

DOCUMENT RESUME

ED 052 668

FL 002 455

AUTHOR Chiu, Rosaline K.  
TITLE Measuring Register Characteristics: A Prerequisite  
for Preparing Advanced Level TESOL Programs.  
PUB DATE 6 Mar 71  
NOTE 23p.; Paper presented at the Fifth Annual TESOL  
Convention, New Orleans, La., March 6, 1971  
EDRS PRICE MF-\$0.65 HC-\$3.29  
DESCRIPTORS Adult Students, \*Advanced Programs, Advanced  
Students, Basic Skills, \*Curriculum Development,  
\*English (Second Language), English Programs,  
\*Instructional Materials, \*Language Instruction,  
Language Learning Levels, Material Development

ABSTRACT

This paper aims at relating the linguistic concept of "register" to the production of TESOL programs, particularly concerning the practical problems of materials selection for advanced level programs. It is the assumption of this paper that this concept is particularly relevant to advanced level programs, when learners are to be able not only to speak or write accurately according to the grammar of English but to speak or write appropriately according to the accepted social and cultural conventions of the language community. The paper reports on the research project conducted by the English Curriculum Section of the Language Bureau, Public Service Commission of Canada, which aims at measuring the register characteristics of a corpus of 1,000 pieces of synchronically sampled administrative correspondence and a sizable corpus of boardroom discussions. It describes how this investigation supplies useful and practical information for the production of advanced level programs. (Author/RL)

ED052668

Measuring Register Characteristics:  
A Prerequisite for Preparing Advanced  
Level TESOL Programs

BY

Rosaline K. Chiu  
Language Bureau  
Public Service Commission  
of Canada

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE  
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE  
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS  
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION  
POSITION OR POLICY.

Paper prepared for  
1971 TESOL Convention  
at New Orleans,  
Louisiana, U.S.A.  
March, 1971.

FL002455



## ABSTRACT

This paper aims at relating the linguistic concept of register to the production of TESOL programs. While linguists concern themselves with how to identify and describe registers that exist within the framework of English, TESOL specialists should work out how to select, grade and present language varieties appropriate to the learners' needs. With the assumption that the linguistic concept of register is particularly relevant to Advanced Level TESOL program, the paper reports on a research project which aims at measuring the register characteristics of a corpus of 1,000 pieces of synchronically sampled administrative correspondence and a sizable corpus of boardroom discussions.

In the field of linguistics, there has been much discussion of intra-lingual varieties in the last ten years on both sides of the Atlantic<sup>1</sup>.

---

<sup>1</sup>Catford, J.C.: "A Linguistic Theory of Translation" O.U.P. 1965, p. 83.

Gumperz, J.J.: "Types of Linguistic Communities" in "Readings in the Sociology of Language". J.A. Fishman ed. Mouton, 1968. p.461.

Stevens, P.D.: "Varieties of English" in "Papers in Language and Language Teaching" P.D. Stevens, ed., O.U.P. 1965, p. 74.

Bernstein, Basil: "Language and Social Class" British Journal of Sociology, Vol. 11, No. 3, 1960, pp. 271-276.

Bernstein, Basil: "Social Class and Linguistic Development" in A. Halsey, J. Floud and C.A. Anderson, eds. "Education Economy and Society" New York, Free Press, 1961, pp. 288-314.

Bernstein, Basil: "Social Structure, Language and Learning" Education Research, 3, 1961, pp. 163-176.

Bernstein, Basil: "Social Class, Linguistic Codes and Grammatical Elements" Language and Speech 5, 1962, pp. 221-240.

Bernstein, Basil: "Elaborated and Restricted Codes: The Social Origins and some consequences" Paper presented to the 1963 Meeting of the American Anthropological Association, San Francisco.

Shuy, Roger W. (ed.): "Social Dialects and Language Learning" Proceedings of the Bloomington, Indiana Conference, National Council of Teachers of English, 1964.

Shuy, Roger W.: "Discovering American Dialects" National Council of Teachers of English, 1967.

Shuy, Roger: "Detroit Speech: Careless or Systematic" Elementary English, May 1968.

Stewart, William A.: "Non-standard Speech and the Teaching of English" Washington D.C., Center for Applied Linguistics, 1964.

---

There have been different frameworks of categorization of intra-lingual varieties presented by different linguists. Most of

them, however, would agree that in one dimension, the variety of language a person uses is determined by who he is. Each speaker has learnt, as his  $L_1$ , a particular variety of the language of his community, and this variety may differ at any or all levels from other varieties of the same language learnt by other speakers as their  $L_1$ . So, he learns to say 'jello' or 'jelly', 'dessert' or 'sweet', 'baby-carriage' or 'pram', 'to do the dishes' or 'to wash up', 'to write an exam' or 'to sit for an exam', 'to have a conflict in his schedule' or 'to have a clash in his time-table', to spell 'colo(u)r' and 'glamo(u)r' without or with a 'u', to say /ei/ /a:/ /lu:tenənt/ /leftənənt/ 'tomato' or 'tomato', 'lieutenant' or 'lieutenant', depending on which side of the Atlantic he is from. Such a variety, identified along this dimension, is called a 'dialect'. For our purpose, 'who he is' means 'where he comes from'. It is usually the region of origin which determines which 'dialect' or 'dialectal variety' of the language a speaker uses. He speaks New York dialect if he comes from New York, Yorkshire dialect if he comes from Yorkshire.

But, for some language communities, England and the United States included, a person's speech is determined not only by the region he comes from but also by the 'social class' he comes from, or the social class he is trying to move into. Apart from the 'geographical' or 'regional' dimension and 'social class' dimension, dialects could also be delimited on the 'temporal' dimension and 'individual' dimension. (See Figure 1.)

Our major concern today is the concept of REGISTER, varieties of a language according to USE. Language varies as its function

varies; it differs in different situations. The label given to a variety of a language according to use is 'register'. It is no good trying to write exactly as we would speak. Very likely, we cannot give a lecture with the kind of language we use with our boyfriend or girlfriend. Gossips over the backyard fence sound very different from VIP's in a boardroom discussion, or so they would have us believe. A page from the Readers' Digest looks very different from a legal document. Each situation, each use, calls for its own appropriate variety. Here, I would like to quote Stevens,<sup>2</sup> illustrations on how the same message could be conveyed

---

<sup>2</sup>Stevens, P.D.: Op. cit. p.85.

---

in five different ways on a five-point scale of formality proposed by Joos<sup>3</sup>

---

<sup>3</sup>Joos, Martin: "The Five Clocks" Bloomington, 1962.

- |               |  |
|---------------|--|
| Frozen style: | Visitors should make their way at once to the upper floor by way of the staircase. |
| Formal:       | Visitors should go upstairs at once.   |
| Consultative: | Would you mind going upstairs, right away, please.                                 |
| Casual:       | Time you all went upstairs, now.   |
| Intimate:     | Up you go, chaps!  |

If you say 'Up you go, chaps!' in a formal situation, your choice of the variety is a wrong one; not that your grammar is inaccurate, but your use of the language is inappropriate. You will be accused of being 'ill-mannered'. If you use the formal style among your friends in a casual social gathering, then undoubtedly you would give the impression of being 'stuffy' or 'behaving strangely', or 'being funny'.

Before we relate the concept of register to TESOL pedagogy, let's look at the framework of language uses to enable us to understand better the functioning of a language in particular situations: who says or writes what, to or for whom, when, in what circumstances.

In this paper, the model adopted is an eclectic one with free adaptations from the models presented by Joos<sup>4</sup>, Halliday, McIntosh and Strevens<sup>5</sup>, Spencer and Gregory<sup>6</sup>, Crystal and Davy<sup>7</sup>, Catford<sup>8</sup>, Ure<sup>9</sup>, and Hill<sup>10</sup>. Registers may vary on three dimensions: field of discourse, mode of discourse and manner of discourse. (See Figure 1.)

---

<sup>4</sup>Joos, Martin: "The Five Clocks" Bloomington, 1962.

Joos, Martin: "The Isolation of Styles" Monograph Series on Language and Linguistics Georgetown University, 12, 1959, pp. 107-113.

<sup>5</sup>Halliday, M.A.K., McIntosh, A, Strevens, P: "The Linguistic Sciences and Language Teaching" London, Longmans, 1964.

<sup>6</sup>Spencer, J. and Gregory, M.: "An approach to the Study of Style" in "Linguistics and Style" London, Longmans, 1964.

Gregory, Michael: "Aspects of varieties differentiation." Journal of Linguistics, Vol. 3, No. 2, Oct., 1967, pp. 177-274.

<sup>7</sup>Crystal, David and Davy, Derek: "Investigating English Style" London, Longmans, 1969.

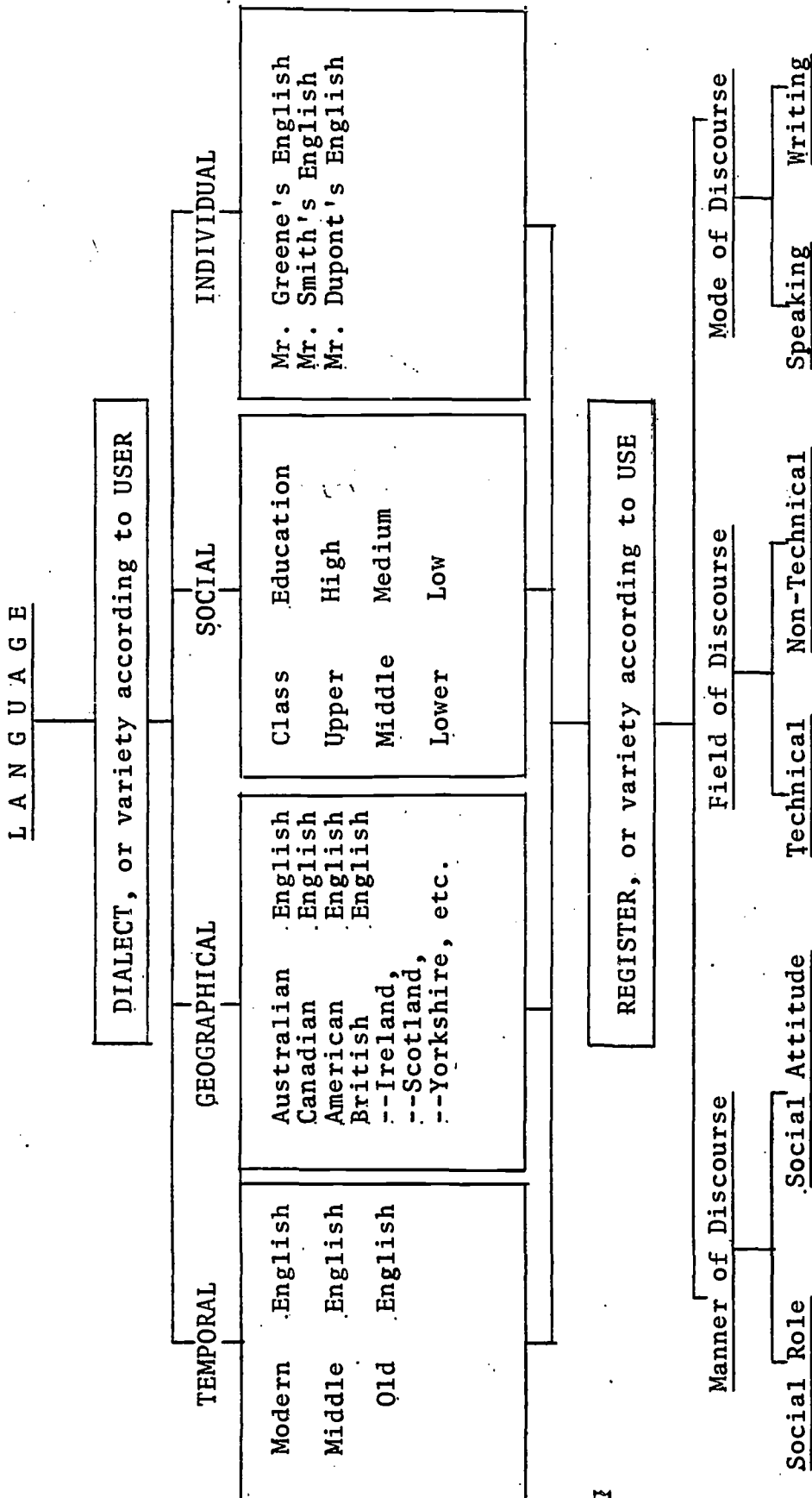


Figure 1



Crystal, David: "New Perspectives for Language Study. 1: Stylistics." English Language Teaching Vol. 24, No. 2, 1970, pp. 99-106.

<sup>8</sup>Catford, J.C.: "A Linguistic Theory of Translation" O.U.P. 1965,

<sup>9</sup>Ure, Jean: "Practical Registers (A)" English Language Teaching, Vol. 23, No. 2, 1969, pp. 107-114, 1969.

Ure, Jean: "Practical Registers (B)" English Language Teaching, Vol. 23, No. 3, 1969, pp. 206-215.

<sup>10</sup>Hill, C.P.: Presented in his series of lectures given to Postgraduate Diploma students in the Teaching of English as a Foreign Language at University of London Institute of Education, 1965-1966.

'Field of discourse' refers to the 'why' of a speaker's use of language -- the purpose the speaker employs language for -- persuasion, discussion, insult (See Figure 2). 'Mode of discourse' is simply the medium, and is primarily a distinction between spoken and written language (See Figure 3). 'Manner of discourse', finally, refers to the relations among the participants in a language activity, which can be further sub-categorized into social rôle and social attitude. It is best treated as a 'cline', with categories such as casual, intimate or polite (See Figure 4).

It is as the product of these three dimensions of categorisation that we can best define and identify register. For example, a talk on linguistics in a university will be in the technical field of discourse, in the mode of speaking, or speaking what is written, and in a polite manner. The same professor, in the common room, a little later, may switch to the field of hockey, conversing mode and casual manner.

Register study thus aims at defining and identifying the linguistic features which are regularly used in recurrent situations,

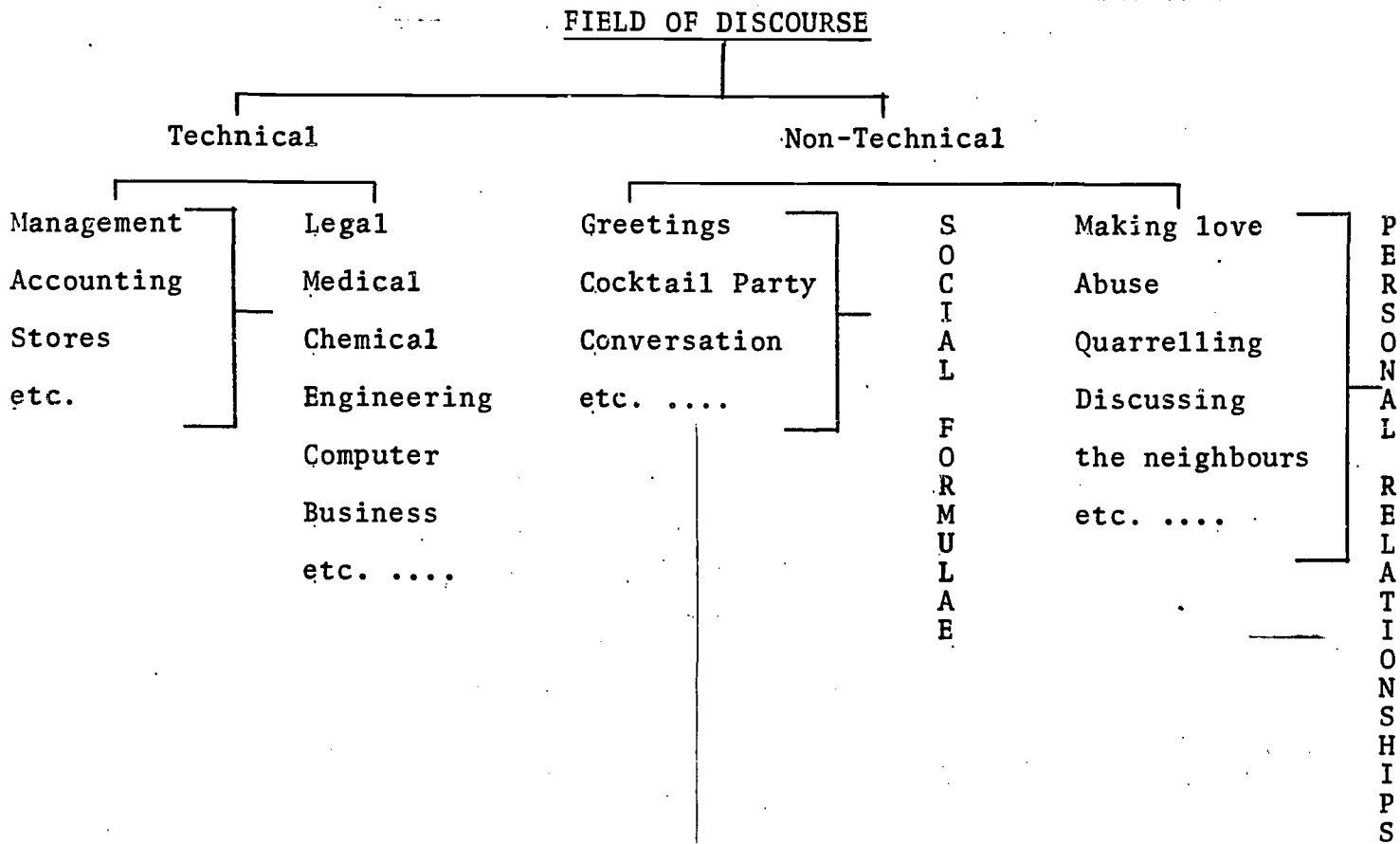


Figure 2

MODE OF DISCOURSE

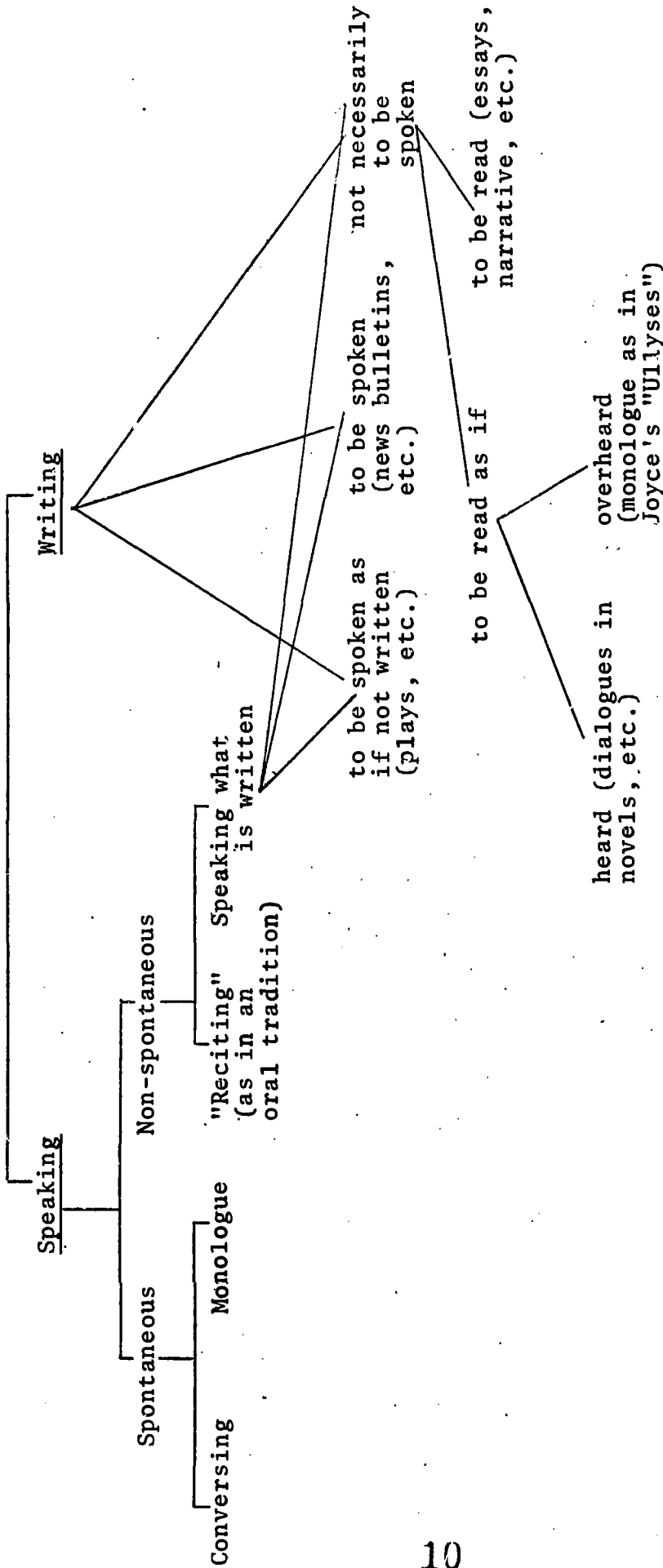
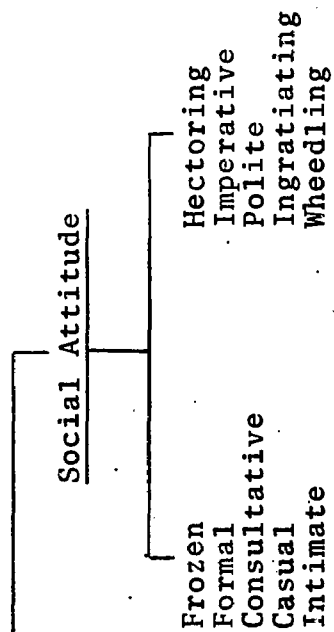


Figure 3

(Adopted from Michael Gregory "Aspects of varieties differentiation" Journal of Linguistics, Vol. 3, No. 2, Oct., 1967, 177-274.)

MANNER OF DISCOURSE



FATHER								
MOTHER								
SON								
DAUGHTER								
TUTOR								
PUPIL								
etc.								

Figure 4

and categorizing the kinds of English that are demarcated in this way. Register characteristics are sets of linguistic forms found to have a regular connection with a particular use of the language in a particular situation, or sets of situations. It does not mean that the connection is necessarily obligatory, but there is enough of a connection to give a native speaker an intuitive feeling of the existence of a link between the language and the situation.

It is the purpose of this paper to propose that a measurement of register characteristics should be considered as a prerequisite for preparing advanced level TESOL programs, that TESOL specialists should work out how to select, grade and present registers within the framework of English according to the needs of the learners. For example, the learner might need to interact with professional colleagues in technical discussion, in committee work, and in formal social situations. We might not be able to cater to the sociolinguistic requirements of any one individual, just as we cannot cater to the psycholinguistic variation in the individual as a language learner; but if a learner typology could be more explicitly formulated, this is likely to bring more effective teaching.

In the English Curriculum Division of the Language Bureau, Public Service Commission of Canada, two of the research projects underway are related to register study. Both projects aim at relating the linguistic concept of situationally differentiated language variety to the preparation of TESOL programs. One project has the specific aim of determining the register characteristics of government administrative writing, for this is a variety of

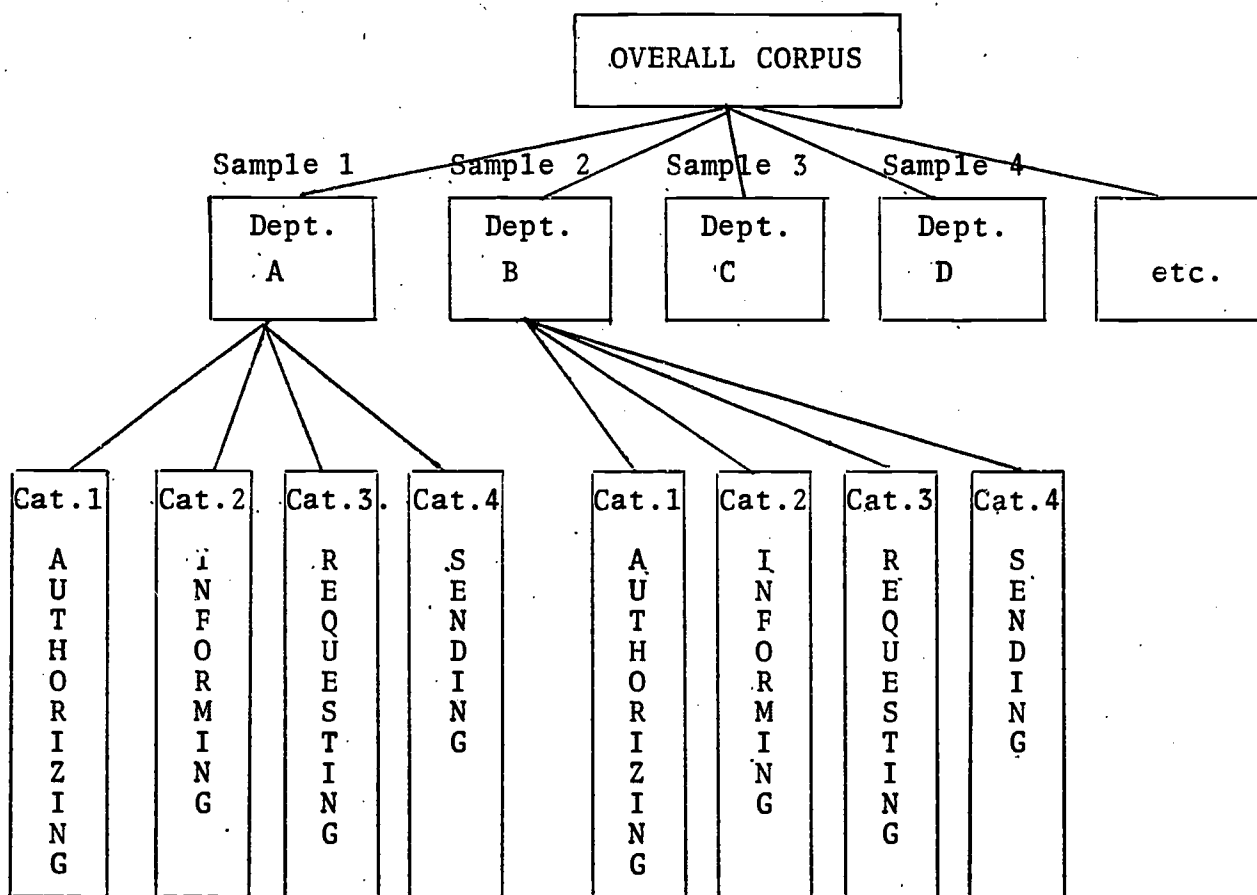
writing which our advanced students, all involved in the public service, are most likely to come across and to have to use. The other is a study of spoken English. One of its aims is to study the register characteristics of topic oriented discussions among high-ranking public servants. We would like to know more about this particular use of English, again believing that this is a variety of English which our students are likely to be confronted with and to use.

In our linguistic study of administrative writing, we have analysed over 1,000 pieces of correspondence supplied to us by various government departments. All documents in the more or less random sample are dated no further back than 1968. Each departmental sample is subdivided into four basic semantic categories as shown in Figure 5. I will not present a detailed description of our reasons for doing so here, but a study of the diagram should give some idea of how the breakdown of samples is organized.

In the first phase of our analysis, we have centered our study on the VERB PHRASE; the term is used to cover segments of language within which statements may be made on the possibilities of co-occurrence of particular verb forms and their relation. So, for phase 1, linguistic analysis is done at 2 levels: lexical and syntactic, both within the verb phrase. The objective of the analysis at this phase is to yield: (1) a list of verbs used in administrative writing, or, to put it more precisely, a frequency list of the lexical verbs; and (2) a list of 'specialized verbs' of the administrative register. This list will be worked out by comparing

Schematically the organization of the corpus is as follows:

SAMPLES:



Semantic Categories according to Purpose:

our frequency list with general word lists previously published by other researchers, West's A General Service List of English Words<sup>11</sup>, Thorndike and Lorge's The Teacher's Word Book of 30,000 Words<sup>12</sup>, and Kučera and Francis' Computational Analysis of Present-day American English<sup>13</sup>, and (3) information on frequency and distribution of different verb structures of the register.

---

<sup>11</sup>West, Michael: "A General Service List of English Words" Longmans, 1936.

<sup>12</sup>Thorndike, E.L. & I. Lorge: "The Teacher's Word Book of 30,000 Words" New York, Columbia University. 1944

<sup>13</sup>Kučera, H. and W.N. Francis,: "Computational Analysis of Present-day American English" Brown University Press, 1967.

---

Verb phrases only are marked off in our raw sample, analysed, structurally coded, and quantified, both by the structural code and lexically. Here are some typical verb phrases:

is amended  
 would appreciate  
 has been requested  
 would appreciate having  
find enclosed

All the underlined verb forms are full verbs, or lexical verbs. Like all other lexical verbs, they are separately recorded and filed. And from this the frequency and range of the lexical verbs of the corpus can be worked out. Frequency is the number of times a word occurs -- say ten, for a verb like consider. It occurs a total of ten times in only seven documents, so its range is seven, but its frequency remains ten.



So, at the lexical level, some of our findings are very interesting. Our corpus consists of slightly less than 140,000 running words. The total number of verbs is 11,323, but the number of different verbs is only 866. Of these 866, the first 15 highest ranking lexical verbs, according to frequency, and excluding BE and HAVE, are the following:

<u>Rank</u>	<u>Lexical Verb</u>	<u>Frequency</u>
1	make	252
2	attach	187
3	enclose	164
	receive	
5	require	159
6	appreciate	148
7	provide	139
8	refer	129
9	forward	123
10	find	120
11	request	118
12	advise	113
	send	
14	take	106
15	give	104

Even at a glance, our list looks very different from other general word lists. If we screen the lexical verbs from Lorge's magazine count<sup>14</sup>, which included nearly 4½ million words from

---

<sup>14</sup>Thorndike, E.L. & I Lorge,: "The Teacher's Word Book of 30,000 Words" New York, Columbia University. 1944.

---

the Saturday Evening Post, Women's Home Companion, Ladies' Home Journal, True Story, and Readers' Digest, the highest frequency verbs will look something like this in descending order of frequency

1. go	6. know	11. think
2. ask	7. get	12. look
3. say	8. see	13. want
4. come	9. take	14. give
5. make	10. like	15. find

When we compare our list with Kucera and Francis' list<sup>15</sup>, which is a compilation of more than 1,000,000 running words,

---

<sup>15</sup>Kučera, H. and W.N. Francis, : "Computational Analysis of Present-day American English" Brown University Press, 1967.

---

spread over fifteen genres of written English, we find that most of our high frequency items come very low in the Kucera and Francis' list:

	<u>Administrative Correspondence</u> (from a corpus of 140,000 run- ning words)	<u>Kucera &amp; Francis</u> (from a corpus of about 1,000,000 running words)
attach	187	44
enclose	164	13
appreciate	148	39
refer	129	108
forward	123	119
request	118	83
advise	113	48
thank	83	82
submit	77	47
acknowledge	40	27
ensure	35	10
amend	27	17

By a process of comparison, we will be able to work out an inventory of 'specialized verbs' of this register of administrative writing.

Let's turn now to the structural level of our analysis. You remember our list of typical verb phrases: 'is amended' is a simple phrase, and so is 'would appreciate' and 'has been requested'. Simple phrases are those that contain only one (form of) full verb, though they may contain a number of (forms of) auxiliaries.

'Would appreciate having' and 'find enclosed' are complex phrases, since they contain more than one (form of a) full verb.

To code our verb phrases for analysis we used a binary distinction grid in which all verb phrases can be quickly coded and catalogued. The grid refers to form and not meaning. Figure 6 is a schematic representation of the coding system.

Say we have a verb phrase like takes, as in 'the trip takes three days'. Takes is unmarked, -- as you can see from the grid: no tense, no mood, no phase, no aspect, no voice.

But take another example: 'would appreciate', as in 'We would appreciate the report by Sunday'. Would marks the phrase in mood, and tense (as opposed to 'will appreciate'), and the form can be uniquely identified as TM in our code.

You can take any verb phrase and identify it on the grid -- and our coding system thus makes it possible to code, catalog and retrieve all of the variety of forms very quickly. It also allows us to computerize the data for more rapid processing, which is now being done in this project.

Here is the distribution of verb phrases.

In our samples 1-10, we have a total of 9,884 verb phrases, of which 16.1% are complex phrases, such as 'became acquainted', 'let me know', 'keep us posted'. Our model does not deal with these complex types, but studies the remaining 83.9% of the total -- the marked and unmarked simple phrases. These total 8,289 verb phrases, phrases such as 'will issue', 'is being prepared', 'are provided', 'proceed'.

Category	Tense	Mood	Phase	Aspect	Voice
Unmarked	non-past	non-modal	non-perfect	non-progressive	non-passive
Marked	past	modal	perfect	progressive	passive
Markers	-D	CAN MAY WILL SHALL etc.	HAVE + -n	BE + -ing	BE + -n
Code for Marked Category	T	M	P	A	V

Figure 6

This gross breakdown shows that of the 8,289 simple phrases, (see Figure 7) 32.1% are unmarked. And 10.0% are marked for tense only -- traditionally simple past.

I should perhaps point out that the figures for the modals, passives, and perfects are composite figures. The figures incorporate items from other categories, so don't try to add them. For example:

I 'would suggest' (Marked in Tense and Mood)

I 'would have suggested' (Marked in Tense, Mood, and Phase)

In these two examples the first item and the second would be calculated for the past and the modal, and the second would be calculated also in the figures for the perfect forms. We are presently beginning to feed all this information into the computer to give us complete breakdowns of each type (something we have had to do manually so far).

The figures are very revealing. And they can help us to define various registers -- will define various registers as our analysis is carried further.

For the sake of comparison I have included here some data from the second project mentioned earlier on verb phrases from a corpus of formal and informal spoken English which was analysed by the same process as I have described above (See Figure 8). And now I think, you can see the importance of these figures more readily. Note the reduced frequency of the unmarked forms, and of the simple past forms when we move to writing from spoken English. But, more important, note the increase in the importance of the modals, passives

	<u>Administrative Writing</u>
Unmarked (Simple Present)	32.1%
Marked in <u>Tense only</u> (Simple Past)	10.0%
Marked in <u>Mood</u> (Modal)	44.2%
Marked in <u>Voice</u> (Passive)	47.3%
Marked in <u>Phase</u> (Perfect)	14.9%
Marked in <u>Aspect</u> (Progressive)	7.5%

Figure 7

	<u>Informal Spoken</u>	<u>Formal Spoken</u>	<u>Administrative Writing</u>
Unmarked (Simple Present)	57.8%	54 %	32.1%
Marked in <u>Tense</u> only (Simple Past)	29.1%	27 %	10.0%
Marked in <u>Mood</u> (Modal)	40.4%	34.4%	44.2%
Marked in <u>Voice</u> (Passive)	9.9%	12.0%	47.3%
Marked in <u>Phase</u> (Perfect)	11.0%	18.7%	14.9%
Marked in <u>Aspect</u> (Progressive)	13.9%	13.7%	7.5%

Figure 8

and perfects. These changes are benchmarks in charting register; they cannot be ignored.

At this stage you may be tempted to say "So what! Just a lot of figures, they don't affect me at all as a teacher." We feel that they do. Indirectly perhaps, through the text materials you may be using. As he progresses in a course, the foreign learner may be expected to be able to speak and write accurately according to the intonation and grammar of English, and he may be expected to have mastered the so-called 'common core' of English, but he may still be lacking in fluency. To reach proficiency in English, he needs to be fluent, and fluency may be measured by his ability to conform and adapt in an appropriate manner to many different sociolinguistic situations. Perhaps, to enable him to achieve this proficiency, TESOL specialists should systematically introduce him to and guide him through these situations, each with its own set of distinctive linguistic features, so that he may develop a sense of linguistic appropriateness, the fluency which corresponds as closely as possible to that of native speakers.

So, our courses will, as our research findings become more refined, reflect these findings, and if our statistical information, when extended over the wider range we plan, stands, our course will, indeed must, look very different from those currently available.