selection", "solidarity", and "combination", and in paradigmatics "specification", "complementarity", and "autonomy". But these functions are defined and applied in somewhat different ways by Hjelmslev and Uldall.

II.3.1. In Hjelmslev's *Prolegomena* the functions are defined by means of the concepts "constant" and "variable". A "constant" is a "functive whose presence is a necessary condition for the presence of the functive to which it has function", i.e. it is presupposed by the other functive. A "variable" is a "functive whose presence is not a necessary condition for the presence of the functive to which it has function"[2], i.e. it is not presupposed by the other functive. In selection (and specification) the function takes place between a constant and a variable, in solidarity (and complementarity) between two constants, and in combination (and autonomy) between two variables.

Uldall does not use the concepts "constant" and "variable", but a different pair: "major" – "minor". These terms are introduced for the first time in Uldall's paper "On Equivalent Relations" 1949[3]. A "major terminal" is here defined as a relate which is equivalent with (i.e. enters into the same relation as) its first degree arrivate (by an arrivate of a unit is understood the chain of which it is a derivate); a relate which is not equivalent with its first degree arrivate is a "minor terminal". In selection (and specification) the function takes place between a major and a minor terminal, in solidarity (and complementarity) between two minor terminals, and in combination (and autonomy) between two major terminals.

In selection Uldall's "major terminal" thus corresponds to Hjelmslev's

[1] U. 3.11.1945.
[2] *OSG*, p. 32.
[3] Bibl. No. 29.

"constant" and his "minor terminal" to Hjelmslev's "variable", whereas in solidarity and combination the correspondence is reversed: here "major" corresponds to "variable" and "minor" to "constant". This difference is due to the fact that Hjelmslev's definitions are based upon the presuppositions between the two terminals, as it appears from their occurrence in other parts of the text, whereas Uldall's are based on the ability of the terminals to occur alone in the same connexion on the preceding derivational level. Uldall had introduced this type of definition already in 1945[1], but at that time without explicit definitions of major and minor terminals.

II.3.2. This is not only a difference of formal definitions. It has various consequences for the analysis. Hjelmslev has often emphasized that the function between the single members of two categories will often be found to be combination, even if the function between the categories is selection or solidarity. Thus the category of consonants selects (by definition) the category of vowels, but there is normally combination between single vowels and consonants. Similarly, in Latin there is solidarity between the categories of number and case, but combination between any particular number and any particular case[2]. According to Uldall's definitions, however, there is not generally combination between the single members of the categories[3]; there is, for instance, selection also between single vowels and consonants, since only the vowel will have the ability of functioning as a syllable alone, and there will be solidarity between the particular case and the particular number since neither can function alone in the same connexion. Uldall's definitions would thus, better than Hjelmslev's own definitions, satisfy Hjelmslev's requirement that the analysis should register as many selections and solidarities as possible, since mere combinations are of less interest[4].

II.3.3. In Hjelmslev's view, however, Uldall's definitions have a serious drawback from a different point of view. Since they are based on functions in definite connexions, they will often be concerned with positional variants, and not with invariants (as mentioned above Uldall did not attach much importance to this distinction), and the same functives may enter into different relations with each other under different conditions, i.e. in respect of different generating connexions. An example may show this:

---

[1] U. 2.11.1945.
[2] e.g. *OSG*, pp. 26 and 77–78.
[3] U. 3.11.1945, and "Notes on the Definition of Direction", undated ms, probably from 1947–48.
[4] *OSG*, p. 75.

According to both Hjelmslev and Uldall there is selection between Latin *sine* and the ablative (*sine* is not found without an ablative); on the other hand, according to Hjelmslev, there is combination between *ab* and the ablative, because *ab* is found elsewhere as a preverb without the ablative[1]. But according to Uldall this is a different connexion, which should not be taken into account; there is selection also between *ab* and the ablative, because *ab* cannot occur in the given connexion without an ablative[2]. Hjelmslev would consider the preposition *ab* as a positional variant.

Uldall is even interested in very restricted classes of variants. One of his examples in *OG* I is the relation between *p* and *l* in English[3]. According to his algebraic notation he considers the relation between *p* and *l* to be a combination in connexion with *-ei* (cf. *play*, *pay*, *lay*, *A*), a solidarity in connexion with *-aant (plant)*, and a selection in connexion with *-ʌv (love)*. Hjelmslev, on the other hand, is only interested in the functions between categories of invariants, and one of his main objections against traditional syntax was that it was concerned with variants[4].

There is an obvious similarity between Uldall's definitions of the three functions and Bloomfield's definitions of endocentric (subordinative and co-ordinative) and exocentric constructions (subordinative constructions corresponding to selections, co-ordinative to combinations, and exocentric to solidarities), and a certain influence cannot be excluded. Paul Diderichsen had, in his syntactic studies, reached exactly the same definitions as Uldall, though independently both of Uldall and of Bloomfield. In a communication in the Linguistic Circle 1947 Diderichsen compares Bloomfield's and Uldall's definitions[5], the latter known from Uldall's correspondence with Hjelmslev, and in his paper "De tre Hovedarter af grammatisk Forbindelse" from 1952[6] he takes up the discussion again. Like Hjelmslev he emphasizes that the glossematic relations, e.g. the selection between vowels and consonants, define classes of invariants by means of their possibility of occurring alone, whereas the description of the relation between "old" and "men" in "old men" as a subordination

[1] *OSG*, pp. 24–25

[2] Uldall Notes (cf. footnote 3, p. XV).

[3] *OG* I, p. 62.

[4] Reports from the meetings of the Glossematic Committee 1950, p. 11 (this report will probably be published in a collection of the selected reports from meetings in the Linguistic Circle 1942–65).

[5] "Klasse, Relation, Helhedstype" 1947, published in Paul Diderichsen *Helhed og Struktur, selected linguistic papers with detailed English summaries*, 1966, pp. 98–115, particularly pp. 104 and 112.

[6] *Festskrift til L.L. Hammerich* 1952, pp. 89–104, reprinted in *Helhed og Struktur*, pp. 192–214.

indicates a relation between specific variants of just these two words[1]. Uldall protested against this interpretation of subordination[2]. As a matter of fact I think that it is necessary to distinguish between three different types of analysis: when in traditional syntax "old" is said to be subordinate to "men" in "old men", this is normally based on a logical or semantic analysis of the relation between these two words in the given construction without taking other occurrences into account, and in this case the functives are single variants (not free variants, but single positional variants, in Hjelmslev's terminology "varieties"), of the words in question. In Uldall's analysis (as in Bloomfield's and Diderichsen's) we are not dealing with single positional variants in the same sense, since e.g. "men" can hardly be said to be the same variant in "the old men came" and in "the men came"; they are the same only in respect of the external relation (with "came"). Here, as in Hjelmslev's glossematics, the analysis is concerned with possibilities of occurrence, not with simple occurrence[3], but whereas the functives in Hjelmslev's analysis are invariants or (generally) categories of invariants, they are, in Uldall's analysis, often categories of variants.

The fact that the functives in Uldall's analysis are not single variants appears also from his terminology. He distinguishes between two sets of terms for syntagmatic and paradigmatic functions: the general terms (corresponding to Hjelmslev's "correlation" and "relation") are "connexion" and "equivalence", and these must be used for functions between single functives, like *p* and *l* in *play*, whereas the terms "relation" and "correlation" have a more restricted application, being used only for functions establishing a category (relations establishing categories of chains, correlations establishing categories of paradigms). The three glossematic functions are defined as types of relations and correlations. "Category" means both in Hjelmslev's and in Uldall's terminology a specific type of paradigmatic class[4], but Hjelmslev defines it as a paradigm with a definite function[5], Uldall as a collection of paradigms

---

[1] *Ibid.* (*Helhed og Struktur*, p. 199 and p. 210).

[2] U. 12.8.1952.

[3] U. 12.8.1952, where he emphasized the difference between actual and possible occurrence, and "Theory II" (an earlier version of OG I, approximately 1948). Here he writes that the functions "give a summary of the possibilities of occurrence of the words involved".

[4] Relations are thus not purely syntagmatic functions, they imply paradigmatics, cf. *OG* I, p. 80.

[5] "La notion de rection", *Acta Linguistica* I, 1939, p. 14, "Essai d'une théorie des morphèmes", *Actes du 4. congrès intern. des ling.*, p. 151. That this function should be a correlation (cf. *OSG*, p. 76 and *Summary*, definition 97) is difficult to reconcile with the examples given and with rule 69 of the *Summary*, according to which relations always take place between categories and correlations between units.

or chains with the same terminals[1]. It is difficult to say precisely what this difference means in practice, but it is at any rate doubtful whether all Uldall's categories would be categories in Hjelmslev's terminology.

II.3.4. This has to do with a further characteristic feature of Uldall's theory, i.e. his use of negatives and negated units. In syntagmatics Uldall distinguishes between positive and negative units: "If two units *ab* and *a* are compared, then *b* is said to be positive in *ab*, negative in *a*, which is now written $a\bar{b}$"[2]. A negative indicates the absence of a particular unit from a particular place, i.e. an unoccupied glossematic place. The number of places is found by counting the number of positive places in the largest possible chain. Thus *play*, *pay*, *lay*, *A* can be written *plei* $p\bar{l}ei$ $\bar{p}lei$ and $\bar{p}\bar{l}ei$. In paradigmatics a distinction is made between asserted and negated units. "A unit which has been registered as a terminal of a given connexion, is said to be asserted in respect of that connexion. Symbol $+$: $(+a)$. $(+b)$." .. "A unit which has been registered as not occurring as a terminal of a given connexion, is said to be negated in respect of that connexion. Symbol $-$: $(-a)$. $(-b)$."[3] The example given is German "auf" with accusative or dative, i.e. "auf".(+a +d), and "um" with accusative only, i.e. "um" .(+a −d). The negation of a unit presupposes that the same unit has been asserted elsewhere in the material.

These concepts, which are not found in Hjelmslev's theory, are used in the definitions of "major-minor" and of the glossematic functions. In *OG* I the functions are defined only in terms of these algebraic notations, but the notations cover the same reality as that described above on the basis of Uldall's definitions in "On Equivalent Relations". For duplex categories (i.e. categories with two members) we get the following definitions: when $+\langle +a\ -b\rangle$, then *a* is called a major correlate, when $-\langle +a\ -b\rangle$, then *a* is called a minor correlate, similarly: when $+a\bar{b}$, then *a* is called a major relate, and when $-a\bar{b}$, then *a* is called a minor relate[4].

In the following we shall use only syntagmatic examples (the categories of paradigms are treated in an exactly parallel manner). The three relations are defined in the following way by means of algebraic formulae: (1) the relation of a duplex category in which $+a\bar{b}\ +\bar{a}b$ is called a combination; if either $+a\bar{b}\ -\bar{a}b$ or $-a\bar{b}+\bar{a}b$, it is called a selection; and if

[1] OG I, p. 59. "Collection" has in the final printed version of OG I replaced "paradigm", one of the very few corrections introduced in the proof from 1952.

[2] OG I, p. 47.

[3] OG I, pp. 49–50.

[4] OG I, pp. 66 and 78.

$-ab$ $-ab$, it is called a solidarity[1]. It will be seen that these formulae correspond exactly to the definitions by means of major and minor relates given above. Now in all these four cases (combination, solidarity, and two types of selection) it is possible to have either $+ab$ or $-ab$. In the first case the relation is called "conjunct", in the second "disjunct"[2]. An example of conjunct combination is *pl* in *play pay lay* $(+ab\ +a\bar{b}\ +\bar{a}b)$, and an example of a disjunct combination would be *pl* in *pad lad* $(-ab\ +a\bar{b}\ +\bar{a}b)$, and similarly for solidarity and selection. The formulae show perhaps more clearly than the definitions given above (II.3.1.) that Uldall does not consider it necessary for the two relates to be found together in order to speak of combination, solidarity, and selection. Hjelmslev's definition also seems to cover the case $-ab$ for combination (a function between two variables, neither of them presupposing the other), but hardly for selections and solidarities. How can we know that *b* presupposes *a* or that they presuppose each other if they are not found together? It is evident that each of the three formulae (1) $-ab\ +a\bar{b}\ +\bar{a}b$, (2) $-ab\ +a\bar{b}\ -\bar{a}b$, and (3) $-ab\ -a\bar{b}\ -\bar{a}b$ can only cover (1) combination, (2) selection, and (3) solidarity, and neither of the other two relations, but it could be maintained that there is no function at all between *a* and *b* in these cases. Uldall was sometimes in doubt himself on this point[3]. Hjelmslev does not give any examples of the type with $-ab$, and his terminology ("combination", "solidarity") seems to indicate that he thought of cases where the terminals do appear together. Here it must, again, be kept in mind that Uldall's categories are much more restricted than Hjelmslev's. Uldall speaks of $(+pl\ +p\bar{l}\ +\bar{p}l).ei$ and $(+pl\ -p\bar{l}\ -\bar{p}l).aant$ as two different categories. For Hjelmslev they would be single paradigms, which might be combined into a category, and only for this more abstract category would he set up the glossematic relations. If none of the paradigms contained $+ab$, Uldall would probably not posit $-ab$, nor any function between *a* and *b*, since a negated unit must be asserted somewhere else in the material (but he has not given any more exact definition of what "somewhere else in the material" means).

Each of the eight duplex categories mentioned above can now be further varied since they may contain $+\bar{a}\bar{b}$ and $-\bar{a}\bar{b}$. The relation of a duplex category in which $+\bar{a}\bar{b}$ is called an "absence" relation, if $-\bar{a}\bar{b}$ it is

---

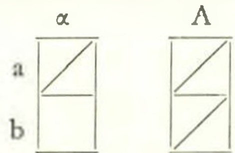
[1] *OG* I, p. 78.

[2] In *OSG*, pp. 34–35, the words "conjunct" and "disjunct" are used in a different sense (= coexistents and alternants).

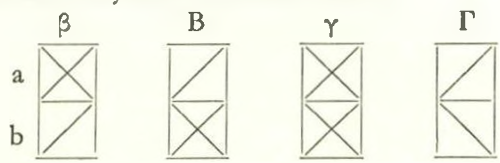
[3] U. 2.7.1948.

called a "presence" relation[1]. A full notation of a duplex category has thus four terms: *ab*, $\bar{a}b$, $a\bar{b}$, and $\bar{a}\bar{b}$, which may all be asserted or negated, and there will thus be 16 different categories in all. In *play pay lay A* $(+pl\ +p\bar{l}+\bar{p}l+\bar{p}\bar{l})$ we have an absence relation (*pl* is absent in *A (ei)*). In *true too rue* $(+tr\ +t\bar{r}\ +\bar{t}r\ -\bar{t}\bar{r})$ we would have a presence relation, since *u:* is not found alone; either *t* or *r* (or something else) must be present in this connexion. This terminology may be somewhat confusing, and, as mentioned by Uldall himself[2], the difference between $+\bar{a}\bar{b}$ and $-\bar{a}\bar{b}$ indicates external relations, i.e. relations between the chain *ab* (e.g. *pl*) and the other terminal of the given connexion (e.g. *ei*), not relations between *a* and *b*. This is probably the main reason for Hjelmslev's objection[3] that there is a confusion of levels in Uldall's algebra.

II.3.5. This very elaborate system (with 16 duplex and 27 triplex categories) had, according to Uldall[4], two main sources: (1) Hjelmslev's three glossematic functions and (2) Hjelmslev's system for participative oppositions as presented in *Cas* I 1935[5]. This system, which was devised for the description of semantic oppositions, but later applied by Hjelmslev in a somewhat different way to the so-called "functival categories", operates with two contrasting zones *a* and *b* (and sometimes a third neutral zone *c*), which may enter into more or less complicated oppositions, the simplest being the opposition with two terms:



which might, for instance, symbolize the opposition between "young" and "old" ("old" may cover the whole field (e.g. "how old are you?") and is therefore symbolized by A, whereas "young" covers only one zone). Uldall started from the system with four terms:



[1] *OG* I, p. 79.
[2] Linguistic Circle 19.9.1950.
[3] *Ibid.*
[4] Earlier version of *OG* I from 1947 or 1948, p. 20.
[5] Louis Hjelmslev, *La Catégorie des Cas* I, 1935, p. 98 ff. and *Lectures*, November 1942.

in which β may cover the whole field, but with "insistence" on *a*, and B may also cover the whole field, but with insistence on *b*, which implies the interpretation that β may have the (principal) variants *a* and *ab*, B the variants *b* and *ab*, γ covers *ab* and Γ either *a* or *b*. Uldall found in this system an analogy[1] to the glossematic functions: β = selection (*a* selected by *b*, thus $a + ab$), B = selection (*b* selected by *a*, thus $b + ab$), γ solidarity *(ab)* and Γ combination (*a* or *b*). "This scheme .... has in fact given rise to the whole theory"[2].

Hjelmslev did not approve of this, partly because he found Uldall's system too complicated to be useful in practical linguistic work[3], partly because (at any rate in 1958) he considered his system of oppositions as an empirical hypothesis which he would not like to include into the theory itself[3].

Besides the indications by means of + and –, Uldall invented a system of arrows which he kept changing during the years 1947–52. The last version, published in *OG* I, is rather different from the preceding version from October 1951. It is thus by no means certain that he would have kept the present version.

II.4. Among the notes from Hjelmslev's and Uldall's discussions in 1951–52 there are some sheets dated 7.11.1951, written by Hjelmslev and containing a list of the definitions in Hjelmslev's *Summary* which should be replaced by Uldall's definitions. Among these are "analysis", "connexion", "equivalence", "relation", "correlation", "selection". The list must be very provisional, since e.g. "selection" is the only one of the glossematic functions mentioned; and it was obviously set up at a time when Hjelmslev still hoped that they could come to an agreement. He must have given up these concessions afterwards, for in the English version of *OSG* (*Prolegomena*) and in the article "Stratification" he has not adopted any of the above-mentioned definitions. Hjelmslev also stuck to their old algebra, which, by the way, is no algebra in the sense that it

---

[1] U. 7.11.1945 and *OG* I ms. 1947 or 1948, p. 21.

[2] *Ibid.* (*OG* I, ms. 1947–48) p. 21b. – One might perhaps also have expected a reference to logical algebra in this place. The general approach is, of course, inspired by this algebra, as in fact appears from Uldall's book. But the concrete algebraic system seems to have been elaborated by Uldall himself rather independently of logical algebra. This appears from his correspondence with Hjelmslev, particularly 1945–50.

[3] Hj. 23.7.1952, and Linguistic Circle 19.9.1950 and 18.2.1958. Hjelmslev's system of participative oppositions is, however, included in the rules and definitions of his *Summary*.

would be possible to manipulate with it (make additions or subtractions or the like), but only a system of symbols.

It is therefore necessary to keep Hjelmslev's and Uldall's theories apart, when evaluating or applying glossematics.

Hjelmslev was probably right in considering Uldall's algebra unnecessarily complicated, but on the other hand Uldall's procedure is no doubt easier to understand and to apply than Hjelmslev's because, although he was aiming at a more abstract theory, he was at the same time in some sense closer to the reality with which most linguists are faced. It is regrettable that Hjelmslev and Uldall did not succeed in uniting their theories, but both of them will certainly remain a source of inspiration to the linguistic world.

ELI FISCHER-JØRGENSEN

GENERAL PRINCIPLES

The status of the humanities[1] *vis-a-vis* the exact sciences has been copiously debated, though no generally accepted conclusion seems to have been reached. It is obvious that there are very considerable differences between the two groups of disciplines, but wherein exactly these differences consist and what is the reason for them are questions which are not only difficult and complicated but calculated to arouse an astonishing amount of emotion.

Nobody would deny that the natural sciences, particularly physics, have reached a higher state of development than the humanities—that they have, to put it bluntly, been more successful. The problem under debate is not whether this is so but why it is, or, more precisely, whether the backwardness of the humanities is due to some inherent—and thus inescapable—factor or if, by changing their methods, the humanities could bring themselves into alignment and become equally successful. Considering the present political and economic state of the world, it is a question of some practical as well as theoretical interest. It is a common, and a reasonable, diagnosis of our current misfortunes that they are due to science having got too far ahead of us, *i.e.* to a discrepancy between the skill with which we control inanimate and the lower reaches of animate nature, and the clumsiness with which we handle human affairs. Some maintain that the root of the evil is moral turpitude; however that may be, it seems indicated that a better understanding of the way human affairs actually work would be a considerable help.

The historians, who have, on the whole, been rather more vocal on the subject than their colleagues in the other branches of the humanities, generally take the line that their material is so unstable, subject to such vacillations and fortuitous changes, that it is quite hopeless to try to

[1] For lack of a better one I shall use this term throughout to refer to those disciplines which deal with the social aspects of human existence. To use "social sciences" would be begging the question, and what else is there?

systematise it. Thus Burckhardt says, in his *Weltgeschichtliche Betrachtungen,* "Die Geschichte ist ja überhaupt die unwissenschaftlichste aller Wissenschaften, nur dass sie viel Wissenswürdiges überliefert. Scharfe Begriffsbestimmungen gehören in die Logik aber nicht in sie, wo alles schwebend und in beständigen Uebergängen und Mischungen existiert. Philosophische und historische Begriffe sind wesentlich verschiedener Art und verschiedenen Ursprungs; jene müssen so fest und geschlossen als möglich, diese so flüssig und offen als möglich gefasst werden."

This is, of course, a possible point of view, and it is understandable that a scholar who feels the very ground shifting under his feet should come to the conclusion that he can do no more than give a simple report, in terms of ordinary common sense, of the events under his observation. Many ethnologists and sociologists, to say nothing of the unfortunate economists, feel the same way about it. Nevertheless, it remains unproved that the fluctuations of which the historian complains are inherent in the material and not in the method brought to bear upon it. There does not seem any reason to believe *a priori* that the material of the exact sciences is in itself more stable than that of history, and yet the scientists have succeeded in applying, even to the study of motion itself, those *fest und geschlossen* concepts which the historian professes himself unable to use.

Burckhardt's argument is essentially another form of the old appeal to the uniqueness and unpredictability of human nature. This is found, for instance, in L. H. Gray's statement that linguistics "is not an exact science in the sense that mathematics and chemistry are exact; the human factor in it is too strong to permit it to be merely mechanical in operation."[1]

If two candles on the dining-table do not burn exactly alike—and they rarely do—it might, analogously, be inferred that the candle factor is too strong to permit physics and chemistry to be merely mechanical in operation. It is, however, also possible and, indeed, usual to try to explain the difference in the performance of the two candles as due to structura differences in the candles themselves or to different external circumstances such as temperature and air-currents. The "candle factor" is thus divested of its mystery and seen for what it is: a combination of quite ordinary, calculable features. Might it not be that Gray's "human factor" could be similarly reduced?

Progress in knowledge has been made only when men were willing to criticise preconceived notions so strongly held that they had never been

---

[1] *Foundations of Language,* New York, 1939, p. 4.

tested. Anatomy, physiology, evolutionary biology were stunted as long as it was a sacred belief that the human body was essentially different from the rest of creation. In the same way the humanities have been prevented from breaking out of the chrysalis by this view of the human race as a sort of cosmic aristocracy, this belief that the human mind is *sui generis*—the one domain in the universe subject to the mysterious force of Free Will and therefore not to be approached in the same spirit or by the same methods as anything else.[1]

But psychology, though itself hardly an exact science as yet, has at least demonstrated that the Free Will is not quite as free as we like to think, that the subject-matter of psychology (one scarcely dare use the word "mind" nowadays) is a product of heredity and environment, and that human behaviour, like plant and animal behaviour, is therefore in principle predictable from a knowledge of the conditioning factors. This product is, no doubt, an extremely complex structure, and it may be that it will never be possible to amass sufficient data on any one specimen to predict behaviour in detail with a high degree of probability (though we must beware of underestimating the contributions of our descendants during the 1,000,000 million years which appear to be the present life expectancy of the earth); but such knowledge is at least theoretically possible even now, and it is therefore no longer reasonable to regard "human nature" as a completely incalculable factor which will inevitably stultify any attempt to make the humanities respectable.

The mystic belief that there are dark corners in Language that defy analysis and can be dealt with only intuitively, through *Sprachgefühl* and total immersion in the *Volksseele,* may have been implanted in our linguists during the years of preoccupation with literature which is nearly everywhere an obligatory accompaniment to the study of language. As Bloomfield puts it, neatly and bitterly, "the discussion of literary values—that is, of the artistic use of language by specially gifted individuals—enjoys general favor as a substitute for the observation of language."[2] The study of literature, as it is traditionally practised, deals with aesthetic values and so encourages a subjective rather than an objective approach.[3]

[1] The recent attempt to introduce similar notions into atomic physics seems to have met with scant success outside the popular press.

[2] *International Encyclopedia of Unified Science,* I. 4, p. 5.

[3] To avoid misunderstanding I would hasten to add that I do not by any means despair of a scientific study of literature, and that I do not wish to be quoted as casting aspersions on the eminently respectable profession of literary criticism.

The argument can be presented in another form which gives it rather more cogency. Lévy-Bruhl found that the thinking of primitive people is not logical but what he calls "prelogical", characterised particularly by its indifference to the Principle of Contradiction: $a\bar{a}=O$, or "you can't eat your cake and have it."[1] This was an extremely important and fruitful discovery, the more so since Lévy-Bruhl was undoubtedly wrong in setting up a fundamental difference between "primitive" and "civilized" man, as a number of writers have pointed out: it is becoming increasingly clear that we are all brothers under our variously coloured skins, and that logical thinking, even among the "civilized", is rather like dancing among horses—a trick which can be taught to some but not to all individuals; which they perform laboriously, with varying degrees of skill; and which even the best of them cannot keep up for very long at a time. It is, in particular, now fairly well established that all languages—and not, as Lévy-Bruhl thought, only "primitive" languages—are based on this participative prelogic.[2] But if language is in itself "illogical", does not that preclude any study of it by scientific, *i.e.* logical methods? I do not think so, because, despite Lévy-Bruhl, the two are not entirely incommensurable; if they had been, indeed, it would hardly have been possible for him—a highly civilized man—to conceive his theory at all. There is nothing to prevent the construction of a theory to accommodate both. And it is, in any case, somewhat rash to believe that the scientific treatment must or does reflect the "real" structure of its object—if only because there is no way of finding out what this "real" structure is.

Toynbee, in his monumental work, *A Study of History*, presents a different argument: "We have empirical knowledge of three different methods of viewing and presenting the objects of our thought, and, among them, the phenomena of human life. The first method is the ascertainment and record of particular 'facts'; the second is the elucidation and formulation of general 'laws' through a process of comparative study; the third is the form of artistic creation and expression known as 'fiction'. We need not doubt that the clear distinction between the techniques of these three methods—a distinction of which we are empirically aware—corresponds to some equally clear distinction between the respective phenomena which are viewed and presented in these different ways. We are not bound, however, to accept without question either the names by

---

[1] Cf. *Les fonctions mentales dans les sociétés inférieures*, Paris, 1928.

[2] Cf. L. Hjelmslev, *Principes de grammaire générale*, Copenhagen, 1928, p. 257 ff., and *La Catégorie des cas*, Aarhus, 1935, p. 102 and *passim*.

which the three techniques are popularly known or the popular anatomy of their respective provinces."[1]

After showing that these three techniques are by no means mutually exclusive, he goes on to say, "...among other differences, they differ in their respective suitability for dealing with 'data' in different quantities. The ascertainment and record of particular facts is all that is either possible or necessary in a field of study where the 'data' happen to be few; the elucidation and formulation of general laws through a process of comparative study is both possible and necessary where the 'data' are too numerous to tabulate but not too numerous to survey. The form of artistic creation and expression known as 'fiction' is the only technique that either can be employed or is worth employing where the 'data' are innumerable.

"Here, as between the three techniques, we have an intrinsic difference of a quantitative order. The techniques differ intrinsically from one another in their utility for handling different quantities of data."[2]

Toynbee proceeds to conclude that history, with no more than 21 civilized societies to deal with, cannot reasonably be asked to do more than it is doing; that anthropology, which has some 650 known primitive societies at its disposal, is in just the right position to employ the scientific method; and that personal relations can be tackled only by the novel and the drama.

This argument, attractive though it seems, leaves Toynbee himself undaunted, since he proceeds forthwith to subject his 21 civilized societies to a process of comparative study with a view to the elucidation and formulation of general laws. And his deeds carry more conviction than his words.

We may concede that there are many more "primitive" than "civilized" societies about which something is known—and, since there are even more languages than primitive societies, the argument would certainly allow us to apply the scientific method in this field—but it has not been proved that the number of data relating to civilized societies is substantially less than the data to be collected from a study of the primitive societies. Other historians complain that they are snowed under with data, for which Toynbee takes them severely to task. A convincing case is made out for choosing what he calls a "society" as the structural unit (a "society" is, or may be, larger than a "nation") which is why he gets only

[1] Vol. I, 2nd ed., reprinted, Oxford, 1945, p. 441.
[2] *ibid.*, p. 452.

twenty-one. What is less immediately acceptable is his treatment of the time-dimension, where a different technique might well produce plenty of data: a society's entire life-cycle (in itself a doubtful analogy) of rise, growth, decline, and fall is not necessarily the most fruitful unit to use. Perhaps there is a complementarity so that, to get enough data, one has to work with smaller units either in time or in space. In any case, what is a datum, and how does one count data? Toynbee himself admits the vagueness of the term by putting it in inverted commas. And what about the natural sciences: have they really got more data than history and fewer than fiction? If the history of a society is truly analogous with the life of one individual, then it is clear that the biologists have the advantage; but to the extent that they have succeeded in reducing the universe to one unit, the physicists seem, on the other hand, to have placed themselves in a much worse position than the historians, with their twenty-one. In what sense can one compare data relating to the Roman Empire with data relating to the hydrogen atom or the eternal triangle? It seems obvious that there must be some connexion between the number, as well as the kind, of data found and the methods employed to find them, so that the comparison is not between three kinds of raw material—which could, in any case, hardly be compared since "raw"="unknown"—but between materials treated by different methods. It is therefore at least possible that the differences noted by Toynbee are inherent in the three methods rather than in the three kinds of material—if, indeed, there really are three kinds of material.

It is by no means certain, or even likely, that a scientific treatment would lead to a distinction between "civilized" and "primitive" societies: the method of science tends to replace differences of kind with differences of degree, and the dividing line between civilization and barbary, which is, at best, pretty vague, seems particularly open to attack. Even personal relations, which invariably reflect cultural patterns, find a place in this totality—in fact, cultural patterns can only be discovered through an analysis of personal relations: a society, whether "civilized" or "primitive", is nothing but a mass of personal relations. It is undoubtedly true that it will never be possible to record and analyse all personal relations, even within a very small society such as a village (and let us, in our private capacity, be thankful for that!) but in this as in other fields there must be a saturation point.[1]

[1] Incidentally, Toynbee quotes Aristotle's remark about the generality of "poetry" as opposed to the particularity of history, and accuses him of "confusing the technique of the Drama and

We may perhaps allow ourselves to derive from this discussion a measure of skepticism with regard to the opinion that certain materials require certain methods to the exclusion of all others.[1] We must, of course agree that generalisation on insufficient grounds is dangerous and apt to lead to false prophecies; we shall have occasion to discuss the validity of induction in general later on. But the difference between a hypothesis based on a large number of instances and one based on a few is only a difference in degree of probability, not a difference in kind. If there were only one society (in Toynbee's sense) available for study, it would still be possible (1) to ascertain and record the facts, (2) to elucidate and formulate (several sets of alternative) general laws, and (3) on the basis of these laws to construct novels and plays. I submit that the choice of technique does not depend on the nature—quantitative or qualitative—of the material but largely, if not entirely, on the nature of the investigator.

What is, in the last resort, the difference between the method of history, which is *par excellence* the method of the humanities, and the method of exact science? Is it simply, as Toynbee seems to imply, that while both ascertain and record their facts in the same way, the exact sciences, finding themselves in possession of a sufficient number of facts, go a step further to elucidate and formulate laws? Could one take the data of the humanities as they are now and make laws out of them, and would that in itself be enough to make the humanities exact?

That is precisely the task which Toynbee has set himself: he surveys the whole body of historical data and tries to find a pattern, a set of laws from which the past history of civilizations could be deduced, and on the basis of which it should be possible to foretell the future in broad outline. It is a bold and magnificent conception, brilliantly carried out in the six vol-

---

the Novel with the technique of Science in order to distinguish them both from the technique of 'History' (so called)". In spite of the prevailing wind it seems to me that Aristotle is entirely justified. Both science and fiction deal with generalities, with that which *can* happen as opposed to that which actually *has* happened. The difference between them is a matter of method of presentation: novelists and dramatists present their stuff syntagmatically, scientists theirs paradigmatically. But the scientist's system has been built from an observed sequence, and the novelist's or dramatist's synthetic sequence has been deduced from a system similarly built from an observed sequence, whether or not that system has been expleticised. Shakespeare's plays and Thackeray's novels contain much the same material as a treatise of psychology. To this may be added the further difference that, in our Western society, things are so organised that it is easier and more profitable to market bad fiction than bad science.

[1] Toynbee is evidently willing to go part of the way with us, since he says *(ibid.,* p. 456) "the differences between the objects of study and between the techniques are intrinsic, invariable, and absolute; the difference in quantities of 'data' is accidental, variable, and relative to the passage of Time." His quantitative criterion is thus not held to be epistemologically decisive.

umes so far published; it is a great methodological advance from the timid chronicling of traditional history; but it still does not look very much like exact science. What, then, is the difference?

Let us attack the problem from the other side and consider what science is and does. Its most striking characteristic, compared with other forms of cognition, is perhaps its great abstractness. The exact sciences do not deal with the whole mass of the observed universe but only with one aspect of it, *viz.* functions,[1] and only quantitative functions at that. To the scientific view the world does not consist of things, or even of "matter", but only of functions between things, the things themselves being regarded merely as points in which functions meet. "Matter" as such is completely ignored, so that the scientific conception of the world is a diagram rather than a picture. The prototype of all scientific statements is "*a* is greater than *b*"; about the *a* and the *b*, as *Dinge an sich,* science has nothing to say. This "greater than" can, of course, have a number of different references, but these are all interrelated in such a way that quality does not enter into the picture at all; as Susan Stebbing puts it (in *Philosophy and the Physicists*) "indeed, 'qualitative physics' may well seem to be a contradiction in terms."

We are accustomed, in ordinary daily life, to consider everything under three separate and distinct aspects: a "thing" exists, it has certain properties, and it performs certain activities. This Aristotelian trichotomy would seem, at first sight, to have a great deal in its favour; in fact, it seems almost to impose itself as a necessity: a thing, for instance a chair, must have some sort of existence[2] before any properties can be ascribed to it, and it must have properties before it can engage in any activities, since we conceive the activities to depend upon the properties. If the chair did not exist, it could not have the property of rigidity, and if it did not have this property, it could not perform the activity of supporting the sitter's weight. As Bertrand Russell has pointed out, these three categories correspond exactly with three of Aristotle's Parts of Speech: that which exists, the substance, the thing, is indicated by a substantive (which is thus rightly so called), the property by an adjective, and the

---

[1] The term "function" is used in the following to mean any dependence whatever, irrespective of its specific nature or the nature of its terminals. "Relation" might be more familiar in this sense but is here reserved for a more restricted purpose. Cf. Propositions 1. and 37. of the algebra, below.

[2] Even when the thing, *e.g.* a dragon, cannot be demonstrated to exist in the physical universe, a sort of courtesy existence is bestowed upon it within the universe of discourse.

activity by a verb, although a good deal of confusion has been caused by the widespread untidiness of constructing substantives from adjectives and verbs, and *vice versa*. In a number of modern languages, English for choice, the parts of speech are less rigidly organised than in Greek, but the Aristotelian scheme still seems reasonable and serviceable. Nevertheless, exact science has found it profitable to abandon this view or, rather, to abstract from it, for even the most rabid physicist could hardly put up with living in the kind of nightmare world conjured up by Eddington.

To begin with, the scientists have obliterated the boundary between property and activity: redness—the traditional example of a property—is explained as consisting in electromagnetic waves of a certain length, *i.e.* an activity. Secondly, what is to be understood by existence as divorced from property-activity? Aristotle imagined that each thing had a sort of ghost which was, so to speak, the possessor of its properties and the performer of its activities. Since his day a great deal of energy and ingenuity has been expended on the exploration of the nature of Being with a capital B—*das Sein*—without any very substantial results, and the pragmatical scientists, never having found anything in a pure state of being, refuse to have anything to do with it, particularly since nothing in science or in any other field of human endeavour (with the possible exception of theology) seems to make this concept necessary. The original three aspects are thus reduced to one, and this one is conceived in terms of functions. Apparent properties and activities are interpreted as functions, and something is deemed to exist if it is the terminal of a function.[1]

---

[1] In disposing of the Aristotelian bath-water we must be careful not to toss out a possible baby as well: there is in the scientific view something which can be regarded as a re-interpretation of Aristotle's "substance", *viz.* the concept of organism, or structure. If the "thing" is only a juncture of functions, it may still possess some individuality by virtue of a specific structure, an architecture of junction, which distinguishes one such organism from another.

The Aristotelian view seems to have, or to have had until recently, some currency among philosophers. Lalande's *Vocabulaire de la philosophie*, Paris, 1938, gives the following definition of *relation* in its logical sense: "si dans une proposition, telle que '*A* est fils de *B*', '*Q* est le quotient de *M* par *N*', on fait abstraction des termes considérés et qu'on n'envisage que la forme du lien qui les unit, celle-ci est appelée *relation*"; "relational propositions" which can be "decomposed" in this way, are distinguished from "predicative propositions", in which "le prédicat... est pensé comme une manière d'être du sujet". In other words, the subject represents a "thing", and the proposition says about it, either that it has an activity or dependence (relational proposition) or else that it has a property (predicative proposition).

In *The Principles of Mathematics* (1903) Bertrand Russell has a very similar division. He first defines "term" as follows: "whatever may be an object of thought, or may occur in any true or false proposition, or can be counted as *one*, I call a *term*"; next, terms are divided into "things", which are terms indicated by proper names (in Russell's special sense), and "con-

The methodological advantages to be derived from concentrating on function to the exclusion of everything else are far-reaching. In such a view the universe is homogeneous: all differences are differences of (measurable) degree, differences of kind vanish completely, and nothing is unique. This naturally makes the scientist's work very much easier than it would otherwise have been, and it permits a generalisation, a unification and simplification throughout the vast domain of exact science which is undreamt of in any other department of human endeavour. It would be premature to say that there is already only one exact science, but it is evident that the development is in that direction.

The humanities, on the other hand, cling to the common-sense, Aristotelian picture of the world. To the historian, the linguist, etc. the data to be ascertained and recorded are still "things", each with its properties and activities, and even when we get so far as to arrange our data in classes, these classes are defined by properties (Russell's "class-concepts") rather than by functions.

This has the corresponding disadvantages. "Things" are always unique: no two "things" are ever exactly alike. In such a view, therefore, the universe is heterogeneous, and all differences are differences of kind, which cannot be measured or even compared but can, in fact, only be ascertained and recorded in terms of concepts as *flüssig und offen* as possible.[1] It is an inevitable consequence of this method that no systematisation is possible, and that such far-reaching generalisations as have been achieved in the exact sciences are entirely out of the question.

A corollary of insisting on the Aristotelian "thing" is that no humanistic discipline is able to achieve autonomy. In linguistics, for instance, we

---

cepts", the terms indicated by all other words; finally, concepts are divided into those indicated by adjectives, which are called "predicates" or "class-concepts", and those indicated by verbs, which "are always or almost always relations". Note that Russell uses linguistic criteria for his classification, though he is obviously not satisfied that they will work in all cases. It is a pity that philosophers *will* pounce upon the Parts of Speech, which is one of the most doubtful legacies from classical grammar. In a later book, *An Inquiry into Meaning and Truth* (1940) Russell states that he holds "things" to be a metaphysical delusion (p. 320) and that "wherever there is, for common sense, a "thing" having the quality *C*, we should say, instead, that *C* itself exists in that place, and that the "thing" is to be replaced by the collection of qualities existing in the place in question. Thus "*C*" becomes a name, not a predicate" (p. 98) In *A History of Western Philosophy* (1945) he says of the Aristotelian concept of "substance" that it is "a metaphysical mistake, due to transference to the world-structure of the structure of sentences composed of a subject and a predicate" (p. 202).

[1] Toynbee, as he says, "returns a soft answer" to those who maintain that "history does not repeat itself" and that all historical "facts" are unique; he ends up in a position intermediate between that of traditional history and that of the exact sciences *(op. cit.,* I, p. 178f.).

have had to put up with a confused jumble of historical, physiological, physical, biological, logical, psychological, philosophical, statistical, and even, on occasion, theological ideas. An autonomous discipline can be built up only by resignation, by being willing to do one thing at a time, by a rigid selection of a set of functions as necessary and sufficient for unambiguous description, *i.e.* by abstraction. You cannot make a map if you insist on bringing in all the hills, valleys, houses, and trees in life size and complete to the last wood-louse.

We have here, I believe, the fundamental difference between the exact sciences and the humanities: the humanities do not analyse their data, or if they do, the analysis is not pushed nearly so far as in the exact sciences; it does not go beyond the "thing" as a unit.

Given that this fundamental difference exists, the question arises whether it would not be possible to effect a *rapprochement* between the two groups of disciplines, thereby making their results commensurable and achieving epistemological unity. Since it would be preposterous to ask the exact sciences to put the clock back, this could only take the form of the humanities giving up "things" in favour of functions, and thus, as I contend, becoming exact.[1] The questions to be answered are, then, whether this is possible and, if so, whether it would be desirable.

The argument that the material of the humanities is unsuitable for treatment by exact methods we may now, perhaps, dismiss. We have examined three pleas in this cause: Burckhardt's that his material is too unstable, Gray's that the "human factor" is too strong, and Toynbee's that his data are insufficient; we found them all inconclusive. The same may be said of another frequently advanced argument, which is the opposite of Toynbee's, and which he naturally does not accept: that the material is prohibitively rich and complex. It is a curious notion that social and linguistic systems, which all normal people are able to master practically, should be too complicated for a scientist to unravel theoretically.

It remains to examine one more possible hindrance, *viz.* the curious fact that the humanities so to speak contain themselves: form part of their own material. The linguist speaks and writes in a language—surprisingly rarely in more than one—thus creating a text which becomes part of the material of his own subject; the sociologist lives in a society, and his life and work thereby become part and parcel of his material;

[1] A "functional" anthropology and a "structural" linguistics already exist, but, like Toynbee's history, they can neither of them be said to have gone the whole hog.

the historian is himself caught up in the stream of history. How enviable, by contrast, is the aloofness of the chemist from his test-tubes, of the geneticist from his fruit-flies! Perhaps it is the quest for an Archimedian point that has attracted so many linguists to exotic languages and so many anthropologists to *ultima Thule*. But the chemist is himself composed of chemicals, and the geneticist is himself the product of genes similar to those under his microscope: they are as much part of their own material as the humanist is of his, and their observation of the material has a disturbing effect, as we now know, comparable to that which may be produced by the linguist's microphone and the camera and notebook of the anthropologist. Far from being a scientific *sine-qua-non,* the Archimedian point thus turns out to be nothing more than a pipe-dream. It is deplorable that we have not got it, but we can at least be comforted by the reflexion that we are no worse off than the others and that they have succeeded in being tolerably exact in spite of this disability.

It should, then, be possible—at least in theory—to make the humanities exact, but is it desirable? Many people do not think so, and for a variety of reasons.

Apart from the motive of vested interests, which is always the same and need not occupy us beyond the mere mention, there is first of all human vanity. If "things" are to be eliminated, it follows that man, who is eminently a "thing"—in fact, the prototype of a "thing"—will be eliminated also. It is a most disagreeable thought that one should be subjected to the indignity of giving up one's individuality, one's Aristotelian ghost, to become a mere meeting-point of abstract functions; that one's civic dignity should be reduced to an algebraic formula. There is altogether something degrading about submitting to analysis; it was probably this feeling, as much as religious prejudice, that prevented the dissection of human cadavers for so long, and a cold analysis of human behaviour seems an even more alarming project, liable to bring to light much that is better left shrouded in decent obscurity. Cf. the reaction to the *Kinsey Report*.

This attitude is somewhat akin to that of the mystics. As Hilton Brown puts it, in his book on Kipling, "...it is no part of the mystic's profession to put things down in black and white; he is never anxious to explain his thoughts to others because he is never anxious to explain his thoughts to himself. If all could be made lucid, aboveboard, straightforward and explicit, something precious would have gone out of it, he feels, which could hardly be restored." It is no good arguing that a scientific analysis does not necessarily destroy its object—that a poem or a love-affair or a religious ceremony can be subjected to such an analysis and still be as good as new: "something precious has gone out of it", and that is that.

If the egocentrics and the mystics object to science as spreading too harsh a light, there are others who complain that science is, on the contrary, too obscure, too unreal. The scientists, they say, devour the familiar, tangible world and leave behind them nothing but a skein of abstract functions, a mathematical cobweb at once abstruse and ethereal, which is no good to anyone else. It is no wonder if a good many people feel like that about it. The scientific picture of the world is, as we have said, only a diagram, curiously thin and unsubstantial and therefore without appeal to those who are temperamentally debarred from seeing the abstract beauty which others have found in it. Mathematic formulae are cold comfort, and a handful of functions may well seem a poor substitute for the acquaintance with solid matter obtainable by other means. People who feel like that about it may well be horrified at the thought of seeing the frontiers of science pushed forward and the humanities invaded by this inhuman method, particularly those—and they seem to be many—who have been driven into the humanities by a dislike for mathematics. The future looks grim for those who are by temperament anti-scientific; we may sympathise with them, but we cannot for their sake give up an experiment which, if it succeeds, will benefit even them in the long run.

Another serious complaint against science is the uncertainty of its results. A well-documented historical fact, such as the date of the battle of Waterloo, seems about as near to absolute truth as it is possible to get, but scientific facts, if they can be called "facts" at all, are not only known to be inaccurate—or approximative, to be more polite—but are furthermore hypothetical. Science, as J. W. N. Sullivan puts it, has adopted the pragmatical criterion of truth, *viz.* success: a theory is acknowledged as true as long as it is the most successful one available; the moment a better one is constructed, the earlier theory is brutally discarded. This makes the whole structure very insecure. Any post may bring the scientist news that everything has now been changed and that nothing, or very little, remains of the truths of yesterday. The thoroughbred scientist takes that sort of thing in his stride, but to many people such insecurity is intolerable.

It is, of course, a delusion that the results obtained by ascertaining and recording are more secure than those reached by the elucidation and formulation of general laws, if only because the very act of ascertaining implies a theory—a particular way of "selecting and grouping in attention."[1] Nor is it necessary to hold that there are two kinds of truth, historical and scientific. A historical "fact", such as the date of the battle of Waterloo, is most usefully regarded as a hypothesis in exactly the same way as any scientific "fact", such as the composition of water, H_2O, a hypothesis which could and would be discarded in favour of another if it should be found to be a misfit in the structure of hypotheses to which it belongs. We go on believing that the battle of Waterloo took place on the 18th of June, 1815, for no other reason than that this is the hypothesis which fits in best so far; a new calendar, or another way of looking at battles, would call for another hypothesis. For "the battle of Waterloo" is itself a construct, the result of a particular "selecting and grouping in attention" on the part of the historian, and not something "given" in an absolute sense. But as long as so much in the humanities depends on opinion, it is probably true that it is easier here than in the sciences to become and to remain an "authority".

The case *for* the exact sciences might be put, briefly, in this way: firstly, the selection of functions to the exclusion of all other aspects of the universe has made it possible to give one comparatively simple explanation of an enormous mass of details that would otherwise have appeared unconnected. This is, if you like, a merely aesthetic achievement, an appeal to an intellectual sense of order which is not highly developed in all individuals; but it is an achievement which is highly valued by some people. Secondly, it works. It has been demonstrated that a knowledge of quantitative functions alone is in itself enough to give the scientist a control over his material which has not been equalled by any other method. It is true that these advantages accrue in full measure only to physics; in biology the method has, so far, proved less fruitful than had been expected, and it seems likely that, as Whitehead suggested, the biologists will have to widen their scope. Still, there is no getting away from the fact that the method is brilliantly successful.

To this we may add another argument. Each human being has his own particular way of looking at the universe, his own particular way of "selecting and grouping in attention"; each of us repeats, in each moment of life, the process described in Genesis of dividing light and darkness, land and water. But co-operation among human beings is possible only when, and to the extent that, these different pictures of the universe coincide.

---

[1] Cf. Angus Sinclair, *The Conditions of Knowing*, London, 1951.

Now science is, among other things, a body of rules for "selecting and grouping in attention", and perhaps its principal social benefit is that it enables those who agree to abide by these rules to work together more closely than under any other comparable set of rules. The results obtained by one scientist can be immediately comprehended and used by all the others, and the chances of misunderstanding are reduced to a minimum. In this respect, too, the humanities are far behind.

Our inquiry so far seems to have shown that the method of the exact sciences leads to greater and more uniform control than those hitherto employed in the humanities, and that there is insufficient ground for believing that the material of the humanities is inherently unsuitable for the application of a similarly exact method. It therefore seems an inescapable duty to make the experiment, and to make it in full scale. Only in this way can we find out whether it is possible to make the humanities exact and, if so, what will come out of it. If the experiment fails, a lot of people will have the satisfaction of saying "I told you so", but otherwise nothing will be lost; if it succeeds, much will be gained.

Let us, then, try to visualise what such a change of method would involve and what an exact humanistic science—or exact humanistic sciences—would be like.

The natural sciences, particularly physics, are based on quantitative functions, as we have noted, and, in fact, owe their existence to mathematics. Sullivan goes so far as to say that "the original elements in modern scientific thought, the new way of thinking, have come from the mathematical sciences. It is in these sciences that common sense has been found most inadequate. The other sciences, as chemistry and biology, have done relatively little in the way of making us acquainted with radically new ideas. It is true that there are some biologists who find the current common-sense outlook inadequate, but a satisfactory set of new ideas has not yet been evolved. The present break-away from long-established habits of thought owes practically nothing to the non-mathematical sciences. Their cultural value is to be found in their facts rather than in their principles."[1] Our first task, therefore, is to inquire whether mathematics is a possible basis for our projected new science.

Toynbee, for one, does not think so: "In the world of action, we know that it is disastrous to treat animals or human beings as though they were stocks and stones. Why should we suppose this treatment to be any less

[1] *Limitations of Science*, Pelican edition, 1938, p. 240.

mistaken in the world of ideas? Why should we suppose that the scientific method of thought—a method which has been devised for thinking about Inanimate Nature—should be applicable to historical thought, which is a study of living creatures and indeed of human beings? . . . We are sufficiently on our guard against the so-called 'Pathetic Fallacy' of imaginatively endowing inanimate objects with life. We now fall victims to the inverse 'Apathetic Fallacy' of treating living creatures as though they were inanimate."[1] This argument occurs in the course of an attack on "the industrialization of historical thought", but it is relevant to our discussion here.

It may be that the conclusion is correct—and certainly, on the evidence available, we could not go so far as to say that it is wrong—but surely the premise of the fundamental difference between inanimate nature and human beings is out of order: it is just the kind of reasoning that Toynbee himself refuses to accept in connexion with "the comparability of the 'facts' encountered in the study of civilizations". If human beings have anything in common with stocks and stones—and why should we suppose that they have not?—is it not most likely to be precisely those quantitative functions which the method of the natural sciences is equipped to deal with? It would obviously be dangerous to assume *a priori* that human social institutions, including language, have no mathematical aspects at all; and if they have mathematical aspects, then those aspects are, in the light of previous experience, best treated mathematically. This Toynbee would, I imagine, be the last to deny; what he must mean is that the essential part of historical thought cannot be mathematical in character. The other things he would relegate to ancillary sciences.

Given that our new science, or sciences, must be limited to functions in order to be exact, the question really is whether control of the kind we envisage can be obtained by a purely quantitative method, *i.e.* whether the functions we have to deal with are quantitative or not.

We are now once again faced with the question whether there is any connexion between the specific nature of a material and the methods applicable to it, though this time the problem presents itself in a different form. It is clear that we must be prepared to find quantitative functions within our material; a beginning has already been made with the study of some of these functions by such disciplines as economics, political science, and, in linguistics, phonometrics, experimental phonetics, and word

---

[1] *Op. cit.*, I, pp. 7—8.

counting. On the other hand it is equally clear that the functions traditionally—and inexactly—studied by history, anthropology, sociology, grammar, etc. are, for the most part, non-quantitative. Although we can therefore not assume that the universe is sharply divided into two parts, one quantitative and the other non-quantitative, yet it may be that there are two inherently different kinds of material, as Toynbee seems to assume: inanimate nature, with a preponderance of quantitative functions, and social phenomena, with a higher proportion of non-quantitative functions. Perhaps there are more than two: the material of biology, for instance, might turn out to have yet another composition.

Very little seems to be known about non-quantitative functions in inanimate nature, perhaps even less than about quantitative functions in language and society, and certainly not enough to prove or disprove the hypothesis tentatively set up above. The natural scientists have confined themselves to quantitative functions, partly because the development of mathematics invited them to do so, and partly, no doubt, because this method has given them the control they want. Whether they have been right in supposing—or acting as if they supposed—that they were thereby exhausting the usefulness of inanimate nature, we have no present means of knowing. However that may be, it is at least possible that the determining factor is the kind of control desired rather than the kind of material to be dealt with. In that case, a non-quantitative treatment of the same material—inanimate nature—would be possible and should lead to another kind of control, which may or may not be worth having. If we can hold as a hypothesis that the choice of quantitative or non-quantitative functions is entirely a matter of the kind of control desired, then it is no longer necessary to assume that the universe is heterogeneous, which seems an advantage on the ground of simplicity. On this hypothesis there would then be two basic sciences, one quantitative and the other non-quantitative, each with the whole of the universe as its field of inquiry.

The existing quantitative studies of the material of the humanities have led to the kind of control expected and desired: predictability of the composition of a given population at a given time in the future, ability to calculate the size of a printed book, or the time needed for a speech, from a manuscript, etc. But the scope of this control is very limited and does not offer possibilities for the development of technologies comparable to those based on the natural sciences. It would appear that ability to predict what changes will, or can be made to, take place in a given society or language, or to calculate what would be the effect of any such change, cannot be obtained by a study of quantitative functions alone. And that is, of course, the kind of control it would be useful to have. In trying to estimate the effect of, say, a new divorce-law, the statesman must take into consideration important non-quantitative functions, which are at present imponderable, although statistics may enter into his calculations as well. Similarly, a linguist designing an artificial language or a new scientific notation, or a judge construing the law (which is also a linguistic activity) is only peripherally concerned with quantitative functions. Further, when a linguist does count or measure, that which he counts or measures is not itself quantitatively defined: the words that the word-counters count are defined, in so far as they are defined at all, in quite different terms.

We have, then, two alternative hypotheses: (1) that the universe is divided into two or more parts differing in respect of the proportion of quantitative to non-quantitative functions, and (2) that the universe is homogeneous, the choice of quantitative or non-quantitative functions as terms of description depending upon the type of control desired. Of these, number two has the advantage of simplicity and seems, on the whole, more promising. Whichever we choose to test, there seems no other way of doing it than by making the experiment of supplementing the natural sciences by a study of non-quantitative functions within their material, and the humanities by a study of both quantitative and non-quantitative functions within theirs. The authors propose, in this work, to turn a modest first sod by attempting to sketch the outline of a non-quantitative science to be applied, in the first instance, to the material of the humanities, particularly language. It is this proposed new science that we call *glossematics*.

As the natural sciences are based on mathematics, so must glossematics be based on a theory of non-quantitative functions. We have noted that the prototype of statements in physics is "*a* is greater than *b*"; the corresponding statement in glossematics is "*a* presupposes *b*", which in a similar way serves to establish an ordered dependence of great generality. This ordered dependence, in glossematics as in physics, occurs in a variety of forms, and the glossematic theory of functions is a calculus of modes of ordered dependence, in which the primitive idea is developed into an algebra. The algebra has been designed as a means of describing a (humanistic) material as a structure of non-quantitative functions complete in itself without the necessity for bringing in definitions derived from other sciences. We have therefore been at pains to develop the algebra to the point of supplying the means of differentiation necessary and sufficient for unambiguous description.

The glossematic algebra owes much to symbolic logic but differs from it, as we shall see, in several respects. These differences, which it has not been found possible to eliminate, appear to be due to the different purposes and different starting-points of the two algebras. Symbolic logic is concerned with the interrelations of classes and of propositions which may be true or false; and the logicians take their propositions, classes, and members of classes for granted without bothering about where they come from. The material of symbolic logic is thus open, unlimited; and the logical approach presupposes an atomistic view of the universe or a preceding analysis which is outside logic itself. Glossematic algebra is concerned with closed structures and presupposes a coherent material the analysis of which is an integral part of glossematics itself; it is not designed to deal with propositions, or with truth and falsehood, and classes emerge, not at the outset, but only after the analysis of the material has supplied something to classify. It should, however, be possible to construct a metatheory from which both can be deduced, and as mathematics is derivable from symbolic logic, all functions—quantitative and non-quantitative—would then be brought under one hat.

The algebra is presented in this book in the form of a series of "propositions" in ordered dependence so that definitions presuppose the primitive propositions in which their defining terms are introduced. Now there are many ways of building up such a system, and many possible propositions for each term; it is, for example, a matter of comparatively free choice which elementary terms are to be left as primitive and which defined, since the whole system serves to define the formally undefined terms. We have tried out a large number of different arrangements with a view to obtaining the greatest possible simplicity and self-consistency, and some of the terms and propositions presented here are therefore different from the earlier versions that have appeared from time to time in other publications. We apologise for any inconvenience caused by this inconsistency and hope to have made up for it by the greater stringency of the final result, which should form a better basis for continued discussion than any of the earlier attempts.

Before going on to a detailed presentation of glossematics it will be convenient to make a rapid survey, in non-technical terms, beginning with the principles on which the theory is founded. This should serve the double purpose of showing the reader at once what sort of theory it is—what we propose to do and, in general outline, how we propose to do it—and of enabling him to see the wood when, later on, he is among the trees. No more than a sketch is intended at this point, and a good deal will have to be taken on trust until the formal exposition can be brought to bear.

The ideal of all scientific description is simplicity. There does not seem to be any absolute necessity why this should be so, and the appeal of simplicity is probably, in the last resort, aesthetic: a simple explanation is just more pleasing than a more complicated one—though evidently not to all of the people all of the time. There is something else that makes simplicity desirable: as it is nearly always possible to think of more than one way of describing anything, and as it is often possible to think of several descriptions that will do equally well, it is necessary to have some criterion for deciding which to choose, and it must, furthermore, be a criterion that can be accepted and applied by everybody. The criterion which has been adopted is simplicity: other things being equal, the simplest possible description is preferred. As a criterion simplicity has the advantage of being objective, and although it is not always easy to apply,[1] it is probably easier than its opposite, the greatest possible complexity, would be.

From simplicity can be derived all the other scientific ideals: objectivity, self-consistency, exhaustiveness. An objective description is simpler than a subjective one because it does not presuppose the personal prejudices or private experiences which enter into subjectivity; an objective description involves only that part of human experience which is, or can be made, accessible to all. A self-consistent description is simpler than one which is not, because contradiction implies more than one set of basic ideas. And an exhaustive description is simpler than one which is not, because any remainder potentially contains and conceals contradiction.

To make these ideals operational we have embodied them in a set of principles, which we have tried to obey in the construction of our theory, and by which we wish our work to be judged.[2]

I. *The Principle of Empiricism:* the description must be self-consistent, exhaustive, and the simplest possible. The three parts of the principle

[1] Cf. H. Spang-Hanssen, *On the Simplicity of Descriptions* in *Recherches structurales,* this series, vol. V, 1949.

[2] These principles were first published in Hjelmslev's *Omkring Sprogteoriens Grundlæggelse* (1943) = *Prolegomena to A Theory of Language* (1953).

are in ordered dependence so that exhaustiveness is subordinate to self-consistency, and simplicity to exhaustiveness.

This principle is, among other things, a definition of scientific truth, which, as we have noted, is somewhat different from ordinary civil truth. The two questions, "is it true that George has bought a house?" and "is it true that the ozone molecule contains three oxygen atoms?" have two different variants of "true"; the first one refers to historical "fact", but the second one is not concerned with fact[1]—it can have no other scientific interpretation than "is O_3 the simplest possible self-consistent and exhaustive description of what is called 'ozone'?" As we have noted earlier, it is possible to generalise the scientific conception of truth, to regard George's purchase of a house as a self-consistent, exhaustive, simple description rather than as a "fact", but that is perhaps not likely to become common practice, although there are considerable advantages to be derived from discussing descriptions instead of "facts".

A scientist may privately believe in absolute truth if he so chooses, *i.e.* he may believe that "God is a mathematician", that the universe possesses an inherent structure which it is the task of science to unveil, and that each scientific discovery is a step forward on the long but finite road; as long as he does not identify any actual description with absolute truth, that is entirely a matter for his own conscience. Or he may believe that there is no absolute truth, no inherent structure, and that science is a projection of the human mind—whatever he may mean by that—on to chaos; in that case he will regard the history of science as an infinite progression of increasingly simple working arrangements of chaos—a sort of endless repetition of the Six Days of Creation. The Principle of Empiricism can be entertained together with either faith, which explains how it is possible for scientists of widely differing creeds to go on working together.

The traditional humanities, dedicated as they are to historical rather than scientific truth, have hitherto paid scant attention to the Principle of Empiricism. Instead, they have set up for themselves what might be called the Principle of Plausibility, though it is rarely made explicit: of two otherwise equally satisfactory explanations the more plausible one shall be preferred. For this they cannot be blamed, as there is hardly any other way open to them. But if we are to establish an exact science, not only must the Principle of Empiricism be adopted, but the Principle of

[1] Leaving out of consideration the possible equivalent "does the textbook say that...?"

Plausibility must be given up. In discussions about the humanities the argument that a hypothesis is "far-fetched" is often enough to kill it on the spot; this would presumably not be so in the case of, say, a physical hypothesis, at least not with a common-sense interpretation of "far-fetched". For the exact sciences are, on the whole, not plausible and never go out of their way to try to be; their explanations are often, from the point of view of common sense, wildly fantastic: cf. most of atomic physics, or even such an old-established theory as that water is made up of two gasses—who, in the humanities, would dare to put forward an explanation as preposterous as that?[1] This difference between the two sets of disciplines is directly traceable to the fundamental difference which we have observed: the Principle of Empiricism goes with functions, the Principle of Plausibility with "things". But even at this level plausibility is a dangerous criterion, for, as Lévy-Bruhl says, "la première règle d'une méthode prudente n'est-elle pas de ne jamais prendre pour démontré ce qui n'est que vraisemblable? Tant d'expériences ont averti les savants que le vraisemblable est rarement le vrai!"[2]

The Principle of Empiricism applies on several levels, to some extent with conflicting results, because the term "description" can be interpreted as referring to either a particular description of a particular object or the general descriptive apparatus. A description of a particular object is exhaustive if it has been carried through until, within the scope of the method, there is no remainder, *i.e.* until the whole of the object has been reduced to a structure of the kind envisaged; it is simple if it describes the object as consisting of as few final resultants as possible while remaining self-consistent and exhaustive. The final resultants are a kind of inevitable remainders, which mark the ultimate limits of the scope of the method: their external functions are known, and serve to differentiate and define them, but as they are by definition unanalysable, nothing can be known of their internal structures, and the Principle therefore demands that they should be reduced to the smallest possible number.

The algebra is a general description from which all particular descriptions, actual or potential, must be deducible. The ultimate test of its

[1] Here is Bernard Shaw on this subject: "In the Middle Ages people believed that the earth was flat, for which they had at least the evidence of their senses: we believe it to be round, not because as many as one percent of us could give the physical reasons for so quaint a belief, but because modern science has convinced us that nothing that is obvious is true, and that everything that is magical, improbable, extraordinary, gigantic, microscopic, heartless, or outrageous is scientific." From the Preface to *St. Joan.*

[2] *Les fonctions mentales dans les sociétés inférieures*, Paris, 1928, p. 11.

exhaustiveness must be by induction from all the particular descriptions deduced from it, but, to the extent that it is general, this induction necessarily remains incomplete, and as the particular descriptions presuppose the algebra, which must therefore be constructed in advance, its exhaustiveness can only be estimated as a probability. The same applies to its simplicity, which must ultimately be tested inductively, by the simplicity of the particular descriptions deduced from it.

Now it will always be possible to simplify any algebra at the expense of its applicability, and from any algebra thus reduced a limited number of particular descriptions can be deduced which are individually simpler than the corresponding particular descriptions deducible from the more general algebra. In other words, any one material, *e. g.* any one language, can be described in a very simple way if the descriptive apparatus, the algebra, is adapted to that one purpose alone; if it is desired, on the other hand, to give uniform descriptions of more than one material, *e.g.* of more than one language, then the descriptive apparatus, and hence each particular description, is likely to be less simple. The reason is obvious: particular descriptions differ as to degree of complexity, and the descriptive apparatus must be equipped to deal with the highest degree of complexity that can be foreseen to come within its scope. As the glossematic algebra is designed to be general—in part, universal—it cannot be claimed that any particular glossematic description is the simplest possible self-consistent, exhaustive description of its object. The compensation for this sacrifice of particular simplicity is the gain in general simplicity which results from getting a large number of uniform descriptions.

The construction of a descriptive apparatus is thus beset with the struggle between two conflicting desires: (1) to make the algebra as general as possible, applicable to as wide a diversity of particular descriptions as possible, which means increasing its power of differentiation and thus its complexity; and (2) to ensure the greatest possible simplicity of particular descriptions. The result is a compromise which must be unendingly tested and revised.

Besides the final resultants of particular descriptions there is another group of unknown elements that must be kept down to a minimum, *viz.* the undefined terms of the algebra. Here, again, we know their external functions, *i.e.* the rôle of these terms in the algebra, but not their internal structures, and the self-consistency of the theory is therefore threatened by an excessive number of undefined terms. An algebra must be constructed with one eye upon the metatheory from which it will eventually be deducible and in relation to which it is, in its turn, a particular description. The greater the number of undefined terms, the greater the burden thrown upon the metatheory, and the greater the probability that the algebra, and all its particular descriptions with it, will have to be revised to make the metatheory self-consistent, exhaustive, and simple.

As we have seen, the purpose of a science is to describe and differentiate the greatest possible number of objects in the simplest possible way, *i.e.* in terms of the smallest possible number of final resultants. This is done by analysis and reduction, *i.e.* by describing the object as composed of a small number of elements each of which may occur a large number of times, and all of which are held together by functions. The classical example of this kind of description is, of course, chemistry. In order to employ this method it is necessary to adopt two working hypotheses: (1) that the object is analysable, and (2) that its components, as found by analysis, can be arranged in a finite number of classes.

(1) This hypothesis is so obvious that probably most of the people who work on that assumption have never bothered to formulate it. It is, however, by no means superfluous to do so, because complete objectivity can only be reached when all assumptions are made explicit. It is this hypothesis that the mystics refuse to subscribe to.

(2) The second working hypothesis may seem to be superfluous and may even be suspected of being designed to pull the wool over the reader's eyes, since the glossematician appears to be free to decide for himself how many classes he will create: the number of classes is surely determined by the algebra, which he has himself made for the purpose. But the glossematician is bound by his principle of Empiricism; if he makes the algebra too narrow in order to get a small number of classes, he runs the risk that it will not furnish exhaustive descriptions, or that it will not be applicable to all the objects he wishes to describe. And in any case, however the algebra is constructed, it remains an assumption that any given object will prove to be self-consistently and exhaustively describable by means of the number of classes provided by the algebra, or by means of any finite number of classes at all. If the classes provided by the algebra prove insufficient, then the algebra stands condemned as unsuitable for the description of that object; if an object should be encountered which cannot be described in terms of a finite number of classes, then science itself is at the end of its tether, since a description in terms of infinite diversity is manifestly impossible. It will be seen that it is necessary to

hold the hypothesis in order to make the attempt at all. It is this hypothesis that the history-does-not-repeat-itself and the human-factor schools refuse to subscribe to.

The description, then, takes the form of a gradual division of the object into smaller and smaller components and a progressive reduction of the number of different components by classification. This is what we call a *procedure*. The procedure is articulated as a series of *operations* in ordered dependence, an operation being a description in accordance with the Principle of Empiricism. Each operation is a description of the results of the preceding operation, and as each operation comprises both analysis and synthesis (classification) this gives rise to both a deduction, *i.e.* a series of analyses in ordered dependence, and an induction, *i.e.* a series of syntheses in ordered dependence.

Now it may happen that it is possible to describe a particular object by means of two or more procedures that lead to equally simple results. In such a case we are bound by the Principle of Empiricism to prefer the simplest procedure, *i.e.* the procedure which has the smallest number of operations. These considerations are summed up in the following Principle:

II. *The Principle of Simplicity:* of two self-consistent and exhaustive descriptions the one that gives the simpler result is preferred. Of two self-consistent and exhaustive descriptions giving equally simple results the one that requires the simpler procedure is preferred.

The synthesis which is carried out in each operation is, of course, a classification in which the "class-concepts" are functions, not "properties". An analysis is the registration of a function and of its terminals, and the synthesis consists in classing all those components together which can be terminals of the same function(s). This leads to a reduction of the number of different components, since all those which are members of all the same relevant classes, no one of them being a member of any relevant class of which all the others are not also members, *i.e.* all those which are equivalent in respect of all relevant functions, are declared structurally identical: separate instances of one and the same element. In this way, and to the extent that our working hypotheses are verified, the number of different components is reduced in each successive operation.

As operations are defined as being in accordance with the Principle of Empiricism, each operation must be exhaustive, *i.e.* it must be continued and, if necessary, repeated until all the results of the preceding operation have been dealt with. This means that in each operation an attempt must be made to analyse each element taken over as a resultant of the preceding operation, and that in each operation an attempt must be made to synthesise the classes handed down from the preceding operation. When the resultants of an operation are heterogeneous, *i.e.* of varying structural complexity, the attempt to analyse them in the following operation will not in all cases be successful. Elements which thus resist analysis are carried over to the next operation.

When the components of an object are heterogeneous, operations must be repeated as many times as may be necessary. This is the case, for example, when the object of the description is a language. The theory leads to the description of a text as composed of four separate but connected parts, or *strata, viz.* the two parts conventionally called *content* and *expression,* each of which consists of two strata which, in the terminology taken over from F. de Saussure, are called *form* and *substance.*

These four are certainly different in kind, from the common-sense point of view: the two central strata, content form and expression form, are "linguistic forms", *i.e.* abstractions, which have never been described in any other terms than functions, and which have often been denied any existence at all. The expression substance varies—it may be speech-sounds, which have been described both physiologically and physically, it may be writing of various kinds, dots and dashes, signal flags, buzzing noises, flashes of light, etc., even dancing;[1] each of them can be described from some non-linguistic point of view. It is this extreme variability that has commended to linguists the opinion that there is a comparatively loose connexion between the central part of language and the expression substance, although it is usually, but not necessarily or always, the case that a change of expression substance goes with a change of expression form, so that the connexion between expression form and content form would appear to be equally loose. By analogy this view is extended to the content substance. This is the most mysterious of the four because the most *mal étudié.* Is it properly a psychological object?—surely no more than the other three: Sapir and Trubetzkoy as well as Saussure and Baudouin de Courtenay have taught us that the expression, too, has a psychological aspect. Well, what then?—philosophical? physiological again (structure of the brain)? or doesn't it exist at all? My own feeling, on this level, is that a strong case can be made out for describing the

[1] Cf. G. K. Chesterton's story, "The Noticeable Conduct of Professor Chadd" in *The Club of Queer Trades.*

content substance as a sort of ethnic philosophy, a *Weltanschauung,* a "climate of opinion", a set of hypotheses or attitudes or beliefs about epistemology, ethics, economics, religion, manners, politics, geography, history, mathematics, the sciences, music, art—the whole of the area which used to be the preserve of philosophy. Otherwise expressed, it is that "culture" which is said to be what is left when you have forgotten what you learnt at school. A court of law is concerned with deciding whether a given extra-linguistic event does or does not fall within "the meaning of the act", *i.e.* with defining the content substance of the words and phrases of the law; every time we open our mouths each of us makes judgements of precisely the same kind—"the meaning of the act" being, now, the set of opinions held by the speaker. A simple utterance like "the dog is asleep" presupposes a whole string of such judgements: that the animal "is" a dog; that it "is" asleep; that sleep is a state which can be entered into by a class of objects including dogs; that the speaker is able and entitled to make statements of this sort; that there is a sufficient reason for making the statement; that the universe is so constituted that "dogs" and "asleep" are reasonable classes to operate with; etc., etc. All these judgements presuppose a body of opinion, and it is this body of opinion which constitutes the content substance. There is, naturally, considerable variation from one group or individual to another, just as, at the other end of the spectrum, there is considerable variation in the sets of speech-sounds employed by different groups or individuals. There may even be variation, in both strata, within the usage of one individual, who may think scientifically at one moment and commonsensically at another, with different contents to his words, just as he may use a set of sounds different from his normal ones for special occasions or just for fun—what Elizabeth Uldall aptly calls "phonetic slang"; think of Sapir's Nootka. This line of enquiry has been opened up by Angus Sinclair in his book, *The Conditions of Knowing,* London, 1951, though—as he would be the first to admit—a great deal more work needs to be done, particularly on the content substances of languages outside the Atlantic group of cultures.

But all these considerations do not concern the glossematic description of languages at all. In the structural view, which we are bound by the Principle of Empiricism to adopt, there are no differences of kind. The reason for describing a text as consisting of four strata is the purely formal one that the components of one stratum cannot be found by analysis of components of any of the others; the strata are, in other words, not mutually conform. A text must therefore be described by means of four separate deductions and inductions—for only the resultants of one deduction can be synthesised in one induction—*i.e.* operations must be carried out four times. The first few operations of the procedure may present no difficulties, the resultants of each operation lending themselves to analysis in the next without the need for repetition. But sooner or later—how soon depends upon the structure of the text—a point will be reached from which it is necessary to carry on in separate deductions; this point is reached in the operation where, for the first time, the units of content and expression, or of form and substance, do not coincide. The larger units of a text (paragraphs, periods, sentences) often coincide in all four strata, though this is not necessarily or always the case; the operations dealing with these levels are therefore not repeated, and the strata are not separated. From there onward—or earlier or later as the case may be—separate deductions have to be employed. In English, for example, the unit "I saw him when he came in" consists of two content-nexus but can be, and usually is, pronounced so as to consist of only one expression-nexus (*i.e.* with one falling intonation); in the operation where the analysis is made it is therefore necessary to recognise that only one part, the content, lends itself to this analysis, while the other, the expression, must be carried over unanalysed to the next operation.[1] The names of the strata are, of course, purely conventional; the glossematic description is, in principle, the same for all four, and only structural differences between them will emerge.

It is for this reason that we do not join in the current effort to analyse "phonemes" into "distinctive features". It does not matter that the "distinctive features" are said to be physical or physiological phenomena and the "phonemes" linguistic—or psychological—"forms": both can be described as terminals of functions and, in glossematics, must be so described. But a "distinctive feature" such as "voice" is not found by analysis of a "phoneme" such as the English m, *e.g.* in miizli ("*measly*"), but by analysis of a larger unit, the whole of miizli, throughout the length of which it extends. In other words, the "distinctive features" are found in a deduction which is separate from that in which "phonemes" emerge, and the two deductions part company in an operation earlier than the one of which "phonemes" are the resultants. A unit like miizli can, and must, be divided both into "phonemes" and into "distinctive features", but the

[1] This example is very considerably simplified.

two divisions are not conform; they belong, so to speak, to two different dimensions. A "phoneme" is generally understood to be a class of variants in the expression which are equivalent in respect of functions with the content, but there is not necessarily any connexion between such a class as a whole and any class of units of "distinctive features"; but each of the members of the class, the variants, is a terminal of a function the other terminal of which is a unit of "distinctive features". The "distinctive features" form classes of their own, and the term "phoneme" would be more reasonably employed on this level—in fact, in something like Daniel Jones's sense of "a family of sounds".

Lack of conformity between form and substance is, of course, no structural necessity; it is possible to imagine objects in the description of which this distinction would not have to be made, but *languages* of this type are probably very rare.

It seems likely that the descriptions of cultures will also call for repetition of operations, but a great deal of work remains to be done before it will be safe to say much about that. My own tentative hypothesis is that this is so, and that, furthermore, the stratum of the content substance is shared by a culture and the language spoken by its nationals. What I have in mind could be illustrated by a diagram shaped something like a pup-tent, with language on one side and culture on the other, each divided into strata. The top stratum common to them both— the ridge-pole, as it were—would be the content substance. A theory on these lines would explain, among other things, the relation between a text and its context of situation. A text, of course, always occurs in a context of situation, from which it can be properly isolated only by analysis; the general practice is, however, to act on the assumption that "we know what a text is" and ignore the context of situation as far as possible. When this is not possible, because of functions across the boundary, the requisite part of the context of situation is translated into a piece of text so that it can be treated as a linguistic context.

To extract a text from its context of situation by proper analysis would clearly require a uniform description of an object formed of text and situation together. In such a description the two would presumably be distinguished as separate strata but not necessarily so as to conform to the common-sense boundary we are used to. The situation is a series of behavioural events and, in fact, must bear the same relation to its culture as a text to its language. The glossematic description of a particular context of situation therefore implies a glossematic description of the whole culture to which it belongs, and a truly functional anthropology (and sociology) thus becomes a linguistic *desideratum*. Conversely, anthropologists and sociologists must be equally interested in a reliable description of languages, since they can no more afford to ignore the texts of their situations than the linguists the situations of their texts.[1]

The delimitation of what Toynbee calls an "intelligible field of study", *i.e.* a suitable object for description, is a thorny problem. The philosophy of science leads to the view that the universe is a continuum or, rather, that the description of the universe is a coherent structure. A description of a particular object is therefore in principle only part of the description of the whole universe and cannot be regarded as definitive until all the pieces have been fitted into their places in the great picture puzzle. To be absolutely sure of the description of, say, an English text, one would therefore have to begin with an analysis of the universe in the first operation of the procedure and descend gradually until the text, or some slightly larger unit comprising it, is reached. In practice that is clearly impossible, and so the investigator is forced to select his object by common sense or by "scientific intuition", which makes his description tentative and of uncertain ultimate validity. But, as Ritchie says, "the predicament of the scientific man is perhaps not so hopeless as it appears. One respect in which he differs from the common-sense practical man is that before tackling any problem he makes explicitly a number of assumptions about the particular situation he is dealing with and always makes his assumptions tentatively so that he can revise them when they turn out wrong or inconvenient. In this way practically all the drawbacks of his position are avoided. He assumes, for instance, that in considering a small portion of the universe he can neglect all the rest. He goes on on this assumption until he finds it is wrong. If it is wrong he looks round and brings another little bit of the universe into his ken, and continues altering his field of observation until his isolated system behaves as though it were really isolated. All the time he is able to leave the whole universe as such severely alone; he gets all the advantages he could have got out of a theory of the universe without the disadvantages."[2] All the advantages, that is, except full security.

---

[1] See in this connexion the works of the late Bronislaw Malinowski, particularly his supplement to *The Meaning of Meaning* by Ogden and Richards, 5th ed., London, 1938, and his *Coral Gardens*, London, 1935.

[2] A. D. Ritchie, *Scientific Method*, London, 1923, pp. 6—7.

Given the view stated above, that the scientific description of the universe must be imagined to be a coherent structure, a continuous network of functions, it seems strange that there should be portions of it which "behave as if they were isolated". This can be explained by the hypothesis that the network is of uneven density: there are islands of comparatively high density surrounded by areas of comparatively low density, and it is these islands in the description of the universe which can be singled out as particular descriptions because the surrounding thin areas give them a certain measure of independence.

The practical device of translating between situation and text—also known from stage-directions in plays—seems to me significant. Ordinary translation from one language to another is generally and, in my opinion, correctly held to be possible because, and in so far as, the content substance is common to both; translation thus becomes a matter of transferring the common substance from one set of forms to another, like moving one's clothes from one chest of drawers to another of different design. But if that is correct, the translation between situation and text implies a common substance of content.

On the number of strata required in the description of a culture it would not be profitable to speculate without the experience of actually working on such descriptions. But perhaps I may be permitted a diffident reference to Thurman Arnold's engaging idea, which does not seem to have received the attention it deserves, that a culture comprises two strata, one consisting of the "real" working of its institutions, and the other of the popular beliefs about them.[1] Arnold makes a good case for his theory, and if it holds, *i.e.* if description in terms of these two strata (with, it may be, one or two others) turns out to be self-consistent, exhaustive, and the simplest possible, then the descriptions of languages and cultures are structurally more similar than common sense would lead one to believe.

In any stratified description there will be two types of units: intrinsic units, which are terminals of functions the other terminal of which belongs to the same stratum, and projected units, which are terminals of functions the other terminal of which belongs to another stratum; both types must be registered and classified in each of the operations in which they occur. It is clear that the different strata of a description do not necessarily all have the same number of effective operations.

---

1 Thurman W. Arnold, *The Folklore of Capitalism*, New Haven, 1937.

Whenever necessary, the whole procedure must be repeated. This happens, again, when the object of description is a language. In the first procedure the text is divided up, as we have seen, through successive analyses, and the resultants classified. But the first deduction is necessarily blind, since an object is amorphous until it has been reduced to a structure through a procedure; whatever we may "know" of the object from other sources must be rigorously excluded, however hard it may be to do so. It is therefore largely a matter of chance in what order the analyses are made; the first analysis of a unit which will ultimately emerge as triplex, *abc,* may be *a(bc)* or *(ab)c:* since the final result cannot be known in advance, there is no means of deciding which is the better order or of keeping the technique uniform. Consequently, it cannot be taken for granted that the resultants of a given operation are comparable to the resultants of an operation bearing the same number in a first procedure describing another language. This inconvenience is mitigated by the introduction of fixed rules for the reckoning of degrees of derivation, but some uncertainty inevitably remains. The second procedure is a description of the text as described by the first procedure; the deduction is therefore no longer blind, and units can be definitively classified in terms of the classes of final resultants, or units thereof, which are represented among their components. Comparisons can now be made with confidence that opposite numbers from descriptions of different languages are truly comparable.

The rules we have been discussing here are summed up in the following Principle:

III. *The Principle of Reduction:* the description takes the form of a procedure. Each operation must be continued or repeated until the description is exhausted, and must lead to the registration of the smallest possible number of resultants.

It follows from what has been said that analysis is not an end in itself but only a means of reducing the number of elements. The object of an operation is to reduce the number of resultants taken over from the preceding operation, and when an operation is reached in which analysis does not lead to further reduction, the procedure comes to an end and, unless the procedure is to be repeated, the description is exhausted. Let us suppose, for instance, that in a certain operation there has been registered a class of four elements, say four moods. In the next operation we may think of a way to analyse these four further by arranging them in a two-dimensional system, like this:

	p	q
r	a	b
s	c	d

so that $a = pr$, $b = qr$, etc. This analysis leads to no reduction since the number of resultants, four, is the same as before, and the analysis should therefore not be made, unless some or all of the four new resultants, p, q, r, s, can be identified with elements already registered, while a, b, c, d cannot, in which case there is a real reduction of the inventory of elements. This rule is introduced to save unprofitable labour—and to prevent quarrels among investigators with different ideas about when to stop. The rule is stated in the following Principle:

IV. *The Principle of Economy:* the procedure must be designed to give the simplest possible result and must be discontinued when no further simplification ensues.

In our discussion of scientific truth we have hitherto left out one important thing, *viz.* the rules of verification, which must now occupy us briefly. There are elaborate rules for ascertaining historical truth, and in chemistry you can verify your hypothesis that water$=H_2O$ by taking two bottles of hydrogen and one of oxygen and making water out of them. In linguistics it is possible to do something similar: if your description of a text is correct, you should be able to deduce from it any number of new texts acceptable to native speakers; in anthropology, likewise, you should be able to deduce from your description new sequences of behaviour acceptable to the natives. Such tests are a great help and—when the result is positive—a great encouragement to the investigator in his toil, but they are not really conclusive. To find the ultimate verification we must go back to Sullivan's "pragmatical criterion of truth", success: a theory is verified to the extent that the laws and predictions made in terms of it turn out to work. Not that a theory is in itself either true or false in any other sense than as judged by the Principle of Empiricism; what can be criticised is its expediency, its power of producing laws and predictions, and the control which can thus be derived from it.

Since it is not always or even often possible to examine all of an object —you can't analyse every drop of water in the world—scientists, like businessmen, have had to make a rule whereby it is considered justifiable to go by fair samples and take it for granted that the unexamined part of the material is uniform with the sample. It will be seen that such a rule is necessary and also that it is dangerous and liable to lead to nasty surprises. We shall need a similar rule and for similar reasons, and we shall have to be careful not to overwork it. There is something else that makes a rule of this kind necessary: it often happens that a certain description is clearly indicated for part of a material while for another part one is in doubt whether to use this same description or another that fits equally well. In such a case it is necessary to come to a decision in order to avoid the fate of starving between two bales of hay, and it follows from the Principle of Empiricism that the decision should be in favour of using the same description if at all possible, since one description is simpler than two. But let no man forget or underestimate the danger of generalisation. The rule is stated in the following Principle:

V. *The Principle of Generalisation:* if an object unambiguously permits a certain description and another object ambiguously permits that same description, then the description is generalised to apply to both objects.

It has been said above that in the procedure deduction precedes induction and acts as a controlling factor in that only those resultants which have been found in one deduction are synthesised in one induction. We shall see later that the two processes are in fact worked together, but in general it is correct to say that the theory is so constructed that induction presupposes deduction, and that the theory is in itself deductive, developed from the general concept of function. This is a general characteristic of scientific theories as opposed to the laws framed in their terms, and the reason for it is, in the last resort, the quest for security, the conclusions of deductive reasoning being the only perfectly certain knowledge there is. By constructing a theory deductively one is at least sure what it is going to lead to. Scientific laws, on the other hand, are built up inductively and are appallingly uncertain, yet all our practically usable knowledge is of this nature, for deduction, though certain, does not lead to control directly but only indirectly, through the laws which it makes possible. The position is, then, that the inductive laws which give us control of the concrete world are themselves controlled by a deductive theory without which no laws could be made. It will be instructive as well as entertaining to quote Ritchie again at this point: "It is an attractive notion that in an investigation we should start with no presuppositions about the state of things to be discovered but with perfectly open minds and a single eye to the facts. There is a fine Baconian smack about it. One thinks of Darwin examining the facts for fifteen years (or whatever the period was) before framing his hypothesis. In fact it is all in the sound English tradition. Nobody can have more respect for the English tradition than I have, so that it must not be thought that I have bowed the knee to any continental Baal when I say that all this is nonsense. Darwin must have had some sort of hypothesis or he would not have known what facts to examine. There were millions of facts and he could not attend to them all. To have an open mind is not the same thing as to have a vacant mind. The vacant mind is like the bottomless pit; no amount of facts will ever fill it. What is absolutely necessary is that the investigator should not allow any hypothesis to give him a bias against the facts. Apart from this the more hypotheses he has the better. I expect Darwin in his account of his work was thinking of Newton's little joke, "Hypotheses non fingo"."[1] Ritchie is directly concerned with another level than the one we are discussing, but his remarks are relevant all the same, for a deductive theory is no more than a very general and very elaborate hypothesis. Just as the ordinary hypothesis selects and limits the "facts" to be examined, so does the theory select and limit the terms in which particular descriptions shall be made and thus the laws that can be framed by induction from the sum of particular descriptions. Without such direction induction becomes wild and unmanageable—if, indeed, it is possible at all.

What is presented in this book is only a theory and thus no more than the foundations of our projected science. If this theory, or another equivalent to it, is widely accepted, and a large number of particular descriptions made within its frame of reference, then the second, inductive part of the new science can be begun: the elucidation and formulation of general laws.

---

[1] *op. cit.*, p. 104.

GLOSSEMATIC ALGEBRA

In establishing our glossematic algebra we shall begin with the concept of function itself:

1. By a *function* is understood any dependence.
   Symbol: φ.

This is, of course, no definition, since the concepts used in it have not been previously either defined or accepted as axioms. *Function*, then, remains as a "primitive idea", *i.e.* a term which, for the time being, must be taken for granted. This is not to say that it is indefinable in an absolute sense: its content is the whole of the algebra, and the establishment of the algebra will serve to define it. The system of definitions is thus self-enclosed and self-supporting, characteristics which such systems share with ordinary dictionaries. It has, however, been designed with a view to being readily absorbed into a wider epistemological system, where some, if not all, of our primitive ideas may be expected to be definable.

Proposition 1. is intended to signalise the adoption of *function* as a technical term—and of φ as an algebraic symbol—and to give a preliminary delimitation of the area to be covered by this term. The words "any dependence" are not technical terms but are taken from the common stock of ordinary language and may safely be accepted at their face value.[1]

As a glossematic term *function*, then, covers considerably more than the ordinary use of the word, not to mention the logico-mathematical sense, which can, however, be derived from it. The OED gives the following definition: "The special kind of activity proper to anything; the mode of action by which it fulfils its purpose."[2] Our use of the term does not imply any judgement as to what is or is not "proper", nor, indeed, anything so specific as "activity", let alone "purpose". In *Prolegomena*

[1] It might have been clearer to say "connexion" instead of "dependence", but "connexion" is introduced as a technical term later on, so that its use here in the *landläufig* sense might give rise to confusion.

[2] Cf. the definition of *fonction* in Lalande's *Vocabulaire de la philosophie*: "rôle propre et caractéristique joué par un organe dans un ensemble dont les parties sont interdépendantes."

the term is used in a slightly more restricted sense ("uniform dependence") but the real, as opposed to the formal, difference is insignificant.

In our view, any function that a "thing" may enter into is part of its characteristics, though not necessarily of its formally defining characteristics. No function can safely be discarded *a priori* as not proper or not characteristic, though it is often possible to show that some functions occur more frequently than others. Statistical value is, however, no certain guide to structural importance, and, particularly in the history of language and other social phenomena, it often happens that what appeared at first to be an insignificant function gradually comes to usurp the first place. The propriety of any given function is, then, not inherent in the material itself but is an evaluation imposed by the theory brought to bear on it: all functions are potentially proper, and every function is proper from some point of view; a change in the point of view, *i.e.* a change of theory, necessarily involves a revaluation. When, in glossematics, we select certain functions to the exclusion of others, it is therefore not through a belief in any absolute superiority of the functions selected but only because these functions alone are relevant to the theory.

2. Anything that enters into a function is called a *functive*. Symbol: F.

3. The functives bound together by a given function are called the *terminals* of that function.

It follows from 2. and 3. that *functive* and *terminal* cover the same ground, since a functive must be the terminal of at least one function, and a terminal is expressly said to be a functive. It is therefore a bit of a luxury to retain both terms, though there is between them a difference of emphasis which may be found to excuse, if not to justify, this departure from strict economy. It is convenient to be able to speak of a functive without reference to any specific function of which it may be a terminal, and, on the other hand, to speak of the terminals of a function with attention focussed on the function rather than on the specific functives which enter into it. But it must be admitted that one could get by without one of these terms.

The term *functive*, it will be seen, is of extremely general application; it says nothing of its referent except that it enters as a terminal into at least one function, and that much is true of anything that one can think of. Since we hold that scientific knowledge is concerned only with function (and with structure, which we shall endeavour to bring under the same hat) it is necessary to have a term of this degree of abstractness: any term of higher specific gravity would tend to prejudice the analysis, and every other term that suggests itself—"thing", "object", "quantity", or whatever—is to some extent tainted with extraneous theoretical affiliations.

Since *functive* is derived from *function*, it follows that a classification of functives presupposes a differentiation of various kinds, or levels, of functions, except that, as a function has two terminals, the registration of any function automatically leads to the registration of two (classes of) functives—with the limiting case that the two may be one and the same. This is expressed by the formula $F_1 \varphi F_2$, *i.e.* "the functive F_1 enters into a function with the functive F_2" or "the function φ has the terminals F_1 and F_2".

It is important not to read into the distinction between *function* and *functive* any more than is contained in the formal explanations of the terms, particularly to resist the temptation to believe that a functive is necessarily more concrete than a function. Both are purely structural concepts, and nothing is said or implied about their material composition. Functions serve to bind functives together and vice versa, like the knots and strings of a net. It may be argued that, having given a distinction with one hand, we now take it away with the other, leaving no criterion for deciding how to apply the two terms. The answer must be that, while it is convenient to have two complementary terms, it is of no importance which is applied to what in any given situation, as long as we are dealing with a single hierarchy. When, however, we come to deal with functions reaching beyond a single hierarchy it is advisable to aim at arranging the application so that opposite numbers in different hierarchies are functives rather than functions or a mixture of both. We may choose to regard *sovereignty* as a complex of functives, and *king* as the functions binding them together, without inconvenience as long as we go no further; but if we extend our scope to include the function between *king* as one terminal and some specific *man* (belonging to a different, biological, hierarchy) as the other, it becomes expedient to reverse the arrangement and regard the *king* as a functive and *sovereignty* as a complex of functions into which he enters as a terminal.

4. By a *functional field* is understood a function together with its terminals. A functional field is said to be *established* by its function.

A functional field is, thus, $F_1 \varphi F_2$. But each or either of the terminals may, itself, be a functional field, so that a functional field may include one or more other functional fields. Consider, by way of illustration, the variation of the English court ceremonial according as the sovereign is a man or a woman, and, to save irrelevant detail, let us say that ceremonial *a* is used under a king, ceremonial *b* under a reigning queen. We then have a function between *a* as one terminal and the functional field *sovereign* φ *man* as the other, and likewise a function between *b* as one terminal and the functional field *sovereign* φ *woman* as the other. Diagrammatically:

$$\boxed{a\ \varphi\ \boxed{\begin{array}{c}\textit{sovereign}\\ \varphi\\ \textit{man}\end{array}}} \qquad \boxed{b\ \varphi\ \boxed{\begin{array}{c}\textit{sovereign}\\ \varphi\\ \textit{woman}\end{array}}}$$

It is obvious that the function is not between *a* and *man*, or between *b* and *woman*, because the ceremonial does not vary with a change of sex in other dignitaries, such as Members of Parliament: it is only when the *sovereign* is a man, *i.e. sovereign* φ *man*, that we get *a*, and it is only when *sovereign* φ *woman* that we get *b*. Nor can *sovereign* be regarded as a terminal in the functions in which *a* and *b* are the other terminals, since such a treatment would fail to explain the conditions under which *a* and *b* occur.

And here is a linguistic example:[1] you hear somebody say [ænt] in a context which makes it clear that he means '*aunt*', and you say to yourself, "American". Here is a function between *American* as one terminal and the functional field '*aunt*' φ [ænt] as the other:

---

[1] It must be made clear, once and for all, that the authors refuse to be held responsible for the ultimate validity of their examples. Examples are given, here and there, not for their own sake, as "facts", but to smoothe the reader's path, to make easier and more enjoyable for him the admittedly sometimes arduous task of following the exposition of the method. Whenever possible, these examples are specially constructed to illustrate the point in question with a minimum of irrelevant detail, partly for the sake of clarity, and partly to avoid recriminations. But there are occasions when it is necessary, or at least preferable, to give an "actual" example. Such "actual" examples are often simplified to suit the needs of the moment—the present one is a case in point—but even when they are not, no guarantee is furnished with them, though every reasonable care is taken. It is a corollary of the method that all conclusions are tentatively held until the description of which they form part has been exhausted, and as examples are unavoidably often chosen from incomplete descriptions, any "actual" example here given may have to be modified in the light of further research upon the language, or whatever it may be, to which it belongs. The reader is therefore warned that the detection of errors of this kind will not entitle him to get his money back. On the other hand, the authors will gratefully accept suggestions for improvements in later editions.

$$\boxed{\textit{American}\ \varphi\ \boxed{\begin{array}{c} \text{‘\textit{aunt}’} \\ \varphi \\ \text{[ænt]} \end{array}}}$$

We cannot regard either [ænt] or ‘*aunt*’ as a terminal in the function with *American*, since [ænt] occurs in other kinds of English as a pronunciation of ‘*ant*’, and ‘*aunt*’ occurs in other kinds of English with the pronunciation [ɑ:nt].

Another instance of a functional field serving as a terminal is furnished by the very common case of two or more functives alternating. A marriage, for example, may be regarded as a function between a man and a woman, who, as terminals of this function, are called *h(usband)* and *w(ife)*: suppose, now, that in a given marriage the husband has been married before: the other terminal is then a functional field consisting of w_1 and w_2 and the function between them. If the wife, too, has been married before, both terminals are functional fields:

$$\boxed{\boxed{\begin{array}{c} h_1 \\ \varphi \\ h_2 \end{array}}\ \ \varphi\ \ \boxed{\begin{array}{c} w_1 \\ \varphi \\ w_2 \end{array}}}$$

It is clear that we shall need techniques for distinguishing between bigamy and alternative marriage and between the case where h_1 and w_1 are now married to each other and the case where they are not and never have been.

Similarly, the English syllables “*see*”, “*saw*”, “*we*”, “*war*” can be regarded as functional fields in the following way (English R. P.): s φ ii, s φ ɔr/ɔ, w φ ii, w φ ɔr/ɔ. These we can now add together so as to obtain one function with functional fields as its terminals:

$$\boxed{\boxed{\begin{array}{c} \text{s} \\ \varphi \\ \text{w} \end{array}}\ \ \varphi\ \ \boxed{\begin{array}{c} \text{ii} \\ \varphi \\ \text{ɔr/ɔ} \end{array}}}$$

It will be seen that the first example presupposes a similar procedure, since its immediate data must be the marriages $h_2 \varphi w_2$, $h_2 \varphi w_1$, $h_1 \varphi w_2$, with or without $h_1 \varphi w_1$.

It may be objected that in these examples the functions alone, not the whole functional fields, should have been regarded as terminals, *i.e.*

$$\begin{array}{ccccc} & & & & \textit{sovereign} \\ & & & & | \\ a & — & \varphi & — & \varphi \\ & & & & | \\ & & & & \textit{man} \end{array} \qquad \begin{array}{ccccc} \mathrm{s} & & & & \mathrm{ii} \\ | & & & & | \\ \varphi & — & \varphi & — & \varphi \\ | & & & & | \\ \mathrm{w} & & & & \text{ɔr/ɔ} \end{array}$$

etc. This is certainly a possible interpretation, but it is less convenient. With the arrangement here advocated, a function can be identified only by its structural position, while a functive nearly always enters into a function with a functive belonging to a different hierarchy and therefore is more easily recognised; it is consequently easier in practice to deal with functional fields than with functions pure and simple.

> 5. By a *functival field* is understood a functive together with the functions into which it enters. A functival field is said to be *established* by its functive.

The functival field is the companion-piece of the functional field, and its possibility as an alternative follows from what has been said about the complementarity of function and functive; it is simply another way of "selecting and grouping in attention". A complex of functions and functives is seen, now, as grouped round a functive instead of round a function: $\varphi_1 \; \mathsf{F} \; \varphi_2$ instead of $\mathsf{F}_1 \; \varphi \; \mathsf{F}_2$, or, to return to our old simile, we choose to view a section of the net as centering on a knot rather than a string; in one case the field is established by a function, in the other by a functive. The formula $\varphi_1 \; \mathsf{F} \; \varphi_2$ serves to illustrate what we have said before, that a functive binds together the functions into which it enters (there may, of course, be more than two) just as a function binds together the functives which enter into it.

The boundaries of the two kinds of field have been left undefined[1] so that any given field may be expanded or contracted to suit varying requirements. Thus a functival field with *king* as its establishing functive can be made to include any desired collection of the functions into which the relevant king enters as a terminal, *e.g.* his constitutional to the exclusion of his ceremonial functions. The investigator is, in other words,

[1] It is, indeed, not possible to construct any such definition at this stage, but that difficulty could be overcome by introducing the two terms later.

left free to define his own sphere of interest, to choose his small portion of the universe, and to adjust it "until it behaves as though it were really isolated".

6. By a *connexion*, or *syntagmatic function*, is understood the function 'both-and'. If *a* and *b* are two functives, their connexion is symbolised by *a.b* or *ab*.

This is another primitive idea, since "both-and" has not been defined and cannot be derived from any of the terms so far established.

There are two kinds of function: that in which the functives occur together, and that in which they alternate. It is the former which we call connexion; the latter we shall presently deal with under the name of *equivalence*. The words "together" and "alternate" must not be taken as referring specifically either to time or to space: we are dealing with functions, which are completely abstract; they may be actualised in time or in space, or they may remain abstract. The distinction may equally well be visualised dynamically, as attraction and repulsion.

Whether any given function is to be regarded as a connexion or an equivalence is, to some extent, a matter of choice, *i.e.* it depends upon the investigator's point of view. Suppose, for instance, that we wish to study the two vocables [haus] and [hu:s]; we may regard them as alternating: a particular speaker, on a particular occasion, says either one or the other (which may be expanded to refer to a particular population at a particular time *or* in a particular place) or we may regard the two vocables as co-existing within an area—the area being, again, definable as temporal or spatial or both.

It is a matter of *Weltanschauung* which of the two kinds of function is to be placed first. It is clear that all scientific cognition begins with chaos. It is, in fact, creation, despite Poincaré, of exactly the kind envisaged by the ancient Jews, *i.e.* not absolute creation, out of nothing, but organisation of chaos, and Genesis is a very beautiful poetic vision of science. But the initial chaos can be conceived either as a compact mass or as a discrete collection of "things" or "events". If you take the former view, your method will be analytical, deductive, and your first type of function will be connexion (which the logicians call "conjunction" or "logical multiplication") since analysis means dividing into parts (functives) seen as the terminals of connexions. The latter view, conversely, leads to induction, the gradual synthesis of individual "things" or "events" into classes on the grounds of their various similarities, and the first kind of function to meet the eye is thus naturally the either-or function (logical "disjunction").

In symbolic logic $a+b$ stands for "either *a* or *b*", which is a little surprising until you realise that it is to be seen from the point of view of the individual: if you combine the contents of two baskets, each individual in the combined collection must have come from one or the other.[1] But since not all classes are mutually exclusive, the logicians have introduced the reading "either *a* or *b* or *both*", and it is to the special case of "both" that they apply the names "conjunction" and "logical multiplication": the logical product *ab* thus applies to those individuals which are members of both class *a* and class *b*. What we mean by *ab* is something different: a functive, *A*, is analysed into two parts, *a* and *b*, which are thus both present together in *A*; the function between *a* and *b* is consequently a connexion, and that is what is meant, in our notation, by *ab* or *a.b*. The logical product indicates that one or more individuals have two properties (the "class-concepts") in common, or that two propositions are both true, in which case truth is the class-concept; the glossematic connexion indicates that two functives are connected in such a way that they occur together or, if you prefer the dynamic interpretation, that they attract each other. Glossematic connexion thus cannot without reservations be identified with logical multiplication, but the two are sufficiently alike to warrant the use of the same notation, and, as we shall see, connexion can be treated algebraically as a kind of multiplication.

Connexions would appear to be of infinite variety, *i.e.* they seem to differ as to the specific nature of the bond between their terminals, and this specific nature of the bond would appear to depend on the specific nature of the functives involved: a marriage is different from the function between table and chair, and an adjective and its substantive are not connected in the same way as, say, two consonants belonging to the same syllable. But it would obviously be a step in the wrong direction to define connexions in terms of their functives, since we are expressly trying to obtain a functional understanding of the universe and avoid metaphysical speculation about different ways of Being. If we confine our attention to connexions found by scientific analysis, it will be seen that the specific nature of the connexions can be satisfactorily accounted for by

---

[1] Applied to propositions, $p+q$ reads "either the proposition p is true or the proposition q is true", which comes to the same thing: the truth is in either one basket or the other.

reference to the branch of science through which they were recognised and, within each branch, to degree of derivation. This may seem to lead us nowhere, since the different branches of science are properly defined by the kinds of connexion they explore, but if we admit the possibility of "unified science" the solution is clear. The different branches of science are, ideally, parts of one vast deduction with the whole universe as its object. Thus fitted together, they can be structurally defined by their places in this hierarchy, their degrees of derivation; and the specific connexions which characterise them will thereby also be defined without recourse to *das Ding an sich*.

Let us examine two different connexions, *e.g.* that between husband and wife and that between two business partners. Such connexions, according to the deductive method, could not be found direct but would be discovered through an analysis of a (minimal) family and a business firm respectively. Now it will be seen that, if each of these analyses were the first one of a separate deduction, we should in both cases have a simple "both-and" function and nothing more: the object of the first analysis of any deduction must necessarily be treated as if it had no connexions ("as if it were really isolated") which means that we have no scientific knowledge about it, although we must certainly have a hunch about its placing in a larger structural context or we should not have chosen to start our analysis at such a point. But if we do not know, as opposed to guess, what a family is or what a firm is, we cannot know anything about their parts except that they are connected, nor can we differentiate between the two connexions. When, however, the two separate analyses are fitted into a larger deduction where both are seen to derive from a common point, the two connexions can be defined in terms of their degrees of derivation, and we shall thus obtain an unambiguous differentiation without drawing on feelings, estimates, or other extra-scientific sources of "knowledge". It is easy enough to understand this as a theoretical principle, but experience shows that it is extremely difficult to impose upon oneself the discipline of not listening to the siren-song of traditional and intuitive "knowledge", to renounce the comfortable feeling that "after all, we know what a preposition is".

This is, of course, not to say that any apparent connexion can be adequately described as a simple both-and function between two otherwise anonymous functives; clearly the conjugal relations between any specific Mr and Mrs X have considerably more to them than that. But a marriage is a complex of connexions, each of which is a simple both-and function defined by its place in the hierarchy, or complex of hierarchies, and we can be certain of a complete definition only if the deduction embraces the whole universe and everything in it.[1]

In glossematic, as in logical, algebra *ab* is the same as *ba*, *i.e.* we are free to write our functives in whatever order seems most convenient, and the order in which functives are written should not be taken as having any significance. This is an algebraic necessity—or at any rate it would, apparently, be extremely difficult to design an algebra into which order enters as a significant feature. Since, as everyone knows, order is utilised as a distinguishing feature in the linguistic expression (though the psychologists are now casting some doubt on the traditional conception of speech as linear) it is obviously necessary to devise some means of differentiating *e.g.* "*fist*" and "*fits*"; but that is a problem which does not concern us here, and will be dealt with later.

Glossematic connexion is what is called *associative*, *i.e.* $a(bc) = (ab)c$.

7. By an *analysis* is understood the registration of a connexion field. This connexion field is called the *object* of the analysis. The terminals of the connexion are called the *resultants* of the analysis.

8. By a *deduction* is understood a series of analyses such that the resultants of each analysis are the objects of the following analyses.

If connexion is a kind of multiplication, analysis is a kind of factorisation: a functive, A, is resolved into a connexion field, $a.b$, *i.e.* it is regarded as consisting entirely of the two functives a and b united by connexion. But where does one get the resultants from, and how does one decide at what point to divide the object of the analysis? About the resultants, it must be remembered, nothing more has been said—and nothing more is necessarily known—than that they are connected and parts of the object, though one or both of them may, of course, be known as resultants of other analyses already made. If it is decided, for instance, to analyse '*boy*'

---

[1] In his book *Power* (London, 1938) Bertrand Russell makes a very good case for the view that social relations are a matter of power. This he conceives as a constant, like the "energy" of physics, transmutable into various forms such as political power, military power, economic power, etc. according to ascertainable laws. It is a popular book, and it is not clear whether Russell visualises power as the content of social relations—substituting, as it were, pipelines for the strings of the net—but such a theory is not necessarily incompatible with the view adopted here, since one can well imagine power as emerging at an early stage of the deduction. The various connexions would, however, be scientifically defined by their positions in the hierarchy, not by their intensional interpretation.

into *'young'*, *'male'*, *'human'*,[1] we may already know *'young'* from *'kid'*, *'cub'*, *'calf'*, etc.; *'male'* from *'man'*, *'bull'*, *'stallion'*; *'human'* from *'man'*, *'woman'*, *'who'*. But in the early stages the dividing line must often be drawn free hand—always in the case of the first analysis of a deduction. The investigator has to take a leap in the dark and must be prepared to climb back again and start all over if his first analysis should turn out to be unsatisfactory, *i.e.* to lead to conflict with the Principle of Simplicity. Once the deduction has been started, he can proceed by successive divisions, *e.g.*

$$\frac{\textit{`boy'}}{\textit{`young'}} = \textit{`male human'}, \qquad \frac{\textit{`boy'}}{\textit{`male'}} = \textit{`young human'};$$

$$\frac{\textit{`boy'}}{\textit{`human'}} = \textit{`young male'}, \qquad \therefore \quad \textit{`boy'} = \textit{`young'} \,.\, \textit{`male'} \,.\, \textit{`human'}$$

In this case it so happens that there are separate English words for the three resultants, but it will be seen that the same analysis could have been made if these words had not been parts of the English vocabulary; we can, for instance, extract *'male'* from *'brother'* by the same process of division, although there is no single English word for the remainder.[2]

If the deduction is continued as far as is possible, the order of its constituent analyses is immaterial to the inventory of final resultants. However, as we shall see under No. 51 and, in more detail, in Part II, the investigator can save himself some re-arrangement by aiming at organising his procedure so as to derive each final resultant from the original object through the maximum number of analyses. It is therefore no advantage to begin by lopping off a single irreducable element, which then has to be "carried" through all the analyses of the deduction until the bitter end. This may look like a counsel of perfection, since neither the final resultants nor the number of analyses can be known until the deduction is exhausted, and so it is: there is no short cut to scientific results. The investigator must proceed by trial and error, guided only by his experience and such "scientific intuition", whatever that is, as he can

---

[1] These elements are to be taken in a functional sense; thus *'young'* does not necessarily mean 'of low age' but rather 'who can be treated as if of low age'. Without this qualification it is impossible to explain how an elder statesman can be referred to as "the old boy".

[2] Apart from the technical term "sibling", specially adapted for the purpose, and that is a constituent of only one variant of *'brother'*, which, of course, covers more than "a male, considered in his relation to another having the same parent(s)".

muster, ever holding himself ready to revise his decisions in the light of later findings.

For examples of partial deductions, see under No. 37.

9. Two connexion fields are said to be *connected* when they have a terminal in common.

By a connexion field is meant a functional field (No. 4) in which the establishing function is a connexion (No. 6). If *a* is connected with *b* which is connected with *c*: *a.b.c*, then *b* is a terminal in two connexion fields, *(a.b)* and *(b.c)*, and we now say that we will regard the two fields as being connected by virtue of *b* belonging to both. This is a simple consequence of our determination to treat both functives and functions as binders. If *b* is common to *(a.b)* and *(b.c)*, and *c* to *(b.c)* and *(c.d)*, then all three fields are connected, and so on *ad infinitum*, as long as the connecting functives really act as connexions: if *b* is a terminal of *(a.b)* only when it is not a terminal of *(b.c)*, it is obviously not legitimate to regard the two fields as being connected; in that case the functive *b* acts as an either-or, not as a both-and binder. In skriim, "*scream*", k is a terminal in both (s.k) and (k.r), which are thereby connected; k is also common to (k.r) in krai, "*cry*" and (k.l) in klei, "*clay*", but the two are mutually exclusive, and (k.r) and (k.l) are therefore not connected fields.

10. By a *chain* is understood a connexion field or an indifferent number of connected connexion fields. Symbol: F^N

11. By a *unit* is understood a single functive, or a chain, functioning as a terminal of a connexion. Symbol: F^n

12. By a *sequence* is understood the totality of connexion fields registered in one deduction.

An analysed utterance, whether spoken or written, is thus a sequence—irrespective of its length, as long as it has been treated in one deduction—and so is any other material which lends itself to this type of description, *e.g.*, and notably, the succession of events known as "the context of situation".

13. If the two units *ab* and *a* are to be compared, then *b* is said to be *positive* in *ab*, *negative* in *a*, which is now written $a\bar{b}$.

The symbol for the negative is taken from one of the many logical notations, but the glossematic negative is not the same as the logical one.

The difference can best be illustrated by an example: if, as we have supposed, '*boy*' is '*y(oung)*'.'*m(ale)*'.'*h(uman)*' then '*child*' is $y.h$, *i. e.*, by comparison, $y.\bar{m}.h$.[1] In logic, that would indicate "a young human who is not male", which is obviously not the intention of the glossematic formula. What we mean is that '*male*' has not been found as a terminal in the chain registered through the analysis of '*child*'; the negative $\bar{m}$ is written only because '*boy*' has been registered as $y.m.h$, in comparison with which '*child*' includes an unoccupied glossematic place. Perhaps it can be made clear in this way: the glossematic negative indicates the absence of a particular unit from a particular chain, not the presence of its opposite. Absence is thus interpreted as a connexion into which the negative of the absent unit enters as a terminal. It follows from this that no separate existence is postulated for negatives and that they should not be included in the inventory.[2]

It will be seen that the glossematic negative is intended to replace the zero which has been widely, and somewhat light-heartedly, used in linguistic literature. Zero is a difficult symbol to operate with in an algebra O has been used in logical algebra in the sense of "nothing", or "the null class") and it has, besides, the disadvantage of being anonymous, *i.e.* of not indicating what it is that has been missed out; $y.\bar{m}.h$ is therefore clearer than $y.O.h$. This anonymity of zero may, on the other hand, seem an advantage when more than two units are to be compared: if '*girl*' is $y.f.h$, then '*child*' is both $y.\bar{m}.h$ and $y.\bar{f}.h$, which could conveniently be indicated by $y.O.h$; but the advantages of making clear what is left out and of writing only one symbol can be combined by choosing an arbitrary symbol, *e. g.* n, to indicate "either m or f": $\bar{n}$ would then denote an empty glossematic place which could be occupied by either m or f. In logical algebra one could write $y\bar{m}\bar{f}h$, but that solution is not possible here because such a formula would indicate more glossematic places than intended.

It is now clear how many negatives should be written in each particular case: there should be one for each unoccupied glossematic place, and the number of glossematic places is found by counting the number of positive units in the largest comparable chain; if there is a chain $abcde$, then ac must be written $a\bar{b}c\bar{d}\bar{e}$. This follows the usage of ordinary daily life: nobody would think of saying, "Smith was not at the office to-day"

[1]Let us not quibble about "Elephant's Child" and such.

[2]This rule, as we shall see, does not hold for projected units: the plural is expressed by $-s$ in "*boys*" and the singular by $-\bar{s}$ in "*boy*"; an inventory of sign expressions must therefore include $-\bar{s}$.

unless Smith could have been expected to be at the office, *i.e.* unless his absence from the office could be compared with his presence there.

For logical multiplication there is a rule, called the Law of Tautology, according to which $aa = a$, but this law is deemed not to hold for "relations", where, for instance, *father* multiplied by *father* yields *grandfather*, not *father*. It is evident that this rule cannot be adopted for glossematic connexion: it would preclude *a priori* the possibility of any unit occurring more than once, or occupying more than one place, in any chain, and connexion between two, or more, identical units is a pattern which is often needed in the description of language.[1] The phenomenon known as grammatical agreement, or concord, must be treated in this way: in '*my father has a pen-knife*' there is a connexion between the singular of '*father*' and the singular of '*has*'. In many languages, also, it is possible to regard long vocoids and contoids as manifestations of chains of identical vowels or consonants; thus the long [t:] of [aimet:u:men] "*I met two men*", clearly manifests the final t of "*met*" in connexion with the initial t of "*two*", and the long [i:] of [fi:t], "*feat*" can be regarded as a manifestation of the chain ii,[2] cp. ai in "*fight*"; short [i] must then be taken to manifest a chain comprising a negative: fit, in comparison with fiit and fait, has an unoccupied place.

It follows from this that glossematic algebra is not subject to the rule $a\bar{a} = O$, which, in logic, is called the Law of Contradiction and reads "nothing is both *a* and not-*a*". In our notation the formula $a\bar{a}$ indicates that *a* is here found alone but that, under other conditions, it occurs in connexion with another *a*.

14. A unit which has been registered as a terminal of a given connexion, is said to be *asserted* in respect of that connexion. Symbol: $+$: $(+a).(+b)$.

The statement that a unit, *a*, enters as a terminal into a connexion, *a.b*, entails the statement that this *a* occurs in the material under consideration, and thus the further statement, or postulate, that it "exists"; in the case of an asserted negative unit $(+\bar{a})$ the postulated existence is once

---

[1] And of other structures: any official holding more than one place, such as the Primate of All England, who is also Archbishop and Bishop of Canterbury, must be represented as being in connexion with himself—and may even have to write letters, perhaps stern letters, to himself.

[2] Without prejudice to the problem of the "semivowels" in English, which is here purposely left out of account.

removed: it is the corresponding positive $(+a)$ which is deemed to exist, the negative in itself only indicates a glossematic place. That is the reason for the term *assertion*, and this is the only sense in which we shall be glossematically concerned with questions of "existence" or "reality". If it cannot be unambiguously stated into what connexion(s) an alleged functive enters as a terminal, then that functive has no glossematic relevance, irrespective of whether it can be said to exist in some other sense or from some other point of view. Nor, if it passes the test, shall we bother about the existence of a unit apart from the functions into which it enters. The Principle of Simplicity enjoins the registration of as few final resultants as consistent with an exhaustive description; it can now be seen that this condition can be fulfilled only when the inventory is limited to asserted positive functives.

Assertion in glossematics corresponds to "truth" in the logic of propositions—algebraically, and in the sense that assertion, like a postulate of truth, entails a statement of validity within the universe of discourse, which, in our case, is a linguistic or a social material. With truth as such we are not concerned. Our task is to give an exhaustive description of a particular structure, and for that purpose a mendacious utterance, or misleading behaviour, is as good material as gospel truth: even the most whopping lie may be couched in the King's English. Truth is a peculiar style of the content to which a speaker may choose to restrict himself, just as he may choose to speak in Alexandrine verse, or to lead a moral life; the glossematic investigator would be ill advised to limit himself to such scanty material.

> 15. A unit which has been registered as not occurring as a terminal of a given connexion, is said to be *negated* in respect of that connexion. Symbol: $-$: $(-a).(+b)$. Assertion and negation are called *paradigmatic functions*.

Negation is thus the opposite of assertion, and the negation of a unit presupposes that that same unit has been asserted elsewhere in the material; without this restriction, there is no end to the number of minusses one could write in respect of any connexion. We shall, in general, only write negations when there is some reason for doing so, *i.e.* when it is important to say explicitly that a particular unit does not occur as a terminal of a particular connexion. Such a statement does not necessarily entail that it is *sprachwidrig* for the unit to occur in this function, but only that such occurrence has not been registered within the material, although the investigator, as he indicates by making the statement, has searched for it. An example will show how negation is used: The German preposition '*auf*' governs either the a(ccusative) or the d(ative), *i.e.* '*auf*'. $(+a\ +d)$; '*um*' governs only the accusative: '*um*'.$(+a)$, which can therefore be written, by comparison, '*um*'.$(+a\ -d)$.

It will be seen that we cannot simplify by making the negative do duty for both: a negated unit, unlike a negative, does not hold any place open in the chain, and since there is no '*um*'.(d), there is no sense in talking about '*um*'.$(\bar{d})$.

The glossematic negation is more like mathematical subtraction than logical negation. Our assertions are like addition with applicate numbers, and we cannot, therefore, subtract anything that is not already there, just as one cannot subtract three apples from four bananas. When we write $a-b$ we seem to be doing just that, but the formula is to be read as short for $a+b-b$: in '*um*'.$(+a-d)$ the dative is not subtracted from the accusative, which would be nonsense, but from the accusative plus dative which we know from the '*auf*' formula. Since a minus thus always implies a plus, it is unnecessary to write the plus.

A minus multiplied (*i.e.* connected) with either a minus or a plus gives minus, as follows from the definition of negation.

Paradigmatic functions are associative, *i.e.* $a+(b+c)=(a+b)+c$, $a-(b-c)=(a-b)-c$, and $a+(b-c)=(a+b)-c$.

16. Two or more units which are asserted, or two or more units which are negated, as one and the same terminal of a given connexion are said to be *equivalent* in respect of that connexion.

It will be seen that "equivalence" is here used in a more abstract sense than in ordinary language,[1] and it may therefore be useful to insert a warning (borrowed, *mutatis mutandis*, from Whitehead and Russell): if any further idea attaches to "equivalence", it is not required here. It should also be borne in mind that there is no such thing as equivalence in general: only equivalence in respect of a specific connexion or specific connexions. We shall say, then, that '*when he came*' is equivalent to '*as he did not come*' in respect of the connexion with '*I went*', since we have both '*when he came I went*' and '*as he did not come I went*', and we shall say that "*pl*", "*p*", and "*l*" are equivalent in respect of the connexion with "*ay*", since we have "*play*", "*pay*", and "*lay*".

[1] Though the OED does give the definition "having the same relative position or function; corresponding".

Equivalence is the glossematic either-or, and neither-nor, function, closely akin, as far as assertions are concerned, to mathematical addition and to logical disjunction, which is also called "addition", and partaking also of the nature of logical equivalence. However, in this notation $a+b$ means "either a or b is the terminal of a given connexion", not, as in logic, "either a or b or both"; "both" (ab) may be equivalent to a and b, as in the case of "*play*" above, but must not be taken to be so unless it is expressly mentioned, since it cannot be taken for granted that there are two glossematic places available in the chain. This is the main difference between logical addition and glossematic equivalence, and it is, as will be appreciated, a very important one.

There is one more feature of the algebraic use of equivalence which must be mentioned here: the formula $(a+b).c$ presupposes ac and bc but not necessarily $ac+bc$: we have '*glad*'.(-'*ness*'+-'*ly*') because of '*gladness*' and '*gladly*', but there is not necessarily any relevant connexion in respect of which '*gladness*' and '*gladly*' are equivalent. Since the method is deductive, the equivalence, or otherwise, of '*gladness*' and '*gladly*' would be registered before either is analysed, so that the danger of making a mistake should be eliminated, but it is just as well to be aware of this point all the same.

The following rules hold: $a+b=b+a$ and $a+a=a$.

17. Two or more units which are equivalent in respect of all relevant connexions are said to be *identical*. Symbol: $=$

If two keys fit the same lock, they are equivalent in respect of their connexion with the lock: either will open the door; if neither fits the lock, they are also equivalent. If opening the door is the only relevant connexion, the two keys are furthermore identical, but if some other connexion is or becomes relevant, *e.g.* that of harmonising in colour or design with the metal-work on the door, and if the two keys are not equivalent in respect of this further connexion (one being brass and the other nickel) then they cease to be identical but remain equivalent in respect of the connexion with the lock.

The question of identity is metaphysically thorny. Are two things ever quite the same? Is one thing ever quite the same at one moment as it was the moment before? I once had a Model T Ford which had most of its parts replaced and the rest considerably battered before I finally sold it for twelve dollars and fifty cents: was that the same car? And if not, what accounts for the continuity which there obviously was between the car I bought and the car I sold two years later? One way out is to declare firmly that all "events" are unique: the Model T at one moment is not the same as the Model T at any other moment, however short the moment elected as a unit of measurement. But if science is to be possible, it is clear that this will not do: one cannot make science out of an infinity of differences; science presupposes recognisable similarities, and similarity implies identity. Our somewhat robust definition of identity has been designed to fulfil this requirement.

There is an optimal level of accuracy for every activity: a difference of a fraction of a millimeter may be vital in the construction of a machine tool but is negligible in cutting cheese for a sandwich; indeed, not only is it possible, but it is necessary to neglect it: a man who insisted on the highest obtainable accuracy in all departments of life would soon find himself starving among his microtomes. In the same way the scientist has to grasp the metaphysical nettle of identity. Two functives may not be absolutely identical, but if they are equivalent in respect of all the connexions in which we are interested—even if only for the moment—then we shall define them as identical, whatever irrelevant differences there may be between them. We may, of course, make mistakes: we may, wittingly or unwittingly, disregard some difference which later turns out to be relevant; in that case we must revise our judgment, but the principle remains the same. In point of fact, all scientific work is based on this view of identity, though definitions vary; if it were not, it could not be done at all.

Our definition of identity is, then, relative, not absolute, and what is legitimately considered identical in one context may have to be treated as not identical in another. The necessary concomitant is the rule that any two functives must be treated as different until they have been proved to be identical. This is an obvious precaution, and it is only mentioned here because one is so often tempted to take a short cut. The procedure is, of course, to exhaust the relevant connexions rather than to try to make a list of all connexions and then eliminate those which are irrelevant.[1]

[1] In *An Inquiry into Meaning and Truth* Bertrand Russell discusses identity in terms of distinguishability; the conclusion is given on p. 107: "We thus arrive at the following statement: I give the name C to the shade of colour that I see at the visual place (θ, φ); I give the name C' to the colour at (θ', φ'). It may be that C and C' are distinguishable; then they are certainly different. It may be that they are indistinguishable, but that there is a colour C'' distinguishable from one but not from the other; in that case also, C and C' are certainly different. Fi-

As a matter of fact, we shall have much more use for the concept of *equality*, which cannot be formally defined until much later (cf. Part II). By this we mean the special equivalence by which two expressions have the same content or two contents the same expression; symbol: =. This is what is needed in the algebra, where, for instance, the two expressions

$$(a+\bar{a}) \cdot (b+\bar{b}) = ab + a\bar{b} + \bar{a}b + \bar{a}\bar{b}$$

are not identical but have the same content.

The lynx-eyed reader will hardly have failed to notice that the last two definitions contain a vicious circle: in No. 16 we tacitly assumed the two instances of b to be identical, and this assumption is a necessary prerequisite for the definition of equivalence; nevertheless, identity is not defined until No. 17, and then in terms of equivalence. See the note on p. 88.

18. Two or more chains are said to *correspond* when their terminals are the same or negatives of the same.

We shall begin with an example which we already know, following the sound pedagogical rule of proceeding from the known to the unknown: from the English units "*play*", "*pay*", "*lay*", "*A*", in phonematic-algebraic notation: plei, $\mathrm{p\bar{l}ei}$, $\mathrm{\bar{p}lei}$, $\mathrm{\bar{p}\bar{l}ei}$, we can infer the equivalences

$$(+\mathrm{p.l} + \mathrm{p.\bar{l}} + \mathrm{\bar{p}.l} + \mathrm{\bar{p}.\bar{l}}).\mathrm{ei}$$

The four units within the brackets are equivalent in respect of the connexion with ei, and according to No. 18 they correspond because their terminals are the same or negatives of the same.

---

nally, it may be that every colour known to me is either distinguishable from both or indistinguishable from both; in that case, C and *C′ may* be identical, *i.e.* "*C*" and "*C′*" may be two names for the same thing. But since I can never know that I have surveyed *all* colours, I can never be sure that *C* and *C′* are identical." It is to avoid lying awake at night, racked by such nerve-shattering uncertainty, that a working definition of identity is necessary. Faced with Russell's problem, we should first ascertain what are the relevant connexions and then find out experimentally whether the two colours are equivalent in respect of all of them. In other words, we should behave in much the same way as any housewife having to buy a reel of cotton *for a given purpose*. And, having established identity within our frame of reference, we, like the housewife, should refuse to bother about the possible existence of irrelevant differences. The belief in absolute identity seems to me a piece of superfluous mysticism—harmless if it is not allowed to prey on the mind, but bearing the seeds of dark and hideous insanity.

Correspondence and equivalence do not always go together. We have the equivalence (+s.t + p.l).ei because of "*stay*" and "*play*", but the connexion fields in s.t and p.l do not correspond since their terminals are different. This is, incidentally, the only sense in which one could speak of the negation of a connexion, and a minus sign in front of a chain can therefore always be taken to indicate the negation of the connexion field, not the connexion alone.

Conversely, '*gladly*' and '*glad*' give the corresponding connexion fields '*glad*'.'*ly*' and '*glad*'.'$\overline{ly}$', but '*glad*$\overline{ly}$' is not equivalent to '*gladly*' in a chain such as "sand for his spinach I'll *gladly* bring".[1]

It sometimes happens that correspondence and equivalence go together in one part of the material but not in another. In a certain school the masters used to be divided into three classes: (a) junior masters, who were not allowed to marry, (b) senior masters, who were allowed but not obliged to marry, and (c) the head master, who was, in fact, always married, though I am not sure that he was obliged to. The three classes were differentiated in other ways, too—by pay, social status, etc., all of which we can sum up in the three symbols *a*, *b*, *c*. Including a few widows who were also about, and letting *m* stand for *master* and *w* for wife, the actual occurrences were as indicated by the following assertions and negations:

$$a.(-mw+m\bar{w}-\bar{m}w)$$
$$b.(+mw+m\bar{w}+\bar{m}w)$$
$$c.(+mw-m\bar{w}+\bar{m}w)^{2}$$

The four chains, $m.w$, $m.\bar{w}$, $\bar{m}.w$, $\bar{m}.\bar{w}$, correspond, since their terminals are the same or negatives of the same, but the assertions and negations are differently distributed in respect of the three connexions with *a*, *b*, and *c*.

And a linguistic example: suppose that there is a language with the consonants p, t, k, s, r, l, that its syllabic themes have the structure $C^{n}VC$, and that the following consonant clusters occur: spr, skl, sp, st, sk, pr, tr, kr, pl, kl; all the consonants occur alone both initially and finally. We then have the following assertions and negations in respect of initial *(i)* and final *(f)* connexion with the vocalic unit:[3]

---

[1] "and tabasco sauce for his teething-ring." From Ogden Nash's *Song to be Sung by the Fathers of Infant Female Children.*

[2] It will be seen that a vacancy would mean $+\bar{m}\bar{w}$, and no vacancy $-\bar{m}\bar{w}$, in each class.

[3] The reader is asked to take the symbols *i* and *f* on trust pending the discussion of the glossematic treatment of order, which belongs in Part II.

$$
\begin{array}{ll}
i. & (+\mathrm{spr}+\mathrm{skl}-\mathrm{spl}-\mathrm{str}-\mathrm{skr} \\
 & +\mathrm{sp}\bar{\mathrm{r}}+\mathrm{sk}\bar{\mathrm{l}}+\mathrm{sp}\bar{\mathrm{l}}+\mathrm{st}\bar{\mathrm{r}}+\mathrm{sk}\bar{\mathrm{r}} \\
 & -\mathrm{s}\bar{\mathrm{p}}\mathrm{r}-\mathrm{s}\bar{\mathrm{k}}\mathrm{l}-\mathrm{s}\bar{\mathrm{p}}\mathrm{l}-\mathrm{s}\bar{\mathrm{t}}\mathrm{r}-\mathrm{s}\bar{\mathrm{k}}\mathrm{r} \\
 & +\mathrm{s}\bar{\mathrm{p}}\bar{\mathrm{r}}+\mathrm{s}\bar{\mathrm{k}}\bar{\mathrm{l}}+\mathrm{s}\bar{\mathrm{p}}\bar{\mathrm{l}}+\mathrm{s}\bar{\mathrm{t}}\bar{\mathrm{r}}+\mathrm{s}\bar{\mathrm{k}}\bar{\mathrm{r}} \\
 & +\bar{\mathrm{s}}\mathrm{pr}+\bar{\mathrm{s}}\mathrm{kl}+\bar{\mathrm{s}}\mathrm{pl}+\bar{\mathrm{s}}\mathrm{tr}+\bar{\mathrm{s}}\mathrm{kr} \\
 & +\bar{\mathrm{s}}\mathrm{p}\bar{\mathrm{r}}+\bar{\mathrm{s}}\mathrm{k}\bar{\mathrm{l}}+\bar{\mathrm{s}}\mathrm{p}\bar{\mathrm{l}}+\bar{\mathrm{s}}\mathrm{t}\bar{\mathrm{r}}+\bar{\mathrm{s}}\mathrm{k}\bar{\mathrm{r}} \\
 & +\bar{\mathrm{s}}\bar{\mathrm{p}}\mathrm{r}+\bar{\mathrm{s}}\bar{\mathrm{k}}\mathrm{l}+\bar{\mathrm{s}}\bar{\mathrm{p}}\mathrm{l}+\bar{\mathrm{s}}\bar{\mathrm{t}}\mathrm{r}+\bar{\mathrm{s}}\bar{\mathrm{k}}\mathrm{r} \\
 & -\bar{\mathrm{s}}\bar{\mathrm{p}}\bar{\mathrm{r}}-\bar{\mathrm{s}}\bar{\mathrm{k}}\bar{\mathrm{l}}-\bar{\mathrm{s}}\bar{\mathrm{p}}\bar{\mathrm{l}}-\bar{\mathrm{s}}\bar{\mathrm{t}}\bar{\mathrm{r}}-\bar{\mathrm{s}}\bar{\mathrm{k}}\bar{\mathrm{r}})
\end{array}
\qquad
\begin{array}{ll}
f. & (-\mathrm{spr}-\mathrm{skl}-\mathrm{spl}-\mathrm{str}-\mathrm{skr} \\
 & -\mathrm{sp}\bar{\mathrm{r}}-\mathrm{sk}\bar{\mathrm{l}}-\mathrm{sp}\bar{\mathrm{l}}-\mathrm{st}\bar{\mathrm{r}}-\mathrm{sk}\bar{\mathrm{r}} \\
 & -\mathrm{s}\bar{\mathrm{p}}\mathrm{r}-\mathrm{s}\bar{\mathrm{k}}\mathrm{l}-\mathrm{s}\bar{\mathrm{p}}\mathrm{l}-\mathrm{s}\bar{\mathrm{t}}\mathrm{r}-\mathrm{s}\bar{\mathrm{k}}\mathrm{r} \\
 & +\mathrm{s}\bar{\mathrm{p}}\bar{\mathrm{r}}+\mathrm{s}\bar{\mathrm{k}}\bar{\mathrm{l}}+\mathrm{s}\bar{\mathrm{p}}\bar{\mathrm{l}}+\mathrm{s}\bar{\mathrm{t}}\bar{\mathrm{r}}+\mathrm{s}\bar{\mathrm{k}}\bar{\mathrm{r}} \\
 & -\bar{\mathrm{s}}\mathrm{pr}-\bar{\mathrm{s}}\mathrm{kl}-\bar{\mathrm{s}}\mathrm{pl}-\bar{\mathrm{s}}\mathrm{tr}-\bar{\mathrm{s}}\mathrm{kr} \\
 & +\bar{\mathrm{s}}\mathrm{p}\bar{\mathrm{r}}+\bar{\mathrm{s}}\mathrm{k}\bar{\mathrm{l}}+\bar{\mathrm{s}}\mathrm{p}\bar{\mathrm{l}}+\bar{\mathrm{s}}\mathrm{t}\bar{\mathrm{r}}+\bar{\mathrm{s}}\mathrm{k}\bar{\mathrm{r}} \\
 & +\bar{\mathrm{s}}\bar{\mathrm{p}}\mathrm{r}+\bar{\mathrm{s}}\bar{\mathrm{k}}\mathrm{l}+\bar{\mathrm{s}}\bar{\mathrm{p}}\mathrm{l}+\bar{\mathrm{s}}\bar{\mathrm{t}}\mathrm{r}+\bar{\mathrm{s}}\bar{\mathrm{k}}\mathrm{r} \\
 & -\bar{\mathrm{s}}\bar{\mathrm{p}}\bar{\mathrm{r}}-\bar{\mathrm{s}}\bar{\mathrm{k}}\bar{\mathrm{l}}-\bar{\mathrm{s}}\bar{\mathrm{p}}\bar{\mathrm{l}}-\bar{\mathrm{s}}\bar{\mathrm{t}}\bar{\mathrm{r}}-\bar{\mathrm{s}}\bar{\mathrm{k}}\bar{\mathrm{r}})
\end{array}
$$

It will be seen that, although the connexion fields in each vertical column correspond, the assertions and negations of the consonant chains vary according as they can or cannot function as terminals in the two outside connexions, with *i* and *f*.

Perhaps this is as good a place as any to make a general remark about this method. Some readers may feel that we have made an unnecessary song and dance about a comparatively simple arrangement of six consonants; "if," we can hear these readers muttering, "if glossematics consists in making half a page full of complicated formulae out of six little consonants, then we will have no part of it." We will at once admit that the example is simple—it was purposely so constructed—and, in itself, could be adequately described with less algebraic apparatus. But there are many much more complicated systems to be described: most languages have more than six consonants, and many have syllabic themes of great complexity. Our aim is to provide one method which can be applied to any material, regardless of degree of complexity. The advantage of one universally applicable method need hardly be pointed out; the corresponding, and unavoidable, disadvantage in very simple cases is, by comparison, negligible. If we have not, so far, illustrated anything very complicated, it is only to save the reader time and trouble, for which he should not be ungrateful.

19. By a *paradigm* is understood a unit which is asserted or negated in respect of a given connexion, together with such other units as may be asserted or negated as one and the same terminal of that connexion. The unit(s) belonging to a paradigm are called its *member(s)*. A paradigm is said to be *established* by the paradigmatic functions of its members and to be *generated* by the connexion in respect of which they are asserted or negated.
Symbol: $\langle\ \rangle$: $\langle +a+b-c\rangle .q$.

A paradigm is thus a class of units which have this in common that they are all associated, by assertion or negation, with a certain terminal of a certain connexion, *i.e.* with a certain glossematic place; a plus indicates ability, a minus inability to occupy that place.

A connexion often generates a paradigm of several members at each of its terminals: we have noted the paradigm $\langle +\mathrm{pl} + \mathrm{p}\bar{\mathrm{l}} + \bar{\mathrm{p}}\mathrm{l} + \bar{\mathrm{p}}\bar{\mathrm{l}} \rangle$ at one terminal of the connexion field pl.ei, but there is another at the other terminal, since, in addition to "*play*", "*pay*", "*lay*", and "*A*", we have also "*ply*", "*pie*", "*lie*", "*I*", and "*plea*", "*pea*", "*lea*", "*E*", *i.e.*

$$\langle +\mathrm{pl} + \mathrm{p}\bar{\mathrm{l}} + \bar{\mathrm{p}}\mathrm{l} + \bar{\mathrm{p}}\bar{\mathrm{l}} \rangle . \langle +\mathrm{ei} + \mathrm{ai} + \mathrm{ii} \rangle$$

The members of the latter paradigm can be further analysed so as to make two paradigms, *i.e.* they can be regarded as occupying two glossematic places each, and we can make the formulation more precise by indicating that a positive member is always present:

$$\langle +\mathrm{e} + \mathrm{a} + \mathrm{i} - \bar{\mathrm{e}} - \bar{\mathrm{a}} - \bar{\mathrm{i}} \rangle . \langle +\mathrm{i} - \bar{\mathrm{i}} \rangle$$

20. By a *synthesis* is understood the registration of a paradigm.

21. By an *induction* is understood a series of syntheses such that the paradigm of one synthesis enters as a member into the paradigm of the following synthesis, etc.

Synthesis is thus the counterpart of analysis, and induction of deduction. The procedure is, upon the registration of each connexion, to open two paradigms and start collecting members for them by determining whether the resultants of other analyses will fit into either or both. There is nothing new or surprising in this—probably every linguist and social scientist of whatever denomination works in this way—but it needs to be put on record all the same, and in our own terminology. After synthesis follows induction, until the process is exhausted.

The result of a deduction followed by induction may be submitted to further deduction and induction, as will be demonstrated in Part II, but as the procedure remains the same, nothing further need be said about it here.

22. Two or more paradigms are said to *correspond* when they have the same members.

This is a simple expansion of *correspondence*, introduced in No. 18, to cover paradigms as well as chains.

23. By the *sum* of two or more corresponding paradigms is understood the totality of their assertions. Symbol: $\langle\ \rangle^+$. Members which are not asserted in any of the paradigms considered, are carried over into the sum as negations.

In our example of the German prepositions and cases we found that the connexion with '*auf*' as one terminal generated the paradigm $\langle$+accusative+dative$\rangle$, and the connexion with '*um*' as one terminal, the paradigm $\langle$+accusative−dative$\rangle$; German has also '*aus*'.$\langle$−accusative+dative$\rangle$. The sum of all three paradigms is $\langle$+accusative+dative$\rangle^+$, according to No. 23, since both the accusative and the dative are asserted at least once in the totality of paradigms considered. If, as we might have done, we had written-in the nominative and the genitive as negated in these paradigms, they would have been transferred as negations to the sum:

$$
\begin{array}{l}
\text{'}auf\text{'}.\langle +a+d-n-g\rangle \\
\text{'}um\text{'}.\langle +a-d-n-g\rangle \\
\text{'}aus\text{'}.\langle -a+d-n-g\rangle \\
\hline
\qquad\langle +a+d-n-g\rangle^+
\end{array}
$$

When two paradigms have one or more members in common, it is always possible to expand one or both, by inserting the appropriate negations, until they correspond, so that a sum can be found; *e. g.*

$$
\begin{array}{rl}
\langle +a+b+c\rangle = & \langle +a+b+c-d-e\rangle \\
\langle +c+d+e\rangle = & \langle -a-b+c+d+e\rangle \\
\hline
\text{sum:} & \langle +a+b+c+d+e\rangle^+
\end{array}
$$

It will be seen that any paradigm is equal to a sum of paradigms, *viz.* to a sum of identical paradigms or, if the paradigm has more than one assertion, to a sum of paradigms with the requisite fewer assertions; *e. g.*

$$
\begin{array}{r}
\langle +a-b-c\rangle \\
\langle +a-b-c\rangle \\
\hline
\langle +a-b-c\rangle = \langle +a-b-c\rangle^+
\end{array}
$$

$$\begin{array}{r} \langle +a-b-c\rangle \\ \langle -a+b-c\rangle \\ \hline \langle +a+b-c\rangle = \langle +a+b-c\rangle^{+} \end{array}$$

When a paradigm has more than two assertions, there are always alternative ways of making up the equal sum, not counting sums of identical paradigms, *e. g.*

$$\begin{array}{cccc} & & & \langle +a-b-c\rangle \\ \langle +a+b-c\rangle & \langle +a+b-c\rangle & \langle +a+b-c\rangle & \langle -a+b-c\rangle \\ \langle +a-b+c\rangle & \langle -a-b+c\rangle & \langle -a+b+c\rangle & \langle -a-b+c\rangle \\ \hline \langle +a+b+c\rangle^{+} = & \langle +a+b+c\rangle^{+} = & \langle +a+b+c\rangle^{+} = & \langle +a+b+c\rangle^{+} \end{array}$$

Sums, as we shall see, are needed for various purposes, especially for making inventories. An inventory is, in fact, the sum of all relevant paradigms.[1]

24. By a *category* is understood a collection of correspondents. Symbol: $\{\}$.

Instead of making a sum of two or more corresponding paradigms, we shall find it convenient, for some purposes, to regard the corresponding paradigms as members of a further class (cf. No. 21), which, as a collection of correspondents, we shall call a *category*. Thus our German case-paradigms, of which, under No. 23, we made a sum, can also be treated as making a category; we then have, so far,

$$\{+\langle +a+d-n-g\rangle + \langle +a-d-n-g\rangle + \langle -a+d-n-g\rangle\}$$

It will be seen that, as there are four plusses and/or minusses in each of the member paradigms, the full category has sixteen members, ranging from $\langle +a+d+n+g\rangle$ to $\langle -a-d-n-g\rangle$, which, of course, are not necessarily all asserted. We shall make it a rule always to operate with full categories, although they take up rather a lot of space, because only in this way can the investigator be sure that no possibility has been overlooked. The technique is, then, immediately upon the registration of any paradigm, to write out a complete category, in which those member paradigms which have not been registered are provisionally negated; whenever, in the course of further work, one of them is registered, its

[1] An inventory of projected units is a simple sum; in an inventory of intrinsic units the negatives are left out of account; cf. Part II.

negation in the category is changed to an assertion. Thus the paradigm $\langle +a+d+n+g \rangle$ would, at the present stage, have to be negated in our category of German cases, but would eventually be registered, and its negation would then be changed to an assertion.

The number of members of a category depends, as we have seen, on the number of members of the constituent paradigms; the mathematical formula is 2^n, where n is the number of members of the constituent paradigms.

By No. 18, two or more chains are said to correspond when their terminals are the same or negatives of the same, and a collection of corresponding chains is thus also a category; cf. the examples given under No. 18. Here, again, we shall make it a rule to operate always with complete categories, so that, for instance, the paradigm $\langle +ab+\bar{a}b \rangle$ must be turned into the category $\{+ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}$, which comprises all the duplex chains that can be made out of a and b and their negatives, the negations being subject to confirmation in the light of further research. Indeed, the registration of the single chain *ab* is, as will be seen, sufficient for setting up a category, in which the three other members will be provisionally negated.

The number of members of a category of chains depends on the number of positive functives involved; the formula is, once more, 2^n, where n is the number of positive functives entering into the chains which are the members of the category. If there are three, a, b, c, the number of members will thus be $2^3 = 8$, *viz.*

$$\{\pm abc\pm ab\bar{c}\pm a\bar{b}c\pm a\bar{b}\bar{c}\pm \bar{a}bc\pm \bar{a}b\bar{c}\pm \bar{a}\bar{b}c\pm \bar{a}\bar{b}\bar{c}\}$$

It is obvious that a category may, itself, be treated as a member of a category of categories: thus the category $\{+ab-a\bar{b}-\bar{a}b+\bar{a}\bar{b}\}$ is one of the sixteen members of a category of categories ranging from $\{+ab+a\bar{b}+\bar{a}b+\bar{a}\bar{b}\}$ to $\{-ab-a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}$. A category of categories we shall call a *category of the second power*. The number of members of a category being 2^n, the number of members of a category of the second power is 2^{2^n}. A category of the second power may, in its turn, be treated as a member of a category of the third power, etc.[1]

[1] I hasten to allay the reader's very proper horror at this appalling prospect by remarking that a continuation of the induction beyond categories of the second power can only very rarely, if ever, be profitable in the description of a single language or culture. It would, on the other hand, obviously be rash to exclude the possibility *a priori,* and it is clear that for comparative work categories of a higher power than two will almost certainly be needed.

25. The sum of the categories generated by all relevant connexions is called the *exhaustive category*. Symbol: $\{\}$

To find the exhaustive category for any given collection of functives, the procedure is to take the sum of all asserted members of the relevant category of the second power. Suppose, for example, that we wish to find the exhaustive category of paradigms for the two functives *a* and *b*. The first thing to do, then, is to examine all the paradigms of which either *a* or *b* or both are members; if one or the other is missing from any of these paradigms, it is registered as negated. Suppose that there are the following paradigms:

$$\begin{array}{ll} (1) & \langle +a+b+c+d \rangle \\ (2) & \langle +a-b+e+f+g \rangle \\ (3) & \langle -a+b+f+h \rangle \\ (4) & \langle +a+c+g+h \rangle \end{array}$$

We can then abstract the following paradigms relevant to our present inquiry:

$$\begin{array}{ll} (1) & \langle +a+b \rangle \\ (2) & \langle +a-b \rangle \\ (3) & \langle -a+b \rangle \\ (4) & \langle +a-b \rangle \end{array}$$

of which (2) and (4) are the same. The next step is to arrange these paradigms in categories, and we must therefore now decide whether the paradigms are all to be registered as members of one category or if they should be distributed over two or more; if only one category is needed, it will, of course, in itself be exhaustive, and there is no reason for a category of the second power. This will depend on circumstances, *i.e.* on the connexions generating the paradigms: if the other terminals of these connexions are homogeneous, there will be only one category, if they are themselves clearly differentiated into separate categories, our paradigms must be distributed into a corresponding number of categories (cf. Part II). We will suppose that, in our present example, (1) and (2) belong to one category and (3) and (4) to separate categories. The asserted members of the category of the second power are then

$$\begin{array}{rl} (1), (2) & +\{+\langle +a+b \rangle+\langle +a-b \rangle-\langle -a+b \rangle-\langle -a-b \rangle\} \\ (3) & +\{-\langle +a+b \rangle-\langle +a-b \rangle+\langle -a+b \rangle-\langle -a-b \rangle\} \\ (4) & +\{-\langle +a+b \rangle+\langle +a-b \rangle-\langle -a+b \rangle-\langle -a-b \rangle\} \end{array}$$

The exhaustive category is the sum of these, *i.e.*

$$\left\{+\langle +a+b\rangle +\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\right\}^{+}$$

The exhaustive category of chains for any given collection of functives is found in a similar way: let the collection of functives be *a*, *b*, *c*, and the categories registered

$$\begin{aligned}
&+\left\{+abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\}\\
&+\left\{-abc-ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\}\\
&+\left\{-abc+ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\}\\
&+\left\{+abc-ab\bar{c}+a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\}
\end{aligned}$$

The exhaustive category is then

$$\left\{+abc+ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\}^{+}$$

It will be seen that, since we make it a rule always to write the members of categories in the same (arbitrary) order, it is sufficient, in the calculations leading to the exhaustive category, to write the plusses and minusses alone, which saves a good deal of time and space.

It follows from the definition of a sum (No. 23) that a category in which all the members are asserted is identical with the exhaustive category, since the sum of an assertion and either a negation or another assertion is an assertion. If such a category is found to be asserted, there is therefore no need for any calculations to find the exhaustive category, though any other members of the category of the second power must, of course, in any case be registered. Our example above, of the English consonants p and l is a case in point: in the category generated by the connexion with ei ("*play*", "*pay*", "*lay*", "*A*") all the members are asserted:

$$\left\{+\mathsf{pl}+\mathsf{p}\bar{\mathsf{l}}+\bar{\mathsf{p}}\mathsf{l}+\bar{\mathsf{p}}\bar{\mathsf{l}}\right\}.\ \mathsf{ei}$$

and this category is therefore identical with the exhaustive category. There are, however, other members of the category of the second power, *e. g.* those which can be abstracted from "*plant*", "*aunt*" and from "*love*":[1]

$$\begin{aligned}
&\left\{+\mathsf{pl}-\mathsf{p}\bar{\mathsf{l}}-\bar{\mathsf{p}}\mathsf{l}+\bar{\mathsf{p}}\bar{\mathsf{l}}\right\}.\ \mathsf{aa/rnt}\\
&\left\{-\mathsf{pl}-\mathsf{p}\bar{\mathsf{l}}+\bar{\mathsf{p}}\mathsf{l}-\bar{\mathsf{p}}\bar{\mathsf{l}}\right\}.\ \mathsf{\Lambda v}
\end{aligned}$$

---

[1] These illustrations have been chosen as immediately intelligible without long explanations; they are, however, of no relevance to the cenematic system of English since, as categories of projected units, they belong to the description of the usage.

26. By *correlation* is understood the functions establishing a category of paradigms. The member paradigms are said to be *correlated*, and their members are called *correlates*.

Since glossematic functions are associative, it is possible to arrange any paradigm so as to consist of simplex paradigms:

$$\langle +a-b+c\rangle = \langle +\langle +a\rangle + \langle -b\rangle + \langle +c\rangle\rangle$$

There are plusses in front of the simplex paradigms because, as paradigms, they are asserted, even though $\langle -b\rangle$ is established by a negation.

A category of paradigms, similarly, can be arranged so as to consist of categories of simplex paradigms, *i.e.* categories with only one correlate:

$$\begin{aligned} &\{+\langle +a+b\rangle + \langle +a-b\rangle - \langle -a+b\rangle - \langle -a-b\rangle\} \\ =&\{+\{+\langle +a\rangle - \langle -a\rangle\} + \{+\langle +b\rangle + \langle -b\rangle\}\} \end{aligned}$$

A category with only one correlate we shall call a *simplex category*.

The exhaustive simplex category is found from the sum of the categories in which the relevant simplex categories occur, *e.g.*

$$\begin{array}{l} \{+\langle +a+b\rangle + \langle +a-b\rangle - \langle -a+b\rangle - \langle -a-b\rangle\} \\ \{-\langle +a+b\rangle - \langle +a-b\rangle + \langle -a+b\rangle - \langle -a-b\rangle\} \\ \hline \{+\langle +a+b\rangle + \langle +a-b\rangle + \langle -a+b\rangle - \langle -a-b\rangle\}^{+} \end{array}$$

$$=\{+\{+\langle +a\rangle + \langle -a\rangle\}^{+} + \{+\langle +b\rangle + \langle -b\rangle\}^{+}\}$$

A simplex category has $2^1=2$ members, *viz.* the paradigms $\langle +a\rangle$ and $\langle -a\rangle$. Its category of the second power has $2^{2^1}=4$ members, *viz.* the categories

$$\begin{array}{l} \{+\langle +a\rangle + \langle -a\rangle\} \\ \{+\langle +a\rangle - \langle -a\rangle\} \\ \{-\langle +a\rangle + \langle -a\rangle\} \\ \{-\langle +a\rangle - \langle -a\rangle\} \end{array}$$

The glossematic value of the symbols is, however, such that the fourth of these categories ("*a* is neither asserted nor negated") would invariably be negated; it can therefore be left out of account, and we can regard the category of simplex categories as having, for all practical purposes, three members.

For the sake of compactness we shall now introduce a special symbol for each of the three simplex categories, as follows:

$$\{+\langle +a\rangle +\langle -a\rangle\} = \sqsupset a \sqsubset$$
$$\{+\langle +a\rangle -\langle -a\rangle\} = \supset a \subset$$
$$\{-\langle +a\rangle +\langle -a\rangle\} = \subset a \supset$$

to which, for the sake of completeness, we can add

$$\{-\langle +a\rangle -\langle -a\rangle\} = \sqsubset a \sqsupset$$

in case there is ever any use for it. For the first and the last term in any formula the symbol need be written only once, *e. g.*

$$\{+\langle +a+b\rangle +\langle +a-b\rangle -\langle -a+b\rangle -\langle -a-b\rangle\} = \{+\{+\langle +a\rangle -\langle -a\rangle\} + \{+\langle +b\rangle +\langle -b\rangle\}\} = a \subset\sqsupset b$$

$$\{+\langle +a+b+c\rangle +\langle +a+b-c\rangle +\langle +a-b+c\rangle +\langle +a-b-c\rangle +\langle -a+b+c\rangle + \langle -a+b-c\rangle +\langle -a-b+c\rangle +\langle -a-b-c\rangle\} = \{+\{+\langle +a\rangle +\langle -a\rangle\} + \{+\langle +b\rangle +\langle -b\rangle\} + \{+\langle +c\rangle +\langle -c\rangle\}\} = a \sqsubset\sqsupset b \sqsubset\sqsupset c$$

As we have seen, the category of simplex categories has, in effect, three members, *i.e.* there are three unitary correlations, *viz.* the three to which we have assigned symbols above. With these symbols we can therefore indicate $3 \times 3 = 9$ binary correlations, $3 \times 3 \times 3 = 27$ tertiary correlations, etc. But the category of duplex categories has $2^{2^2} = 16$ members, of which one, the category in which all the members are negated, has no glossematic application and can, in any case, be indicated by the formula $a \sqsupset\sqsubset b$; there are therefore six duplex categories which cannot be indicated by any simple arrangement of our three symbols. This difficulty can be overcome by treating the six missing categories as sums of categories which *can* be indicated: we saw, under No. 23, that any paradigm is equal to a sum of two or more paradigms, and we shall now make use of this fact. Thus the category

$$\{+\langle +a+b\rangle +\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\}$$

cannot be indicated by any simple arrangement of the correlation symbols, but it is equal to four alternative sums of categories which can be so indicated, *viz.*

$$\begin{array}{ll}
\{+\langle +a+b\rangle +\langle +a-b\rangle -\langle -a+b\rangle -\langle -a-b\rangle\} & = a \subset\sqsupset b \\
\{+\langle +a+b\rangle -\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\} & = a \sqsubset\supset b \\
\hline
\{+\langle +a+b\rangle +\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\}^{+} &
\end{array}$$

$$\begin{array}{ll}
\{+\langle+a+b\rangle+\langle+a-b\rangle-\langle-a+b\rangle-\langle-a-b\rangle\} & =a\subset\sqsupset b \\
\{-\langle+a+b\rangle-\langle+a-b\rangle+\langle-a+b\rangle-\langle-a-b\rangle\} & =a\supset\supset b \\
\hline
\{+\langle+a+b\rangle+\langle+a-b\rangle+\langle-a+b\rangle-\langle-a-b\rangle\}^{+} &
\end{array}$$

$$\begin{array}{ll}
\{+\langle+a+b\rangle-\langle+a-b\rangle+\langle-a+b\rangle-\langle-a-b\rangle\} & =a\sqsubset\supset b \\
\{-\langle+a+b\rangle+\langle+a-b\rangle-\langle-a+b\rangle-\langle-a-b\rangle\} & =a\subset\subset b \\
\hline
\{+\langle+a+b\rangle+\langle+a-b\rangle+\langle-a+b\rangle-\langle-a-b\rangle\}^{+} &
\end{array}$$

$$\begin{array}{ll}
\{+\langle+a+b\rangle-\langle+a-b\rangle-\langle-a+b\rangle-\langle-a-b\rangle\} & =a\subset\supset b \\
\{-\langle+a+b\rangle+\langle+a-b\rangle-\langle-a+b\rangle-\langle-a-b\rangle\} & =a\subset\subset b \\
\{-\langle+a+b\rangle-\langle+a-b\rangle+\langle-a+b\rangle-\langle-a-b\rangle\} & =a\supset\supset b \\
\hline
\{+\langle+a+b\rangle+\langle+a-b\rangle+\langle-a\quad\rangle-\langle-a-b\rangle\}^{+} &
\end{array}$$

We can therefore write

$\{a\subset\sqsupset+\sqsubset\supset b\}$, $\{a\subset\sqsupset+\supset\supset b\}$, $\{a\sqsubset\supset+\subset\subset b\}$, or $\{a\subset\supset+\subset\subset+\supset\supset b\}$ or, more compactly, $a{\substack{\subset\sqsupset\\ \sqsubset\supset}}b$, $a{\substack{\subset\sqsupset\\ \supset\supset}}b$, $a{\substack{\sqsubset\supset\\ \subset\subset}}b$, or $a{\substack{\subset\supset\\ \subset\subset\\ \supset\supset}}b$.

If there is nothing to indicate which of the possible sums is preferable, we shall make it a rule to choose the solution which has the smallest number of correlations and the greatest number of assertions; in this case, therefore, $a{\substack{\subset\sqsupset\\ \sqsubset\supset}}b$.

Table I. shows the 16 members of the category of duplex categories together with their correlations; for the six sums the solutions have been chosen in accordance with the rule just stated. Table II. gives the 27 triplex categories which can be indicated by simple arrangements of the three correlation symbols; the rest of the 255 effective members of the category of triplex categories can be stated in terms of sums of two or more of these. For the sake of compactness the member paradigms have been left out; they are always written in the same order, so that, *e. g.* category 1. in Table II represents

$$\{+\langle+a+b+c\rangle+\langle+a+b-c\rangle+\langle+a-b+c\rangle+\langle+a-b-c\rangle+\\ \langle-a+b+c\rangle+\langle-a+b-c\rangle+\langle-a-b+c\rangle+\langle-a-b-c\rangle\}$$

27. The correlation of a duplex category in which $\{+\langle+a-b\rangle+\langle-a+b\rangle\}$ is called an *autonomy* (1, 2, 9, 10).

28. The correlation of a duplex category in which $\{+\langle +a-b\rangle - \langle -a+b\rangle\}$ or $\{-\langle +a-b\rangle + \langle -a+b\rangle\}$ is called a *specification* (3, 4, 5, 6, 11, 12, 13, 14).
29. The correlation of a duplex category in which $\{-\langle +a-b\rangle - \langle -a+b\rangle\}$ is called a *complementarity* (7, 8, 15).
30. When, in a duplex category, $+\langle +a-b\rangle$, then a is called a *major correlate*.
31. When, in a duplex category, $-\langle +a-b\rangle$, then a is called a *minor correlate*.

An autonomy thus has two major correlates, a specification has one major and one minor, and a complementarity has two minor correlates.

32. The correlation of a duplex category in which $\{+\langle +a+b\rangle\}$ is called *conjunct* (1—8).
33. The correlation of a duplex category in which $\{-\langle +a+b\rangle\}$ is called *disjunct* (9—15).
34. The correlation of a duplex category in which $\{+\langle -a-b\rangle\}$ is called an *absence correlation* (odd numbers).
35. The correlation of a duplex category in which $\{-\langle -a-b\rangle\}$ is called a *presence correlation* (even numbers).

These definitions are purely terminological and hardly need any comments; the numbers in brackets refer to Table I. The reason for introducing special terms for binary correlations is that duplex categories play a particularly important rôle in the description of languages and other systems: the paradigmatic definition of each functive in the inventory consists in the totality of the binary correlations into which it enters.

The idea is, in accordance with our general principle, to give a purely functional definition of all units, *i.e.* to define them in terms of relative dependence. *Complementarity* is the correlation in which there is the highest degree of mutual dependence: neither correlate occurs as a member of any relevant paradigm[1] of which the other is not also a member. In

[1] If all paradigms were taken into consideration, there would be no complementarities or specifications, since each unit necessarily occurs as a simplex paradigm in respect of at least one connexion: otherwise it could not be registered as a separate unit at all. In stratified descriptions, such as those of languages, the final resultants of separate strata will often be projected units, some of which form simplex paradigms only in respect of extrinsic connexions; in order to obtain the greatest possible differentiation, correlations are therefore based on intrinsic connexions alone.

TABLE I.

1. $\{+\langle +a+b\rangle +\langle +a-b\rangle +\langle -a+b\rangle +\langle -a-b\rangle\} = a\Subset\Supset b$

2. $\{+\langle +a+b\rangle +\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\} = a\substack{\subset\,\Supset\\ \Subset\,\supset}b$

3. $\{+\langle +a+b\rangle +\langle +a-b\rangle -\langle -a+b\rangle +\langle -a-b\rangle\} = a\substack{\Subset\,\subset\\ \subset\,\Supset}b$

4. $\{+\langle +a+b\rangle +\langle +a-b\rangle -\langle -a+b\rangle -\langle -a-b\rangle\} = a\subset\Supset b$

5. $\{+\langle +a+b\rangle -\langle +a-b\rangle +\langle -a+b\rangle +\langle -a-b\rangle\} = a\substack{\Subset\,\supset\\ \supset\,\Supset}b$

6. $\{+\langle +a+b\rangle -\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\} = a\Subset\supset b$

7. $\{+\langle +a+b\rangle -\langle +a-b\rangle -\langle -a+b\rangle +\langle -a-b\rangle\} = a\substack{\subset\,\supset\\ \supset\,\subset}b$

8. $\{+\langle +a+b\rangle -\langle +a-b\rangle -\langle -a+b\rangle -\langle -a-b\rangle\} = a\subset\supset b$

9. $\{-\langle +a+b\rangle +\langle +a-b\rangle +\langle -a+b\rangle +\langle -a-b\rangle\} = a\substack{\Subset\,\subset\\ \supset\,\Supset}b$

10. $\{-\langle +a+b\rangle +\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\} = a\substack{\subset\,\subset\\ \supset\,\supset}b$

11. $\{-\langle +a+b\rangle +\langle +a-b\rangle -\langle -a+b\rangle +\langle -a-b\rangle\} = a\Subset\subset b$

12. $\{-\langle +a+b\rangle +\langle +a-b\rangle -\langle -a+b\rangle -\langle -a-b\rangle\} = a\subset\subset b$

13. $\{-\langle +a+b\rangle -\langle +a-b\rangle +\langle -a+b\rangle +\langle -a-b\rangle\} = a\supset\Supset b$

14. $\{-\langle +a+b\rangle -\langle +a-b\rangle +\langle -a+b\rangle -\langle -a-b\rangle\} = a\supset\supset b$

15. $\{-\langle +a+b\rangle -\langle +a-b\rangle -\langle -a+b\rangle +\langle -a-b\rangle\} = a\supset\subset b$

(16. $\{-\langle +a+b\rangle -\langle +a-b\rangle -\langle -a+b\rangle -\langle -a-b\rangle\} = a\Supset\Subset b$)

TABLE II.

1. $\{+\,+\,+\,+\,+\,+\,+\,+\} = a\Subset\Supset b\Subset\Supset c$

2. $\{+\,-\,+\,-\,+\,-\,+\,-\} = a\Subset\Supset b\Subset\supset c$

3. $\{-\,+\,-\,+\,-\,+\,-\,+\} = a\Subset\Supset b\Subset\subset c$

4. $\{+\,+\,-\,-\,+\,+\,-\,-\} = a\Subset\supset b\subset\Supset c$

5. $\{+\,-\,-\,-\,+\,-\,-\,-\} = a\Subset\supset b\subset\supset c$

6. $\{-\,+\,-\,-\,-\,+\,-\,-\} = a\Subset\supset b\subset\subset c$

7. $\{-\,-\,+\,+\,-\,-\,+\,+\} = a\Subset\subset b\supset\Supset c$

8. $\{-\,-\,+\,-\,-\,-\,+\,-\} = a\Subset\subset b\supset\supset c$

9. $\{-\,-\,-\,+\,-\,-\,-\,+\} = a\Subset\subset b\supset\subset c$

10. $\{+\,+\,+\,+\,-\,-\,-\,-\} = a\subset\Supset b\Subset\Supset c$

11. $\{+\,-\,+\,-\,-\,-\,-\,-\} = a\subset\Supset b\Subset\supset c$

12. $\{-\,+\,-\,+\,-\,-\,-\,-\} = a\subset\Supset b\Subset\subset c$

13. $\{+\,+\,-\,-\,-\,-\,-\,-\} = a\subset\supset b\subset\Supset c$

14. $\{+\,-\,-\,-\,-\,-\,-\,-\} = a\subset\supset b\subset\supset c$

15. $\{-\,+\,-\,-\,-\,-\,-\,-\} = a\subset\supset b\subset\subset c$

16. $\{-\,-\,+\,+\,-\,-\,-\,-\} = a\subset\subset b\supset\Supset c$

17. $\{-\,-\,+\,-\,-\,-\,-\,-\} = a\subset\subset b\supset\supset c$

18. $\{-\,-\,-\,+\,-\,-\,-\,-\} = a\subset\subset b\supset\subset c$

19. $\{-\,-\,-\,-\,+\,+\,+\,+\} = a\supset\Supset b\Subset\Supset c$

20. $\{-\,-\,-\,-\,+\,-\,+\,-\} = a\supset\Supset b\Subset\supset c$

21. $\{-\,-\,-\,-\,-\,+\,-\,+\} = a\supset\Supset b\Subset\subset c$

22. $\{-\,-\,-\,-\,+\,+\,-\,-\} = a\supset\supset b\subset\Supset c$

23. $\{-\,-\,-\,-\,+\,-\,-\,-\} = a\supset\supset b\subset\supset c$

24. $\{-\,-\,-\,-\,-\,+\,-\,-\} = a\supset\supset b\subset\subset c$

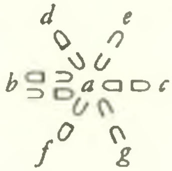
25. $\{-\,-\,-\,-\,-\,-\,+\,+\} = a\supset\subset b\supset\Supset c$

26. $\{-\,-\,-\,-\,-\,-\,+\,-\} = a\supset\subset b\supset\supset c$

27. $\{-\,-\,-\,-\,-\,-\,-\,+\} = a\supset\subset b\supset\subset c$

autonomies the correlates are paradigmatically independent of each other, each occurring as a member of at least one paradigm of which the other is not a member. And in *specifications* the dependence is unilateral.

Since there are 15 different binary correlations, and each functive may enter into any number, depending on the structure of the description to which it belongs, it will be seen that the theory offers considerable possibilities of differentiation—there need be no fear that the high degree of abstraction will lead to difficulties in that respect. The paradigmatic definition of a functive will then look, in principle, something like this:



In addition, there are other possibilities, of which we can give no more than a hint at this stage. With the categories comprising only a single asserted paradigm (8, 12, 14, 15) of course no more can be done, but the other, more complex categories lend themselves to statistical treatment as a supplement to the glossematic description, which opens up further prospects of differentiation. This line of inquiry will be of particular interest in the study of historical changes because it furnishes a technique for describing the gradual replacement of one pattern by another. Suppose, for example, that we have the three functives *a*, *b*, *c* with the following correlations:

$$\{+\langle +a+b\rangle+\langle +a-b\rangle+\langle -a+b\rangle+\langle -a-b\rangle\}=a\sqsubset\sqsupset b$$
$$\{+\langle +a+c\rangle+\langle +a-c\rangle+\langle -a+c\rangle+\langle -a-c\rangle\}=a\sqsubset\sqsupset c$$
$$\{-\langle +b+c\rangle+\langle +b-c\rangle+\langle -b+c\rangle+\langle -b-c\rangle\}=b\genfrac{}{}{0pt}{}{\sqcap\,\sqsubset}{\sqsupset\,\sqsubset} c$$

b and *c* then have exactly the same paradigmatic definitions, but a differentiation can be obtained if it is possible to show that there is a difference in the proportions of the numbers of connexions generating the four member-paradigms in each case: *e. g.* in $a\sqsubset\sqsupset b$ the proportions might be 12:6:4:3[1] and in $a\sqsubset\sqsupset c$ 3:4:12:3. If, now, instead of $a\sqsubset\sqsupset b$ and $a\sqsubset\sqsupset c$ we have two specimens of $a\sqsubset\sqsupset b$, culled from different stages in the history of, for instance, a language, then a difference of statistical emphasis

[1] *I.e.* (a multiple of) 12 connexions generating the paradigm $\langle +a+b\rangle$, (a multiple of) 6 generating $\langle +a-b\rangle$, etc.

may give a hint about the direction in which the language is developing. It is at least possible that an accumulation of evidence of this kind may lead to a better understanding of the mechanism of historical change, perhaps even to the formulation of general laws.

36. By a *set* is understood a correlated paradigm the correlates of which are a functive and its negative.

For any functive, *a*, the category of sets has $2^2 = 4$ members, *viz.* the paradigms $\langle +a+\bar{a} \rangle$, $\langle +a-\bar{a} \rangle$, $\langle -a+\bar{a} \rangle$, and $\langle -a-\bar{a} \rangle$. The category of the second power has $2^{2^2} = 16$ members; cf. Table I, which is equal to the category of the second power for the sets of *a* when $b = \bar{a}$. An exhaustive category of sets can be found in the usual way but, as we have remarked, is of no relevance except in the case of projected units.

The correlations of the asserted category or categories of sets depend, of course, on the connexions into which the sets enter; if, for instance, the total occurrences of sets of *a* are

$$\langle +a+\bar{a} \rangle \,.\, \langle -b+\bar{b} \rangle = \{ -ab+a\bar{b}-\bar{a}b+\bar{a}\bar{b} \}$$
$$\langle +a-\bar{a} \rangle \,.\, \langle +c+\bar{c} \rangle = \{ +ac+a\bar{c}-\bar{a}c-\bar{a}\bar{c} \}$$

then there will be asserted, according to circumstances, either one category of sets of *a*, *viz.*

$$\{ +\langle +a+\bar{a} \rangle + \langle +a-\bar{a} \rangle - \langle -a+\bar{a} \rangle - \langle -a-\bar{a} \rangle \} = a \subset\supset \bar{a}$$

or two categories, *viz.*

$$\{ +\langle +a+\bar{a} \rangle - \langle +a-\bar{a} \rangle - \langle -a+\bar{a} \rangle - \langle -a-\bar{a} \rangle \} = a \subset \supset \bar{a}$$
$$\{ -\langle +a+\bar{a} \rangle + \langle +a-\bar{a} \rangle - \langle -a+\bar{a} \rangle - \langle -a-\bar{a} \rangle \} = a \subset \subset \bar{a}$$

The decision as to whether the two asserted sets of *a* should be placed in one or in two categories depends on the glossematic definitions of *b* and *c*; for criteria see Part II.

In logical algebra the formula $a.(b+\bar{b})$ is held to be universally valid, because it says that all the members of the class *a* must either be members of the class *b* or members of the class not-*b*. In glossematic algebra, it will be remembered, the symbols have a different value, and the corresponding formula, $a.\langle +b+\bar{b} \rangle$ is of course not universally valid: it asserts $+a$, $+b$, and $+\bar{b}$ and is therefore only valid if the chains *ab* and $a\bar{b}$ have been registered as occurring in the description under consideration.

The fourth set, $\langle -a-\bar{a} \rangle$, which reads "neither *a* nor not-*a*" may seem, at first sight, to contain a contradiction, but a moment's reflexion will show that, given the special glossematic value of the symbols, there is no contradiction: "neither a nor not-$\bar{a}$" amounts to the negation of a glossematic place, *i.e.* "there is no place here which could be either occupied by a or left unoccupied". We shall, however, find very little practical use for this set, since its assertion as a terminal of a connexion naturally leads to nothing but negated chains, whatever the sets of the other functive(s) entering into the connexion.

In order to save space we shall now assign a symbol to each set, as follows:[1]

$$\langle +a+\bar{a}\rangle = \mapsto a \mapsfrom$$
$$\langle +a-\bar{a}\rangle = \rightarrow a \leftarrow$$
$$\langle -a+\bar{a}\rangle = \leftarrow a \rightarrow$$
$$\langle -a-\bar{a}\rangle = \mapsfrom a \mapsto$$

It will be seen that these symbols are similar in principle to the correlation symbols introduced under No. 26 above; it is hoped that this similarity will make both notations easier to learn. For the first and the last term in any formula the arrows need be written only once, *e. g.*

$$\langle +a+\bar{a}\rangle . \langle +b-\bar{b}\rangle . \langle -c+\bar{c}\rangle = a \mapsfrom \rightarrow b \leftarrow \leftarrow c$$
$$\langle +a+\bar{a}\rangle . \langle -b+\bar{b}\rangle = a \mapsfrom \leftarrow b$$

It will be seen that the sum of $\rightarrow a \leftarrow$ and $\leftarrow a \rightarrow$ is $\mapsto a \mapsfrom$, and that $\mapsto a \mapsfrom + \rightarrow a \leftarrow = \mapsto a \mapsfrom$ and $\mapsto a \mapsfrom + \leftarrow a \rightarrow = \mapsto a \mapsfrom$; cf.

$$\langle +a+\bar{a}\rangle \; . \langle +b-\bar{b}\rangle \; = \{+ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\} \; = a \mapsfrom \rightarrow b$$
$$\langle +a+\bar{a}\rangle \; . \langle -b+\bar{b}\rangle \; = \{-ab+a\bar{b}-\bar{a}b+\bar{a}\bar{b}\} \; = a \mapsfrom \leftarrow b$$

$$= \langle +a+\bar{a}\rangle^{+} . \langle +b+\bar{b}\rangle^{+} = \{+ab+a\bar{b}+\bar{a}b+\bar{a}\bar{b}\}^{+} = a \mapsfrom \mapsto b$$

$$\langle +a-\bar{a}\rangle \; . \langle +b+\bar{b}\rangle \; = \{+ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\} \; = a \leftarrow \mapsto b$$
$$\langle -a+\bar{a}\rangle \; . \langle +b+\bar{b}\rangle \; = \{-ab-a\bar{b}+\bar{a}b+\bar{a}\bar{b}\} \; = a \rightarrow \mapsto b$$

$$= \langle +a+\bar{a}\rangle^{+} . \langle +b+\bar{b}\rangle^{+} = \{+ab+a\bar{b}+\bar{a}b+\bar{a}\bar{b}\}^{+} = a \mapsfrom \mapsto b$$

---

[1] The arrow-notation here presented is the outcome of considerable experimentation. Many shapes, combinations, and values of arrows have been tried out and discarded as insufficient, or clumsy, or suitable only for duplex chains, or for a variety of other reasons. Two earlier notations may be inspected, if anyone is interested, in my article, "On Equivalent Relations" in *Recherches structurales*, this series, vol. V.

$$
\begin{array}{llll}
\langle +a-\bar a\rangle & .\langle +b-\bar b\rangle & =\{+ab-a\bar b-\bar a b-\bar a\bar b\} & = a \leftarrow\;\rightarrow b \\
\langle +a-\bar a\rangle & .\langle -b+\bar b\rangle & =\{-ab+a\bar b-\bar a b-\bar a\bar b\} & = a \leftarrow\;\leftarrow b \\
\langle -a+\bar a\rangle & .\langle +b-\bar b\rangle & =\{-ab-a\bar b+\bar a b-\bar a\bar b\} & = a \rightarrow\;\rightarrow b \\
\langle -a+\bar a\rangle & .\langle -b+\bar b\rangle & =\{-ab-a\bar b-\bar a b+\bar a\bar b\} & = a \rightarrow\;\leftarrow b \\
\hline
\end{array}
$$

$$
=\langle +a+\bar a\rangle^{+}.\langle +b+\bar b\rangle^{+}=\{+ab+a\bar b+\bar a b+\bar a\bar b\}^{+}=a \mapsfrom\;\mapsto b
$$

However, this does not mean that we can simply write $\mapsto a \mapsfrom$ whenever there are two asserted sets in the category, for the constituent sets may not be equivalent in respect of all the connexions into which they separately enter. If, for instance, we have

$$
\{+abc-ab\bar c-a\bar bc-a\bar b\bar c-\bar abc-\bar ab\bar c-\bar a\bar bc+\bar a\bar b\bar c\}
$$

which can be analysed into

$$
\begin{aligned}
(a\leftarrow\;\rightarrow b).c &= +abc-a\bar bc-\bar abc-\bar a\bar bc = a\leftarrow\;\rightarrow b\leftarrow\;\rightarrow c \\
(a\rightarrow\;\leftarrow b).\bar c &= -ab\bar c-a\bar b\bar c-\bar ab\bar c+\bar a\bar b\bar c = a\rightarrow\;\leftarrow b\rightarrow\;\leftarrow c
\end{aligned}
$$

then, although $\rightarrow a\leftarrow + \leftarrow a\rightarrow = \mapsto a\mapsfrom$, and $\rightarrow b\leftarrow + \leftarrow b\rightarrow = \mapsto b\mapsfrom$, and $\rightarrow c\leftarrow + \leftarrow c\rightarrow = \mapsto c\mapsfrom$, we cannot write $a\mapsfrom\;\mapsto b\mapsfrom\;\mapsto c$, because the result would be

$$
\{+abc+ab\bar c+a\bar bc+a\bar b\bar c+\bar abc+\bar ab\bar c+\bar a\bar bc+\bar a\bar b\bar c\}
$$

which, in addition to $a\leftarrow\rightarrow b\leftarrow\rightarrow c$ and $a\rightarrow\leftarrow b\rightarrow\leftarrow c$, would presuppose also $a\leftarrow\rightarrow b\leftarrow\leftarrow c$, $a\leftarrow\leftarrow b\rightarrow\rightarrow c$, $a\leftarrow\leftarrow b\rightarrow\leftarrow c$, $a\rightarrow\rightarrow b\leftarrow\rightarrow c$, $a\rightarrow\rightarrow b\leftarrow\leftarrow c$, and $a\rightarrow\leftarrow b\rightarrow\rightarrow c$. We must therefore write $a\substack{\leftarrow\,\rightarrow\\ \rightarrow\,\leftarrow}b\substack{\leftarrow\,\rightarrow\\ \rightarrow\,\leftarrow}c$, and the categories of sets will be

$$
\begin{aligned}
\{-\langle +a+\bar a\rangle+\langle +a-\bar a\rangle+\langle -a+\bar a\rangle-\langle -a-\bar a\rangle\} &= a\substack{\subset\\ \supset}\substack{\subset\\ \supset}\bar a \\
\{-\langle +b+\bar b\rangle+\langle +b-\bar b\rangle+\langle -b+\bar b\rangle-\langle -b-\bar b\rangle\} &= b\substack{\subset\\ \supset}\substack{\subset\\ \supset}\bar b \\
\{-\langle +c+\bar c\rangle+\langle +c-\bar c\rangle+\langle -c+\bar c\rangle-\langle -c-\bar c\rangle\} &= c\substack{\subset\\ \supset}\substack{\subset\\ \supset}\bar c
\end{aligned}
$$

37. By *relation* is understood the functions establishing a category of chains. The constituent chains are said to be *related*, and the positive functives of a relation are called its *relates*.[1]

---

[1] The term "relation" has been used in previous publications, *e.g.* in Hjelmslev's *Prolegomena* and in my article "On Equivalent Relations", in the sense in which we now use "connexion". We hope that this change of terminology will not cause undue hardship, and that the discriminating reader will sympathise with our difficulties in finding suitable terms. "Relation", heaven help us, is used in yet another sense in logic.

It will be seen that a category of chains can always be analysed so as to emerge as the product of a connexion between sets or as the sum of two or more such products; *e.g.*

$$\{+abc+ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\}=a\leftarrow\ \mapsto b \mapsfrom\ \mapsto c$$

$$\{+ab+a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}= \left| \begin{array}{ll} \begin{array}{l} \{+ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \mapsto b \\ \{+ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}=a\mapsfrom\ \rightarrow b \end{array} & (1) \\[2ex] \begin{array}{l} \{+ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \mapsto b \\ \{-ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}=a\rightarrow\ \rightarrow b \end{array} & (2) \\[2ex] \begin{array}{l} \{+ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}=a\mapsfrom\ \rightarrow b \\ \{-ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \leftarrow b \end{array} & (3) \\[2ex] \begin{array}{l} \{+ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \mapsto b \\ \{-ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \leftarrow b \\ \{-ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}=a\rightarrow\ \rightarrow b \end{array} & (4) \\[2ex] \begin{array}{l} \{+ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}=a\mapsfrom\ \rightarrow b \\ \{-ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \leftarrow b \\ \{-ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}=a\rightarrow\ \rightarrow b \end{array} & (5) \\[2ex] \begin{array}{l} \{+ab-a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \rightarrow b \\ \{-ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\}=a\leftarrow\ \leftarrow b \\ \{-ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\}=a\rightarrow\ \rightarrow b \end{array} & (6) \end{array} \right.$$

It is therefore possible to state all relations in terms of sets, *viz.* either as simple connexions between sets or, when no simple connexion will yield the category required, as sums of connexions between sets. Here, again, whenever there is a free choice of alternative sums, we shall make it a rule to prefer the solution which involves the smallest number of sets and yields the greatest number of assertions; in the example above, the preferred solution is therefore (1), $a \genfrac{}{}{0pt}{}{\leftarrow}{\mapsfrom} \genfrac{}{}{0pt}{}{\mapsto}{\rightarrow} b$.

Table III shows the 16 members of the category of categories of duplex chains together with their relations indicated in terms of sets; the solutions for the six categories which have to be treated as sums are in accordance with the rule just stated. Table IV gives the 27 members of the category of categories of triplex chains which result from simple connexions between sets, excluding the sets with no asserted members; the

TABLE III.

1. $\left\{+ab+a\bar{b}+\bar{a}b+\bar{a}\bar{b}\right\} = a \mapsfrom \mapsto b$
2. $\left\{+ab+a\bar{b}+\bar{a}b-\bar{a}\bar{b}\right\} = a \substack{\mapsfrom \\ \leftarrow} \substack{\rightarrow \\ \mapsto} b$
3. $\left\{+ab+a\bar{b}-\bar{a}b+\bar{a}\bar{b}\right\} = a \substack{\mapsfrom \\ \leftarrow} \substack{\leftarrow \\ \mapsto} b$
4. $\left\{+ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\right\} = a \leftarrow \mapsto b$
5. $\left\{+ab-a\bar{b}+\bar{a}b+\bar{a}\bar{b}\right\} = a \substack{\mapsfrom \\ \rightarrow} \substack{\rightarrow \\ \mapsto} b$
6. $\left\{+ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\right\} = a \mapsfrom \rightarrow b$
7. $\left\{+ab-a\bar{b}-\bar{a}b+\bar{a}\bar{b}\right\} = a \substack{\leftarrow \\ \rightarrow} \substack{\rightarrow \\ \leftarrow} b$
8. $\left\{+ab-a\bar{b}-\bar{a}b-\bar{a}\bar{b}\right\} = a \leftarrow \rightarrow b$
9. $\left\{-ab+a\bar{b}+\bar{a}b+\bar{a}\bar{b}\right\} = a \substack{\mapsfrom \\ \rightarrow} \substack{\leftarrow \\ \mapsto} b$
10. $\left\{-ab+a\bar{b}+\bar{a}b-\bar{a}\bar{b}\right\} = a \substack{\leftarrow \\ \rightarrow} \substack{\leftarrow \\ \rightarrow} b$
11. $\left\{-ab+a\bar{b}-\bar{a}b+\bar{a}\bar{b}\right\} = a \mapsfrom \leftarrow b$
12. $\left\{-ab+a\bar{b}-\bar{a}b-\bar{a}\bar{b}\right\} = a \leftarrow \leftarrow b$
13. $\left\{-ab-a\bar{b}+\bar{a}b+\bar{a}\bar{b}\right\} = a \rightarrow \mapsto b$
14. $\left\{-ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\right\} = a \rightarrow \rightarrow b$
15. $\left\{-ab-a\bar{b}-\bar{a}b+\bar{a}\bar{b}\right\} = a \rightarrow \leftarrow b$
16. $\left\{-ab-a\bar{b}-\bar{a}b-\bar{a}\bar{b}\right\} = a \mapsto \mapsfrom b$

TABLE IV.

1. $\left\{+abc+ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}+\bar{a}bc+\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \mapsto b \mapsfrom \mapsto c$
2. $\left\{+abc-ab\bar{c}+a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \mapsto b \mapsfrom \rightarrow c$
3. $\left\{-abc+ab\bar{c}-a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \mapsto b \mapsfrom \leftarrow c$
4. $\left\{+abc+ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \rightarrow b \leftarrow \mapsto c$
5. $\left\{+abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \rightarrow b \leftarrow \rightarrow c$
6. $\left\{-abc+ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \rightarrow b \leftarrow \leftarrow c$
7. $\left\{-abc-ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \leftarrow b \rightarrow \mapsto c$
8. $\left\{-abc-ab\bar{c}+a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \leftarrow b \rightarrow \rightarrow c$
9. $\left\{-abc-ab\bar{c}-a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \mapsfrom \leftarrow b \rightarrow \leftarrow c$
10. $\left\{+abc+ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \mapsto b \mapsfrom \mapsto c$
11. $\left\{+abc-ab\bar{c}+a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \mapsto b \mapsfrom \rightarrow c$
12. $\left\{-abc+ab\bar{c}-a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \mapsto b \mapsfrom \leftarrow c$
13. $\left\{+abc+ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \rightarrow b \leftarrow \mapsto c$
14. $\left\{+abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \rightarrow b \leftarrow \rightarrow c$
15. $\left\{-abc+ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \rightarrow b \leftarrow \leftarrow c$
16. $\left\{-abc-ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \leftarrow b \rightarrow \mapsto c$
17. $\left\{-abc-ab\bar{c}+a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \leftarrow b \rightarrow \rightarrow c$
18. $\left\{-abc-ab\bar{c}-a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \leftarrow \leftarrow b \rightarrow \leftarrow c$
19. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc+\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \mapsto b \mapsfrom \mapsto c$
20. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \mapsto b \mapsfrom \rightarrow c$
21. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \mapsto b \mapsfrom \leftarrow c$
22. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \rightarrow b \leftarrow \mapsto c$
23. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \rightarrow b \leftarrow \rightarrow c$
24. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \rightarrow b \leftarrow \leftarrow c$
25. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \leftarrow b \rightarrow \mapsto c$
26. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \leftarrow b \rightarrow \rightarrow c$
27. $\left\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\} = a \rightarrow \leftarrow b \rightarrow \leftarrow c$

rest of the 255 triplex categories comprising at least one assertion can be found as sums of two or more of the 27, *e.g.*

$$
\begin{aligned}
10.\ &\{+abc+ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}-\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\} = && a \leftarrow \mapsto b \mapsfrom \mapsto c\\
3.\ &\{-abc+ab\bar{c}-a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\} = && a \mapsfrom \mapsto b \mapsfrom \leftarrow c
\end{aligned}
$$

$$\{+abc+ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\}^{+} = a \substack{\leftarrow \\ \mapsfrom} \substack{\mapsto \\ \mapsto} b \substack{\mapsfrom \\ \mapsfrom} \substack{\mapsto \\ \leftarrow} c$$

$$
\begin{aligned}
3.\ &\{-abc+ab\bar{c}-a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc+\bar{a}b\bar{c}-\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\} = && a \mapsfrom \mapsto b \mapsfrom \leftarrow c\\
7.\ &\{-abc-ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}-\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\} = && a \mapsfrom \leftarrow b \rightarrow \mapsto c\\
19.\ &\{-abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc+\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\} = && a \rightarrow \mapsto b \mapsfrom \mapsto c
\end{aligned}
$$

$$\{-abc+ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}+\bar{a}bc+\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\}^{+} = a \substack{\mapsfrom \\ \mapsfrom \\ \rightarrow} \substack{\mapsto \\ \leftarrow \\ \mapsto} b \substack{\mapsfrom \\ \rightarrow \\ \mapsfrom} \substack{\leftarrow \\ \mapsto \\ \mapsto} c$$

The constituent sets of any binary or tertiary relation can easily be found from the tables; for more complex relations—pending the publication of a "Glossematician's Vademecum"—the fixed order of writing gives a simple clue: if there are assertions in both halves of a category of chains, whatever its complexity, then both $+a$ and $+\bar{a}$, *i.e.* either $\mapsto a \mapsfrom$ or $\substack{\rightarrow \\ \leftarrow} a \substack{\leftarrow \\ \rightarrow}$; if all the assertions are concentrated in the first half, then $\rightarrow a \leftarrow$; if in the second half, then $\leftarrow a \rightarrow$. If, within each half, there are assertions in both halves, then $\mapsto b \mapsfrom$ or $\substack{\rightarrow \\ \leftarrow} b \substack{\leftarrow \\ \rightarrow}$; if only within the first half, then $\rightarrow b \leftarrow$, if only within the second half, then $\leftarrow b \rightarrow$; etc., etc. *E. g.*

$$
\begin{gathered}
\{+abcd+abc\bar{d}-ab\bar{c}d-ab\bar{c}\bar{d}+a\bar{b}cd+a\bar{b}c\bar{d}-a\bar{b}\bar{c}d-a\bar{b}\bar{c}\bar{d}\\
-\bar{a}bcd-\bar{a}bc\bar{d}-\bar{a}b\bar{c}d-\bar{a}b\bar{c}\bar{d}+\bar{a}\bar{b}cd-\bar{a}\bar{b}c\bar{d}+\bar{a}\bar{b}\bar{c}d-\bar{a}\bar{b}\bar{c}\bar{d}\}
\end{gathered}
$$

There are assertions in both halves, so $+a$ and $+\bar{a}$, but we cannot tell yet whether $\mapsto a \mapsfrom$ or $\substack{\rightarrow \\ \leftarrow} a \substack{\leftarrow \\ \rightarrow}$; in the first half (1—8) there are assertions in both halves (1—4 and 5—8), so $+b$ and $+\bar{b}$, but in the second half (9—16) there are assertions only in the second half (13—16), so that here $\langle -b+\bar{b} \rangle$; it is therefore clear that the sets, so far, are $a \substack{\leftarrow \\ \rightarrow} \substack{\mapsto \\ \leftarrow} b$. The next division shows $\langle +c-\bar{c} \rangle$ for 1—8 and $\langle +c+\bar{c} \rangle$ for 9—16; therefore $a \substack{\leftarrow \\ \rightarrow} \substack{\mapsto \\ \leftarrow} b \substack{\mapsfrom \\ \rightarrow} \substack{\rightarrow \\ \mapsto} c$. And finally $\langle +d+\bar{d} \rangle$ in 1—8, $\langle +d-\bar{d} \rangle$ in 9—16. The sets for the whole relation are thus

$$a \substack{\leftarrow \\ \rightarrow} \substack{\mapsto \\ \leftarrow} b \substack{\mapsfrom \\ \rightarrow} \substack{\rightarrow \\ \mapsto} c \substack{\leftarrow \\ \mapsfrom} \substack{\mapsto \\ \rightarrow} d$$

We shall now demonstrate the *modus operandi* on an "actual" example—the piece of English text, '*George went to the cinema while Mary had her hair* *curled*'.[1] This piece of text we shall call A, and we shall make the following assumptions about it: (1) that A is a resultant of an analysis registering the connexion $A.B$, and (2) that $\langle +A+\bar{A}\rangle . B$. Our first analysis is $A=a.b$: ('*George went to the cinema*').('*while Mary had her hair curled*'), because a is the largest part of A found separately elsewhere in the material in a comparable function. The category of chains generated by the connexion with B is then

$$\{+ab+a\bar{b}-\bar{a}b+\bar{a}\bar{b}\}=a \overset{\leftarrow\!\shortmid\;\longleftarrow}{\underset{\longleftarrow\;\longmapsto}{}} b \text{ (Table III, no. 3)}$$

This amounts to the statement that a without b, and nothing at all $(\bar{a}\bar{b})$, but not b without a, are equivalent to ab in respect of the connexion with B, *i.e.* that in the wider text from which our piece has been taken, '*George went to the cinema*' and nothing, but not '*while Mary had her hair curled*' can occur as the same terminal of the same connexion (of which B is the other terminal) as '*George went to the cinema while Mary had her hair curled*'. As English is a living language, we will assume this wider text to consist of the utterances of an informant—some carefully selected native speaker, over sixty, unspoilt, and with a reasonably complete set of teeth—and our material is thus confined to what he or she is willing and able to write, since we have chosen traditional orthography.

Next, we shall analyse $b=c.d$: ('*while*').('*Mary had her hair curled*'), again because d is the largest part of b which can be identified with anything occurring separately elsewhere in the material. Of the four members of the category of chains for $c.d$ we already know two, *viz.* $cd=b$ and $\bar{c}\bar{d}=\bar{b}$; these can therefore be inserted into the formula without further ado:

$$+acd+a\bar{c}\bar{d}-\bar{a}cd+\bar{a}\bar{c}\bar{d}$$

The two others, $c\bar{d}$ and $\bar{c}d$, are new and unknown, and their connexions with a and with $\bar{a}$ must now be tested to see whether the chains $ac\bar{d}$, $a\bar{c}d$, $\bar{a}c\bar{d}$, and $\bar{a}\bar{c}d$ are asserted or negated as terminals of the connexion with B. The result is as follows:

$$-ac\bar{d}+a\bar{c}d-\bar{a}c\bar{d}+\bar{a}\bar{c}d$$

[1] The example is given with all reservations as to its ultimate validity. In order to get the greatest possible freedom of combination, and to avoid irrelevant difficulties, I have chosen for this content analysis a text expressed in traditional orthography and without punctuation marks. Any phonematic expression, with intonation, stress, and pauses (which, in "real life", would naturally also have to be analysed) would considerably restrict the possibilities of combination, besides needing a great deal of explanation, and we are here primarily concerned with the mechanics of the technique.

and the whole category of *a.c.d* in respect of *.B* is thus

$$\left\{+ab \quad +a\bar{b} \quad -\bar{a}b \quad +\bar{a}\bar{b}\right\} = a\substack{\leftarrow\\ \mapsfrom}\substack{\mapsto\\ \leftarrow}b$$

$$\left\{+acd-ac\bar{d}+a\bar{c}d+a\bar{c}\bar{d}-\bar{a}cd-\bar{a}c\bar{d}+\bar{a}\bar{c}d+\bar{a}\bar{c}\bar{d}\right\} = a\substack{\mapsfrom\\ \leftarrow}\substack{\leftarrow\\ \rightarrow}c\substack{\rightarrow\\ \leftarrow}\substack{\mapsto\\ \rightarrow}d$$

This category is equal to the sum of Nos. 7 and 11 and also to the sum of Nos. 7 and 14 in Table IV; the former solution gives the greater number of assertions, but the sum of 7 and 14 is nevertheless preferable here because, with $\substack{\rightarrow\\ \leftarrow}c\substack{\leftarrow\\ \rightarrow}$, it shows more clearly that it is the presence or absence of *c* (*'while'*) that decides the relation between *a* and *d*.

The deduction-induction is continued in the same way, one analysis following after another, and, at each step, the registration of a category of chains with its relation stated in terms of sets. But there would be no point in going on with it here; a more detailed *Leitfaden* is given in Part II.

Here, with the same reservations, is an example from a somewhat different language, *viz.* Southern Maidu:

'hɔjjam	kawi	nɛɛm	huukɛ	sijɛɛn	hɛɛsɛm	huukum	mɨ	thɛm	wɔɔnɔtʃɛ	
early	time	big	chief	making	old	chief	his	father	when	died

thɛɛbɛjɛ	huukjatɔm	wɔɔnɔhɔm	huukɛ	mɨ	thɛhɛ'
youth	made chief	dead	chief's	his	son

i.e. "in the old days, when (they) created a big chief, when the old chief, his father, died, (the Maidus) made a young man chief, the dead chief's son". It will be seen that there is apparently nothing in the interlinear translation corresponding to the words in brackets in the free translation, and the question now arises on what grounds these additions have been made, and whether a corresponding expansion of the Maidu text would be justified or desirable. The second bracket ("the Maidus") has been taken from the context of situation: the text is part of an account of the procedure for creating a new chief, dictated to me by a Maidu Indian; these accounts always refer to his own people and often, but—as the example shows—not always, do so explicity by means of the term nisɛnaan, which means something like "our people". It is clear that, in abstracting the text from its context of situation, this item must be retained and translated from situation into language; without it, the text is not intelligible. A study of similar texts shows that the expression unit nisɛnaanim could be inserted after hɔjjam kawi without materially altering the content of the text, *i.e.* it is probable that, with a less definite context of situation, the informant would himself have made this addition. But for our structural purposes it is not necessary to make even such a comparatively safe guess: all that is needed is a subjective case ('-m') to act as a terminal in two connexions, (1) with the subjunctive mood of '**sijεεn**' and (2) with the indicative of '**huukjatɔm**'. This addition we shall make because the description of Maidu is thereby considerably simplified, but we must, of course, make a note to the effect that, under certain conditions to be described, the subject is placed outside the text, in the context of situation.

The first bracket ("they") is not really an addition, since the ending of **sijεεn** indicates that this verb has the same subject as '**huukjatɔm**'; the only way to render this in English is by inserting the appropriate personal pronoun, but the Maidu text is, on this point, complete in itself.

Our first analysis will divide the text after '**wɔɔnɔtʃε**', and we shall call the two parts a and b as usual. The category is then as follows:

$$\left\{+ab-a\bar{b}+\bar{a}b-\bar{a}\bar{b}\right\}=a \leftarrow\!\dashv\!\rightarrow b$$

The first chain, ab, is equal to the whole text and is therefore asserted; a without b does not occur anywhere in the wider text, and $a\bar{b}$ is therefore negated; b without a, on the other hand, does occur, so $\bar{a}b$ is asserted;[1] our text is presupposed by the continuation of the account (not printed here), therefore $-\bar{a}\bar{b}$.

The second analysis is $a=c.d$: ('**hɔjjam kawi nεεm huukε sijεεn**'). ('**hεεsεm huukum mɨ thεm wɔɔnɔtʃε**'). Of the eight members of the category of $b.c.d$ four are known, *viz.*

$$+bcd+b\bar{c}\bar{d}-\bar{b}cd-\bar{b}\bar{c}d$$

Of the four new chains two, $bc\bar{d}$ and $b\bar{c}d$, are asserted, the other two, $\bar{b}c\bar{d}$ and $\bar{b}\bar{c}\bar{d}$, negated. The category is, then

$$\left\{+bcd+bc\bar{d}+b\bar{c}d+b\bar{c}\bar{d}-\bar{b}cd-\bar{b}c\bar{d}-\bar{b}\bar{c}d-\bar{b}\bar{c}\bar{d}\right\}=b \leftarrow\!\mapsto c \leftarrow\!\dashv\!\mapsto d$$

[1] If **nisεnaanim** had been manifested in the sequence immediately after **hɔjjam kawi**, $\bar{a}b$ would have been negated too, but that is not the case in our text, and although we have imported a subjective case, it would not be justifiable to assign to it a specific place in the order of the sequence, since the context of situation must be taken to have an equal bearing on all parts of the text in a case like this. However, the whole problem of the context of situation is *mal étudié*.

Third analysis: $b = e.f$: ('thεεbεjε huukjatɔm').('wɔɔnɔhɔm huukε mɨ thεhε'). The category of $c.d.e.f$ has sixteen members, of which eight are known from the preceding operation, *viz.*

$$+cdef-cd\bar{e}\bar{f}+c\bar{d}ef-c\bar{d}\bar{e}\bar{f}+\bar{c}def-\bar{c}d\bar{e}\bar{f}+\bar{c}\bar{d}ef-\bar{c}\bar{d}\bar{e}\bar{f}$$

The whole category is

$$\{+cdef+cde\bar{f}-cd\bar{e}f-cd\bar{e}\bar{f}+c\bar{d}ef+c\bar{d}e\bar{f}-c\bar{d}\bar{e}f-c\bar{d}\bar{e}\bar{f}$$
$$+\bar{c}def+\bar{c}de\bar{f}-\bar{c}d\bar{e}f-\bar{c}d\bar{e}\bar{f}+\bar{c}\bar{d}ef+\bar{c}\bar{d}e\bar{f}-\bar{c}\bar{d}\bar{e}f-\bar{c}\bar{d}\bar{e}\bar{f}\} = c \leftarrow\!\dashv \mapsto d \leftarrow\!\dashv \rightarrow e \leftarrow \mapsto f$$

A comparison of this category with the preceding one shows that it is not the whole of *b* which must be positive but only that part of it which we have called *e* ('thεεbεjε huukjatɔm X-m').

Further analysis would show that the terminals of the selection (cf. No. 39) $c \leftarrow\!\dashv \rightarrow e$ are, not the whole of *c* and *e*, but the subjunctive mood, close contact of *c* ('-n') and the indicative mood of *e* ('tɔm', which also comprises the distant past tense); similarly, the selection $d \leftarrow\!\dashv \rightarrow e$ is carried by the subjunctive mood, loose contact of *d* ('-tʃε')[1] and the indicative of *e*; and the selection $e \leftarrow \mapsto f$ has as its terminals the objective case of 'thεεbεjε' in *e* and the objective case of 'thεhε' in *f*.[2]

38. The relation of a duplex category in which $+a\bar{b}+\bar{a}b$ is called a *combination* (1, 2, 9, 10).

39. The relation of a duplex category in which either (1) $+a\bar{b}-\bar{a}b$ or (2) $-a\bar{b}+\bar{a}b$ is called a *selection* (3, 4, 11, 12; 5, 6, 13, 14). In (1) *a* is said to be *selected*, *b selecting*; in (2) *a* is said to be *selecting*, *b selected*.

40. The relation of a duplex category in which $-a\bar{b}-\bar{a}b$ is called a *solidarity* (7, 8, 15, 16).

41. When, in a duplex category, $+a\bar{b}$, then *a* is called a *major relate*.

42. When, in a duplex category, $-a\bar{b}$, then *a* is called a *minor relate*.

A combination thus has two major relates, a selection has one major (the selected) and one minor (the selecting), and a solidarity has two minor relates.

---

[1] The subject is the junction 'hεεsεm huukum mɨ thεm'.

[2] The second procedure will identify our text as comprising three nexus, *viz.* *c*, *d*, and *ef*; the relations between the moods of *c*, *d*, and *e* are thus heteronexual, while the relation between the cases of *e* and *f* is homonexual.

43. The relation of a duplex category in which $+ab$ is called *conjunct* (1—8).

44. The relation of a duplex category in which $-ab$ is called *disjunct* (9—16).

45. The relation of a duplex category in which $+\bar{a}b$ is called an *absence relation* (odd numbers).

46. The relation of a duplex catogory in which $-\bar{a}b$ is called a *presence relation* (even numbers).

These definitions are parallel to Nos. 27—35 and are designed for a similar purpose, *viz.* the syntagmatic definition of units in terms of binary relations. These binary relations, moreover, are needed for the classification of glossematic elements in terms of which the units of the second procedure are defined; cf. Part II. *Solidarity* is the relation of greatest mutual dependence, neither relate occuring without the presence of the other; the relation registered as a result of the first analysis of a description is necessarily a conjunct presence solidarity: since the object of the first analysis is the whole material, neither relate has any possibility of being encountered without the other. *Combination* is a loose association in which each relate can occur without the other. And in *selections* the dependence is unilateral, one, but not the other, of the two relates occurring alone, *i.e.* connected with the negative of the other. Each relation is, of course, conditioned by the generating connexion in respect of which the members of the category are asserted or negated; it is therefore possible, as we have seen, for two given functives to enter into different, alternative, relations with each other under different conditions, *i.e.* in respect of different generating connexions. The exhaustive binary relation (cf. No. 25) in each case gives a summary of mutual relations within the sequence as a whole.

The same possibility exists here as with the correlations of getting further differentiation by statistical means, and of describing trends in historical development in terms of increasing or decreasing syntagmatic association. This method also lends itself to the differentiation of styles, or dialects, between which there are no glossematic differences. It is, for instance, a common phenomenon that two styles within one and the same (European) language differ by one making much greater use of relative clauses than the other; in both, the relation between principal and relative clauses would be a conjunct selection, probably an absence relation, $a \begin{smallmatrix}\mapsfrom & \leftarrow \\ \leftarrow & \mapsto\end{smallmatrix} b$, where *a* is the principal clause, and the difference would come out as a greater proportion in one style than in the other of *ab* to $a\breve{b}$.

It is, however, always preferable to give a glossematic rather than a statistical differentiation whenever possible. For example, a comparison of greeting-habits would lead to the conclusion that Danes take off their hats much more frequently than Englishmen—a statistical material could easily be provided by stationing observers in Ostergade and in Piccadilly, preferably in the winter. But how much more satisfactory to learn that a Dane will take off his hat when he meets either a woman or another man that he knows—and each time he meets the same person, however often it may be—while the Englishman will uncover only when given permission to do so by a nod from a lady. Statistics is the last resort, to be used only when no stone is left unturned and no avenue unexplored.

47. By a *system* is understood the totality of correlation fields generated by the connexions of one sequence.

48. By a *hierarchy* is understood a sequence together with its system.

System and sequence are, so to speak, two dimensions of the same thing—the thing which we call a hierarchy. A sequence is, by No. 12, the totality of connexion fields registered in one deduction. Each connexion generates a paradigm at each of its terminals, and each paradigm is a member of a category; each unit is consequently a correlate. The sequence is a deductive description of the material as found, in terms of both-and functions, or connexions; the system is an inductive description of that same material in terms of either-or functions, or correlations, from which the original sequence, and an indefinite number of other sequences, can in turn be deduced. Relations, it will be seen, are hierarchical functions, at once both-and, syntagmatical, as they refer to the structure of chains, and either-or, paradigmatical, as they refer to categories. Each functive is thus described both syntagmatically and paradigmatically, and the hierarchy is established by the identity of the functives of its sequence with those of its system. While system and sequence are functional fields, the hierarchy is thus a functival field.

As we have seen, the determination of a given function as both-and or either-or is, in the last resort, arbitrary, a matter of expedience. When, nevertheless, we emphasise the distinction between sequence and system, it is because the decision, once made, is fundamental and binding. The two kinds of function are complementary, in the ordinary sense of the word, and mutually exclusive in this way: a given function can be interpreted as either one or the other but not as both, and the choice is binding for the treatment of all the functions of the hierarchy to which it belongs; the decision is therefore not so much whether a single function is to be regarded as syntagmatic or paradigmatic, as whether a given material is to be treated as a sequence or as a system. On this decision depends the application of the procedure.

The definition of a system is based on that of a sequence because of the peculiar interdependence between syntagmatic and paradigmatic functions: the paradigmatic functions are associated with glossematic places ("either *a* or *b* in a given glossematic place") and a glossematic place is created by a syntagmatic function. Since two given functives may enter into both kinds of function with each other, the interdependence may be very intricate, and in the course of descriptive work the investigator constantly has to go back and forth from one to the other. This can best be illustrated by an example:

Southern Maidu has 23 modal units,[1] equivalent in respect of certain connexions which need not be detailed here. These twenty-three units thus form one paradigm, but it is convenient, in the first instance, to treat them as forming three: ⟨*A*⟩ nine units which are not immediately analysable;[2] ⟨*B*⟩ five units made up of three signs, as follows:

-naa-kha	potential-interrogative
-naa-tʃɛj	potential-mirative
-naa	potential
-kha	interrogative
-tʃɛj	mirative

Since the interrogative and the mirative do not connect, ⟨*B*⟩ is most conveniently treated as two categories of chains:

$$\left\{+pi+p\bar{i}+\bar{p}i-\bar{p}\bar{i}\right\}=p\substack{\leftarrow\!\dashv\;\rightarrow\\ \leftarrow\;\mapsto}i$$
$$\left\{+pm+p\bar{m}+\bar{p}m-\bar{p}\bar{m}\right\}=p\substack{\leftarrow\!\dashv\;\rightarrow\\ \leftarrow\;\mapsto}m$$

⟨*C*⟩ nine units made up of five signs, as follows:

---

[1] "Modal" is here used in the traditional sense, referring to the content substance.

[2] ⟨+ indicative + constative + imperative + desiderative + voluntative + permissive + persuasive + hortative + dubitative⟩.

1	2	3	4		
bi	sa	n		*abc*	prohibitive-sociative-subjunctive
bi		ε		*ac*	prohibitive-subjunctive
bi				*a*	prohibitive
	sa	n		*bc*	sociative-subjunctive
		n/ε		*c*	subjunctive[1]
	ε	ε	thε	*cde*	intentional-subjunctive-concessive
	ε	n/ε		*cd*	intentional-subjunctive
		ε	thε	*ce*	subjunctive-concessive
			thε	*e*	concessive

What we referred to before as one glossematic place now appears as four sign places (the four columns of endings) but since no unit has more than three positive functives, and as the prohibitive and the sociative (a and b) do not connect with the intentional or the concessive (d and e) this paradigm can be treated as two categories of triplex chains, *i.e.* we can regard these modal units as taking up three glossematic places each. The two categories are:

$$\{+abc-ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}+\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\}=a\underset{\leftarrow}{\mapsfrom}\underset{\leftarrow}{\mapsto}b\underset{\rightarrow}{\mapsfrom}\underset{\mapsto}{\rightarrow}c$$
$$\{+cde+cd\bar{e}+c\bar{d}e+c\bar{d}\bar{e}-\bar{c}de-\bar{c}d\bar{e}+\bar{c}\bar{d}e-\bar{c}\bar{d}\bar{e}\}=c\underset{\mapsfrom}{\leftarrow}\underset{\leftarrow}{\mapsto}d\underset{\rightarrow}{\mapsfrom}\underset{\rightarrow}{\mapsto}e$$

These two categories have the same structure: $\underset{\mapsto}{\rightarrow}c\underset{\mapsfrom}{\leftarrow}$ in both; $\underset{\rightarrow}{\mapsto}a\underset{\leftarrow}{\mapsfrom}$ in one, $\underset{\rightarrow}{\mapsto}e\underset{\leftarrow}{\mapsfrom}$ in the other; $\underset{\leftarrow}{\mapsto}b\underset{\rightarrow}{\mapsfrom}$ in one, $\underset{\leftarrow}{\mapsto}d\underset{\rightarrow}{\mapsfrom}$ in the other. The assertions and negations are differently distributed only because of our conventional habit of writing the letters in alphabetical order; if we change this, the distribution is the same:

$$\{+abc-ab\bar{c}+a\bar{b}c+a\bar{b}\bar{c}+\bar{a}bc-\bar{a}b\bar{c}+\bar{a}\bar{b}c-\bar{a}\bar{b}\bar{c}\}=a\underset{\leftarrow}{\mapsfrom}\underset{\leftarrow}{\mapsto}b\underset{\rightarrow}{\mapsfrom}\underset{\mapsto}{\rightarrow}c$$
$$\{+edc-ed\bar{c}+e\bar{d}c+e\bar{d}\bar{c}+\bar{e}dc-\bar{e}d\bar{c}+\bar{e}\bar{d}c-\bar{e}\bar{d}\bar{c}\}=e\underset{\leftarrow}{\mapsfrom}\underset{\leftarrow}{\mapsto}d\underset{\rightarrow}{\mapsfrom}\underset{\mapsto}{\rightarrow}c$$

But since a and b do not connect with d and e, a is not equivalent to e, nor b to d, although, as we know, abc is equivalent to cde, etc.

Having found that the nine units of $\langle C\rangle$ can be described as triplex chains, we must now, in accordance with the Principle of Generalisation, go back and re-examine the nine units of $\langle A\rangle$ and the five units of $\langle B\rangle$ to see if they cannot also be described as triplex chains, *i.e.* as taking up the same number of glossematic places. $\langle B\rangle$ we described earlier on as

---

[1] The subjunctive is expressed by the two alternative endings n and ε, which also express close contact ("same subject") and loose contact ("different subject") respectively.

two categories of duplex chains, but it is easily transformed into a single category of triplex chains, like this:

$$\left\{ -pmi + pm\bar{i} + p\bar{m}i + p\bar{m}\bar{i} - \bar{p}mi + \bar{p}m\bar{i} + \bar{p}\bar{m}i - \bar{p}\bar{m}\bar{i} \right\} = p \overset{\leftarrow|\rightarrow}{\underset{\leftarrow|\rightarrow}{\leftarrow|\leftarrow}} m \overset{\leftarrow\leftarrow}{\underset{\leftarrow|\leftarrow}{\rightarrow\rightarrow}} i$$

There is no chain with three positive functives, and we should not have thought of treating ⟨*B*⟩ in this way if it had not been suggested by the structure of ⟨*C*⟩; however, categories in which the first member is negated have been provided precisely for such cases.

With ⟨*A*⟩ it is more difficult, since the expression gives no clue to the content analysis; this does not make the analysis impossible—any more than it is impossible to analyse '*went*' by analogy with '*want-ed*' etc.—but it does make the analysis less convincing to those who subscribe to the philosophy of naive realism. The advantage of continuing the deduction beyond the minimal sign is that we should thus be able to reduce the inventory: we have already reduced the nine members of ⟨*C*⟩ to an inventory of five and the five members of ⟨*B*⟩ to an inventory of three, and we may look forward with some confidence to obtaining similar benefits from the analysis of the nine members of ⟨*A*⟩. It is, however, not possible to carry through an analysis of this kind until Part II, because it involves syncretism, defectivation, and manifestation – functions which have not yet been defined or explained. At this stage we can therefore do no more than point out the possibilities to be explored, without being able to give criteria for the adoption or rejection of hypotheses. The first thing to be done is to see whether any of the members of ⟨*A*⟩ can be identified with the negated members of the three categories already set up, *viz.* $ab\bar{c}$, $\bar{a}b\bar{c}$, $\bar{c}de$, $\bar{c}d\bar{e}$, pmi, $\bar{p}mi$; any members of ⟨*A*⟩ which cannot be accommodated in this way—and that may, of course, be all nine—must then be treated otherwise. The next possibility is that some or all of the remaining members of ⟨*A*⟩ can be identified with hitherto unrealised chains made up of the functives already registered, *i.e.* *a.d.e*, *b.d.e*, *c.m.i*, etc.; if this could be done, we should need only the eight functives already registered to describe the whole twenty-three units—with the possibility still left open of yet further reduction by analysis. If, after that, any members of ⟨*A*⟩ remain unaccounted for—and it may, again, be all nine —a sufficient number of new functives have to be registered to provide the requisite number of triplex chains, and the question will then arise whether any or all of these new functives can be identified with functives already registered elsewhere in the material.

In this way sequence and system are made to throw light on each other, a continued mutual adjustment designed to produce a description in accordance with our principles.

49. By *derivation* is understood the function between a unit and the chain in which it is a terminal. Symbol $F^n \triangleleft F^N = F^N \triangleright F^n$

50. By the *derivates* of a functive is understood its derivational parts and the derivational parts of its derivational parts, etc.

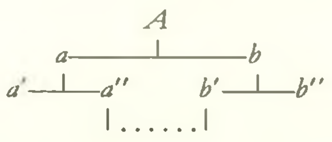
51. By the *degree of a derivate* is understood the maximum number of chains through which it can be derived from a given chain. The degree is symbolised by the appropriate number above the derivation sign, *e. g.* $abc \overset{1}{\triangleright} a$.

52. By the *arrivate(s)* of a unit is understood the chain(s) of which it is a derivate.

53. By the *degree of an arrivate* is understood the maximum number of chains through which the relevant derivate is derived from it.

54. By the *degree of a function* is understood the degree of its terminals in respect of their nearest common arrivate.

Derivation is essentially the same kind of function as that on which a family tree is based: the derivates are descendants, and their degree is counted from the original sequence (the founder of the family) or from any other arrivate (ancestor) that may be made relevant; the arrivates are ancestors, and their degree is counted the other way, from whichever derivate is the centre of interest; the degree of function is a device for calculating collateral relationship through the nearest common ancestor.



If A is the first functive analysed, then a and b are first-degree derivates, and a', a'', b', and b'' are second-degree derivates; a is the first-degree arrivate of a' and a'', and A is their second-degree arrivate. The functions between a and b and between a' and a'', b' and b'' are first-degree functions, and the function between a'' and b' is a second-degree function, since the nearest common arrivate of its terminals, A, is twice removed.

The rule about counting degrees through the maximum number of chains has been made in order to ensure uniformity of treatment: without some such device the degrees of derivates, and therefore of arrivates and of functions as well, would depend on the order of analyses, and this order is in principle fortuitous since analysis is our first means of acquiring knowledge about a material. This can best be shown by an example: the English sentence

$$\overset{a}{\textit{`when}}\ |\ \overset{b}{\textit{he came}}\ |\ \overset{c}{\textit{I went'}}$$

can be analysed as indicated, in two successive analyses, but these two analyses can be ordered in two different ways, *viz.* either (1) $a(bc)$ or (2) $(ab)c$. In case (1) a appears as an immediate derivate of abc, while b and c are derived from abc through bc; in case (2), on the other hand, c is the immediate derivate, and a and b are derived from abc through ab. Furthermore, since the category of $a.b.c$ is

$$\left\{+abc-ab\bar{c}-a\bar{b}c-a\bar{b}\bar{c}+\bar{a}bc+\bar{a}b\bar{c}+\bar{a}\bar{b}c+\bar{a}\bar{b}\bar{c}\right\}=a\,\substack{\rightarrow\,\mapsto\\ \leftarrow\!\mid\,\rightarrow}\,b\,\substack{\leftarrow\!\mid\,\mapsto\\ \longleftrightarrow}\,c$$

both b $(=\bar{a}b\bar{c})$ and c $(=\bar{a}\bar{b}c)$ are encountered in the material without derivation from abc at all. Unless we make up our minds, once and for all, how to deal with problems of this sort, we shall therefore be in doubt at every step and run the risk of perpetrating an inconsistent description. The rule enables us to decide that b and c must be derived from abc and that, like a, they are second-degree derivates of abc. The rule is, admittedly, arbitrary: the effect of dispelling doubt could be obtained equally well by the opposite rule, of always counting through the minimum number of chains. But as the object of reckoning derivation is to afford a further means of differentiation, it is an advantage to have as many degrees available as possible.

Any given unit, say c in the example above, may of course be a derivate of a large number of alternative chains, which are, in that respect, equivalent: '*(as soon) as he came, I went*', '*when she sang, I went*', etc. We might therefore envisage a derivational definition of each unit in the shape of a catalogue of all its equivalent arrivates of each degree. Such a method applied to a reasonably ample material would be extremely laborious: there are enormous numbers of low-degree chains, and since the establishing relation in these regions of the hierarchy is apt to be combination, the amount of differentiation obtained would hardly repay the toil. What is needed for the statement of final results is not so much a technique for indicating that '*I went*' is derived specifically from '*when he came, I went*' (although that must be a necessary intermediate stage) as some way of determining the *order* of chain from which it is derived, without reference to the particular components thereof. This, as we shall see in Part II, is supplied by the second procedure, where units are reclassified and graded, and chains are assigned to orders by the grades instead of by the individualities of their components.

The algebra we have presented here, in Part I, is universal, *i.e.* its application is not confined to materials of any particular kind, and it is thus not specifically linguistic, or even humanistic, in scope or character, though our main purpose in designing it has been to provide for the description of linguistic and other humanistic materials. Its aim is to furnish a calculus of non-quantitative functions, and its application to a material is intended to result in a description of that material in terms of relations, correlations, and derivations. In accordance with the Principle of Reduction the description takes the form of a procedure, which, as we have seen, first of all reduces the material to a sequence of units strung together by connexions; the connexions entered into by each unit are compared, and registered in the form of binary relations; the paradigms generated by each connexion are collected together into categories, and each unit receives a paradigmatic definition registered in the form of binary correlations; finally the place of each unit in the hierarchy is defined in terms of its derivational degree, which makes differentiation possible even though the same relations and correlations recur on each level.

The algebra is a calculus of glossematic possibilities, and the description of a material consists in indicating which of these possibilities are realised in that material, so that any completed description is immediately ready for comparison with any other completed description in terms of the relation of each to the general calculus; this will tend, in some cases, to make the description more cumbersome than if the descriptive apparatus had been designed with only that one material in mind, but this disadvantage is easily outweighed by the advantage of not having to recast any description for comparative purposes.

The sequence of this book contains the details of the procedure and describes the (general but not universal) algebra of structures comprising more than one hierarchy, *i.e.* of stratified descriptions—another calculus of functions designed to provide a means of self-consistent, exhaustive, and simple descriptions of those materials to which it is applicable. This further albegra is thus also defined in general terms and not as specifically linguistic or humanistic, although, again, its primary purpose is to provide for the description of linguistic and other humanistic materials. It may be that its scope will prove to be wider than we can now foresee.

NOTE

The circle in Nos. 16 and 17 is in so far only apparent as complete identity is not required in the second terminal. Let there be two connexion fields, *a.b* and *c.b′*; if *b* and *b′* are equivalent in respect of the connexion with *a*, that is $a.(+b+b')$ and also in respect of the connexion with *c*, that is $c.(+b+b')$, then *a* and *c* are equivalent irrespective of whether *b* and *b′* can ultimately be proved to be equivalent in respect of *all* relevant functions.

It often happens, however, that *b* and *b′* cannot be demonstrated to be equivalent at all, because they cannot be brought into the same glossematic place. If, for example, you wish to establish the equivalence of i: and ɑ: in English on the grounds that they both occur with f in f_1i: and f_2ɑ:, then you need to show that f_1 and f_2 are identical or, at least, equivalent in respect of the two connexions; but nowhere in the material will you find f_1ɑ: or f_2i:. It is true that, by judicious cutting and splicing of sound-film or magnetic tape, you can replace f_1 by f_2 and vice versa, but this would make you dependent on a sort of plebiscite of native speakers to decide whether a significant disturbance had been created by the replacement—a highly distasteful situation, since there is no guarantee of a unanimous vote.

The principle of complementary distribution, whereby f_1 and f_2 can be identified if it can be shown that they never occur in the same surroundings, does not provide a solution, since it fails to meet the argument that all events are unique, so that everything is in complementary distribution: the principle presupposes recognisable difference, which implies the concept of identity.

In everyday life we go on the assumption that two events are instances of one and the same unless there is some special reason to suppose that they are different: when a man wakes up in the morning, he takes it for granted that his bed is the one he got into last night and that he, himself, is the same person as the one who went to bed the night before. This assumption is philosophically unjustifiable and, indeed, often leads to error; it is also a practical necessity and so fundamental that, when its application proves fallacious—when the bed is different after all—we are seized with panic comparable to that induced by an earthquake.

I can see no way out but to admit in all humility that this assumption is necessary even in science. The best we can do is to make it explicit and to be constantly aware of its inherent danger. Our definition of identity can then be seen to be what it really is: a definition of the cases in which the assumption of identity can safely be made.

In our example, above, the equivalence of i: and ɑ: is thus established on the assumption that f_1 and f_2 are identical. This can do no harm as long as we remember that there is a hypothesis involved and are willing to drop that hypothesis, and any conclusion based upon it, if it becomes necessary to do so.

INDEX OF TECHNICAL TERMS

Italics indicate number of page

PUBLICATIONS BY H. J. ULDALL

1. Det danske r. *Danske Studier*, 1928, pp. 172–75.
2. Danish (text in phonetic ranscription). *Le Maître Phonétique* 23, 1928, pp. 56–57.
3. A Suggestion. *Le Maître Phonétique* 24, 1928, pp. 73–74.
4. West Norwegian. *Le Maître Phonétique* 26, 1929, p. 20.
5. Review of M. G. Mori, The Pronunciation of Japanese (1929). *Le Maître Phonétique* 30, 1930, p.35.
6. Review of Pablo Wernstedt, Libro para la enseñanza del Aleman. *Le Maître Phonétique* 30, 1930, p. 35.
7. Review of Olof Gjerdman, Studier över de sörmlandska stadmålens kvalitative ljudlära II. *Le Maître Phonétique* 30, 1930, p. 36.
8. Dansk (text in phonetic transcription). *Le Maître Phonétique* 30, 1930, pp. 37–38.
9. Maidu. *Le Maître Phonétique* 33, 1931, pp. 8–9.
10. Pomo. *Le Maître Phonétique* 37, 1932, pp. 5–6 (with Jaime de Angulo).
11. The Perception of Speech. *Le Maître Phonétique* 41, 1933, pp. 3–5.
12. A Sketch of Achumawi Phonetics. *International Journal of American Linguistics* 8, 1933, pp.73–77.
13. A Danish Phonetic Reader, University of London Press, 1933.
14. Eskimo. *Le Maître Phonétique* 47, 1934, pp. 76–78 (with A. C. Lawrenson and William Thalbitzer).
15. A Note on Vowel-length in American English. *Le Maître Phonétique* 48, 1934, pp. 97–98.
16. Maidu, et kalifornisk sprog. *Bull. du Cercle Linguistique de Copenhague* I, 1935, pp. 6–8.
17. The Phonematics of Danish. *Proceedings of the Second Intern. Congr. of Phon. Sciences*, London 1935 (1936), pp. 54–57.
18. Louis Hjelmslev and HJU, Synopsis of an Outline of Glossematics. Aarhus, 1936.
19. Review of J. Saxe, Bernard Shaw's Phonetics. A Comparative Study of Cockney Sound-Changes. *Gymnasieskolen* 2, 1937, p. 26.
20. Speech and Writing. (Summary). *Congrès international des sciences anthropologiques et ethnologiques*, 1938, p. 374.
21. On the Structural Interpretation of Diphthongs. *Proceedings of the Third Intern. Congr. of Phon. Sciences*. Ghent 1938, pp. 272–276.
22. Review of D. M. Beach, The Phonetics of the Hottentot Language. *Africa* 12, 1939, pp. 369–372.
23. The General Rules of English (R. A. Close with H. J. Uldall). Institute of English Studies, Athens 1940. Translated into Greek by Patroclus Eliascos.
24. Speech and Writing. *Acta Linguistica* IV, 1944, pp. 11–16, reprinted in *Readings in Linguistics* II, ed. E. Hamp, F. W. Householder, and R. Austerlitz, Chicago (1966).
25. The English Vowels. Pamphlet to accompany filmstrip No. 86 produced by the British Council, 1945.
26. Danish (Dansk). *Le Maître Phonétique* 85, 1946, pp. 3–4.
27. The English Vowels. Practice Lists. The British Council, Buenos Aires (1947).
28. Notes on the English Tenses. *English Language Teaching* 2, 1948, pp. 122–128 and 147–153.

Reprinted in *E. L. T. Selections* 1, articles from *English Language Testing*, ed. W. R. Lee, London 1967.

29. On Equivalent Relations. *Recherches Structurales 1949. Travaux du Cercle Linguistique de Copenhague* V, pp. 71–76.
30. Maidu Phonetics. *International Journal of American Linguistics* 20, 1954, pp. 8–16.
31. Hermann Vinterberg og C. A. Bodelsen, under medvirkning af H. J. Uldall og B. Kjaerulff Nielsen, *Dansk-Engelsk Ordbog*. København 1954.
32. Louis Hjelmslev and HJU, *Outline of Glossematics, A Study in the Methodology of the Humanities with Special Reference to Linguistics. Part I. General Theory* by H. J. Uldall. *Travaux du Cercle Linguistique de Copenhague* X, 1957. Russian translation: *Osnovy glossematiki. Novoye v Lingvistike* I, Moscow, 1960.
33. On the Preparation of a Text. *Archivum Linguisticum* 11, 1959, pp. 1–17.
34. HJU and William Shipley, *Nisenan Texts and Dictionary, University of California Publications in Linguistics*, vol. 46, 1966.