te reo

1965

Proceedings of the Linguistic Society of New Zealand

8

TE REO IS PUBLISHED BY THE LINGUISTIC SOCIETY OF NEW ZEALAND, AND SUPPLIED FREE TO MEMBERS OF THE SOCIETY. THE SERIES OF **TE REO MONOGRAPHS** IS ALSO AVAILABLE AT SPECIAL RATES TO MEMBERS. MEMBERSHIP RATES ARE AS FOLLOWS (IN NEW ZEALAND CURRENCY):

AUCKLAND RESIDENTS:	£1. 0. 0
STUDENTS:	5. 0
ALL OTHERS:	10. 0

IN THE PUBLICATION OF ARTICLES IN **TE REO**, PREFERENCE WILL BE GIVEN TO ARTICLES SUBMITTED BY MEMBERS. MANUSCRIPTS SHOULD BE SENT TO THE HONORARY SECRETARY, LINGUISTIC SOCIETY OF NEW ZEALAND, C/O UNIVERSITY, BOX 2175, AUCKLAND. THEY SHOULD BE TYPED ON QUARTO PAPER, ON ONE SIDE OF THE PAPER ONLY, AND DOUBLE-SPACED. PREPARATION OF THE COPY TO BE PRINTED IS MUCH EASIER IF THERE ARE NO HANDWRITTEN ADDITIONS. ALL WORDS UNDERLINED WILL BE PRINTED IN ITALICS. FOOTNOTES SHOULD BE NUMBERED CONSECUTIVELY THROUGHOUT THE TEXT, AND TYPED SEPARATELY ON AN ADDITIONAL SHEET. REFERENCES GIVEN IN THE FOOTNOTES SHOULD GIVE ONLY THE SURNAME OF THE AUTHOR, THE DATE OF PUBLICATION AND THE PAGE NUMBER:

[1] Smith, 1954: 249.

WHERE REFERENCES ARE MADE TO TWO OR MORE WORKS PUBLISHED BY THE SAME AUTHOR IN THE SAME YEAR, THEY SHOULD BE DISTINGUISHED BY THE ADDITION OF *a*, *b*, ETC. TO THE YEAR:

[1] Smith, 1954a: 249.
[2] Smith, 1954b: 492.

ON A FURTHER ADDITIONAL SHEET SHOULD BE GIVEN A COMPLETE BIBLIOGRAPHY OF ALL THE REFERENCES CITED, ARRANGED IN ALPHABETICAL ORDER OF AUTHORS, AND WHERE SEVERAL WORKS BY THE SAME AUTHOR ARE INCLUDED, IN CHRONOLOGICAL ORDER. THE PLACE OF PUBLICATION AND THE NAME OF THE PUBLISHER SHOULD BE QUOTED FOR BOOKS. THE TITLES OF BOOKS AND OF PERIODICALS SHOULD BE UNDERLINED; THE TITLES OF ARTICLES PUBLISHED IN PERIODICALS SHOULD BE PLACED IN INVERTED COMMAS. EXAMPLES OF THE PRINTED VERSION:

Smith, X.Y., 1954a. "The Phonemes of Kerguelen." *Journal of the Phonemic Society*. 49: 249-257.
---, ---. 1954b. *A Grammar and Vocabulary of Kerguelen.* Kerguelen, Smith and Smith.

AUTHORS OF ARTICLES PUBLISHED RECEIVE 20 FREE REPRINTS. IF MORE ARE DESIRED, THE AUTHOR SHOULD COMMUNICATE WITH THE HONORARY SECRETARY.

BACK ISSUES OF **TE REO** ARE 7S.6D. EACH FOR VOLS. 1-5; 10S. FROM VOL. 6.

ALL SUBSCRIPTION RATES AND PRICES INCLUDE POSTAGE AT 2ND-CLASS MAIL RATES. PAYMENTS FROM WITHIN NEW ZEALAND SHOULD INCLUDE EXCHANGE. PAYMENTS FROM OUTSIDE NEW ZEALAND SHOULD INCLUDE BANK CHARGES AND SHOULD BE DRAWN ON AUCKLAND. THE SOCIETY HAS ACCOUNTS AT THE AUCKLAND SAVINGS BANK (NO. 18603) AND AT THE AUSTRALIA AND NEW ZEALAND BANK (NO. 86-72-418).

TE REO

Proceedings of the Linguistic Society of New Zealand

Volume 8 **1965**

Published for the Linguistic Society of New Zealand by the Wellington Branch, Victoria University of Wellington.

MEETINGS HELD IN 1965

AUCKLAND BRANCH

April 28 — Mr. C. C. Bowley read a paper on the structure of the verbal group in English.

July 20 — Mr. A. K. Pawley read a paper entitled "Pruning the Polynesian Family Tree", in which he reviewed evidence for the subgrouping of the Polynesian languages.

September 8 — Mr. F. R. Wyatt read a paper on mechanical translation.*

October 28 — The Annual General Meeting was followed by the reading of a paper from Dr. O. Blixen (Montevideo) in which the honorific terms of Uvean were compared with those of other languages of Western Polynesia.

Auckland Officers for 1966

President:	Professor F. S. Scott
Vice-Presidents:	Dr. B. G. Biggs; Dr. A. W. H. West
Hon. Secretary:	Dr. W. J. Kirkness
Hon. Treasurer:	Mr. W. O. Droescher
Hon. Auditor:	Mr. J. E. Field
Committee:	Mr. C. C. Bowley; Mr. J. C. Corne; Mr. A. K. Pawley; Mr. D. S. Walsh.

WELLINGTON BRANCH

The following papers were read during the 1965 year:

March 25 — Professor C. J. Adcock *Some Requirements for an International Auxiliary Language*.

June 9 — Professor H. A. Murray *The Stylistics of Satire*.

July 29 — Professor G. E. Hughes *Speaker-Committing Verbs*.

September 9 — Mr. Colin Bowley, Auckland University, *Multiple Syntagmatic Relations in English Clause Structure*.

October 21 — Annual General Meeting followed by a language-film evening and discussion.

Wellington Officers for 1966

President:	Professor L. F. Brosnahan
Vice-Presidents:	Mrs. Helene L. Woolston; Mrs. J. G. Gordon
Hon. Secretary:	Mr. H. W. Orsman
Hon. Treasurer:	Mr. R. D. Hughes
Committee:	Mr. D. A. Cooke; Mr. R. L. Fountain; Mr. R. W. Kerr; Dr. D. G. McArthur.

NATIONAL COMMITTEE

The National Annual General Meeting of the Society was held on 8 December, 1965 in Auckland. The following officers were elected to the National Committee for 1966:

Patron:	Sir Gilbert Archey
President:	Professor L. F. Brosnahan
Hon. Secretary:	Dr. W. J. Kirkness
Hon. Treasurer:	Mr. W. O. Droescher
Hon. Auditor:	Mr. J. E. Field
Committee:	Dr. B. G. Biggs; Dr. D. G. McArthur; Mr. H. W. Orsman; Mr. D. S. Walsh (editor).

Following the National Annual General Meeting Mr. C. C. Bowley read a paper on Multiple Syntagmatic Relations in English Clause Structure.*

* Published in this issue.

CORRIGENDA

Volume 6 (1963) Supplement, or Reprint No. 1, Français régional de l'Indo-Pacifique.

p.14, line 1: instead of |e|, read |E|

p.16, Indications bibliographiques, read: Hollyman 1962a.

p.18, under |Œ| read: |fIIœl| filleul: |fIIœ œl| filleule

p.18, under Corrélation nasale, read [vẽ]~[veŋ]; and in the following line:
En SFF, on a [ṽC]~[ṽnC]~[ṽnÇ]

p.24, in line 5 of the paragraph on |E| read: |pikɛ| piquet, piquait

p.25, line 1, read: En SFF: [u]~[u:]

p.29, line 9 of text, read: et des diverses variantes

Volume 7 (1964)

p.2, footnote 4, line 3, for low u, ô, k read; back u, ô, o.

p.8, 9 lines from bottom, read: jop 'horsefly'

p.8, 8 lines from bottom, read: chu'p 'cockroach'

p.11, 6 lines from bottom, for mouse, deer read: mouse deer

p.11, 3 lines from bottom, for runn- read: run-

p.11, 2 lines from bottom, for -ing read: -ning

p.13, 7 lines from bottom, for lêng read: lênh

p.15, footnote 1, read: Hua Thi Them

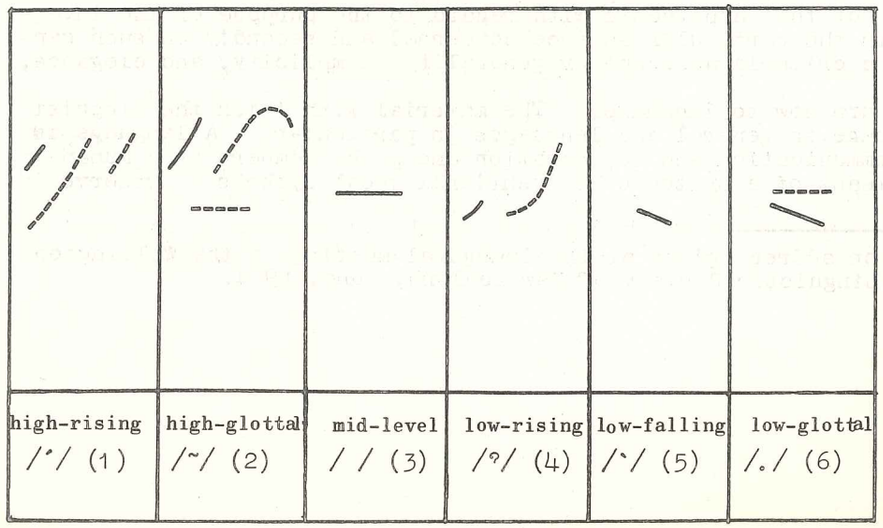
p.16, the diagram "Figure 1. Chart of Tones" should be as printed below.

p.18, line 11, instead of ɬak 'to wash', read: tak 'to wash'

p.37, line 11 of text, read: to say [ɻip] for Dieb

p.41, the last triad in the oral consonant correlation should be kh, k, ɣ

Figure 1. Chart of Tones



THE SCIENCE OF LINGUISTICS [1]

L. F. Brosnahan

(Victoria University of Wellington)

In briefest form, linguistics may be described as the scientific study of languages. The term 'study' needs little further discussion here - we understand by it the application of the intellectual faculty to the material which is to be studied. The qualification 'scientific', however, is a little more complex. When we say that a study is 'scientific', what in fact do we mean? It seems to me that what characterises scientific study from what I might call 'unscientific' study is that it is devoted primarily to the analysis and investigation of the material concerned in a search for features of order or of regularity therein, that is, to the uncovering or the establishment of organization and patterning in that material. There is a further process - the classification of the analysed material to display the patterns so arrived at, but this is a subsequent operation, even though there may well be a sort of 'feed-back' from the display attempts, a feed-back which suggests new features of order in the material.

The procedure which the scientist adopts is basically one of generalization and hypothesis with regard to the material: that is, he creates ideas or hypotheses as to the nature of the features of order and patterning which may exist within the material, and subsequently tests these hypotheses by applying them to the material. Scientists themselves frequently believe that their progress depends on analysis and classification of the material and a quasi-logical extrapolation of it, but in practice advance in a science is now held to come primarily from those guesses or hunches or intuitions with regard to the material, which we have here called hypotheses.

The safeguards erected by scientists for science are basically of two types. The first is that of publication: the scientist publishes his hypotheses and the results of his testing so that they are available, in theory at least, to all other scientists who are interested in the same mass of material. The second safeguard is repetition: published accounts of scientific work are ideally in such form that the work may be repeated by other scientists as a check upon its validity.

The evaluation of a hypothesis is based first on the usefulness or convenience of that hypothesis with regard to the purpose of the investigation within the particular science concerned and secondly on such general scientific criteria as scope or generality, simplicity, and elegance.

To turn now to language. The material with which the linguist deals is language in general and languages in particular. A language is a method of communication and co-operation among the members of a human community by means of a system of conventional vocal symbols. Observe

---

[1] The text of an address given at the inaugural meeting of the Wellington Branch of the Linguistic Society of New Zealand, June, 1964.

that in this definition, language is clearly conceived as a possession of the community. A language is a system which is, in a very clear sense, available to all the members of a community, but is not possessed in its entirety by any single one of them. The speaker in his speaking makes reference to the system of the language and the hearer refers what he hears to the same system and the correspondence of the features referred to in the system constitutes the communication between the two. Speech, then, is the individual's using of the system of language. It is concrete and real in the sense that what is spoken exists in the world of energy, time and space. Language, on the other hand, is an abstract system which exists in the minds of and by virtue of the speakers as a community but is nevertheless, in a very clear sense, superordinate to them as individuals.

Language is primarily a system of symbols. It is the symbol upon which all language depends. A symbol is a structure of two parts; one part consists of a perceptible entity, something which can be perceived by one or other of the modes of the human sensory apparatus. In language, the material from which this part of the symbol is constructed is sound and the stretch of sound from which a symbol is composed is known as its <u>form</u>. The other part of the symbol is an idea, a concept, a segment of experience: in linguistics this is the <u>meaning</u>. The two parts of the symbol, the form and the meaning, are linked by a relationship of association which enables either of these, if occurring in consciousness, to call up the other. This relationship of association is a purely conventional and arbitrary one, which must be learned separately for each symbol during the course of apprenticeship to the language.

Since the study of language must focus on the symbol, we might expect two major branches of linguistics; in fact, however, that branch which concerns itself with the form of the symbol is itself divided into two sections. The form of the linguistic symbol shows a dual patterning, and each of these patterns is the province of a separate branch of linguistics. The whole of the study of language thus consists of three branches - two parts concerned with the form of the symbol and one part with the meaning.

The first branch is the study of form at the lower level. This branch is named <u>phonology</u>. The subject matter of this branch is the sounds, stresses and pitches in which the forms of language are manifested by the processes of articulation. It also includes the significant units at this level, the phonemes and the prosodies. These units bear significance by their opposition to one another - |p| is not |b| is not |f| etc. - and it is this function of being different which renders them significant units of form as far as the whole language is concerned. They do not, as phonemes or prosodies, carry meaning.

The second branch of linguistics is that which deals with form at the higher level where meaning is involved; this branch is termed <u>grammar</u>. The forms at this level are patterned sequences of phonemes and prosodies which are associated with meaning. The units here are morphemes, the minimum meaningful forms of the language, words, the minimum sentence-forming units, and the constructions, the more complex meaningful units of the language. Grammar is itself frequently subdivided into the two sections of <u>morphology</u> which deals with the simpler meaningful units, the

morphemes, words, and syntax which deals with the ways in which these units are put together to form more complex units.

The third branch of language is the study of meaning: this is semantics, the branch in which linguistics has made least progress. The difficulties are mainly those which are apparently inherent in the study of meaning and which have bedevilled philosophy and metaphysics from the beginning of human inquiry. The meanings which are associated with the forms of the language are concepts, ideas, relations and such like, which are derived from the experience of the community, either directly or indirectly, and in an abstract or generalised form associated with the forms. Strictly speaking, the meaning which each member of a speech community associates with each form or construction derives from his own individual experience and hence is unique and individual to him alone, but language as a system of communication can operate only as a result of a contract social, a general convergence and agreement among the members of that community with regard to the meaning which they associate with each of the forms. As far as language is concerned, then, the meaning of a particular form is a sort of highest common factor of the individual meanings associated with that form by each of the users of the language.

The study of language is one of respectable antiquity and yet it does not date, as far as we know, from anything like the period of the beginning or the origin of language. Recent work in the study of human evolution suggests that human or human-like men have been constructing tools, and hence (we assume) speaking, for a very large proportion of the last million years, but the beginning of the study of human language seems hardly to be older than a few thousand years. In the first millenium B.C., in three separate areas, South-eastern Europe, India, and China, separate peoples independently seem to have reached that stage of critical detachment at which they were able to take an interest in the nature and function of language. In the West, from Greek times, through the Roman grammarians, the early Fathers of the Church, and the scholastic philosophers of the medieval period, the study of language was kept alive. In the later Middle Ages it formed an integral part of the studies of the newly developing universities. With the Renaissance and world exploration came new intellectual interests in the nature and role of language and a wider range of languages for investigation, but it was not until near the end of the eighteenth century that reasoned and objective study began to replace what seems to us today interesting, but very largely subjective, speculation with regard to the nature and origin of language.

The modern study of language dates from the very end of the eighteenth century and its florescence in the nineteenth century must be regarded as one of the great triumphs of the human mind. The realization of the relatedness of languages, and of development and change in languages through time was a remarkable achievement and quite fit to rank in the humanities alongside the Darwinian concept of evolution in the biological sciences. In the twentieth century, the emphasis in linguistics has shifted from the historical to the descriptive. It has been realised that progress in the understanding, the nature and the functioning of language and in the understanding of its evolution and history is dependent upon improvement in the accuracy, precision and completeness of our methods of describing language.

From the study of many languages both as they appear at points in time and as they have developed over long periods of their known or inferred history, the linguist has come to two main conclusions. The first is that language is an orderly and organised phenomenon. The concept of language as a system is basic to modern linguistics. A language does not consist of an aggregation of words, or of strings of units in use, but of a complex of integrated and interlocking systems, and each part of a system holds its position by virtue of its contrast with every other part. Language is a system of oppositions, as it has been said, and these oppositions are amenable to scientific study and to objective description.

The second main conclusion is that languages may be related and these relationships are of different types. The first type of relationship is typological, that is, related to the structural system or type of the language. For example, two languages may show similarities with regard to corresponding features of one or more systems. As a simple example, we note that the syllable structure in Maori of New Zealand and in Yoruba of West Africa is very similar: the syllable consists of either a vowel or of a single consonant followed by a vowel, no clusters of consonants occur initially in the syllable and no consonants occur at the end of the syllable. A second example, at the grammatical level, is the similarity between Latin and Turkish in the case system of the noun. In both these languages, the noun has six cases, five of them closely corresponding with regard to their functional significance within their own grammatical systems. Relationships of this nature are said to be typological.

The second type of relationship between languages is with the linguists have come to call by the rather unfortunate term genetic; that is, related as to origin. Two languages are said to be in genetic relationship if they have ever been one and the same language; thus, the English language of the twentieth century and the English language of the tenth century are in genetic relationship. French and Spanish of today are also genetically related, both being modern developments or descendants of the same language, Latin, and Icelandic and Bengali likewise, being developments from Indo-European. In the same way, Hebrew, Arabic, Amharic and Hausa have been shown to be genetically related, as developments from a proto-Semitic language; Twi of West Africa, Zulu of South Africa, and Swahili of East Africa are genetically related descendants of a proto-Bantu ancestor and so on.

The third type of relationship between language is a relationship derived from contact and inter-influencing. Whenever speakers of two languages, or even of two forms of one language, come into contact with one another, a bridge is opened for borrowing from one to the other. Few languages in human history have existed for a long period in complete isolation, and most languages bear the marks of contact with others. In some, such marks are but traces, as in a few words such as boomerang, corroboree English reveals its contact with the autochthonous languages of Australia; in others the borrowing and influencing has been extensive and long-continued, as in Albanian, in which the borrowings from Greek, Romance and Slavonic have almost swamped the native words in the vocabulary. A further interesting reflection of this type of relationship, a type of semantic borrowing, is found in Europe in the widespread occurrence of forms of similar meaning as a farewell: "Au revoir", "A rivederci", "Auf wiedersehen", "Viszontlátásra".

The Tasks of Linguistics: The first ask of linguistics is a simple one; it is to record and describe the languages of the world and in so doing to develop and test improved methods and techniques of such description. It is today clear to the linguist that the continued development of the understanding of the nature and functioning of language is dependent upon improvement of methods of describing the material of this science.

The second task of linguistics is the expanding and deepening of our understanding of language. We can usefully divide this into two parts. In the first place, the linguist is concerned with what language is as language, what its nature is, what its units and systems are and what the relations between these are, how it functions and how it develops and what the relations between languages are. In the second place, he is concerned with language and the rest of human life and culture, and how language affects and is affected by other aspects of human activity. Language impinges on and penetrates all human thinking, all human doing and all human activity, and the linguist's interest in and study of language in this role will depend upon co-operation and assistance from a large range of specialists.

Consider briefly some of the important aspects. Language is in an important sense a vocal phenomenon produced by activity of the human nervous system and the human vocal apparatus. For the understanding of this aspect of language, the linguist requires the co-operation of the anatomist, the physiologist and the neurologist. Then, processed into speech, language emerges from the speaker in the form of information-bearing sound-wave signals, and for the study of these the linguist must rely on the help of the acoustician, physicist, mathematician and communications engineer. Again, language is a complex series of activities in the brain, of perception and of thinking, and for the understanding of this aspect the co-operation of the psychologist and the philosopher is necessary. Language is also something which the individual acquires during the period of his growth and development; it is part of the process of socialisation and of the development of his personality, and for the understanding of language in this aspect, the linguist requires assistance from the pediatrician, the child psychologist, the educationist and the social psychologist. And, further, since the course of this development is not always untroubled, we may need to call on the speech pathologist, the logopaedist, the speech therapist and in exteme cases even the psycho-analyst, for their contribution to the understanding of language.

Language is also a social phenomenon. It is indeed the main means by which the interaction of human beings and hence human society, as we understand it, are rendered possible. A language is the chief medium in which a community stores the knowledge which it has of itself, its history, and its environment, and the medium by which a community transmits its knowledge and wisdom from generation to generation. Language is also the primary means by which peoples interact with peoples, and by which ideas are conveyed from society to society - we need only think of the Renaissance in Europe in the sixteenth century or the impact of the West on African cultures in the twentieth century - with results that have changed and are changing the world as we know it. For the study of this aspect of language, we must call on the historian, geographer, anthropologist, sociologist, and ethnographer.

Thirdly, but of equal importance, language is also a biological

phenomenon. It is the most characteristic property of our species, that property which separates us most clearly from all other living things and that property which more than any other makes us human. It is a property, however, which is based on and depends on bodily structures and bodily processes which are transmitted from generation to generation by the mechanism of heredity. As a property of a biological species, it has been exposed through millennia, indeed through the course of all human history, to the influence of natural selection: we need have little doubt that a great deal of the perfection of our present-day languages is due to evolutionary advantages which improvement in language and language ability bestowed on certain humanoid and human groups in the struggle for existence within a specific evolutionary niche. For the study of this aspect of language, we must call on the biologist, the geneticist and the evolutionist.

This does not exhaust the list of aspects of language. Language is also the medium of professional use of the writer, the journalist, the poet, the statesman, the demagogue, the advertising man, the public relations officer; and it is a medium of daily use to all of us within human society. Some of the linguistic understanding of language drives and will continue to derive from the use of the ordinary user. In short, linguistics is that branch of human activity which is concerned with the study of language as the most overtly and characteristically human property, as the indispensable basis of human life and culture as we know it. and with the study of languages as the differing and changing manifestation of this property; but the manifold nature of the object of its study is such that linguistics needs and acknowledges its need for the co-operation and assistance of a wide range of scholars, scientists and users.

MULTIPLE SYNTAGMATIC RELATIONS IN ENGLISH CLAUSE STRUCTURE

C. C. Bowley

(University of Auckland)

This paper[1] deals with the problem of describing certain kinds of clause in English which until comparatively recently have received little searching consideration from linguists.[2] Any native speaker of English would agree that clauses such as

(1) The man is able to go.

(2) The man is easy to please.

(3) The man is clumsy to fall.

while superficially similar, are in fact grammatically dissimilar. It is the two-fold task of the linguist (a) to devise formal tests which will enable him to discover the differences which exist intuitively for the native speaker, and (b) to devise appropriate structural descriptions embodying the significant relations revealed by his textual operations.

This paper falls into three sections. Section I outlines the theory and method of description which has been applied. Section II discusses some past and current solutions to the descriptive problem. Section III suggests how the two tasks outlined above can be fulfilled.

I

A grammatical description of English consists of a set of descriptive categories which are related to each other and to English sentences in certain defined ways. "Clause", "group", "word", "subject", "predicator", etc. are all descriptive categories in English. The nature of these descriptive categories and their relations to each other are in turn specified in the general theory of grammar.

Four kinds of descriptive categories are needed to account for all the grammatical features of a particular natural language such as English. They are class, structure, unit and system. For the purposes of this paper I shall restrict myself to a brief and partial consideration of the categories of class, structure and unit.

Units are stretches of language that carry patterns. English needs five units in all to account for the kinds of pattern that are carried by different stretches. These are the units: sentence, clause, group, word and morpheme. The units are related rigorously to one another in fixed relations. The relations are fixed so that sentences consist of clauses, clauses of groups, and so on down to the smallest unit, the morpheme. This paper will be mainly concerned with the ranks of sentence, clause and group.

---

[1] This paper was first read at the Wellington Branch of the Linguistic Society of New Zealand, September, 1965. I am indebted for many of the ideas in this paper to M. A. K. Halliday and J. Sinclair whose lectures on Modern English Grammar I attended in 1961-62.

[2] Some well-known grammars contain little or no consideration of these clauses. See, for example, Francis 1958, Fries 1957, Long 1961, Sledd 1959.

Each unit is expounded by a set of different structures. A structure is a string of one or more elements. An element is a place in structure where one grouping of the items belonging to the rank next below operates. The structure of the English sentence, for example, can be expressed by means of the elements α and β. α is the place in structure where operates the independent class of the clause: β is the element at which operates the dependent class of the clause, that is clauses which do not usually occur alone as sentences. In

He'll be welcome if he comes.

"if he comes" is an exponent of a dependent clause, while "he'll be welcome" is an exponent of an independent clause. This formal item, therefore, is an exponent of the structure αβ. Primary classes are defined by reference to elements of structure. At each different element of a given structure operates a different class of the items of the rank next below. At the elements of the sentence, for example, as we have just seen, operate the classes of the clause.

Here is a brief sketch of the structures and classes of the sentence, clause and group which will be referred to in Section III.

Sentence:

elements of structure: α, β

Clause:

classes: independent (at α)

dependent (at β)

elements of structure: Subject (S), Predicator (P), Complement (C), Adjunct (A)

The subject is the place in the clause at which operates the nominal group immediately preceding the Predicator.

S P

e.g. ||| John | went.|||[3]

The Predicator is the place at which the verbal group operates. The Complement is the place at which operates the nominal group which normally follows the Predicator. The Adjunct is the position at which the Adverbial Group operates.

A S P C A

e.g. ||| In the morning | John | took | his brother | to school.|||

Group:

Classes: Nominal (N.G.), Verbal (V.G.), Adverbial (A.G.).

The group structure that is chiefly relevant in this paper is that for the nominal group. This structure has elements Modifier (M), Head (H), Qualifier (Q), e.g.

M H Q

The man next door

The need for separate class and structure descriptions may not perhaps be immediately obvious. Both are needed because their roles are different. Structural descriptions are needed in order to account for syntagmatic or chain relations while class descriptions are needed to account for choice or paradigmatic relations. If we consider these clauses:

(1) John painted the house green.

(2) John saw Peter this morning.

(3) John painted the house next door.

The description by group classes of these items consists of the sequence N.G. + V.G. + N.G. + N.G. Any native speaker of English will realize that there are structural differences in the above set of clauses which are not revealed in the class description. What this describes is the class of the items and the sequence of occurrence of the classes but it does not reveal anything about the relations of the classes to each other. That the relations of the classes to each other can be quite different is revealed by applying a permutation operation to the above sentences. They become:

The house was painted green by John.

Peter was seen by John this morning.

The house next door was painted by John.

The class description does not reveal that "next door" is dependent upon "the house" nor that "this morning" differs from "green" in its range of positions. This information about syntagmatic relations of classes, is conveyed in the grammar by means of the structural description. Thus each of (1), (2) and (3) is an exponent of a different structure, being SPCC, SPCA and SPC respectively. The difference in structure is therefore represented either as a difference in the number of elements or as a difference in the choice of elements.

In (3) occurs the grammatical feature which has been called rankshift by Halliday.[4] Rankshift is the occurrence of a unit in the structure of a unit which is of the same or lower rank. "The house next door" is a nominal group operating at Complement. Its own structure is

M H Q
the house [next door]

However, "next door" can also participate directly in the structure of a clause as exponent of Adjunct, as in

A P A
||| Next door | was | an old house.|||

Therefore, in "the house next door", since 'next door' is part of "the house next door" which is itself a group item, it is rankshifted.

Rankshift of one kind or another is a frequent phenomenon in English. It may involve a clause rankshifted within a clause: e.g.

S P C
||| [[Hunting rabbits]] | used to be | fun.|||

---

[3]The following conventions are adopted in parsing grammatical items:

|||	=	sentence boundary
||	=	clause boundary
|	=	group boundary
[[ ]]	=	rankshifted clause
[ ]	=	rankshifted group
( )	=	group interpolated within a discontinuous group

[4]For rankshift, see Halliday 1961.

or a clause within a group: e.g.

M H Q
the man [[who came]]

In this section I have given a partial outline of a taxonomic grammar. Such a grammar consists of a number of different ranks. For each different rank there exists a set of classes and of structures with clearly defined relations between classes and structures. Taxonomic grammars account for syntagmatic relations by means of structure, paradigmatic relations by means of class. Each unambiguous grammatical item is described by being assigned to one of the set of structures for each rank, while each ambiguous grammatical item is described by being assigned to two or more of the set of structures for each rank at which it is ambiguous.

II

Utterances such as:

(1) "The man is able to go."

(2) "The man is easy to please."

(3) "The man is clumsy to fall".

are properly distinguishable at the ranks of clause and of sentence. They certainly cannot be distinguished at any rank lower than the clause since they are all instances of the same sequence of group classes:

N.G. + V.G. + N.G. + V.G.

The internal composition of the groups in each position is the same for each utterance. There is a difference in the syntagmatic relations of the groups which, in a taxonomic grammar, will be reflected in the choice of different structures at higher ranks. The number and choice of elements will have to be clarified.

Jespersen described the difference between utterances such as (1) and (2) in non-grammatical terms by appealing to contextual categories of logical subject and object. In a discussion of "the path is easy to find" which is analogous to "the man is easy to please" he states:

> If we analyze this sentence logically, we see that it is not so much the path that is easy, as the finding of the path; if we transcribe: "It is easy to find the path", the subject in the first instance is "it", which is representative of the following infinitive + object "to find the path". Notionally, we may therefore say that in "the path is easy to find" "the path" is at once a subject and an object, but grammatically "the path" is treated as the sole subject and "to find" as an appendage to "easy".[5]

Because of his logico-grammatical dichotomy, Jepersen is presumably able to distinguish the two sentences:

(1) He is able to go.

and (2) He is easy to please.

only in notional terms "He" is the logical subject of "is" in (1), but in (2) it is not only logical subject to "is" but logical object to "to please".

---

[5]Jespersen 1928: 215

Little attention was focused on the descriptive problem posed by these utterances in the grammars of many American linguists. In part this is due to the limitations of their descriptive procedures. They tend arbitrarily to limit their textual operations to substitution within a frame. Such a procedure is useful only for establishing paradigmatic relations but it does not contribute to the explanation of syntagmatic relations.

The most interesting recent description of these utterances is to be found in the writings of the transformative-generative grammarians. Chomsky (1964), for example, explains that to achieve descriptive adequacy a grammar would have to show that in "John is easy to please", "John" is the direct object of "please" while in "John is able to please" "John" is "the logical subject" of "please". The differences in the syntagmatic relations of the groups are handled in the work of Chomsky and Lees chiefly by means of transformational rules.[6] Chomsky has indeed claimed that it is impossible to handle the descriptive problem in any other way:

> In cases of this sort, the taxonomic model of generative grammar discussed above (or any of its variants) cannot achieve the level of descriptive adequacy, since information of this kind cannot be represented in the Phrase-Marker that it provides as the full structural description on the syntactic level. The transformational model does, however, make grammars available that can supply structural information of this sort, and therefore can, in this case at least, achieve the level of descriptive adequacy.[7]

It is not possible here to discuss at length the question of whether or not a non-transformative-generative grammar can give an adequate description of such utterances. Chomsky's claim is acceptable if, indeed, the characterization that he gives to phrase-structure grammars (Cf., inter alia, Chomsky 1964: 11) were an adequate characterization of all non-transformative taxonomic grammars. Chomsky has received strong support from Postal (Cf. Postal 1964) who has attempted to demonstrate that all taxonomic grammars, including Halliday's one outlined in Section I, are all adequately included within the framework of the phrase-structure characterization for taxonomic grammars. This reduction of Halliday's taxonomic grammar is accomplished only at the price of over-simplification, distortion and misinterpretation of some crucial features of Halliday's grammatical theory. Moreover, more trivially perhaps, Postal's attempt to belittle Halliday's theory as unformalized and perhaps unformalizable overlooks Dixon's previously published formalized version of Halliday's theory (Cf. Dixon 1963).

What is crucial in the strategy of Postal and Chomsky is the exclusion of certain linguistic aspects from the phrase-structure component of grammar and their inclusion via the transformational component. What are excluded from the phrase-structure component are important syntagmatic relations of language. The omission of these reduces the phrase-structure grammar to a classification device. The phrase-structure grammar turns out to be an impoverished version of a taxonomic grammar. In fact, any taxonomic grammar must incorporate information not merely about paradigmatic relations (the class identification of an item) but about syntagmatic relations (the structural identification of an item). The crucial role that structure plays in a taxonomic grammar has been underrated and even dismissed by Postal as superfluous. The truth of the matter seems to be that whereas a transformative-generative

---

[6] See Chomsky 1957, 1962 and 1964; Lees 1963, 1964.

[7] Chomsky 1964: 34

grammar shows syntagmatic relations between classes chiefly by transformation rules and by configuration, a taxonomic grammar shows them by structure.[8] Both structure and transformation rules are grammatical mechanisms (within different grammatical theories) for explicating or characterizing syntagmatic relations in language.

Since it is the function of structure to represent, in an abstract way, the syntagmatic relations of language, it is not sufficient to say, as I did earlier, that elements of structure are places in a unit where operate different classes of the unit below. The elements of the clause, for example, are not just places in the clause where group classes operate, they are terminal points in dyadic relations. For example, the utterance:

‖John | came.‖

which is a realization of the group-class sequence:

N.G. + V.G.

has a clause structure description, SP. In this structure S not only represents a place but it also represents a terminal in the subject-predicator relation, of "John" and "came". Likewise in

"John saw the man"

whose clause representation is SPC, "saw" is a terminal in two dyadic relations; it is a terminal in a subject-predicator relation and in a predicator-complement relation. The two different relations are indicated in this structure by the two different symbols S and C. The symbol P is neutral in respect to information about dyadic relations since the information about the dual relations of the element which it represents is incorporated in the symbols chosen to mark the places which are the other terminals of the relations. Each different dyadic relation need be distinctively indicated only at one of th two places which are its terminals.

Normally, then, a structure will represent syntagmatic relations by a sequence of unitary symbols some of which, such as S or C, indicate dyadic relations, others of which, such a s P, carry neutral information. It is not the case, however, that where place in structure is a terminal in two dyadic relations that the position can always be represented by a single symbol. There are some clause structures in English which must have a double symbol for one of their elements otherwise information about important syntagmatic relations will not be incorporated in the structural description. In other words, there are cases where, like P, an element is a terminal in two different syntagmatic relations but where, unlike P, the element is not neutral with respect to information about the relations but its symbolic representation must indicate the existence of two different relations. Failure to incorporate information about such multiple syntagmatic relations in structural description leads directly to a vindication of Chomsky's claim that taxonomic grammars are incapable of revealing the differences in the kind of clauses illustrated at the beginning of this paper.

Elements of structure are either unitary or dual. A unitary element is a terminal in a singly dyadic relation; a dual element is a terminal in two dyadic relations. All elements, whether unitary or dual, are representated in structural description, by a symbol. All unitary elements are represented by a single symbol, e.g. the element Subject is represented by the symbol S. Dual elements are represented either by a single symbol (e.g. the element Predicator is represented by P) or by a double symbol (illustrated in Section III).

---

[8] For a similar view see Belasco 1964: 77.

III

The first step in the description of such sentences as:

(1) The man is able to come.

(2) The man is easy to please.

(3) The man is clumsy to fall.

is the suggestion of discovery procedures. Certain tests will have to be proposed in order to indicate what structural differences are present. The tests cannot be of the type which note the presence or absence of morphological features since in respect to morphology the sentences are the same. Nor does a test which examines the behaviour of the three sentences in substitution operations lead to any revealing differences. The most illuminating tests are those which examine the behaviour of the sentences under permutation operations. A permutation operation will change the order of occurrence of classes, at the same time deleting or adding certain items.

The second step in structural description is to interpret the results of the permutation test so as to give maximum representation in the structural description to the inter-group relations revealed by the tests. The tests therefore turn out to have a dual function. They not only disclose differences but they also suggest the nature of the differences. The incorporation of this information in structure, however, cannot be simply determined by the test, since it involves other considerations as well. Since the structure may consist of a string of one or more unitary elements or of a string of unitary and dual elements, its structural representation will consist of a sequence of single symbols or of single and double symbols. If

(1) The man is able to go.

(2) The man is easy to please.

(3) The man is clumsy to fall.

are structurally different, it will be necessary to provide two different tests. It will be sufficient to propose one test to distinguish (1) and (2) from (3); and a second to distinguish (1) from (2).

Test 1: $N.G. + V.G. + N.G._2 + V.G._2 \rightarrow$

$it + V.G. + N.G._2 + of + N.G. + V.G._2$ [9]

Applying this test to (1), (2), and (3) we get

It is able of the man to go.

It is easy of the man to please.

(4) It is clumsy of the man to fall.

Only the last of these transforms is acceptable to native speakers, thereby distinguishing (3) from (1) and (2). This test not only reveals that (3) is different from (1) and (2) but it also provides information about group relations in (3). In (4) "of the man" can be deleted, thus showing the independence of "the man" and "to fall" - there can be no SP relation between them. Likewise the occurrence of "of the man" between "clumsy" and "to fall" shows the independence of these groups. Finally, "to fall" is substitutable for "it" to give the transform:

[9] The symbols + and → are to be interpreted as "concatenated with" and "transformed to" respectively.

To fall is clumsy of the man.

Since this particular substitution is considered elsewhere in the grammar to be a crucial test for certain kinds of dependent clauses, the appropriate conclusion is that in

"The man is clumsy to fall."

"to fall" is to be treated not as a constituent of the clause but as the constituent of a sentence whose other component is "the man is clumsy". The correct analysis for this sentence would be

α β — sentence structure

||| S P C || P ||| — clause structures

Test 2: $N.G. + V.G. + N.G._2 + V.G._2 \rightarrow$

$V.G._2 + N.G. + V.G. + N.G._2$

The application of this test to (1), (2) and (3) gives the transforms:

To go the man is able.

To please the man is easy.

To fall the man is clumsy.

Thus only the transform of (2) is acceptable to a native speaker, serving to distinguish (2) from (1). The structural analysis of

S P C

P C

||| [[to please | the man]] is | easy.|||

reveals that "the man" is a complement to "to please" as part of a rankshifted clause. This syntagmatic dependence of "the man" on "to please" is also present in the source sentence. Thus in "the man is easy to please", "the man" is in subject relation to "is" since it is the nominal group that immediately precedes the verbal group. Likewise it is complement to "to please". The fact that the complement precedes its predicator in no way debars it from being a complement since this sequence is already allowed for in certain kinds of thematic clause in English: e.g.

C S P A C S P

|||Mary | we | like,| but | John | we | dislike.|||

"The man" is therefore a terminal in two dyadic relations both of which must be explicitly indicated in the symbol in this initial position in structure if they are to be structurally represented. It is necessary, therefore, to incorporate a double symbol, C/S, in the structural representation where S indicates that the class operating at this position is a terminal in a relation with the left-most verbal group, while the C indicates that it is also a terminal in a relation with the right-most verbal group. The structural analysis will therefore be

C/S P C P

||| The man | is | easy | to please.|||

To say that "the man" is both S and C is purely a grammatical

description. In particular it has nothing to do with logic. The categories of subject and complement can be fully explicated within grammar without resorting to logical notions. The discovery that "the man" is an exponent in both relations gives us a grammatical fact and not a logical fact. Indeed it is clear that since the notion of logical subject is dependent upon the notion of grammatical subject and not vice versa, to attempt to define the latter in terms of the former leads to circularity.

The provision of a unique structure for

"The man is easy to please."

makes it plain that the dual relation of "the man" is sui generis and in no way dependent upon the transform for its existence. It seems necessary to stress this since some linguists seem rather confused about the relation of a sentence to a transform. Lees, for example, writes: "There is no reasonable way to construe certain sentence types other than as permuted, elided, or embedded versions of source sentences".[10] It is odd to claim that the explanation of a sentence is that it is a version of another sentence. The explanation of a given sentence is to be found in the grammatical description assigned to that sentence. A sentence and its transform can be explained independently of each other.

Since "The man is able to go" had negative results in both tests, it might appear that they will not indicate the crucial syntagmatic relations. However the fact that it is not acceptable to say:

"It is able of the man to go."

is an indication of the close dependence of "to go" and "able". Their relation is probably similar to that of "ready" and "to serve"in

"It is ready to serve." [11]

In this sentence it is not possible to substitute "to serve" for "it" to give

"To serve is ready."

thereby directly revealing the dependence of "to serve" upon "ready". This dependence is recognized in structural description by means of nominal group structure. Thus the structural analysis of

It is ready [[to serve]]

if	S	P	C	Cl. Structure
		H	Q	N.G. Structure

The syntagmatic likeness of "the man is able to go " can be indicated by giving it the same structural description, that is, S P C.

The consideration of these three sentences reveals therefore that it is possible in a taxonomic grammar to represent simply the difference between the above kinds of sentence by means of minimal differences in the number and choice of elements. This structural description is both economical in choice of symbol and non-redundant since it incorporates

---

[10] See Lees 1964: 142

[11] For a discussion of some of the problems connected with sentences like "It is ready to serve", see Lees 1963: 77 ff.

facts of grammatical relations which would not elsewhere be represented in the grammar.

The recognition of three different structures has paradigmatic consequences. There are three sub-classes of adjective distinguished according to their potentiality of occurrence at C in these different structures.[12]

(1) Subclass occurring at H in S P C with marked infinitive as qualifier:

Like able, are accustomed, anxious, apt, ashamed, bound, delighted, eager, entitled, fit, free, glad, happy, inclined, keen, liable, likely, loath, pleased, qualified, ready, reluctant, slow, sorry, willing, etc.[13]

(2) Subclass at H in ||| S P C || P |||

Like clumsy, are ambitious, careful, mad, sweet, wise, wrong, etc.

(3) Subclass at H in C/S P C P

Like easy, are deadly, delectable, delightful, excellent, horrible, impatient, pleasant, right, strange, sweet, useful, etc.

The tests used to differentiate the exponents of these different structures are not necessarily, of course, the only possible tests that might be used but they are ones that seem to provide results most consistent with the native speaker's grammatical intuitions. An independent validation of them is provided by their ability to provide a formal analogue for the intuition of grammatical ambiguity. Consider for example:

He was strange to watch.

By test (1) it becomes

It was strange of him to watch.

By test (2) it becomes

To watch him was strange.

Everytime both tests are applicable, ambiguity is present. "He was strange to watch" must be provided with two structural descriptions since it is ambiguous: C/S P C P and S P C || P
The fact that the tests can provide an explanation for empirical facts beyond those for which they were originally proposed provides them with an independent validation. Other examples of ambiguity are:

She is delightful to watch

He is excellent to choose.

She is sweet to remember

The recognition of multiple syntagmatic relations will also serve to help characterize the difference between the following set of sentences:

---

[12] The adjective is a class of the word defined by its potentiality of operation at position E (Epithet) in Nominal Group structure.

[13] Some of these adjectives usually require the infinitive qualifier e.g. "bound", "apt", "liable", "likely", etc.

(4) John chose the gift to send.

(5) John is the man to send.

(6) They chose John to go.

The syntagmatic differences between the three sentences are revealed if a passive transformation test is applied. (4) becomes

The gift to send was chosen by John.

This test will not provide acceptable transforms for (5) or (6). (5) can be differentiated from (6) be a permutation test:

$$N.G. + V.G. + N.G._2 \rightarrow N.G._2 + V.G. + N.G.$$

The man to send is John.

but not

John to go chose they.

The significance of these operational tests is that they reveal the close dependence of the marked infinitive in (4) and (5) upon the preceding nominal group. These two sentences can therefore be described by means of the structure S P C. The description is however inadequate for (6) since it fails to reveal the close syntagmatic relations of "John" with both "chose" and with "to go". The relation between "John" and "to go" is comparable to that in "For John to go is wrong" which is analyzed as S[p[[S P]]] P C. It therefore seems plausible to give as an analysis of "they chose John to go", the structure S P $^{S}/_{C}$ P where the uppermost letter indicates the terminal relation with the rightmost predicator and the lower one indicates the terminal with the leftmost predicator.

It is possible, of course, to have longer complex clauses of this type: e.g.,

He wanted Peter to help John do his homework. S P $^{S}/_{C}$ P $^{S}/_{C}$ P...

The verbal group after the dual element is not restricted to a marked infinitive as this last example shows. The possibilities include all the non-finite verb forms:[14]

(1) unmarked infinitive:

I saw him go

We let him do it.

(2) present participle:

We saw them going.

We caught them trying to get in.

(3) past participle:

They wanted him killed.

They wanted it kept.

[14] For illustrations of these possibilities see Nida 1960: 125 ff.

(4) marked infinitive:

They allowed him to choose.

They wanted him to stop.

Ambiguity is just as possible with these sentences as it was with the earlier types.

(6) "John saw the man waiting outside" may become either

The man waiting outside was seen by John,

or The man was seen, waiting outside.

The first transformation requires the structural description for (6) to be S P C with "waiting outside" a rankshifted clause, while the second transformation requires that the structural description be S P $^{S}/_{C}$ P.

These multiple syntagmatic relations realized in structure as the dual elements $^{S}/_{C}$ or $^{C}/_{S}$ prove useful in description elsewhere. The difference, for example, between

(7) John is easy for us to please. and

(8) John is eager for us to please.

can be accounted for by the structures $^{C}/_{S}$ P C P, and S P C respectively. "For us" is treated as an exponent of an element of group structure and not of the clause because of its close syntagmatic dependence upon "easy" as evident in the transforms:

It is easy for us to please John.

To please John is easy for us.

The significant structural differences between

(9) John is a person who is easy to please. and

(10) John is a person who is eager to please.

can likewise be handled in the same way. Their respective structures will be

(11) John is a person [[who is easy to please]]
S/C P C P

(12) John is a person [[who is eager ‖to please]]
S P C P

It is perhaps sufficient to suggest the analogy of

He is easy to please.

and He is eager to please.

by way of justification for this analysis without suggesting any specific permutation test or tests to yield the difference.

The kind of multiple syntagmatic relations I have been dealing with are relations of group classes described through clause structures. There are other kinds of dual syntagmatic relations which cut across the rank-scale since the class of the group involved enters into syntagmatic relation both at clause and group ranks simultaneously. Strictly speak-

ing the analysis of these relations lies outside the scope of this paper but since the phenomena are partially similar to the ones I have been describing and since they do concern the clause I will deal with them as well.

Consider the two clauses:

(13) This room is easy to clean.

(14) This room is easy to work in.

The first clause is the complex type whose structure can be described as S/C P C P. This structural description is appropriate because each of the group classes represented in the clause is operating directly in the structure of the clause. Each of "this room", "is", "easy" and "to clean" is operating at a place in clause structure. By virtue of this there is no obstacle in regarding the double symbol for the element at which "the room" is operating as symbolizing the dual relations which "the room" is entering into with other groups. It is for this reason, of course, that one has to regard "to clean" not as a rankshifted verbal group like "to do" in "the thing to do", which is operating in the structure of a group at position Q, but as a group which operates directly in the structure of a clause.

The analysis of

This room is easy to work in.

is different. It is possible to show that "This room" has dual syntagmatic relations: it is both subject to "is" and also part of the prepositional phrase "in the room" as the following transform reveals:

It is easy to work in this room.

However, it is not possible to handle these syntagmatic relations by having a dual symbol to represent them in clause structure. The reason is the one that I gave before: as a group in subject relation with "is", "this room" is operating in the structure of the clause; as a group which is rankshifted within a discontinuous prepositional phrase - "this room ... in", - "this room" is operating in the structure of a group.

There is thus no problem in representing the group relations of "this room". At clause rank it is simply represented by S. But one can't legitimately use some such symbol as S/A to have a dual symbol since this would conflate relations which occur at different ranks in the grammar. There is, however, a problem in representing in structural terms the relation of "this room" to "in" within a discontinuous prepositional phrase. The usual way to represent discontinuity of exponents of an element can be illustrated in this example:

||| The man | was (certainly) trying.|||

Here the exponent of Adjunct ("certainly") is interpolated within the discontinuous exponent of a predicator ("was...trying"). The discontinuity can be indicated in clause structure like this:

S P A

where the curved arrow indicates that the exponent of A occurs somewhere within the exponent of P. This solution will not apply to our clause

This room is easy to work in

unless in some such fashion as:

(S P C P) A

However, the representation of this second relation of "This room" does not seem crucial in description: it seems convenient to just present the structure as S P C P A.

A similar problem of representation occurs with clauses such as

(15) The man was taken care of by them.

(16) This must be paid attention to by us.

which are obviously related to the transforms:

They took care of the man.

and

We must pay attention to this.

But though "the man" and "this" are presupposed by "of" and "to" as part of the exponents of prepositional phrase, this is also not an instance of a dual element in clause structure.

The solutions to the descriptive problems dealt with in this paper do not differ greatly from those found in some older grammars. There are, of course, differences in the methods by which they were solved and in the form in which they were presented. Recently much more has been learnt about the kind of criteria which are relevant in making grammatical descriptions. Operations such as substitution, deletion, and permutation are applied to a test to enable the linguist to discover the likenesses and differences in language form. There is no appeal to meaning as a criterion for establishing categories though meaning is always presupposed. The formal discoveries are described by means of abstract categories of class, system, structure and unit. Finally the use of symbols as a shorthand representation of categories enables the linguist to detect inconsistencies and ad hoc solutions, thus reducing undesirable descriptive features often masked in the past by the use of anecdotal methods of presentation.

BIBLIOGRAPHY

Belasco, S. 1964. "Tagmemics and Transformational Grammar." Linguistics 10: 5-16

Chomsky, N. 1957. Syntactic Structures. The Hague, Mouton.

1962. "A Transformational Approach to Syntax." Proceedings of the Third Texas Conference on Problems of Linguistic Analysis in English. Austin: 124-58

1964. Current Issues in Linguistic Theory. The Hague, Mouton.

Dixon, R. M. W. 1963. "A Logical Statement of Grammatical Theory." Language, 39: 654-658

Francis, W. N. 1958. The Structure of American English. New York, Ronald Press.

Fries, C. C. 1957. The Structure of English. London, Longmans.

Halliday, M.A.K. 1961. "Categories of the Theory of Grammar." Word, 17: 241-92

Jespersen, O. 1928. A Modern English Grammar on historical principles. Part III, Vol. 2. London, George Allen & Unwin Ltd.

Lees, R. B. 1963. The Grammar of English Nominalizations. The Hague, Mouton.

1964. "Transformational Grammars and the Fries Framework." Readings in Applied English Linguistics. (ed.) Allen New York, Appleton-Century-Crofts. Second Edition: 137-146.

Long, R. B. 1961. The Sentence and its Parts. Chicago, University Press.

Nida, E. 1960. A Synopsis of English Syntax. Norman, Oklahoma, Summer Institute of Linguistics (Linguistic Series No. 4).

Postal, P. 1964. Constituent Structure: A Study of Contemporary Models of Syntactic Description. The Hague, Mouton.

Sledd, J. 1959. A Short Introduction to English Grammar. Chicago, Scott, Foresman.

Strang, B. 1962. Modern English Structure. London, Arnold.

CLAUSE AND SENTENCE TYPES IN MNONG RƠLƠM

EVANGELINE BLOOD

(Summer Institute of Linguistics)

0. Introduction
1. Clause Types
2. Clause Functions
3. Clause Expansion
4. Sentence Types

0. This paper presents the analysis of clause and sentence types in Mnong Rơlơm,[1] a Mon-Khmer language of Vietnam. The data studied for this analysis were a number of texts, which included legends, narratives of current events, and descriptions of customs, together with some elicited material containing sentence types not in the texts.

1. Clauses in Mnong Rơlơm are grammatical units of predication. There are eight clause types: transitive, object complement transitive, intransitive, stative, fused, nominal equative, adjectival equative, and quotative. The formulae for these types follow:

---

[1]Mnong Rơlơm (also known as Rlâm or Lâm) is a dialect of Mnong, a language of the Mon-Khmer family. It is spoken in Lac Thien district, Darlac province, South Vietnam. The data for this paper were collected during our residence among the Mnong Rơlơm people from May 1960 to December 1962.
I would like to acknowledge the helpful suggestions of Viola Waterhouse, Darlene Bee, and David Thomas of the Summer Institute of Linguistics, and of Sue Albright of the Unevangelized Fields Mission; also the help of my husband, Henry, with his knowledge of the language.

Orthographic symbols used in Mnong Rơlơm represent the following phonemes: (at bilabial, alveolar, palatal, and velar positions) voiceless stops p, t, c, k; voiced stops b, d, j, g; preglottalized stops ɓ, ɗ, ɗj; prenasalized stops mb, nd, nj, ngg; nasals m, n, ñ, ng; lateral l, trill r; semivowels w, y; fricatives s (-ih), h; glottal (- between vowels, ˘ word final).

Front vowels: i high, î mid-high, îa (-ie-, -ia) glided, ê mid, e low. Central: u' mid-high short, â mid short, o' mid long, ă low short, a low long. Back: u high long, ŭ mid-high short, û mid-high long, ûa (-uo-, -ua) glided, ô mid long, o low (short).

In this paper the hyphen is used between the syllables of words or between the words in compounds.

Transitive ± S + P_1 + O

Object Complement Transitive + S + P_{1a} + O + C_1

Intransitive + S + P_2

Stative + mau + S ± P ± O

Fused ± S + P_1 + O/S + P ± O ± C_1

Nominal Equative + S + P_3 + C_2

Adjectival Equative + S + M

Quotative + S + P_4 + Q

(S = Subject, P = Predicate, O = Object, C = Complement, M = Modifier, Q = Quotation)[2]

1.1 The transitive type with P_1 predicate (verbs like see, hit, etc.) has obligatory object. kan so sau 'he saw dog'. In clauses with O-S-P order the object is in focus. bo' kan han lai nai so (pl. masc. there already they saw) 'they already saw them'.

1.2 The object complement transitive type with P_{1a} predicate (filled by verbs nan 'to name' or kuol 'to name or call') has obligatory object and complement. nai nan kan han Y Pal 'people call him there Y Pal'. The object is kan han 'him there', and Y Pal, a man's name, is the object complement.

1.3 In the intransitive clause the predicate P_2 is filled by verbs that do not take an object. añ plŏ 'I go-home'.

1.4 The stative clause type requires mau 'is, there is' with a subject, optional predicate, and optional object following mau. mau pay nau yat gong 'there-are three people hit gong'. mau bal bri mơng uon añ (there-are pl. forest in village mine) 'there are forest people in my village'.

1.5 In the fused clause type the obligatory object of the first predicate is simultaneously the subject of the second predicate. The second predicate can be simple transitive, object-complement transitive, or intransitive. nai phung bo' kan mhŏ djoh ta nai (they fear pl. him do bad to people) 'they fear they will do bad to people'.

---

[2]A complete analysis of word classes has not yet been undertaken, but the following general statements can be made about the fillers of subject, object predicate and modifier slots. Noun expressions manifest the subject, complement, and object tagmemes, and are further identified as those forms that can occur in the frame ta ___. Ta has many meanings, such as 'to, for, with, by, at, pertaining to'. The minimum form of a noun expression is a noun or pronoun. Verb expressions manifest the predicate tagmeme and are further identified as those forms that can occur in the frames wih ____ sưt, or ____ un. (wih, sưt 'again', un 'thus help'.) The minimum form of a verb expression is a verb. Adjectives manifest the modifier tagmeme in adjectival equative clauses, and are further identified as those forms other than nouns and verbs that can occur in the frame ____ ngăn. (ngăn 'very')

1.6 The nominal equative clause with optional P_3 filled by a copulative verb (njing 'is', blah, ndrơm blah, ndrơm ta 'is same as') always has a noun or pronoun as the complement. añ ndrơm blah e 'I am-same-as you', jau-dak-duh u-ur go'-gu (sorcerer-water-hot women only) 'midwives are women only'.

1.7 The adjectival equative clause has obligatory M filled by an adjective. The usual order is subject followed by modifier, but this order may be reversed, usually to mark emphasis on the condition. kan ndrong (he rich) 'he is rich', ar ngan añ (tired very I) 'I am very tired'.

1.8 The quotative clause has P_4 filled with verb lah 'say' or lup 'ask'. The obligatory quotation includes anything from one word up to a number of sentences. The subject is an animate noun or a pronoun. nai lah, hi ay so gưt 'they say, we do not know'.

2. Clauses may be either dependent or independent. The quotative and fused types occur independently most of the time. Dependent clauses are marked by one of eight introducers (to' or dî 'if, when', cieng bi 'in order to', koyuadah or koyuahan 'because', lai or lai le 'already, finished', tui or tui si 'according to', mhay 'as soon as') and must co-occur with an independent clause. To', dî, koyuahan, lai, lai le, and mhay introduce dependent clauses almost always preceding the independent clauses. The dependent clauses introduced by to' or dî indicate time or condition. Koyuadah, cieng bi, and koyuahan indicate purpose. Tui and tui si mark condition. Lai, lai le, and mhay mark time.

3. Clauses may be expanded by inclusion of location, time, instrument, and indirect object constituents.[3]

Most location constituents are introduced by the prepositions ta 'to, at, in', bak 'with', dơng bơ 'from', mơng 'in', and tang 'on'. Direction words such as nam 'south', buk 'north', ien 'there', to 'there far', au 'here', and nouns and pronouns are the head words in phrases with ta, dơng bo', and mơng. Only nouns and pronouns are the head words in phrases with tang 'on' and bak 'with'. The location constituents usually follow the predicate in an intransitive clause or follow the object in a transitive clause. They may also occur sentence initial.

Among the time constituents are words denoting definite time periods such as mơ-mang 'night', mho 'evening', nam 'year', khe 'month'; words denoting passage of time such as jok ta ien 'after a while', luor 'first or before', yau lak 'long ago', ejai 'during'; and expressions with the prepositions truh ta 'until' and dơng bơ 'from' with some of the time period words and yau and lak.

The time constituents usually precede the subject, or precede the predicate if there is no subject, but also may follow the object or complement.

The instrument constituents are introduced by ta 'by, with' and usually follow the predicate in an intransitive clause or the object in a transitive clause. The distinction between instrument, location, and indirect object is not always clear. Instruments never occur clause initial, and

---

[3]There are other minor constituents which expand clauses, such as manner, which are not treated in this paper.

they usually co-occur with verbs such as sreh 'cut', geh 'to stick', chok 'gore', pañ 'shoot'.

The indirect object is always a noun or pronoun introduced by the preposition ta 'to, for'. It follows the predicate in an intransitive clause or the object in a transitive clause. It is distinguishable from time constituents in that time constituents usually stand clause initial; it is distinguishable internally from location constituents only when a location word is in the head of the phrase. It usually co-occurs with verbs such as lah 'say', an 'give', mhŏ 'make, do'.

These constituents are illustrated by the following sentences:

Jok ta ien nai yat gong mơng hih

Time S P O Loc.

After a while they hit gongs in the house.

Nai puol dih ta kan

S P O I.O.

People hang the cotton for him.

4. The sentence in Mnong Rơlơm is defined as a syntactic unit which potentially occurs alone as a complete utterance. The predominant sentence structure is a single clause, since an independent clause can be a simple sentence. Sau kap kan 'The dog bit him.' The potential clause combinations are:

a) Two or more independent clauses optionally joined by conjunction lai-han 'and' or bia-dah 'but' make a compound sentence. De sok kuon sau, de lang truoñ su, lai-han de hao mơng tơm go-gier. 'She takes the dog, she wraps it in a loin cloth, and she climbs the go-gier tree.'

b) One independent clause together with at least one dependent clause make a complex sentence. Jok ta ien bal han njoh, kơyuadah ar. 'After a while they rest, because they are tired.'

c) At least two independent clauses together with at least one dependent clause make a compound-complex sentence. To' nai trûh ta hih kan, mau ok bal bri guk jum-dar hih kan, mau bar nau bal bri sak cơn mơng hih kan. 'When they arrived at his house, there were many Communists staying around his house, and there were two Communists who went up into his house.'

4.1 There are four sentence types: declarative, interrogative, vocative, and imperative.

4.1.1 The declarative sentence type consists of those sentences simply stating facts or declaring action, and they occur without any special marker. Hi tong nai yat gong. 'We hear them hit the gongs.'

4.1.2 Interrogative sentences are marked by one or more of a class of interrogative words, with or without final rising intonation. A number of these indicators almost always appear sentence initial, such as pơp or pơp ien 'what, pơp mhŏ 'how, why', nih tơm 'who, whose', tuk or ta ien tuk 'where', tuk ta nang 'what day', tuk ta with a noun of location such as village, house or mountain 'where'. Frequently the word ya 'is that so, yes?' occurs sentence-finally with rising intonation, but it can also occur right after the preposition in question. E sak ya? 'Do you go?'

Dơi ya ăn sak ta Uon Thuot? 'Can I go to Banmethuot?' Nih tơm sau deh kuon? 'Whose dog bore pups?'

4.1.3 Vocative sentences consist of a proper name or kinship name optionally preceded by wo' 'hey!' and optionally followed by the corresponding second person pronoun — me masculine or e feminine. The pronoun occurs only when wo' also occurs; wơ, however, may occur without the pronoun. Wo' Muom me! or Wo' Muom! 'Hey Muom!'

4.1.4 The imperative sentences include negative and positive commands. The negative commands are marked by the words ay rau 'don't' or man 'don't' preceding the subject or predicate. If a pronoun is present it is always a second person pronoun. Ay rau ndroh kuon añ hau. 'Don't forsake my child.' Man e pah-rak añ. 'Don't you curse me.'

The positive commands are marked by sentence-initial second person pronoun, or a proper name and a second person pronoun. Muom me sok ba han. (Muom you get rice there.) 'Muom get that rice.' Inclusive commands are marked by the occurrence initially and optionally finally of jŏ' 'let'. Jŏ' ban sak jŏ'. (let us go let) 'Let's go.'

4.2 The four sentence types are further divided by dependence and independence.[4] The dependent sentences presuppose the situation or the content of a previous sentence for interpretation. Independent sentences are unambiguous as they stand.

The dependent sentences are of two types: responsive and sequential.

4.2.1 The dependent response sentences are answers to questions and are not complete clauses. Me plŏ ya? I. 'Are you going home? Yes.' The Yes is the dependent response sentence.

4.2.2 Sequential sentences are dependent on the preceding sentence or sentences in that they (1) build on the information given, or (2) follow in a logical sequence. These are marked by the following words occurring initial: tô-mô 'for example', lai-han 'and', blah ien or kơt blah ien 'so, like that', bia-dah 'but' (if there is only one clause in the sentence), and han 'then'. Bia-dah mih eh sak hui hai. 'But rain it comes always.'

---

4 Viola Waterhouse, 'Independent and Dependent Sentences', IJAL 29: 45-54 (1963).

HALĂNG VERB PHRASE

James S. Cooper

0. Introduction

1. Preverb
   - 1.1 Stative
   - 1.2 Negative
   - 1.3 Dependent

2. Main Verb
   - 2.1 Reduplicative
   - 2.2 Close Knit
   - 2.3 Directional
   - 2.4 Coordinate
   - 2.5 Objective

0. Introduction. We will consider the Halăng[1] verb phrase (for the purposes of this paper) to consist of the main verb and elements occurring adjacent to the main verb which cannot be separated from the main verb by clause-level elements such as subject, locative or instrumental. The verb phrase under consideration is specifically that found in independent declarative clauses, which accounts for a large majority of the verb phrases in text.[2]

1. Preverb. Preverbs are not obligatory to the verb phrase but serve to modify the main verb. For convenience of description we have divided the preverbs into stative, negative and dependent preverbs. The stative preverbs can function as main verbs but negative and dependent preverbs cannot.

---

[1]Halăng is a Mon-Khmer language spoken in Kontum Province of the Republic of Viet Nam. Kơyong, a nearby dialect, is mutually intelligible with Halăng. There are an estimated 10,000 Halăng people. The language data for this paper was gathered over a period of two years, beginning in March 1963. Six months were spent living in the village of Plei Khôk Hơnar, a 'new life' hamlet 15 kilometers west of Kontum City.

[2]During the analysis of this paper there was extensive use of a concordance of 57,000 words of text in Halăng made on the IBM 1410 computer at the University of Oklahoma by the Linguistic Information Retrieval Project of the Summer Institute of Linguistics and the University of Oklahoma Research Institute, and sponsored by Grant GS-270 of the National Science Foundation. I am indebted to David D. Thomas for help in analysis and writing of this paper and to Dĭn, my Halăng language teacher, who has been the source of much of the text material used in this paper.

1.1 Stative Preverbs. The set of words which function as stative preverbs occur frequently in Halăng, because the same set of words can function as preverbs, as main verbs, or to introduce dependent clauses. The usual order for the stative preverb is: jìang[3] 'become, complete', ay 'continue, be', ʔlo ʔ 'know, be, experience', e 'have, be', wa ʔ 'going to, about to', jah 'acquire, result'. ʔlo ʔ can precede ay and can both precede and follow e.

1.11 jìang 'become, complete' can not be preceded by any of the other stative preverbs. It has been found in combinations with ay, ʔlo ʔ, and jah.

jìang ay Aw jìang ay waaw. 'I become continue understand'
'I have come to understand.'

jìang ʔlo ʔ Koan jìang ʔlo ʔ hateh. 'child become know talk'
'The child has come to be able to talk.'

jìang jah Ger jìang jah dàk. 'he become result come-up'
'He has come to the point of having come up.'

1.12 ay appears frequently in text material because it can function as a main verb, both stative and active, with the meaning of 'to live', or as a preverb with the meaning of 'to continue, to be'. ay has been found with other stative preverbs in the following combinations:

jìang ay see 1.11

ay ʔlo ʔ Bò ʔ klia ay ʔlo ʔ cha. 'grandfather tiger continue know eat'
'The tiger still eats (people).'

ay wa ʔ ʔloang ay wa ʔ takliḥ. 'tree continue about-to fall'
'The tree is still about to fall.'

ʔlo ʔ ay jah Gawì ʔlo ʔ ay jah mut. 'they know still complete enter'
'They have still been able to enter.'

1.13 ʔlo ʔ is the only preverb which can occur twice in one verb phrase, coming before and after e or the negative ʔbe. It may also come before or after ay. As an active verb it has the meaning 'to know', as a stative verb the meaning of 'be', and as a preverb the meaning of 'be, experience'. It has been found in the following combinations.

jìang ʔlo ʔ see 1.11

ay ʔlo ʔ see 1.12

ʔlo ʔ ay jah see 1.12

ʔlo ʔ e or e ʔlo ʔ

Bloy ʔlo ʔ e ʔnha ʔ un. or Bloy e ʔlo ʔ ʔnha ʔ un.
'Bloy know have bring fire' 'Bloy have know bring fire'
'Bloy has brought fire.'

ʔlo ʔ e ʔlo ʔ wa ʔ

Bloy ʔlo ʔ e ʔlo ʔ wa ʔ blah yòon.
'Bloy be result be about-to fight we-exclusive'
'Bloy is always about to fight us.'

---

[3]For a description of the Halăng sound system see James & Nancy Cooper "Halăng Phonemes" in Mon-Khmer Studies II (Saigon 1965). Non-conventional orthography used in this paper is ` for breathy voice quality, double vowel for long vowel, a following a vowel for neutral off-glide, nh for ñ.

ʔlo? waʔ

Ger ʔlo? waʔ go? aw. 'he be about-to beat me'
'He is about to beat me.'

ʔlo? jah Aw ʔlo? jah ke. 'I be result see'
'I have been able to see (it).'

ʔlo? negative ʔlo? (for negative see 1.3)

Katap ʔlo? ʔbe ʔlo? chen. 'egg be not experience cook'
'The egg is not becoming cooked.'

1.14 e 'have, be' may be used as a main verb, a stative preverb, or to introduce dependent clauses with the meaning 'if there are' or 'who?'.

e ʔlo? or ʔlo? e see 1.13

e ʔlo? jah

Aw e ʔlo? jah danùh. 'I have experience result poor'
'I have become poor.'

ʔlo? e ʔlo? waʔ see 1.13

e jah Aw e jah chiw. 'I have result go'
'I have gotten to go.'

1.15 waʔ as a preverb means 'about to, going to' and as a main verb has the meaning of 'to want'.

ay waʔ see 1.12

ʔlo? waʔ see 1.13

ʔlo? e ʔlo? waʔ see 1.13

waʔ jah

Bòʔ klia waʔ jah kap. 'grandfather tiger about-to result bite'
'The tiger almost bit (her).'

1.16 jah as a preverb has the meaning of 'acquire, result', as a main verb 'to have', and in introducing dependent clauses the meaning of 'if'.

jìang jah see 1.11

ʔlo? jah see 1.13

ʔlo? ay jah see 1.12

e ʔlo? jah see 1.14

waʔ jah see 1.15

Since all of the stative preverbs can occur as main verbs with very similar meanings, it has been difficult to decide in every instance whether a word is a preverb or a main verb. A possible alternate solution would be to include all the stative preverbs under the classification of main verb in an objective relationship (see 2.5). This solution was not adopted because it is not possible to expand the stative preverb and the main verb by repeating the subject or supplying a different subject between the stative preverb and the main verb.

1.2 Negative Preverbs. Negative preverbs consist of ʔbe 'not', maʔ 'don't', and tàam 'not yet'. ʔbe can be used to negate parts of speech other than verbs.

1.21 ʔbe 'not, no' is the most common of the negatives. It can negate stative preverbs or main verbs, and it comes directly before the item being negated. It has been found before all stative preverbs except jìang and ay.

Aw ay ʔbe tùaʔ. 'I continue not believe'
'I still don't believe.'

Yòon jìang ʔbe ʔloʔ waaw. 'we-excl. become not be understand'
'We came to the point of not being able to understand.'

1.22 maʔ 'don't' is used with subjects which include the second person: àih 2 sg., ìh 2 sg. honorific, ʔni 2 pl., bri 2 dual, bà 1 dual incl., and hèey 1 pl. incl. maʔ is used with the negative tàam (see 1.23) and before the stative preverb ʔloʔ.

Maʔ ʔloʔ blah. 'don't experience fight'
'Don't fight.'

maʔ used in clauses which are conditional or contrary to fact has the meaning of 'not'.

Jah gàwʔ maʔ phoam aw, aw dah ʔbe siit.
'if he not help me, I rather not complete'
'If he had not helped me, I (probably) would not have finished.'

1.23 tàam 'not yet' follows the distribution pattern of ʔbe in that it is found before all stative preverbs except jiang and ay. It may also occur with the negative maʔ.

Mìh maʔ tàam goʔ aw. 'uncle don't yet beat me'
'Don't beat me yet, uncle.'

1.3 Dependent Preverbs. Dependent preverbs can not occur as main verbs and do not occur unless a main verb is present. The dependent preverbs are maʔ emphatic, khom 'to continue to try', and huung 'to want'.

1.31 maʔ emphatic[4] is homophonous with the negative maʔ 'don't' but receives a lighter stress than maʔ negative.

Dadam klan maʔ lòon ger. 'young-boy snake emphatic swallow him'
'The snake boy swallowed him.'

Aw maʔ choat tuuih. 'I emphatic be-stabbed sharp-stick'
'I was stabbed with a sharp stick.'

The only other preverb which co-occurs with maʔ emphatic is waʔ.

Aw maʔ waʔ tek. 'I emphatic want sell'
'I want to sell (it) very much.'

1.32 khom 'to continue to try' is found before main verbs and the stative preverbs ay, waʔ and jah.[5]

---

[4] maʔ emphatic also occurs before words of the stative verbs class functioning as adverbs following the main verb.

[5] khom is found following the main verb with the closely related meaning of 'continue'. In this position it is usually repeated three times and is often used in conjunction with other words such as nèh, lah, and ʔmòyʔ.

Ger khom blah ʔni. 'he continue-to-try fight you'
'He will continue to try to fight you.'

Aw khom ay waʔ tek. or Aw ay khom waʔ tek.
'I try continue going sell' 'I continue try going sell'
'I am still going to try to sell.'

1.33 huung 'to want' immediately precedes the main verb and can not be separated from the main verb by other words except in cases where strong emphasis is indicated.

Chaw huung cha mok. 'grandchild want eat parched-rice'
'Grandchild wants to eat parched rice.'

Kay ʔloʔ huung chiw koh ʔloang. 'he be want go cut wood'
'He is wanting to go cut wood.'

huung can occur following other preverbs ʔbe, ʔloʔ, and ay.

2. Main Verb. Main verbs include words such as kuuy 'to sleep', cha 'to eat', liam 'to be good, beautiful', haklàng 'to be heavy', chiw 'to go', and braayʔ 'to be tired'.

All of the verbs in a sentence usually occur together. Subjects, objects, and indirect objects which may occur between verbs are omitted if the context is able to supply the meaning. Within the main verb there is a complex of verbs which can best be described by showing the relationships within the complex. The relationship may be reduplicative, close knit, directional, coordinate, or objective. There may be more than one relationship in the main verb complex.

2.1 Reduplicative Relationship. In a reduplicative relationship the second verb can not occur alone. The second verb has no meaning of itself but serves only to modify the meaning of the first verb.

2.11 The most common type of reduplication consists of replacing the vowel and final consonant (if one occurs) with ayʔ. Reduplication of this type conveys the meaning of 'to do reluctantly'.

Aw cha chayʔ ròʔ. 'I eat meat.' (Someone else doesn't want it or it may be rotten.)

Aw kuuy kayʔ. I sleep.' (on the ground, because there is no place else to sleep.)

Aw sìʔ sayʔ. 'I go home.' (by myself because no one would go with me.)

Aw hùum hay . 'I take a bath.' (The water is not very clean.)

2.12 Another type of reduplication consists of replacing the vowel with the vowel àa. Reduplication of this type indicates a reluctance and someone's insistence upon an action.

Cha chà kà àih. 'eat referring-to you' 'Eat it!' (There is not very much and you keep insisting.)

Kuuy kàay kà àih. 'sleep referring-to you' 'Go sleep by yourself.' (I don't want to.)

Aw sìʔ sàaʔ dòh. 'I return final-particle' 'I go home reluctantly (because you insist).'

Aw hùum hàam dòh. 'I bathe reluctantly.' (Someone insists).

2.2 Close Knit Relationship. Verbs in a close knit relationship can occur independently as a main verb but when used together they cannot be reversed in order or separated by other words. They are stereotyped phrases where the two near synonyms together have just one meaning and indicate just one action.

Aw chong cha pang ìh. 'I eat with you.' (chong 'to eat rice', cha 'to eat other than rice')

Aw cha kap. 'I eat.' (cha 'to eat other than rice', kap 'to bite')

Aw rak ke àih. 'I take care of you.' (rak 'care for', ke 'watch')

A number of verbs occur only in the second position, knit to a variety of verbs in the first position. These include words such as dèh 'to throw away', ngòoy 'to play', am 'to give', and chuʔ 'to put'.

Aw koh dèh àih. 'I destroy you.' (koh 'to cut', dèh 'to throw away')

Aw peng dèh àih. 'I kill you.' (peng 'to shoot', dèh 'to throw away')

2.3 Directional Relationship. With few exceptions, verbs of direction or motion precede other main verbs. Nothing can come between the directional verb and the verb which follows. Directional verbs include words such as: chiw 'go', lah 'come', loh 'go out', dang 'go look for', dàk 'ascend', jur 'descend', lay 'go after', wal 'return', weh 'turn aside', sìʔ 'return'.

Hèey chiw hùum. 'We go bathe.'

Aw lah chong pang ìh. 'I come eat with you.'

Two or more directional verbs can precede another verb.

ʔni chiw dàk cha. 'You go ascend eat.'

Aw lah chiw tan ìh. 'I arrive go tie you.'

2.4 Coordinate Relationship. Coordinate verbs show separate actions or separate aspects of the same action. Verbs in this relationship share the same subject and object.

Ger bùh oak cha ka. 'He roasted cleaned ate the fish.'

Pataw ruup bùh ʔmaang pahaang giawʔ àih. 'Rich-man (will) capture pound beat roast cook-on-a-stick you.'

Kataam yok cha ger. 'Crab got ate her.'

2.5 Objective Relationship. In an objective relationship one verb is the object of another verb. This construction may be considered as going beyond the verb phrase because the verbs can be expanded into separate clauses having their own subjects and objects, but the expansion often includes information which does not need to be said or which the speaker does not want to say. In effect, the expansion changes the intent of the clause.

Verbs which take other verbs as objects fall into the general categories of emotion, perception, saying or doing. Of emotion: maat 'to want', wil 'to worry', ʔyuʔ 'to be afraid'. Of perception: saw 'to see', nhàng 'to hear, feel', hapo 'to dream'. Of saying: roay 'to tell', patho 'to teach', chun 'to command', takèey 'to tell'. Of doing: chuʔ 'to put', che 'to do, pretend', ʔnhaʔ 'to bring', dagum 'to help', kaw 'to wait',

pakra 'to repair', yok 'to get', lòoy 'to stop'.

Bò patho tìam màk. 'grandfather teach make knife' 'Grandfather teaches making knives.'

Aw hapo jaal. 'I dream fish-with-a-net.' 'I dreamed someone was fishing with a net.'

Aw nhàng waar chiw. 'I hear say go'
'I heard (someone) say that (he) went.'

Ger che patho pakra hanaat. 'he pretend teach repair gun'
'He pretends to teach how to repair guns.'

The subjects of the verbs may be the same or different, but the second subject seldom occurs in the clause.

Aw ʔyuʔ ʔbe teal chiw. 'I afraid not able go'
'I am afraid that (I, you, or he) is not able to go.'

Aw ʔnhaʔ hùum. 'I bring (you) and (we) bathe.'

Aw yok ay kà aw. 'I get live referring-to me'
'I get (you or someone) and (you or someone) lives with me.'

AN ARTICLE OF FAITH

H. V. George

(English Language Institute, Victoria University)

Linguisticians often describe their work as scientific. I am a Popper man in this respect, and feel that to qualify as scientific a statement about language (like a statement about the physical world) must lead directly to prediction that is experimentally verifiable or falsifiable.[1] From this point of view, many premises of contemporary linguistics are articles of faith. This paper describes one article of faith put in the form of a prediction, experiments to verify to falsify the prediction, and some implications of the findings.

The article of faith is the assertion that in language the relationship between sign and meaning is arbitrary.

It is true that Bloomfield, after Jespersen, recognizes such facts as that many words in English beginning with sl represent varieties either of chopping and striking movements or of physical or mental "muddiness". These facts are noted, like onomatopoeia, as exceptional to the postulate that the meaning attached to a group of sounds is not intrinsic but socially determined.

The evidence for the postulate, or more precisely, the way in which the evidence is considered, is of methodological interest, and will be considered separately. Let us now reduce to experimental terms the statement under reference.

With an initial consonant, a vowel and a final consonant we have a simple formation pattern for an English word. Using l initially, varying the vowel, and taking final p, for instance, we form a series of "words": lip lep lap lop lup leap lape lipe ... and so on. Half of the words so formed are institutionalized, or "actual" words, and the others, lep lup lape ... I will call "contingent". The latter constitute material one can experiment with.

Taking for granted that actual words have actual meanings, there are three questions: May contingent words have actual meanings? Do actual words have contingent meanings? Do contingent words have contingent meanings? The third question is esoteric, and I have no experimental approach to it; the other two questions can be approached experimentally, and the first is considered here.

The question, May contingent words have actual meanings? cannot be put to experimental proof naively, by asking people, "What does lep mean?", for lep is a priori distinguished as not institutionalized, as not

---

[1] Popper 1935.

(yet) actual. We are by definition dealing not with actual actual meaning but with potential actual meaning.

Now from time to time, a contingent word does become an actual word, and the occurrence of actualization offers a clue to an experimental approach. Here are three actualizations: Klim smog spiv.

The condition for such actualization is the coming into existence, or perception, of a new designatum. Most actualizations are of unknown provenance. However smog is the sm of smoke and the og of fog, while Klim is milk backwards; and if these formation procedures were the whole facts of even one type of actualization we would have formulas for conjuring the contingent into the actual. While it has to be admitted that the unlikely Mho has established itself - at least as a written word - on the klim formula, I for one hardly imagine that niaw or eniw will come to life when this substance too is retailed as a powder. My prejudice is with the poet:

> Twin are the gates of sleep, of horn and polished ivory.
> The true shades issue from the one, and from the other
> Wraiths.

Here indeed we have the basis of experiment. It is said that the relationship between sign and meaning is artibtrary and socially determined. We predict, therefore, that when, for a designatum which as yet has no socially accorded sign, systematically chosen alternative contingent signs are proposed, no feature of any one such sign can make it more acceptable for representation of the designatum than any feature of any other sign; therefore, that when several people are invited to select one of the signs proposed for representation of the designatum, sounds associating randomly with meaning and no particular sign being preferred, the votes for each artificial word will be, within statistical limits, equal.

Experiment 1

Eight designata were proposed. Here is a specimen:

> The month: Having 12 months with different totals of days is very awkward. Having 13 months of four weeks each is a much better idea. A name is required for a thirteenth month to come between May and June in the present calendar.

Ten contingent words (Alternatives a - j in Table 1) were selected to compete for representation of each designatum. The raw materials of the contingent words were obtained by breaking down to their constituent sounds the actual signs for adjacent designata, those of other months for the specimen cited above.[2]

The form of the experiment was explained to a group of graduate students at the Central Institute of English in Hyderabad, India, then a tape-recording was played, consisting of designatum definition followed by the ten contingent words read from the first to the tenth then repeated from the tenth to the first. Each student followed the reading on his

[2] For the month, the ten contingent words were a) tɔsba, b) bevrou, c) mu:pem, d) sedʒbi, e) ækmou, f) dilti, g) nepdu:, h) vlæsi, i) ɔlsem, j) rɔnta.

voting paper showing the words in I.P.A. script, and ticked his preference if he had one. This is how the voting went:

Alternative:	Designatum: 1	2	3	4	5	6	7	8
a	2	8	1	7	4	3	9	8
b	4	2	1	2	2	5	5	3
c	[25]	10	[25]	5	9	[16]	4	2
d	5	3	4	8	-	2	1	4
e	3	6	1	4	[15]	8	2	5
f	3	1	1	1	[17]	7	[21]	6
g	-	5	[15]	2	3	8	-	9
h	1	3	2	4	-	-	5	5
i	2	7	-	-	1	3	2	8
j	4	5	2	[17]	1	1	1	-

TABLE 1

The prediction that there would be substantially the same number of votes for each word was fulfilled with respect to Designata 2 and 8, and unfulfilled with respect to Designata 1, 3, 4, 5, 6 and 7. The postulate leading to this prediction must be held not verified.

What I take to be a similar experiment, though performed with a different purpose, is cursorily described in "Studies on Thought and Speech Problems" by R. G. Natadze.[3] The outcome appears to have been similar too, and one of the conclusions is that the result "seems to us to constitute a powerful argument against the mechanistic explanation in terms of association, which describes the naming process as a simple association between the attribute and the name for it."

Now when we observe, for instance, that Alternative c was accorded 25 votes for representation of Designatum 1, no other Alternative gaining more than 5 votes, we may not conclude that Alternative c is particularly suited to representation of Designatum 1. Indeed, the Soviet psychologist's interest is the processes by which the "objectively indifferent (my italics) sound complex" is "appreciated by the subject as naming a particular attribute". We conclude only that Alternative c and the other popularly selected words are intrinsically more generally acceptable for admission into the language. To discover whether Alternative c, rather than any

[3] Natadze 1957: 306-8

other of the popular words, is particularly suited to representation of Designatum 1, a further experiment is needed.

Experiment 2

A colleague from Osmania University, Hyderabad, undertook this experiment with a group of undergraduates. Each undergraduate was provided with a simple description of Experiment 1, a list, in "ordinary" writing and in alphabetical order, of the eight popular words, (Designatum) 1 (Alternative) c, 3c, 3g, 4j, 5e, 5f, 6c, 7f, from Table 1[4], and the original definitions for the six Designata 1, 3, 4, 5, 6 and 7, with a note that for two of the definitions two words had been found equally appropriate. The student then matched each definition with the word or words he thought most likely to have been chosen for it.

Assuming that the relationship between sign and meaning is arbitrary, we predict that, since no feature of any of the signs can make it more acceptable for representation of a particular designatum than any other sign, the votes for all the artificial words as representatives of particular designata will be, within statistical limits, equal. This is how the voting went:

Alternative:	Designatum:					
	1	3	4	5	6	7
1 c	[9]	8	5	2	8	9
3 c	7	[19]	3	2	2	2
3 g	1	[16]	4	4	5	4
4 j	2	-	[28]	1	6	2
5 e	9	1	4	[19]	4	-
5 f	3	3	-	[23]	9	3
6 c	15	1	2	3	[9]	5
7 f	2	4	1	1	9	[25]

The prediction was fulfilled with respect to Alternative 1c, while Alternative 6c, previously selected as best representing Designatum 6, has been voted more appropriate for representation of Designatum 1. The prediction was unfulfilled with respect to Alternatives 3c, 3g, 4j, 5e, 5f, and 7f, each of which was assigned to the designatum which, in Experiment 1, it had been selected to represent; and the postulate upon which the prediction was based is therefore falsified.

[4] namely, abnoss, dilrima, ma, moopem, neach, stobe, tave, ziferi

The reader may reasonably hesitate to accept this last statement as a statement of fact, based as it is on limited experiment in particular circumstances. As far as the object of this modest paper is concerned, distinction between article of faith and scientific statement, it is indeed not at all necessary to consider the fact established. Rather the contrary, for the statement "the postulate upon which the prediction was based is therefore falsified" is scientific (in the Popper sense) simply because it is itself falsifiable (or verifiable), by experiment.

Having undermined, if not destroyed, an article of faith, one is curious to examine the means by which the article of faith was substantiated. We may go back to De Saussure for denial of a relationship, other than the socially conferred one, between sound and meaning, and for what have become typical evidence and assertion. The evidence is that what is b-ŏ-f on one side of a frontier is o-k-s on the other, the assertion, that the meaning of "sister" might just as well be represented by no matter what (De Saussure's words) other sequence of sounds.[5]

However, there is no practical or logical obligation to assume that, in order for there to be a substantive relationship between sound and meaning, there must be one sound sequence per designatum throughout the world. The existence of such a relationship is entirely compatible with representation of one designatum by b-ŏ-f, o-k-s and indeed as many sequences of sound as we actually find. One would be wrong to conclude, from the external differences between pine needle, oak leaf and banana leaf that there could be only arbitrary relationships between leaf and leaf function.

The methodological error with respect to "evidence" is that reference of a postulate to observation "proves" (De Saussure's word) the postulate. It shows that postulate and evidence are not incompatible; and that is all. It is the particular inference from the evidence, which De Saussure takes for granted, that requires verification; for an observation may be equally compatible with a number of inferences.

It is perhaps always tempting to feel that what is asserted positively in well defined terms and with reference to observation is more "scientific" than something vaguely stated; it may be, and may not. It would not occur to us to call Alexander Pope's line, "The sound must seem an echo to the sense" scientific, yet methodologically Pope compares favourably with De Saussure, for Pope proceeds to demonstration of what he means, for verification on our part; whereas De Saussure's assertion that the meaning of "sister" could just as well be represented by no matter what other sequence of sounds is not experimentally verifiable or falsifiable, and so far as one can judge was not made with this criterion in mind: it is, in fact, an unscientific kind of statement.

The word "scientific" is used both as a classifying, and, more popularly, as a descriptive adjective. In the first (and more "scientific") use, there is understood, I think, a methodological element absent from much linguistic study. Perhaps necessarily, for language, like life of which it is a part, is messy rather than orderly, and its behaviour only partially, often only statistically predictable. The problem of the relationship between sound and meaning is a messy problem. Languages, with

[5] De Saussure 1957: 100, 159.

the peoples that use them, have evolved, each developing its own "cut"[6], to which - however difficult to define - old and new acoustical phenomena adopt. As a language's words have altered, so presumably its speakers' feeling of appropriateness of sound for meaning has altered too, keeping fairly well in step. There is no chance of "sister" getting represented by no matter what other sequence of sounds, since, it seems, sounds are not indifferent counters in sets of permutations. Insisting on the "mathematics called linguistics", our work may be the less, not the more, scientific; for one opposite of "mentalism" is "fundamentalism", with a register of articles of faith.

---

6 Martinet 1961: preface.

POPPER, K., 1935, Logik der Forschung (The Logic of Scientific Discovery, London, Hutchinson, 1959)

NATADZE, R.G., 1957, "Studies on Thought and Speech Problems", in Ed. O'CONNOR, N.O., 1961, Recent Soviet Psychology, Pergamon Press, 306-308.

DE SAUSSURE, F. Cours de Linguistique Générale, Paris, Payot.

MARTINET, A. 1961, A Functional View of Language, Oxford. Clarendon Press.

PROGRAMMING A REMEDIAL PRONUNCIATION COURSE*

Patricia Quaife

(University of Auckland)

If we accept Robert Lado's three point definition of programmed learning as: "learning by means of materials that break up the task into minimal steps, requiring an active response for each step, and providing an immediate check on the correctness of the responses"[1] then it is not appropriate at the present moment to give the label "programme" to any phonetics course, whether remedial or introductory. When every language laboratory booth is provided with both an intonation counter and a modified spectograph, this may be the case, but until that time I prefer to use the more exact, if clumsier expression - "systematic phonetics course". However, in accordance with the title of this paper, I intend to use the verb "programme" meaning organising a corpus of remedial phonetic material in such a way that the task of re-learning is broken up into minimal steps, with each step requiring an active response, thus fulfilling the first two requirements of Lado's definition. I hope to show in due course why his third requirement cannot be fulfilled in a phonetics course.

Programming a remedial phonetics course for any foreign language presents a number of problems, due firstly to the nature of language and secondly to the previously acquired foreign language habits of the students. Both these points need clarification. Few would argue with the following broad working definition of language as a system of organised sounds emitted by human beings for purposes of communication. The important point being that it is a system of sounds that is at the heart of human language, while the complex arrangements and permutations of sounds into structures and lexical items may be classed as forms rather than substance. Without substance (sounds) there can be no form. The American linguist Gleason illuminates this with his statement that "a speaking knowledge of a language...requires very close to a one hundred percent control of the phonology and control of from fifty to ninety percent of the grammar, while one can frequently do a great deal with one percent or even less of the vocabulary."[2] Jakobson prefers the notion of an ascending ladder of liberty where the degree of liberty as regards phonemes is nil, is circumscribed for the phonetic realization of these phonemes in words, increases slightly for the formation of sentences, and increases enormously for the combination of sentences into larger units.[3]

It is also profitable here to look briefly at native language learning habits, and at adult verbal behaviour. While the average child has full control of the sound system of his native language by the age of six or seven, his control of the grammatical system goes on increasing until he is eleven or twelve, and acquisition of vocabulary items continues for life. For the adult native speaker who has been using the sound system of his language automatically for x number of years, speech characteristics have become part of his own personality, so strongly have they been ingrained and reinforced by habit. Everyone has had the experience of being at loss for a word, has rephrased or restructured a sentence without relinguishing his control of sounds. The most naive of listeners to a non-native speaker speaking, say, English is immediately aware of the slightest non-conformation to the English sound system, which he is inclined to stricture globally as a "funny accent" or a foreign accent.

---

*Revised text of a paper given at the University of Auckland Language Laboratory Workshop, 25th August, 1965.

[1]Lado 1964: 220 [2]Gleason 1961: 343 [3]Jakobson 1963: 47

Having established and made habitual the use of a particular sound system over a number of years, the adolescent and adult learner of a second language will automatically transfer all his previously acquired speech habits into the foreign language. He will hear and reproduce the foreign sounds in terms of his own system, he will transfer the grammatical structures from one language to another, and all the "meanings" and connotations of native language vocabulary items will be transferred to the foreign language form.

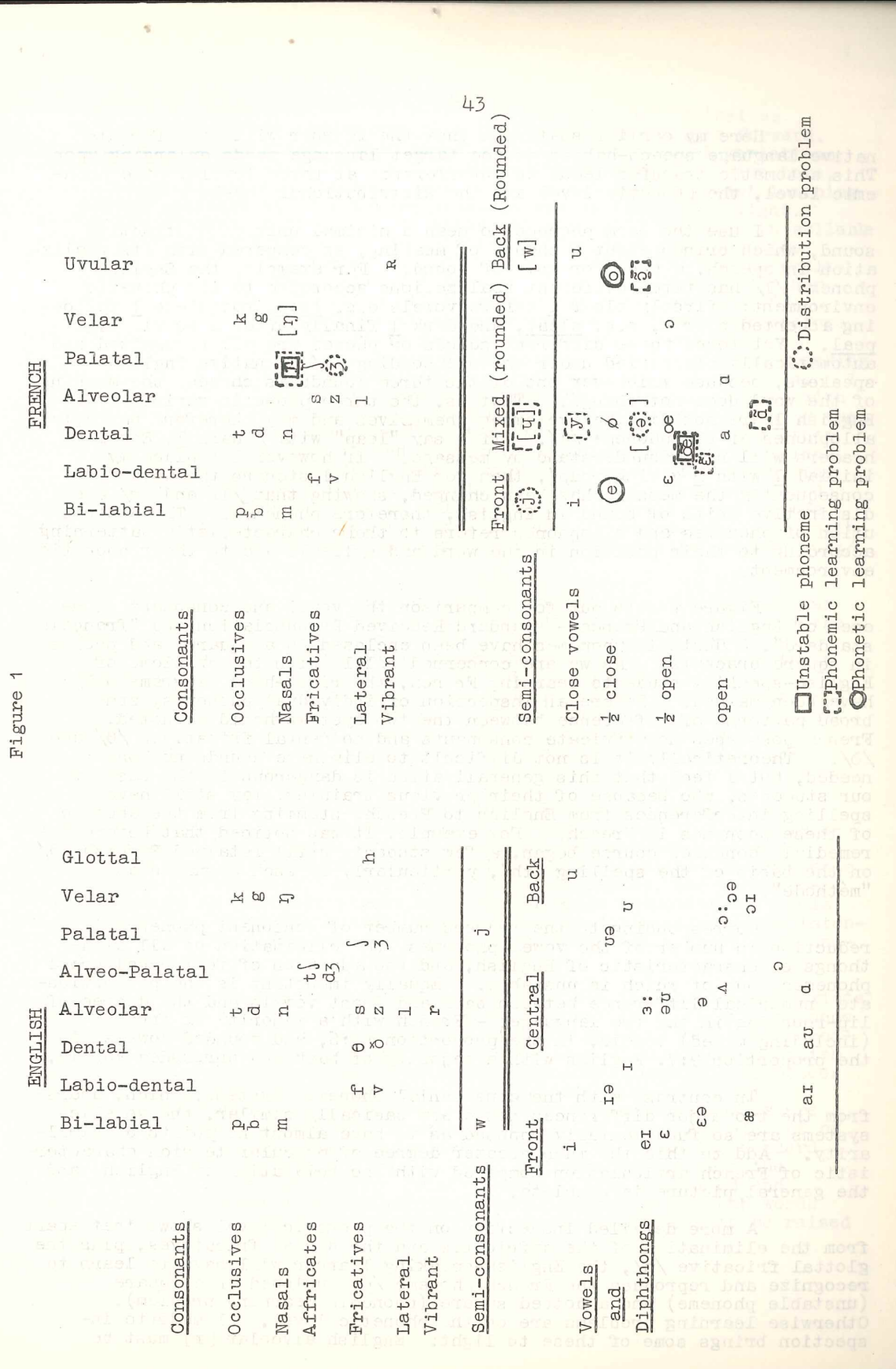
The problem stated in general terms, we must proceed to a specific case - the retraining of English-speaking New Zealand students of French, who have had between four and six years teaching by the traditional grammar/translation method, usually with non-native teachers. The vast majority of these students have developed their reading and writing skills in French to a point far in advance of their listening and speaking skills. This order of development of foreign language skills is diametrically opposed to that put forward by the proponents of the audio-lingual (or audio-visual) method - listening, speaking, reading, writing. It is outside the scope of this paper to discuss the principal tenets of this method, but it must be emphasized that the above order stems from a particular view of language, of language-learning psychology, and of student aims. It should also be remembered that we are not dealing with beginning students, but with university students who have entered upon the academic study of French as a discipline as well as a means of communication.

With this retardation of listening and speaking skills, we have the two-fold problem of training students in aural comprehension and discrimination, and in the production of acceptable French sounds. The latter aspect of the problem is increased by the formation and reinforcement of unacceptable speech habits over their previous four to six years of study. These students may now feel that they are taking a step backwards unless fresh motivation and new goals be instilled in them. It is important that they be made aware that their former standard is not acceptable to an informed non-native speaker (the teacher), and equally unacceptable to a native speaker.

The object of the remedial phonetics course is the deceptively simple one of improving the pronunciation, and the communication powers of the student, by exposing him to carefully graded materials, recorded by native speakers, for comprehension and repetition so that his powers of discimination and aural reporduction will be brought up to a level comparable with his reading and writing skills. It is illusory and unrealistic to expect his powers of speaking to equal his reading and writing skills. Again a quick look at the native situation is helpful. We all recognize and are familiar with many written vocabulary items which we do not, and in some cases may be unable to produce in speech. To a much smaller extent, this statement is also true of some grammatical structures which in their aural and written forms may be part of our passive reservoir, but not available for automatic production. I stress this point, because it seems to me that many applied linguists have created for the average student an artificially high standard of performance in the foreign language, expecting "near-native fluency", while at the same time misunderstanding the nature of fluency. Native fluency certainly involves automatic control of the sound system, but does not of necessity imply rapid, faultless speech, and is not proof against fumbling for words, backtracking, slowness of articulation, lack of vocal punctuation, and structures as yet unrecorded by linguists.

To predict, and where possible to minimise transfers and interference from the native language to the target language, a systematic comparison of the sound systems and prosodic features of the two languages should be undertaken. This analysis, together with a list of the frequency of occurrence of sounds in the target language, will determine the order of presentation of the material.

Figure 1



Here my earlier statement that the learner will transfer his native language speech-habits to the target language needs enlarging upon. This automatic transfer leads to interference at three levels: the phonemic level, the phonetic level and the distributional level.

I use the term phoneme to mean a minimal unit of distinctive sound, which brings about a change of meaning, as compared with its realization in speech, a phone or unit of sound. For example, the English phoneme /l/ has three different realizations according to its phonetic environment: firstly clear l before vowels e.g. leap, voiceless l following accented p or k, e.g. pleat, and dark l finally after a vowel, e.g. peal. Yet these three different sounds or phones are all recognized and automatically classified under the one heading /l/ by native English speakers, because whichever one of the three sounds is chosen, the meaning of the word does not change. That is, the three phonetic varieties of English l are not distinctive among themselves and may therefore be called allophones of the phoneme /l/. If I say "leap" with a dark l, English hearers will not misunderstand my message. If however I replace my initial l with r giving "reap", then for English listeners the word and consequently the message has been changed, showing that /l/ and /r/ are distinctive units of sound in English, therefore phonemes. The distribution of phonemes and allophones refers to their characteristic patterning according to their position in the word and syllable and to their phonetic environment.

Figure 1 sets out for comparison the vowel and consonant phonemes of English and French - standard Received Pronunciation and "français standard". Unstable phonemes have been enclosed in a square, and phones in square brackets. As we are concerned solely with the problems of English-speaking students learning French, the right-hand diagrams only have been marked. Before an inspection of individual phonemes, some broad patterns of difference between the two systems should be noted. French possesses no affricate consonants and no dental fricatives /θ/ and /ð/. Theoretically it is not difficult to eliminate sounds no longer needed, but I feel that this generalization is dangerous in the case of our students, who because of their previous training, may still have spelling interferences from English to French, stemming from the absence of these phonemes in French. For example, it was noticed that before the remedial phonetics course began, a few students still retained English /θ/ on the basis of the spelling "th", particularly in such cognates as "méthode".

Corresponding to the reduced number of consonant phonemes is a reduction in number of the vowel phonemes, the elimination of all diphthongs so characteristic of English, and the addition of four nasal vowel phonemes, one of which is unstable. Equally important is the proportionate numerical difference between back and front vowels and the degree of lip-rounding in the two languages - French with a majority of front (including mixed) vowels, in the proportion 10:6, and rounded vowels, in the proportion 9:7, English with a majority of back and unrounded vowels.

In contrast with the consonantal phonemic systems, which, apart from the two major differences noted are basically similar, the vocalic systems are so fundamentally opposed as to have almost no points of similarity. Add to this the much greater degree of muscular tension characteristic of French articulation compared with the relaxation of English, and the general picture is complete.

A more detailed inspection on the phonemic level shows that apart from the elimination of the affricates and the dental fricatives, plus the glottal fricative /h/, the English-speaking learner will have to learn to recognize and reproduce the French phoneme /ɲ/ enclosed in a square (unstable phoneme) and a dotted square (phonemic learning problem). Otherwise learning problems are on the phonetic level. Immediate inspection brings some of these to light: English alveolar [r] must be

replaced by the French uvular [ʀ] while English alveolar occlusives [t, d, n] become dental occlusives with the tongue tip further forward. The diagram does not reveal other phonetic and distributional variations which require a more detailed analysis than it is possible to give here, but I propose to mention the main ones, so that all the important learning problems and subsequent programming problems may be brought to light. Voiceless occlusive stops [p, t, k] which are strongly aspirated in syllable initial position in English are unaspirated in French - e.g. cognates English "pain" [pʻeɪn] : French "paine" [pɛn]. French [l] has no positional variations comparable with the English clear [l]/ dark [l], other than devoicing in final position after [k] and [p] e.g. "mon oncle" - [mɔnõkl̥]. It is clear, with front vowel resonance. Students therefore have to make a conscious effort to produce the correct sound after a vowel. French [ʒ] has minor problems of distribution for English-speaking students in that English [ʒ] does not appear in word initial position. I quote from Lado[4]: "English speakers will transfer their /ʒ/ phoneme with its limitation into French and will thus have difficulty with learning the word initial /ʒ/ in that language." It seems to me that Lado and other writers on the subject have tended to magnify such distribution problems, which, unless they also involve unfamiliar phonetic variants, are by far the most readily solved. Although English /ʒ/ does not appear in word initial position, it does occur in syllable initial position within the word in a number of cases, e.g. measure, leisure, treasure, etc., and with such similarity of environment it is not unduly difficult to produce a /ʒ/ in word initial position.

We are confronted with far greater variation and therefore far more linguistic interference when we consider the vocalic systems for the two languages. I include for convenience the semi-vowels (or semi-consonants) in this category, even though they function as consonants in the utterance. Vowels enclosed in a dotted square - the four nasals, plus /y/ and [ɥ][5] represent phonemic learning problems: those circled represent important phonetic learning problems, those unmarked represent minor phonetic learning problems, while /j/ and the phone [ə] (dotted circle) represent distribution problems, the latter with a high degree of spelling interference.

Prosodic features characteristic of French and English also differ fundamentally, so that interference is not limited to the segmentals. French with a majority of open syllables has a syllable-timed rhythm, with the phenomena of enchainement and liaison within the rhythmic group, while English with a majority of closed syllables, has a phrase-timed rhythm, plus phonemic stress, and considerable reduction of vowels in non-tonic positiion to neutral [ə] or [ɪ]. Predominantly rising intonation of French similarly contrasts with the predominantly falling intonation of English.

Because no adequate description of the New Zealand dialect exists, I was forced to reproduce the standard English vowel and consonant phoneme diagrams. While the consonantal system does not appear to have been modified to a significant degree, the opposite is true of the vocalic system and some anecdotal remarks about New Zealand/standard English differences may help to pinpoint some of the corresponding interferences and to justify the programming of the course. Most so-called pure vowels are liable to be diphthongised, e.g. beat = [bəit] while the first element of the diphthongs [eɪ] and [ou] is considerably more open than in standard English, e.g. [eɪ]→[æɪ] and [ou]→[ʌu], that is further away from the French pure vowels [e] and [o]. Front vowels tend to be raised and nasalised particularly when followed by a nasal consonant in such words as "pan", "pen". Central vowels e.g. [ɜ:] as in "bird" tend to be raised

---

[4]Lado 1957: 17.

[5]See p.48 below.

with lip-rounding, giving a sound nearer to the French [ø] than standard English [ɜ:] .

Until fairly recently, it was accepted that the most systematic way of teaching pronunciation was to deal with phonemic, phonetic and distributional problems in that order, with special study of prosodic features fitted in when and where possible. The relative frequency of sounds has now entered as another criterion by which order of presentation may be worked out - a criterion which does not always accord with the first criterion of phonemic and phonetic interference. On figure 2 are set out the comparative frequencies of occurrence of French sounds, (not phonemes) as determined by two French phoneticians working in the United States, Léon[6] and Delattre.[7] English consonant figures according to the English phonetician Gimson[8] are given for comparison.

Of French consonant phonemes, only /ɲ/ has to be acquired by the English-speaking student. Yet its frequency is so low as to hardly justify special study before a number of other features. It is interesting to note here that in her Exercises systématiques de la langue française Monique Léon omits /ɲ/ entirely, presumably because the phonetic realization of [nj] is acceptable. Similarly, of the vowel phonemes, /œ̃/ which is phonetically very close to /ɛ̃/, is third to bottom (Léon) and second to bottom (Delattre) on the frequency lists. The functional yield of /œ̃/: /ɛ̃/ is limited to five or six minimal pairs indicating that the teaching of /œ̃/ may safely be omitted at a preliminary stage.

To these criteria, must be added the particular needs and circumstances of students, the materials available, and the planned remedial course should take all of these factors into account. In this particular case, the most important features of French pronunciation had, this year, to be telescoped into a nine-week language laboratory course for all students from first to third year inclusive. I will describe the programming of this in some detail, giving reasons for the choices made:

Lesson 1 - vowels [e, ø, o]

Lesson 2 - close vowels [i, y, u]

Lesson 3 - semi-consonants [j, ɥ, w]

Lesson 4 - nasal vowels

Lessons 5 and 6 - enchainement and liaison

Lesson 7 - e mute, elimination, retention and "groupes figés"

Lesson 8 - tonic accent, rhythmic groups, introduction to intonation.

Lesson 9 - intonation with several rhythmic groups, interrogative intonation.

The course began with the pure vowels [e] and [o], strictly speaking phonetic problems, but equally difficult as some phonemic problems for New Zealand students (see above) who have to change the phonetic nature of their well-established diphthong [eɪ] or [æɪ]. There was the additional problem of[e] : [i] discrimination for some students who heard French [e] as [i], due to the considerably lower starting point of the first element of their own diphthong. The very high frequency of occurrence of [e] (8.14% - Delattre and 6.5% - Léon) was another reason for its high place in the list of priorities. [ø] was included to complete the horizontal contrast of the half-close vowels.

[6] Léon, P. & M. 1964: 42. [7] Delattre 1964a: 89 (Vowels); Delattre 1964b: 180 (Consonants). [8] Gimson 1962: 214.

Figure 2

Comparative frequency of occurrence of French sounds

Vowels	Léon %	Vowels	Delattre %		
a	8.1	e	8.14		
e	6.5	a	7.04		
i	5.6	i	5.23		
ɛ	5.3	ə	3.21		
ə	4.9	ɑ̃	3.20		
ɑ̃	3.3	ɛ	2.83		
u	2.7	u	2.70		
y	2.0	ɔ	2.13		
ɔ̃	2.0	y	1.98		
o	1.7	ɔ̃	1.62		
ɔ	1.5	o	1.10		
ɛ̃	1.4	ɛ̃	1.03		
ø	0.6	œ	0.76		
œ̃	0.5	ø	0.72		
œ	0.3	œ̃	0.44		
ɑ	0.2	ɑ	0.01		
Consonants		**Consonants**		**Eng. cons.**	**Gimson %**
r	6.9	r	8.67	n	7.58
l	6.8	l	6.14	t	6.09
s	5.8	t	5.59	d	4.89
t	4.5	s	5.06	s	4.81
k	4.5	p	4.60	l	3.66
p	4.3	d	4.18	ð	3.56
d	3.5	k	3.67	m	3.22
m	3.4	m	3.46	k	3.09
n	2.8	n	3.02	r	2.91
v	2.4	v	2.57	w	2.81
ʒ	1.7	j	1.86	z	2.46
f	1.3	ʒ	1.67	v	2.00
b	1.2	f	1.48	b	1.97
j	1.0	z	1.35	f	1.79
w	0.9	w	1.33	p	1.78
ɥ	0.7	b	1.31	h	1.46
z	0.6	g	0.65	ŋ	1.15
ʃ	0.5	ʃ	0.57	g	1.05
g	0.3	ɥ	0.49	ʃ	0.96
ɲ	0.1	ɲ	0.15	j	0.88
				dz	0.60
				tʃ	0.41
				θ	0.37
				ʒ	0.10

In such a short course it was not possible to include the vertical contrasts of the middle vowels: [e:ɛ, ø:œ, o:ɔ]. With a few exceptions, the yield of these pairs is extremely low in modern French, and students experience difficulty in production only with the [ø:œ] contrast, which, as Léon[9] says, is truly distinctive only in the artificial case of isolated words such as "jeune/jeune" - [ʒœn:ʒøn]. For practical purposes, [e:ɛ, ø:œ] and [o:ɔ] are regarded as being in complementary distribution.

Lesson 2 introduced the first of the new phonemes proper: front rounded close vowel [y] and the contrasting close vowels, front unrounded [i] and back rounded[u]. The logical follow-up was the corresponding semi-consonants with the difficult distinction [ɥ:w] and the minor distributional problem of [j] which in English does not appear in final position.

The four nasal vowels presented in lesson 4, with the stress on the [ɛ̃:ɑ̃:ɔ̃] distinction, completed the vowel section of the course.

It will be apparent that this analysis (particularly that on which lesson 3 is based) departs slightly from the orthodox phonemic solution. Gimson[10] notes that the phonemic analysis and solution of a language depend to a certain extent on the purpose of the analysis, e.g. purely linguistic, pedagogical etc. For example, [ɥ] and [w] being in complementary distribution with /y/ and /u/ respectively, they are rightly regarded as being allophones of /y/ and /u/. But just as the student has to learn to distinguish between /y/ and /u/, so he has to learn to distinguish between [ɥ] and [w] in such pairs as "lui:Louis, buée:bouée" - [lɥi:lwi], [bɥe:bwe]. For this reason [ɥ] has been classed as a phonemic learning problem. For an English-speaking student, [ɥ] involves both a new contrast and the production of a new sound.

The remaining five lessons concentrated on the most important prosodic features of French as compared with English.

Within each lesson, an attempt has been made to fulfil the first two conditions of Lado's definition of a programme. Students are required to make an active response, by repeating after the model, the word, phrase or sentence, and according to the available material,[11] the learning process has been cut up into minimal steps, progressing from easier to more difficult material.

The following examples come from lesson 3 (semi-consonants). [j] is presented first as the problems it raises are minor compared with those raised by the opposition [ɥ:w].

a) [j] in word initial position, which is familiar:
hier, y a-t-il, hiatus, yoyo, etc.

b) intervocalic [j] in syllable initial position within the word. Similar environment to a):
billet, payé, merveilleux, voyage, etc.

c) [j] in final position - an unfamiliar position for English speakers - and preceded by different vowels:

[9] Léon, P. n.d.: 24

[10] Gimson 1962: 45-46

[11] Material used for lesson 3 comes from M.Léon's Exercises systématiques de prononciation française, (2 vols.) plus three long-playing gramophone records. Exercises a), b), and c) come from vol.1, p.59, exercises d) and e) from vol.1 p.55-6. The programming described in this paper does not follow the order found in the Exercises, which are not designed solely for English-speaking students of French.

fille, bille, quille
soleil, merveille, pareil
taille, maille, paille
mouille, fouille, houille
seuil, oeil, Auteuil etc.

d) opposition [ɥ:w] introduced in minimal pairs for comparison. Subsidiary difficulty of unfamiliar consonant clusters: [lw, nw, bw]
oui:huit, Louis:lui, mouée:muée, etc.

e) short sentences concentrating on the unfamiliar [ɥ]
je suis étudiant, je suis chimiste, etc.
c'est lui qui parle, c'est lui qui écrit, etc.
il faudrait qu'il puisse partir, (finir, venir, etc.)

A twenty or twenty-one lesson first-year remedial course could be programmed as follows:

Lesson 1 - Tonic accent and rhythmic group theory.

Lesson 2 - Declarative intonation.

Lesson 3 - Interrogative intonation.

Lessons 4-9 - vowels and semi-consonants as in shorter course, with the addition of the open-close variation [e:ɛ, ø:œ, o:ɔ] and the complementary distribution of [i:j, y:ɥ, u:w].

Lesson 10 - Nasal vowels continued. Oral:nasal opposition.

Lessons 11-12 - Enchainement and liaison (particularly as applied to nasals).

Lessons 13-14 - e mute.

Lesson 15 - Opposition e mute: [a] (type: il te dit:it t'a dit) and sequence: consonant + r or l, + e mute. (type: mon oncle).

Lesson 16 - [r].

Lesson 17 - [r] and [l].

Lesson 18 - Initial [p, t, k] (non-aspirated).

Lesson 19 - Geminated consonants.

Lessons 20-21 - More advanced intonation.

In this fuller course, the general prosodic framework is given first, lessons on nasal vowels, e mute and intonation have been increased, and lessons on [r] and [l] (the two most frequent French consonants, and for English speakers the most difficult), [p, t, k] (the next most frequent after s) have been incorporated.

In coming years it will also be necessary to programme in depth, that is provide more advanced tapes for students, who although having done a first-year course, still have faults of pronunciation, but who are not going to be content with simple repetition. The majority of these more advanced exercises need a good knowledge of intonation, as the student is required to change either the form or the intonation, (or both) of the sentences proposed, which are also designed to practise particular sounds.

A few random examples of such exercises,[12] which are mostly of the four-phase type: stimulus, response, reinforcement or correction, student repetition, follow:

a) to practice [j] in various phonetic environments, the student is required to answer each question, beginning with "oui" and repeating the terms of the question:

Vous aimez travailler? Oui, vous aimez travailler!
Vous aimez mieux travailler? Oui, vous aimez mieux travailler!
C'est un vieil appareil? Oui, c'est un vieil appareil!

b) to practice [ɛ̃] the student is required to begin each sentence with "il vient" and to end with the indication given, modifying the intonation at the same time:

bientôt	Il vient bientôt.
demain matin	Il vient demain matin.
de moins en moins	Il vient de moins en moins.

c) to practice [w], the student is required to add each new element to the sentence, again modifying the intonation appropriately:

Je me demande (pourquoi)
Je me demande pourquoi (je le vois)
Je me demande pourquoi je le vois (ce soir)
Je me demande pourquoi je le vois ce soir (pour la troisième fois)
Je me demand pourquoi je le vois ce soir pour la troisième fois.

In spite of the fact that these exercises are performed in the language laboratory with facilities for hearing responses both immediately through the activated headphones, and later when playing back the tape, a phonetics course by its nature cannot fulfil Lado's third requirement of a programme - provision of an immediate check on the correctness of the responses. If we take as a criterion of the correctness of a response, complete acceptability to a native speaker, the number of students achieving this would be as low as three to five percent. Obviously criteria of right and wrong are inapplicable. By successive approximations the student can be encouraged to come closer and closer to the goal of acceptability, but without a good deal of auditory experience of French, he will usually be unaware of his improvement or lack of improvement. At best, he may have a general impression that his attempt is not acceptable, without being able to say why and how.[13] For this reason, opinions are divided on the desirability of the language laboratory in this situation. F. Marty says: "The language laboratory is not a good solution because progress is unpredictable for these features of pronunciation."[14] With classes of ten students or fewer, his conclusion may be justified, but in most institutions today it is simply not relevant. Léon holds the opposite view, and his wife Monique Léon has created comprehensive exercises designed for foreign students of French and to be performed where possible in the language laboratory.

The laboratory is of value in the remedial situation where each student is subjected to hearing native speakers of French, and so gradually fixing the meaningful distinctions and correct sounds in his auditory

[12] These exercises were designed at the Centre de Linguistique Appliquée, Université de Besançon, and are not available commercially.

[13] see Rivers 1964: 155-6. [14] Marty 1960: 121.

memory. For a sound to be correctly produced, it must be correctly perceived; and to be correctly perceived and produced, intensive individual practice is necessary. Although of value, the laboratory is not, in this situation, a teaching machine and the monitoring of the class becomes of prime importance. Without careful monitoring, students will reward with their inner approval, their unacceptable responses, and thus reinforce these responses before the goal of acceptability to the teacher is reached. When this occurs, the unacceptable responses will become entrenched and not susceptible to improvement.[15]

While the laboratory will continue to be of value, especially with large numbers of students, for the maximum exploitation of phonetics courses, it seems that criteria for the organization and programming of such courses may be radically changed. I have tried to show that the criterion of phonemic difference does not always correlate with that of frequency. It has recently been put forward that the reactions of linguistically naive native speakers to the pronunciation of their language by non-native speakers should also be taken into account, and that those features of mis-pronunciation to which they react most strongly should be corrected first, regardless of phonemic, phonetic or frequency considerations. Some linguists may regard this as far too unscientific, but then human beings do not react scientifically, nor do they always speak scientifically. Delattre couches a similar concept in more formal language when he speaks of the "characteristic auditory impression of a language."[16] In his opinion, "the following consonants should be emphasized in decreasing order of importance"[17] (and I give his own frequency figures for comparison):

[r] (8.67), [ʒ] (1.67), and the two consonants with the lowest frequency of all: [ɥ] (0.49), and [ɲ] (0.15).

There seems to be no lack of conflicting criteria in the field of applied phonetics, without taking into account individual features of each teaching situation. Whatever the criteria finally proved to be the most productive of results, any new factors which will enable students to acquire an acceptable pronunciation as economically as possible will be welcome.

Postscript. A digital computer[18] at the language laboratory booth, as developed by Harlan Lane of the Behavioural Analysis Laboratory, University of Michigan, can feed back to the student accurate information as to the closeness of approximation of his performance to that of the model (pitch, intensity and speed only)[19] but cannot tell the student how to improve his performance further.

[15] Rivers 1964: 53. [16] Delattre 1964a: 88. [17] Delattre 1964b: 181.

[18] SAID - Speech Auto-Instructional Device. [19] van Teslaar 1965: 91-2.

BIBLIOGRAPHY

Delattre, P., 1964a. "Comparing the vocalic features of English, German, Spanish and French." IRAL II, 2: 71-97.

Delattre, P., 1964b. "Comparing the consonantal features of English, German, Spanish and French." IRAL II, 3: 155-203.

Gimson, A. C., 1962. An Introduction to the Pronunciation of English. London, Arnold.

Gleason, H.A. jr., 1961. An Introduction to Descriptive Linguistics. Rev.ed. New York, Holt, Rinehart & Winston.

Hollyman, K.J., 1964. "A quantitative method of predicting areas of phonemic interference." Te Reo 7: 36-43.

Jakobson, R., 1963. Essais de linguistique générale. Paris, Ed. de Minuit.

Lado, R., 1957. Linguistics Across Cultures: Applied Linguistics for Language Teachers. Ann Arbor, University of Michigan Press.

Lado, R., 1964. Language Teaching: A Scientific Approach. New York, McGraw-Hill.

Léon, P., n.d. Aide mémoire d'orthoépie. Besançon, Université de Besançon.

Léon, P. & M., 1964. Introduction à la phonétique corrective. Paris, Hachette et Larousse.

Léon, M., 1964. Exercises systématiques de prononciation française. 2 fasc. Paris, Hachette et Larousse. 3 long-playing gramophone records.

Marty, F., 1960. Language Laboratory Learning. Wellesley, Mass., Audio-Visual Publications.

Rivers, W.M., 1964. The Psychologist and the Foreign-language Teacher. Chicago & London, University of Chicago Press.

van Teslaar, A.P., 1965. "Learning new sound systems: problems and prospects." IRAL III, 2: 79-93.

THE PROGRESS AND TECHNIQUES OF MECHANICAL TRANSLATION*

F. R. Wyatt

Introduction

If we can properly regard language as one of man's inventions, then it must have been one of the earliest that he made. No subsequent invention can be considered as having greater fundamental importance. Language is man's tool for the storage and dissemination of concepts. Language is essential for the preliminary organization of any complex group activity.

With the technological progress that language has made possible, concepts can be transmitted over greater distances than before and to larger numbers of people. Written language allows the storage of information for subsequent generations, but the written word can be read only by the literate proportion of the world's population. Radio broadcasts, on the other hand, can be heard in every part of the world and do not demand literacy on the part of the listener.

However, the spread of knowledge and understanding among the peoples of the world is now restricted, not so much by the physical nature of the communication channels used as by the nature of language itself. The work of translation and interpretation is probably the greatest obstacle.

Because of the number of languages in everyday use [c.3,000] and the manual labour involved, it is currently impractical to translate all the scientific[1] and technical[2] papers published each year into all the languages man uses.

As an example, the Commonwealth Agricultural Bureau receives annually more than 70,000 scientific papers. This frustrating situation has led some of those people concerned with information - in the widest sense - to look for a way to take the work of translation out of the hands of men and women and put it into the chromium-plated claws of the machine.

Hitherto, much of the writing on the progress of mechanical translation has been at the level of the scientific report. The technical nature of these reports has tended to restrict their readership to those having a direct interest in the subjects discussed. This essay, whilst not aimed at the layman, is an attempt to provide a readable abstraction of these papers for those whose time and/or training permit of no deeper study. Perhaps the experts will forgive me if I have made what

---

*This paper is based largely upon the reports and papers referred to herein. The writer is most grateful to various members of staff of the University of Auckland for their encouragement and for the opportunities to discuss with them the topics reviewed, and particularly grateful to W. O. Droescher for the loan of many reports.

[1]A recently-published list of scientific periodicals includes 59,404 entries, some of which admittedly have now ceased publication. However, it is probably safe to estimate that more scientific and technical periodicals have newly appeared since 1960 than those which disappeared in the preceding sixty years. See World List of Scientific Periodicals.

[2]A British publisher's catalogue listed 8,800 technical books as being in print in November, 1964. See Technical Books in Print.

seems to them some gross over-simplifications. The complexity of the material is not only due to the complexity of language in a general sense. The work of planning the beginnings of a computer system requires re-examination of many ideas which have been previously regarded as self-evident truths. As an example, it was once thought that automatic translation equipment designed for working from just one language to one other language would be more efficient than a multi-lingual machine. This idea was queried by the Russian linguist N. D. Andreyev and now is generally believed invalid.

Early History

Machine translation research at the University of Washington, Seattle, began in November, 1949. (The term "machine translation" has gradually given way to "mechanical translation" and the abbreviation "MT" is commonly used in the literature.) Research began with close cooperation between linguists and computer engineers to produce a pilot-model translating machine. This machine was planned to test a number of input recognition procedures devised by Dr Erwin Reifler. Later work led to a draft programme for use with a computer, envisaged as having so large a storage capacity that each entry in the memory would store not only lexical equivalents for input words, but also control-symbols for operating the machine and editing-symbols for modification of the output text. In May 1956, the study of large-capacity rapid-access memory technique for translation of Russian into English was the subject of a contract with the International Telemeter Corporation of Los Angeles. Like the earlier research, this was a joint effort of the linguists headed by Dr Reifler, and the engineers headed by Professor W. R. Hill. The linguist responsible for the special study of Russian language required for this task was Dr L. R. Micklesen. In the report[3] on this project came the first reference to the proposed use of a rotating optical disk as a memory device.

The first Mechanical Translation Conference was held at the Massachusetts Institute of Technology in June 1952. This was apparently the first time that the individuals working on the problems of MT had met to discuss their differing approaches. Some of these men were not linguists, but electronic computer engineers, and it was a principal conclusion of the conference that for certain types of input information, MT had become a real engineering possibility.

In March 1954, the Department of Modern Languages at Massachusetts Institute of Technology published[4] the first journal devoted entirely to the subject of mechanical translation. The journal appeared three times each year and provided a well-coordinated display of the problems and progress of researchers in mechanical translation from many parts of the world. The first doctoral thesis on MT, entitled "A Study for the Design of an Automatic Dictionary" was presented by A. G. Oettinger at Harvard in April 1954. The first book in the field was Machine Translation of Languages, edited by W. N. Locke and A. D. Booth, published[5] in 1955. This is a collection of fourteen essays by some of the pioneer workers.

In Russia, as elsewhere in the early years, reliance was placed on sophisticated word-for-word translation procedures, with grammatical rules applied chiefly as an aid in producing better constructions in the output language. However, in the reports[6] of the Fourth International

[3] University of Washington News Service 1956: 31

[4] See Mechanical Translation

[5] See Machine Translation of Languages

[6] Rozentsveig 1958: 97

Congress of Slavicists, September 1958, it became clear that the group working at the Steklov Mathematical Institute of the Academy of Sciences had developed their procedure so that a thorough formal syntactic analysis of the input text preceded any attempt at translation. Moreover, the group working at Leningrad University under N. D. Andreyev was attempting to evolve a procedure whereby methods of analysis and synthesis were completely independent, the sole link being a logical system of symbols, serving as an interlingua or pivot-language.

In May 1959 research by the Linguistics Research Center, University of Texas, was begun under the directorship of Dr W. P. Lehmann. This work[7] was directed towards the development of two inter-related computer systems: one for automatic translation and the other for supporting research in linguistics.

The Current Decade

During August 1960, a report[8] was published by M. I. T., presenting a simple, mechanized model for sentence production. From this work, its author V. H. Yngve, predicted rules of syntax which are thought to be applicable to all human languages.

Related work is going on at various universities around the world, notably California, Bonn, Milan and Kyoto, Japan. The "N.A.T.O. Advanced Study Institute on Automatic Translation" was held in Venice in July 1962. Reports indicate that university centres in France are concentrating more upon automatic documentation than translation by machine.

Since November 1962, Arts graduates at the Faculté des Lettres et Sciences Humaines at Nancy have been able to take an option for the Certificate in Progamming Technology. This comprises a short course, including practical work with a machine, followed by an examination.

Most of the material published in France deals with the developments of linguistic theory which are necessitated by the acceptance of machine translation as a practical reality. [9, 10, 11] A few authors refer to specific problems and offer partial solutions (e.g. Ronsse[12]), but I have yet to discover any reports concerning the application of these theoretical studies in the field; that is to say the use of existing computing machines for automatic translation trials, or the development of special computing machinery for the embodiment of the routines and sub-routines variously suggested.

A symposium on the current status of research at the Linguistics Research Center (University of Texas) reviewed techniques and progress towards the goal of fully-automatic translation, as at June 1963. The major problems brought out by the symposium were:

the vast quantity of preparatory work required before any trials could be attempted;

the complexity of the concepts used by the linguists and programmers; and

the difficulty of establishing a thoroughly satisfactory mathematical analysis of linguistic structures.

[7] See Quarterly Progress Reports

[8] Yngve 1960: 445

[9] Marthaler 1964: 12 [10] A.T.A.L.A. 1963: 78 [11] Coyaud 1963: 51

[12] Ronsse 1963: 10

Structure

A structure is a clearly-defined whole built from one or usually more parts. In written language the quanta are morphemes, and the smallest possible structure which can be built to stand without support from other similar structures is a sentence.

At first sight, it would seem reasonable to aim at translating a sentence at a time. A word-by-word translation leads to unresolved ambiguities and a loss of information arising from omissions and redundancies. Phrase-by-phrase translation is much better and for many purposes is good enough, but re-consideration of long or complex sentences which have been translated in this manner shows the need for re-casting some sentences in order to achieve the type of structure normally acceptable in the output language. A further practical point is that a machine can be easily programmed to recognize the beginning and ending of a sentence. Recognition of the boundaries of parts of a sentence is admittedly one function of a machine designed for automatic translation, but the designer of such a machine would not be satisfied with an outpouring of well-formed phrases having no sentence-type boundaries.

On further consideration then, it seems that a better plan would be to analyse the input language a phrase at a time, and perhaps in smaller units, while synthesizing the output language a sentence at a time.

Computer Meta-Languages

Before a comparison of sentence structures in more than one language can be made an automatic process which can be handled by a machine, a meta-language has to be devised to serve as a means of communication between the very complex human language systems and the very simple bivalent language used within the computer. The use of the term "meta-language" is not intended to refer to the sequences of elementary symbols having one of only two possible values which all digital computers use internally. The use of binary symbols to illustrate the stage-by-stage working of the computer circuits is really the province of the mathematician or the electronics engineer, and in any case, the language within the computer is much the same whether it is being used as a translation machine or for road-traffic control. Instead, the term meta-language is used to cover every intermediate level of symbolism employed to convey the intelligence of a sentence. As an illustration, let us look at the sentence "THE MAN IN THE BOAT SWORE LOUDLY" as the computer would.

A lexical analysis would give:- definite article, noun singular masculine, preposition, definite article, noun singular neuter, verb past tense, adverb, punctuation full stop. The machine will read, interpret and encode this analysis as a set of symbols in strict order.

The system outlined in the diagram (fig.1) incorporates a matching-device which I have labelled "comparator". The incoming phrase is held for examination in a temporary store. Successive words are fed one at a time to one of the comparator input terminals. Into the other comparator input terminal is fed the content of the complete lexical store, one word at a time, until a match is obtained. Immediately a match is obtained, a recognition signal is fed from the comparator to the lexical store to halt the flow of lexical items. The arrival of the recognition signal also sets off another circuit which provides a read-out of the information recorded in the lexical entry alongside the matching word. In the example provided for illustration purposes, the incoming phrase contains the word "BOAT". When the matching word "BOAT" is disgorged by the lexical store, the comparator halts the flow and demands the associated information. After a short pause, during which the information "N" is recorded elsewhere in the machine, the output of the temporary store is reset so that the next word of the stored phrase is fed to the comparator circuit and the search cycle recommences.

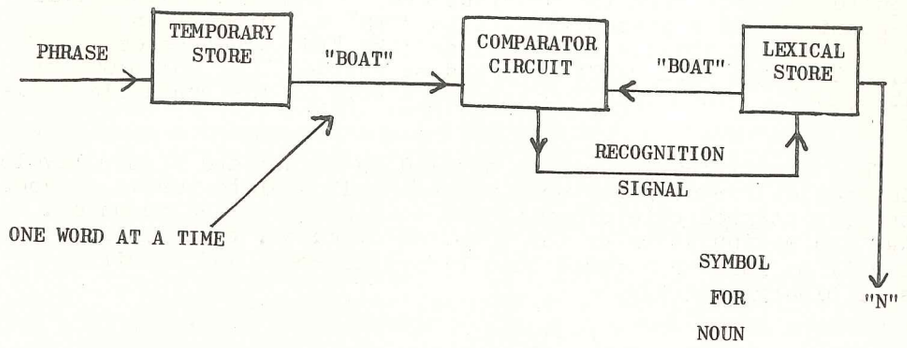


Figure 1 - Lexical Analysis Schematic

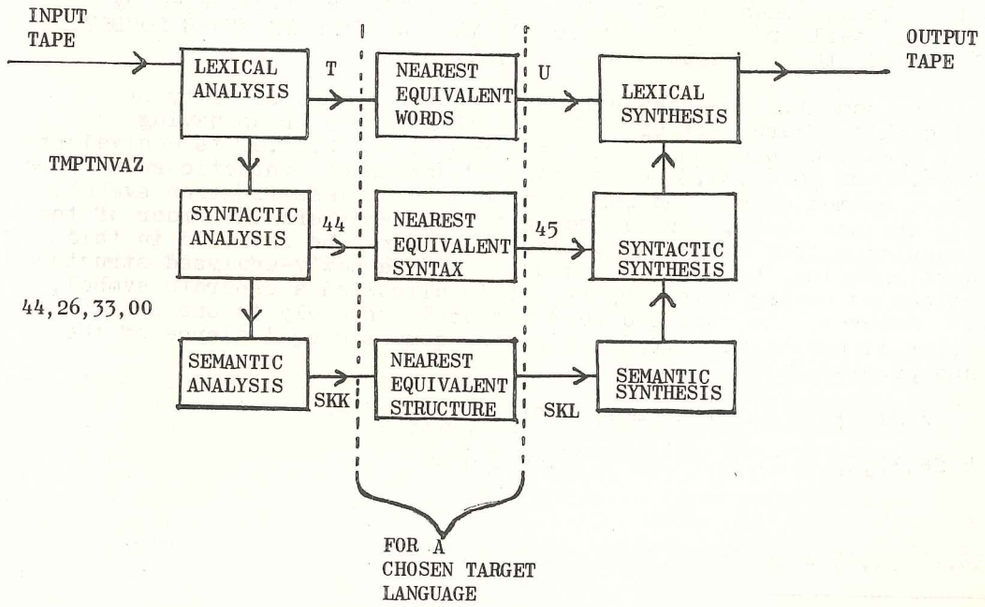


Figure 2 - Translation Schematic

Now, this code might give all definite articles the symbol T. Every singular masculine noun might be given the symbol M, and so on. Thus the sentence could be encoded "T+M+P+T+N+V+A+Z", or, more simply, "TMPTNVAZ". Then TMPTNVAZ becomes a word of the lexical meta-language.

We can take this example a stage further. At the syntactic level, one can say that "PTN" is acting adjectivally in that it qualifies "TM", and of course, "V" is modified by "A". This analysis will be read and enclosed as a further set of symbols in strict order. This code would be quite distinct from the previous one; perhaps all adjectival phrases beginning with a preposition, like "PTN" would be given the symbol "26", and noun article combinations like "TM" might become "44" and so on. Thus the word TMPTNVAZ could be encoded "44, 26+33+00" or simply "44263300". Then 44263300 becomes a word of the syntactic meta-language.

This stratification can be carried to higher and higher levels to provide codings for complete noun phrases and complete verbal phrases, through to subject/predicate ordering and complete sentence formulae. For storage and manipulation by the computer circuits, the complete sentences will be further encoded into binary numerals and electric analogues of binary numerals.

Analysis

The above discourse on computer meta-languages has underlined the stratification of this approach to input analysis. Stratification is on the basis of functional equivalence, so that lexical analysis involves the recognition and encoding of items having equivalent function at the lexical level, while syntactic analysis is seen as the recognition and encoding of items having equivalent function at the syntactic stratum.

Semantic analysis has been defined by Dr Tosh[13] as the process of identifying semantically equivalent subsets of syntactic rules and placing them in a common class. This process may be illustrated by using again my earlier example, where THE MAN IN THE BOAT SWORE LOUDLY became TMPTNVAZ, to be followed by 44,26,33,00.

The semantic analyzer circuit examines the possibility of identifying this ordered set of symbols with a larger group having equivalent semantic functions. In the example 44,26,33,00 is equivalent to 33,44,26,00 or to 26,44,33,00 so these "subsets of syntactic rules" are placed "in a common class" and may be allocated a classification symbol, say SKK. One can easily find examples in English where the order of the symbols resulting from syntactic analysis could not be shuffled in this way without changing the semantic value. When a newly-analysed structure is recognised as having this property it is allocated a separate symbol, say SJJ. However, the rule SJJ need not apply uniquely to one subset of rules, as it may be used to describe the semantic equivalence of the active and passive forms.

Expressing all this more neatly,

SKK = 44,26,33,00 = 33,44,26,00 = 26,44,33,00

but

SJJ = 55,27,34,00 ≠ 34,55,27,00 ≠ 27,55,34,00

[13] Tosh 1963: 66.

however

SJJ = 55,27,34,00 = 27,66,34,55,00.

Clearly, at this level, one can compare structures built from one language with structures built from another.

As the input analyzers deliver a stratified analysis to the sentence synthesizing circuits, (fig. 2), then the only additional information needed by the computer to allow translation to proceed, is a list of the target languages in which the output is to be expressed.

Synthesis

In order to synthesize a sentence at a time, there has to exist a table or list of permissible sentence structures, or alternatively a quasi-mathematical statement of the type:

$$S = (x\ y\ z) \neq (x\ z\ y) \neq (y\ x\ z)$$

which attempts to restrict the constituents of a particular sentence-type and their order. Each sentence produced by the output machinery then has to fit one of the mathematical models, or be one of those listed in the tables. It is clearly more economical of storage facilities to prepare a hierarchy of mathematical rules which covers the synthesis of many possible forms, than to tabulate every known variant. Testing and subsequent modification of the mathematical rules for synthesis should be based upon sampling of the output texts and constructive criticism by a native speaker of the output language.

The procedure for language synthesis will then be:

1. The input analysis will have dictated a particular sentence in the output language; but a computer meta-language, having symbols which describe all possible sentence structures in the chosen output language, will have allowed a selection which results in a formula for the new sentence differing minimally from the sentence structure dictated by the input analysis.

2. The embryonic sentence will now exist as a set of slots into which lexical equivalents of nouns, noun phrases, verbs, verb phrases, adjectives, adverbs etc. may be fitted. The information for this process is provided mainly by the usual automatic dictionary working on the input material; but, when ambiguity threatens, criteria of compatibility and probability are employed in much the same way as the human translator would employ them. Acceptable word order is already ensured by the output mechanism which is under the control of the sentence formula.

Ambiguity

The resolution of ambiguities arising in the input analysis procedure is sometimes seen as the problem of paramount difficulty. Firstly one has to be quite clear about the types of ambiguity which may be encountered. In poetry, for instance, an ambiguity may be quite intentional. In such contexts, the two possible meanings of a word or phrase should be carried right through the translation process and appear in the output. Metaphors in prose are probably best treated in the same way. The unintended doubt which arises from the use of equivocal words, wording or phrasing, can only be resolved if one knows what the author had in mind at the time of writing.

The human translator tackles the problem by finding unambiguous synonyms for alternative possibilities and then weighing the relative

probabilities of juxtaposition. The probability of two events co-occurring can be expressed as the product of their individual probabilities of occurrence. Statistical analysis of suitable texts allows calculation of the individual probabilities of occurrence of the chosen synonyms. The computer can be programmed to use conditional probability or probability in the general sense depending upon whether the operator considers the text to be a specialized treatise or more general reading. In this way, the output will contain the most likely translation in that context.

It is necessary to emphasise the difference between ambiguities which arise through the use of equivocal words and which may be resolved by assessing the probabilities of juxtaposition as outlined above, and those ambiguities which arise through the use of equivocal words and which may not be resolved by this technique in its simplest form. As an example, "WE ARE RETURNING REFUGEES" can have one of two alternative syntactical analyses, depending upon whether the verbal part is intended to be transitive or intransitive. This dilemma may be viewed as an unintended lexical ambiguity resulting from the use of "returning" to convey either "sending back" or "coming back", but the solution proposed above would at first sight seem unhelpful as the juxtaposing words are equally compatible in either sense. The problem is indeed insoluble unless one is given more information about the ambient context. In the light shed by surrounding sentences, the decision would not be painfully difficult.

This trend of thought precipitates a tentative hypothesis of stratified ambiguity, such that insoluble ambiguities at lexical level could be resolved at syntactic level, while ambiguities insoluble at syntactical level could be resolved at semantic level. Thus, to return to this example, let us suppose that the structure SKM is the semantic analysis of "WE ARE SENDING BACK REFUGEES." Then SKM can be said to have a statistical probability P_M. The surrounding structures, say SBB and SMJ, also have probabilities, P_B and P_J respectively. Thenthe probability of these structures co-occurring with SKM will be proportional to the products $P_M P_B$ and $P_M P_J$ respectively. If the analysis of "WE ARE COMING BACK REFUGEES" was SQQ and this had a probability P_Q, then the products would be $P_Q P_B$ and $P_Q P_J$. The computer would be programmed to accept the pairs with the highest net probability and in making this choice, would have selected by implication either SKM or SQQ. Again, the output text will contain the most likely translation.

The above-described synthesis procedure defines a sentence of the output language which is lexically equivalent, grammatically acceptable and semantically probable. Current researches will one day show whether this sort of translation is generally satisfactory.

Conclusion

Since a written language is an empirical code, it has natural redundancy and the rules of encoding and decoding may be phrased in a number of equivalent ways. The system of decoding, chosen at the outset by the man responsible for the design philosophy of the machine, will determine the efficiency of the input analysis techniques; and the system of encoding will control the efficacy of the output language synthesis.

The resolution of ambiguities by a strictly formal method based on probability theory must improve the overall accuracy of the translation process.

As Dr Lehmann has said:[14] "If we have learned anything from

[14] Lehmann 1963: 5

research in machine translation, it is that our knowledge of language must be vastly deepened. Even from a theoretical point of view, work in machine translation is of tremendous importance for linguistics and other social sciences. It provides us with our first opportunity of testing models of language, of verifying linguistic theories in ways comparable with those available to physical and biological scientists."

BIBLIOGRAPHY

A.T.A.L.A. 1963 "le mot en traduction automatique et en linguistique appliquée." La Traduction Automatique 4: 78

Coyaud, M. 1963 "Quelques problèmes de constructions d'un langage formalisé sémantique." La Traduction Automatique 4: 51

Lehmann, W. P. 1963 "Machine Translation: An Introduction." Symposium on Current Status of Research. Linguistic Research Center, University of Texas, Austin.

Machine Translation of Languages edited by W.N.Locke and A.D. Booth, published jointly by The Technology Press of M.I.T. (Cambridge), John Wiley and Sons (New York) and Chapman and Hall (London), 1955.

Marthaler, M. P. 1964 "Considérations sur l'analyse automatique du Langage, base de la Traduction mécanique et du Traitement automatique de messages écrits." La Traduction Automatique 5: 12

Mechanical Translation 1954 - The Massachusetts Institute of Technology, Cambridge.

Quarterly Progress Reports 1959 - Linguistics Research Center, University of Texas, Austin.

Ronsse, F. 1963 "Locutions françaises issues de Mots Russes uniques." La Traduction Automatique 4: 10

Rozentsveig, V. Y. 1958 "The Work on Machine Translation in the Soviet Union, Fourth International Congress of Slavicists Reports, Sept. 1958". Mechanical Translation 5: 97

Technical Books in Print, 1964 J. Whitaker and Sons, London.

Tosh, L. W. 1963 "Linguistic Research" Symposium on Current Status of Research. Linguistics Research Center, University of Texas, Austin.

University of Washington News Service 1956 "News". Mechanical Translation 3: 31

World List of Scientific Periodicals (1900-1960) 4th Edition 1965. Butterworth, London.

Yngve, V. H. 1960 "A Model and an Hypothesis for Language Structure" Proc. American Philosophical Society 104: 445.

ENQUÊTE SUR LE FRANÇAIS RÉGIONAL DU CANADA

J. C. Corne

(University of Auckland)

Enquêteurs: Mlles Jeanette Glasgow, Alison Glenie, Gayleen Verrall; MM. J. C. Corne, I. M. Wilson. Détermination des phonèmes par J. Glasgow et J. C. Corne.

Date: 1965.

Témoins: Mlle A., née au Manitoba, enfance et adolescence à St Eustace, domicilié depuis à Winnipeg; langue maternelle: le français; parle anglais couramment; âgée d'environ vingt ans; parents et grandparents nés au Canada.

M. E., étudiant, âgé d'une vingtaine d'années; né au Canada, de parents canadiens; langue maternelle: l'anglais; parle français et allemand; éducation à Montréal et à Toronto; en Nlle-Zélande depuis cinq mois et demi.

Mme I., née au Canada, de parents canadiens; langues maternelles: l'anglais et le français; éducation à Winnipeg; en Nlle-Zélande depuis sept mois et demi.

Mme O., née au Canada, de parents canadiens; langue maternelle: le français; ne parle anglais que depuis les quatre ans et demi qu'elle est en Nlle-Zélande; a passé quelques mois en France; éducation au Canada, près de Québec.

M. U., né au Canada, père acadien, mère canadienne; langue maternelle: le français; parle anglais (appris aux USA) depuis cinq ans et demi; en Nlle-Zélande depuis quatre ans et demi; a passé quelques mois en France; éducation au Canada, près de Québec.

SIGLES ET SIGNES

C	phonème consonantique
Ø	consonne zéro
IM	initiale du mot
IS	initiale de la syllabe
FM	finale de mot
FS	finale de la syllabe
SFF	syllabe finale fermée
SFO	syllabe finale ouverte
SNF	syllabe non-finale
| |	orthographe phonologique
[ ]	orthographe phonétique
⌈ ⌉ ~ [ ]	variantes facultatives citées dans l'ordre diminuant de leur fréquence
+	suivi de

PHONÈMES VOCALIQUES

Corrélation Orale

I y u
E œ O
A

|I|: A la FM, |i|:|j|, p.ex. |pEi|:|pEj| pays:paye.

-à l'IS, devant V, [j], p.ex. |tRAvAIõ|=[tRavajõ] travaillons; mais à l'intervocalique |II|=[ij], p.ex. |tRavaIIõ|=[tRavaijõ] travaillions.

-dans le contexte C+|I|+V, |I|=[i^j], p.ex. [ʃaRpatsi^{j}e] charpentier, [pi^{j}e] pied; mais lorsque C = |l| ou |R|, |I|=[j], p.ex.[sulje] soulier.

-dans le contexte CC+|I|+V, |I|=[i^j], p.ex. [pRi^{j}e] prier.

En SFF, |I|=[ɪ], p.ex. [sɪs] six,[pɪk] pique, [ɥɪl] huile, mais quand -C = |R|, |I|=[i], p.ex.[RiR] rire.

En SNFO, |I| = [i]~[ɪ], p.ex. [sivɪsm]~[sɪvɪsm] civisme.

En SNFF, |I|=[ɪ], p.ex. [dɪskaRte] discarter.

|E|: En SFO, |e|:|ɛ|, p.ex.,|Ete|:|Etɛ| été : était.

En SFF, |ɛ|:|ɛɛ|, p.ex., |mɛtR|:|mɛɛtR| mettre : mètre, maître;
|ɛɛ| se réalise [$ɛ^I$]~[a^I];
|ɛR| à la finale se réalise [$ɛ^I$R]~[ɛR].

En SNF, le timbre de la réalisation de l'archiphonème n'est pas pertinent, cf. |dEbyte| = [dɛbyte]~[debyte] débuter; [gɛte]~[gɛte]~[gete] gaîté.
La réalisation [køl] pour |kEl|, signalée dans SPFC 1903, n'a pas été notée.

A : Pour tous les témoins sauf A., on trouve partout l'opposition |a|:|ɑ|, p.ex.|pat|:|pɑt| patte : pâte, |taʃe|:|tɑʃe| tacher : tâcher,
|a| = [a]~[ʌ]
|ɑ| = [ɑ]~[å]
Pour le témoin A., on a le système suivant:

En SFO, |ɑ|, p.ex.,|bɑ| = [bʌ] ~ [bɑ] bas;

En SFF, |a|:|ɑ|, réalisé [ʌ]:[ɑ], p.ex. dans |pat|:|pɑt|, mais dans |taʃ|, |a| se réalise [æ], probablement à cause de l'influence de la réalisation de |a| dans|taʃe| = [tæʃe];

En SNF, |a|:|ɑ|, réalisé [æ]:[ʌ], p.ex.[tæʃe] :[tʌʃe].
Pour tous les témoins, dans le contexte C+|u|+|A|, on ne trouve que|a|, réalisé[a]~[ʌ], p.ex. |luA| = [lwa]~[lwʌ] loi; mais quand cette syllabe est fermée par une consonne: CuAc, la réalisation de |A| change pour les témoins E et I, devenant [ɛ:]~[$ɛ^{I}$]~[a], dans p.ex. |buAt| boîte, |muAn| moine.

|u|: A l'IS devant V, et dans les contextes CuAC,CuA, on a toujours [w], p.ex. |ui| = [wi] oui, |buAt| = [bwat] boîte, |luA| = [lwa] loi.
Ailleurs, on ne trouve que [u], p.ex. |lue| = [lue] louer.

|y|: A l'IS devant V, et dans les contextes Cyi, CCyi, CyiC, on ne trouve que [ɥ], p.ex., [kɥijɛʀ] cuiller, [fʀɥi] fruit, [kõdɥit] conduite.
Ailleurs, il n'y a que [y], p.ex. |bye|=[bye] buée, |lyœʀ|=[lyœʀ] lueur.

|o|: En SFO, toujours |o|, p.ex. [bo] beau, [so] sot, [kaʃo] cachot.
En SFF, on a l'opposition |o|:|ɔ|, p.ex. |sot|:|sɔt| saute : sotte, |bot|:|bɔt| boat:botte.
En SNF, quatre témoins sur cinq utilisent l'opposition |o|:|ɔ|, mais rarement et dans des paires suspectes comme |bote|:|bɔte| beauté: botté, où, à cause du conte du Chat botté, la prononciation [bɔte] peut très bien ne pas représenter la pratique courante. La réalisation du |O| en SNF serait donc plutôt [ɔ]~[o]~[ŏ], p. ex. [ŏtostɔp]~[otostɔp]~[ŏtostɔp] autostop, [dɔne] donner.
|O| suivi de |-z| ou |-ʒ| se réalise [o]; suivi de |ʀ| de la même syllabe, [ɔ] sauf chez le témoin A qui dit en SFF [$ɔ^{o}$], et en SNF [ŏ], p.ex. [pŏʀte] porter, [$mɔ^{o}ʀ$] mort.

|œ|: En SFF:[œ]~[ø]~[ɶ]~[ɛ] , p.ex., [ʒœn]~[ʒøn]~[ʒɛn] jeune, jeûne.
En SNF: [ø], p.ex. [dɛʒøne] déjeuner.
Devant |ʀ| de la même syllabe: [œ]~[ɶ]~[ø] , p.ex. [otœʀ] ~[otɶʀ] ~ [otøʀ] auteur.
Ailleurs: [ø], p.ex. [øʀø] heureux.

N'a pas été observée la désonorisation des voyelles |i|,|y|,|u|, signalée par Gendron 1959.

Corrélation nasale

ẽ œ̃ õ
ã

|ẽ| = [ẽ]; |œ̃| = [œ̃]; |õ| = [ẽ].

|ã| se réalise principalement [ã], parfois [ɑ̃], p.ex. [ʃɑ̃te] chanter, |ãfãs|=[ãfãs] enfance, |ãfõs|=[ãfõs] enfonce.

Il y a tendance à toujours dénasaliser les voyelles lorsqu'à la liaison on prononce la consonne nasale: [bɔnarijẽ] bon à rien, [mwajɛnaʒ] moyen âge.

PHONEMES CONSONANTIQUES

L'inventaire consonantique se compose de la corrélation proportionnelle:

p	f	t	s	ʃ	k
b	v	d	z	ʒ	g
m		n		ɲ	ŋ

et des phonèmes hors système |l|, |ʀ|, |ø| et |h|.

|t|: devant les voyelles |y|, |i|, |e|, |t|=[ts], p.ex. [ymɪdzɪtse] humidité, [tsjɛd] tiède.

|d|: devant les voyelles |y|, |i|, chez les Montréalais (surtout le témoin I), |d| = [dz], p.ex. [dzy] du, [kõdzɥɪt] conduite.

|ɲ|: à l'intervocalique, se distingue bien de |nj|, p.ex. [laɲɛl] l'agnelle, [lanjɛl] la nielle.
à la finale, se confond avec |n| et |ŋ|: |siɲ|=[sɪŋ]~[sɪn] signe.

|ŋ|: ne se trouve guère qu'à la finale, surtout dans des mots empruntés à l'anglais.

|ʀ|: se réalise généralement [ʀ], mais parfois [ɹ], p.ex. [kilpaɹt] qu'il parte, [kɹobɑɹ] crowbar.

|ø|: se réalise en soixante-dix pour cent des contextes et des cas étudiés comme en français parisien. En neuf pour cent des cas, |ø| n'existe pas du tout, alors qu'il existe en français parisien, p.ex.[lezɔʀd [lezɔʀdœvʀ]. En huit pour cent des cas, |ø| du francais est remplacé par |h|, et en de faibles pourcentages |ø| se réalise par un hiatus, ou [ʔ] ou [zh] ou [zʔ], p.ex. |lEØERitie| peut se réaliser [lɛɛʀitsje], [lɛhɛʀitsje], [lɛzhɛʀitsje], [lɛzʔɛʀitsje], [lɛzɛʀitsje]. Il est à noter qu'il n'y a aucune homogénéité: un même témoin peut réaliser |ø| par un ou plusieurs des procédés indiqués. Celui qui dira [lɛzʔɛʀitsje], dira [lɛzhɪstwaʀ] et [ãʔo], [lɛkao] et [lɛzom]. La palatalisation de |k|, |g| n'a pas été observée, ni les réalisations [x] et [h] de |ʒ| signalées par Charbonneau 1957: 14.

INDICATIONS BIBLIOGRAPHIQUES

Charbonneau, R., 1957. "La spirantisation du |ʒ|", in JCLA 3,1:14, et 3,2: 71.

Dulong, G., 1953. Atlas linguistique du Canada français: questionnaire. Québec, Université Laval.

Gendron, J.D., 1959. "Désonorisation des voyelles en franco-canadien", in JCLA 5,2: 99.

Ham, E.B., 1959. "Linguistic Lode in Rural Quebec", in Romance Philology 11.

Martinet, A., 1945. La Prononciation du français contemporain. Paris, Droz.

_____ 1954. La Description phonologique. Genève, Droz.

_____ 1961. Eléments de linguistiqu générale. Paris, Colin.

Société du Parler Français au Canada, 1903. L'origine et le parler des Canadiens-Français. Paris.

_____ 1930. Glossaire du parler français au Canada. Québec.

Tougas, G. "Quelques aspects peu étudiés du français canadien", in JCLA 2,2: 60.

The Linguistic Society of New Zealand

The Linguistic Society of New Zealand was founded in Auckland in 1958, in order to "promote and pursue the scientific description and study of the evolution and structure of languages". Meetings are held in Auckland and Wellington, and papers are presented by New Zealand members and overseas visitors. The interests of members have no regional restriction, but particular attention is given to indigenous and European languages in the Pacific. In addition to papers by members and visitors, research papers by graduate students are usually presented at a meeting held late in the academic year.

The Society's proceedings and other contributed articles are published in the journal *Te Reo* (see inside back cover).

TE REO MONOGRAPHS

K.J. Hollyman: *A Checklist of Oceanic Languages (Melanesia, Micronesia, New Guinea, Polynesia)*. 1960. 32pp., double columns 4to. Mimeographed. o.p.

A. Capell: *The Polynesian Language of Mae (Emwae), New Hebrides*. 1962. 55pp., Crown 4to. Xerographed. o.p.

A.G. Haudricourt: *La Langue des Nenemas et des Nigoumak (Dialectes de Poum et de Koumac, Nlle-Caledonie)*. 1963. 85pp., 5 maps, Crown 4to. Xerographed. 13s. (Members: 8s.).

D.S. Walsh and Bruce Biggs: *Proto-Polynesian Word List. I*. 1966, (in press). Crown 4to. Xerographed.

TE REO REPRINTS

K.J. Hollyman: *Le Francais regional de l'Indo-Pacifique. Essais de phonologie*. 1964. 32pp. Crown 4to. Xerographed 6s. (Members: 4s.). (Reprinted from *Te Reo* 6).

David D. Thomas (ed.): *Papers on Four Vietnamese Languages*. 1966. Crown 4to. Xerographed. 6s. (members: 4s.) (Reprinted from *Te Reo* 7 and *Te Reo* 8).

All enquiries should be addressed to:

The Secretary,
Linguistic Society of New Zealand,
c/o University of Auckland,
P.O. Box 2175,
Auckland, New Zealand