National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

The Preference Formula:

Production Versus Prediction

Jamie Baird

A Thesis

in

The Department

of

Applied Linguistics

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts at
Concordia University
Montreal, Quebec, Canada

October 1990

© Jamie Baird, 1990

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-64660-8

Canada

ABSTRACT

The Preference Formula:
Production Versus Prediction

Jamie Baird

This thesis focuses on the discrepancy between what ESL teachers teach and what native speakers of English actually say when expressing preference --likes and dislikes.

Eighteen native speakers of English, divided into two groups according to social class, were interviewed in order to ascertain how preferences were expressed. The speakers were presented with a number of situations in which they described their feelings towards single items (objects, persons, ideas), towards two items, and towards groups of items. From this taped conversation, 1300 preference utterances were gathered and analyzed in terms of both the frequency of their occurrence and the situations in which they occurred.

Ten ESL teachers made predictions as to the language which would be elicited in the above corpus, and four textbooks were investigated for the same information.

Results show discrepancies between the predictions and the actual speech;

the language used was not the language teachers predicted. First,teachers predicted that comparisons with both elements being compared (*This* is better than *that*) would be used frequently, but they were not; second, teachers predicted that the situations and topics would greatly influence the language used, but they did not. Third, teachers did not predict that the language would vary according to social class, but it did.

In addition, it was found that the entire corpus of 1300 utterances could be reduced to five main structural types.

Finally, the implications of the results of this study are discussed in relation to pedagogy and materials writing.

Acknowledgements

I would like to thank Dr. Elizabeth Gatbonton for all her guidance and assistance in the production of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Do ESL teachers teach what people actually say, or only what the teachers themselves think people say? Can we rely on intuition for our teaching material? Tillit and Bruder (1985) in *Speaking Naturally* give a list of compliment phrases commonly used in English (p.70). Heading the list is "I would like to compliment you on...". But do English speakers often use that particular expression for complimenting, or are there other sentences more commonly used? According to Nessa Wolfson and Joan Manes in *The Compliment Formula* (1981) less than 2.8% of compliments would follow the *Speaking Naturally* pattern while 53.6% of compliments fall within a single syntactic category of <NP + {is/looks} + (really) + ADJ>. Why then do textbooks include *I would like to compliment you on...* when it is neither as common nor as simple as *You look nice*, or *That is pretty*? Wolfson (1983) states (p.117):

> ... although native speakers are able to recognize intuitively and respond appropriately to speech acts ..., they are not in a position to describe how such interactions are patterned. Intuitions are judgements of accuracy and appropriateness and as such they are a useful tool. However, when native speakers are asked to describe what they or others would say in a given situation, their responses do not always coincide with observed speech behaviour.... Examples of speech acts which are based on introspection rather than on actual speech, and explanations of sociolinguistic rules

> which depend on native speaker intuitions rather than on close analysis of real data, are frequently unreliable. It follows from this that the best way in which the sociolinguistic patterns of a target language can be analyzed and made available to language learners is through empirically based descriptive analyses.

Support for Wolfson's claim that native speaker intuitions may not be reliable predictors of actual language use comes from both sociolinguistic studies (Blom and Gumperz, 1972; Borkin and Reinhart, 1978) and corpus linguistic research into spoken language (Renouf, 1986). Renouf states (p.184):

> Our data indicates (sic) that the linguistic choices made by our subjects in response to the tasks prescribed are not those which would be predicted by theorists or materials writers in the field of English as a foreign language. This is important, since it means that, at present, learners are not being given the means to achieve classroom tasks in a way that accords with native-speaker practice, because language specialists are typically not taking account of authentic language behaviour.

Renouf (et al, 1986) gathered data through a series of language generating tasks specifically designed to elicit certain lexical realizations, verb tenses, classes of adverbs, etc. She then contrasted her expectations with resulting language. She found, for example, that

1. For the task, FREE TIME, requesting "Find out three things that your partner is planning to do when he/she next has some free time" (p.164), she predicted the use of *I'll*, *I'm going to*, and *I'm planning to*, but the actual response

was *I've got to, I want to, I would like to,* and *I'd like to.* Thus, she predicted use of future tenses but elicited present tenses and modals.

2. For A TYPICAL DAY, she predicted the occurrence of adverbs of frequency combined with the use of the simple present tense to express habitual action. Surely contrary to any ESL teacher or materials writer's intuitions, native speakers of English did not generally use the simple present tense to describe habitual actions. In fact, modals were used, almost exclusively; and the simple present tense was used very infrequently.

3. For DAILY ROUTINES, she predicted the same language as for A TYPICAL DAY, but the resulting language differed greatly. The speakers often used the present continuous to describe habitual actions.

4. For FAVOURITES, she predicted *my favourite x is, I like x best, I prefer x,* and she predicted that full syntax would be used. But none of the expected forms were elicited in the data. Renouf suggests that the speakers were responding to the forms in the instruction card rather than to the task itself; thus, for FAVOURITES, her results are an artifact of the data and not a reliable language sample.

Clearly, there are great differences between the intuitively predicted language use and the actual language used. Some of these false intuitions may be accounted for

as resulting from the method of elicitation, artifacts of the data-gathering procedure, but many may not. If, as Renouf's data suggest and as Manes and Wolfson believe, intuition cannot be relied upon, empirical studies must be conducted in order to ascertain just exactly what people do say.

The Manes and Wolfson study (1981) not only suggests that intuitions may be unreliable, but also that there may be recurring semantic, lexical, or syntactic patterns of utterances which are associated with certain language functions. According to Manes and Wolfson, these patterns may be reduced to what they call formulas. In their analysis of the data for the compliment formula, they suggest that due to "the combination of a restricted semantic set and an even more highly restricted set of syntactic structures" (p.123), they are not dealing solely with frequency of occurrence, but with actual formulas. Compliments are "highly structured formulas which can be adapted with minimal effort to a wide variety of situations" (p.123). They can either be used in whole or in part: either as entire chunks of memorized sequences or as elements slotted into certain categories. Manes and Wolfson believe that, at least for compliments (and invitations), frequently occurring syntactic and semantic patterns can be found and that these patterns can be reduced to formulas

representing a large portion of language use in a variety of situations.

Other research seems to support this notion of formulaic speech; to date, language samples have been gathered, analyses have been done, patterns have been found, and formulas of varying degrees of accuracy and comprehensiveness have been proposed for the following:

-compliments (Manes & Wolfson, 1981)
-compliment responses (Pomerantz, 1978)
-invitations (Wolfson, D'Amico-Reisner, Huber, 1983)
-scolding (Reisner, 1984)
-disapproval (D'Amico-Reisner, 1983)
-forms of address (Ervin-Tripp, 1969; Wolfson & Manes, 1979)
-directives (Ervin-Tripp, 1976; Holmes, 1983)
-conversational strategy signals "gambits" (Keller, 1981)
-apologies (Borkin & Reinhart, 1978; Olshtain & Cohen, 1983)
-telephone call openings (Schegloff, 1968).

Manes and Wolfson (1981) gathered data, isolated patterns of frequency, and proposed formulas to account for the patterns. They primarily addressed the question *What do we say?* However, for this project, I attempted not only to gather data from a number of sources, to isolate patterns, and to reduce these patterns to a series of hypothetical formulas, but also to compare the patterns and

formulas with teacher intuitions and textbook information on the expression of preferences in English. I attempted to answer not only the question *What do we say*? but also the question *Do we teach what we say*?

Specifically, this project, considers how preferences, likes and dislikes, in English are indicated. Do we say:

I prefer salty to sweet food, or

I like salty food better than sweet food, or

I would rather eat salty food than sweet, or

Salty food is better than sweet, or

I like salty food more than sweet, or

I love salty food, and I hate sweet food, or

I adore salty food, but detest sweet food, or

This is good, this is bad, or

do different people use different forms at different times in different situations? Perhaps we say none of the above. What exactly do we say most often, and what should we teach our ESL students to say? How closely does our intuition match the actual language that we use?

Simply then, in order to ascertain exactly what we say, a number of English native speakers talked about what they liked and what they did not like. A number of ESL teachers given the same directions guessed/predicted what the speakers would say. The actual language was compared to the predictions to see how well the teachers predicted.

This project examines how native speakers of English express preferences. The components are:

1. to gather data -language samples- from several sources,
2. to analyze the data in order to isolate patterns which may then be reduced to formulas in the language,
3. to survey ten ESL teachers to find what their intuitions and ESL experience tell them that native speakers say (what they think we say) and, therefore what they teach, when expressing preference,
4. to investigate four ESL texts, and
5. to compare the formulas (2) found in the data (1) with the forms presented in ESL texts (4), and in teachers' intuitions (3).
6. to discuss the significance, if any, of the formulas, the intuitions, and the comparisons between the two.

1.1 HYPOTHESES

There were two working hypotheses on which this research was based: the first hypothesis was that the language predicted and taught by the teachers and included in textbooks would differ from the actual language used by native speakers to express preference; i.e. it was expected

that teachers and textbooks would not accurately describe what people actually said; if this proved to be true, it meant that ESL students were studying and learning language which was not most commonly used by native speakers. Obviously, it would be preferable to teach and learn language which is used, authentic living language. It would be both interesting and useful to ascertain the similarity between what is used and what is taught. So, the first question was: *Are students being introduced to what is used?* And the expected answer was *No, not entirely.*

The second hypothesis was that the language elicited in the sample would be reducible to a set of formulas of the Manes and Wolfson (1983) style. The questions were: *Can operational definitions for formulas be devised?* and *Can the language in the corpus be reduced to a series of formulas according to those definitions?* The expected answer was *Yes.*

CHAPTER II

METHOD

In order to investigate the spoken language of preferences, teachers' intuitions, and textbook writings, it was necessary, first, to enlist subjects to create a body of language containing samples of preference language, second, to gather intuitions from ESL teachers, and, third, to examine ESL texts to be able to compare the material presented with both the intuitions and the actual spoken language. Accordingly, this chapter will be presented in the following main categories:

2.1: the **subjects** for both the corpus and the predictions,

2.2: the **procedures** for gathering the language sample as well as for gathering the predictions,

2.3: the **materials** involved for the corpus, the predictions, and the texts, and

2.4: the methods of **analyses** of these data.

2.1 SUBJECTS

A total of twenty-five subjects participated in the study. All were native speakers of Canadian English, all were born in Canada; fourteen had spent most of their lives in Quebec and Ontario, nine in B.C., one in New Brunswick, and one in Saskatchewan. All subjects had studied French; but only eight stated that they were fluent; thirteen spoke a little French, and four claimed that they were completely monolingual in English.

The subjects were categorized into one of three groups, Working-class, Middle-class, or Teachers. The divisions for social classes are not clear as, in *The Vertical Mosaic* (1965, p.10), Porter states "Class boundaries are drawn arbitrarily at points on the scales or indices...". But among sociologists, there seems to be some consensus on the criteria for description and analysis of social class. Porter (1965, p.10) states

> The most commonly used objective criteria of class are income, occupation, property ownership, and education, all of which are ways of expressing objective economic differences among members of the society.... Occupational class categories based on different degrees of skill such as professional, managerial, clerical, semi-skilled, unskilled, manual, or non-manual are economic categories which correspond to the structure of work in the economic system.... The more formal education a person has, the higher his skill level and the higher his income.

Accordingly, Blishen's 1958 occupationally-based scale was used to determine the class of the subjects for the study. In Blishen's scale, social level is determined entirely by occupation but formulated by taking into account both the educational requirements and the salary scale for the occupation in question. For the subjects in this project, in cases in which the occupations of the subjects do not correspond with their educational levels, the other factors mentioned by Porter --income, education, occupation, and property ownership-- were considered. Also, the two women who did not have occupations outside the home, i.e. the housewives, were classified according to their husbands' occupations, according to Blishen's 1967 statement "on the assumption that the family's social status is dependent upon the occupation of the husband..." (p. 42).

According to Blishen's scale, teachers do not, in fact, constitute a social class but are merely members of a larger social group, the Middle-class. However, as some (three) of the Teachers participated in both the corpus and the predictions, it was necessary to distinguish them from the Middle-class group. By maintaining the distinction between Teachers and other Middle-class speakers, the accuracy of the Teachers' predictions could easily be determined by comparing the predictions with output in the corpus. Consequently, all teachers were classified as Teacher-class

regardless of other factors: education, property ownership, or income. Thus, the group of twenty-five can be broken down into:

Working Class:

The seven subjects classified as the working class group (W-class) ranged in age from thirty-seven to sixty-five with an average age of fifty. The educational level ranged from grade eight to twelve, average of eleven, followed in some cases by vocational training. Only one subject was employed and earning a salary at the time of the sampling. The group was composed of two males and five females. Their occupations were: waitress, photographer, salesman, nurse, nurse, busdriver, and housewife. All seven provided data for the corpus.

Middle Class:

The six subjects in the middle class group (M-class) ranged in education from high school graduation to eight years of university studies; they ranged in age from twenty-six to sixty-five, average of forty; three were female and three were male. Their occupations were: graphic artist, writer, publishing company owner, housewife, post office clerk, and electrician. Salaries ranged from $40,000 to $65,000. All six provided data for the corpus.

ESL Teachers:

The twelve ESL teachers (T-class) ranged in age from thirty to forty-five, average of thirty-six; they ranged in ESL teaching experience from two years to twenty-one years. One of the teachers had also taught elementary school, and one had taught Hebrew and French for five years. Their education ranged from four to more than six years of university. (Details of the subjects are found in appendices 1 and 2.) Seven Teachers provided only prediction data; two Teachers provided only corpus data; and three of the Teachers provided both corpus and prediction data. Thus, eighteen subjects provided data for the corpus and ten for the predictions.

2.2 PROCEDURE

CORPUS DATA-GATHERING

The eighteen speakers were interviewed in pairs. All of the pairs of speakers in the corpus were well-known to each other, being either spouses, roommates, or friends, and all of the pairs were well-known to the researcher. Six of the interviews were conducted in the speakers' own homes, and three were held in the home of the researcher. In each case, the speakers were comfortable with their surroundings. The subjects sat around their kitchen tables with the instruction sheets on the table in front of them. The

conversations were taped on a Sony Tape Recorder with a miniature remote microphone set in the middle of the table. Before beginning taping, the researcher read out a prepared statement of instructions to the subjects to explain what was expected of them. (See appendix 3 for statement.)

The speakers read the questionnaires and responded to them. They were encouraged to talk about those questions which interested them and to skip those which did not. Each interview/ discussion lasted until all the questions had been either discussed or discarded. The researcher gave exactly the same instructions to each pair of speakers. The speakers were assured that their names would not be made public at any time, so they felt free to talk/voice opinions on any subject.

The average conversation/taping time was approximately one and one/half hours, resulting in a total of thirteen and one/half hours of taped conversation.

Originally, it was intended to gather a language corpus generated by Canadian speakers and to compare this with the language contained in the London-Lund Corpus of English Conversation. However, as all the subjects were Canadian, including both the corpus speakers and the predictors, and as the texts used were either Canadian or American, the introduction of other factors --English spoken by highly

educated Britishers-- seemed extraneous. Consequently, the L-L corpus was not investigated for this project.

PREDICTION DATA-GATHERING

The ten teachers from whom the prediction data were gathered were interviewed either on the phone or in person. They were presented with a four-part questionnaire designed specifically to record their comments and predictions as to:

a) the language they thought would be elicited in the corpus,

b) the language they believed that they actually used, and

c) the language they taught.

As was done for the corpus, the instructions for the predictions were purposely left vague so as not to suggest any preference forms; e.g., the words *prefer, like. good, love,* etc. did not occur anywhere. In addition, the teachers did not know the subject of the research until after the interview was concluded, thus minimizing interference/contamination by preconception. Once the instructions had been read, the teachers stated their ideas regarding all the language to be elicited, and the information was recorded on the questionnaire. All the relevant preference predictions were incorporated into the data for comparison with the actual corpus to determine how

accurately the teachers did predict. The questionnaire will be discussed further in the Materials section. Also, see appendix 5 for complete prediction questionnaire.

ESL TEXT ANALYSIS

A number of ESL texts were examined to discover which preference forms were presented, in which context, and in which order as targets of study. Textbooks considered for inclusion in this research had to meet several conditions: they were to be Canadian texts, thus describing Canadian language use; they were to be recent; they were to mention register or level of formality to make necessary comparisons with the language elicited in the corpus; they were to deal with oral/spoken English; and they were to clearly label the function of expressing preference or likes/dislikes, want/desire, or some comparable classification. The information in the texts regarding preference, or like/dislike, or the equivalent, was then extracted and recorded along with the context in which it occurred in the text. In addition, in order to more clearly compare the language of the corpus and the language of the texts, any information contained in the text on register, use, frequency, sociolinguistic rules and appropriateness, functions, etc., was included in this research.

2.3 MATERIALS

Two sets of questionnaires were used in this study: one set was used to elicit the corpus data from the W-class, M-class, and T-class speakers; the second set was used to elicit the prediction data from the T-class predictors.

CORPUS QUESTIONNAIRE

The three-page/three-part Corpus Questionnaire was specifically designed to elicit a variety of responses showing feelings (preferences, likes, dislikes, etc.) about the items involved; thus, a variety of subject matters and question types were presented. Ten Category A questions demanded that the speaker make a choice from among three or more items and then discuss the choice. Ten Category B questions had only two items to compare and contrast, and Ten Category C questions had only a single item to discuss.

Sample of Category A question:

> *Discuss the following.*
>
> *On the following page you will see a list of movies currently playing in Montreal. Choose one. Then suggest to your partner that you go to that particular movie and tell him/her why you chose it. If your partner chooses another movie, try to convince him/her to go to yours.*

Sample of Category B question:

> *Compare and Contrast the Following. How do you*

feel about them. Please discuss with your partner.

1. Bourasssa----------Parizeau
2. Bill 101----------Bill 178

Sample of Category C question:

How do you feel about the following? Discuss with your partner.

1. Visits to the dentist.
2. Chocolate cake.

(See appendix 6 for full Corpus Questionaire.)

Items of common interest and common knowledge were chosen. As the purpose of this research was to discover how people express likes and dislikes, both desirable, pleasant items such as chocolate cake and undesirable, painful items such as abortion and the dentist were incorporated into the questions. Also, both innocuous and highly controversial items were included to see if the speakers would use the same language -- structures and lexis -- for stongly emotional subjects as for weakly emotional ones. For example, the speakers discussed muffins versus croissants, and Vander Zalm vs. Barrett (or Bourassa vs. Parizeau if the interview were being conducted in Montreal). The depth of feeling for the muffins was not expected to be as great as for the politicians. Therefore, the type of language elicited might vary.

All three Categories included questions on food, politics, and geographical locations, so the subject matter

of the sections did not differ greatly, only the number of choices in each section differed. Therefore, if the three categories elicited different language, it might be because the number of items in each section varied and not because the subject matter of the individual questions varied.

In order not to contaminate the elicited data (as occurred with Renouf's 1986 data on favourites), care was taken that the expected preference words such as *prefer, like, enjoy, love, hate,* etc. were included neither in the directions nor in the questions themselves. Instead, less directive, less semantically loaded terms such as "How do you feel about...?" and "Suggest to your partner that you..." were used.

TEACHER PREDICTION QUESTIONNAIRE

The Prediction Questionnaire was divided into four parts: the first part asked for predictions as to the types of language which would be elicited by the three types of questions in the actual corpus questionnaire; part two asked for the teachers' opinion regarding personal use of preference forms; part three asked for pedagogical priorities regarding preferences; and part four surreptitiously created a true language choice.

Part I:

In Part I, the teachers were asked the following

question:

> *Part I:*
> *If the following questions were given to native speakers of Canadian English, what language do you think would most likely be elicited? (For example, which verbs, nouns, adjectives, structures, notions, functions, etc.) What language would a native speaker use? What would he/she say?*

The teacher was then presented with the first question from the corpus questionnaire and asked to respond to it.

Part II:

In order to discern what the teachers believed to be their own most common way of expressing preference, a list of forms (taken from my own intuitions) was presented.

> *Part II: If you wanted to tell someone that you believed one thing superior to another, what forms of English would you most likely use? What would/do you most commonly use?*
>
> *Would/do you say:*
>
> *I prefer x to y,*
> *I like x better than y, I would rather x than y,*
> *I love x, but I hate y,*
> *X is good. Y is bad.*
> *X is good. Y is not as good as X.*
> *X is better than Y.*
> *While (although) fond of x, I feel less than admiration for y.*
> *I adore x. I detest y.*
> *Or something else entirely.*

Teachers then discussed their choices. They stated which of the forms were generally acceptable to or generally used by them, which forms were rarely or never used by them, and which other forms they might use. In every case, the

teachers were told that they need not choose any of the sentences if they did not personally use them or if they believed other forms more useful and/or appropriate. In several cases, they listed the forms most used by them in order of preference, but this hierarchical description of frequency of use was not suggested by the researcher. The predictor's role was to predict what language would be elicited by the speakers in the corpus and to state his/her own choices of language items but not to predict what others would say outside the context of the elicitation task.

Part III:

The third part of the predictions sheet was the question: *What would/do you teach your ESL students regarding preferences?* The teachers were asked to describe which of the preference forms, if any, they would teach, in what order, and to what level of student. This section of the questionnaire was intended to discover if the teachers teach ESL students according to what their intuitions tell them are the most frequent and utilitarian forms, or if the ESL teachers teach what they think is simplest and easiest to learn, or if they teach willy-nilly, or if they teach only according to texts, or if they are even aware of what they teach. If the predictions in Part I, the personal language choices in Part II, and the teaching beliefs in Part III all closely resembled each other, the teacher could

clearly be seen to teach what his intuition told him was correct, what he believed to be common spoken English language. The next question remained, however: How closely would these intuitive predictions match real language use, the language generated in the corpus?

Part IV:

Part IV was a trick question on the part of the researcher. At the end of each of the interviews for teacher predictions, the researcher conversationally stated "My wife and I are considering going to Toronto for a holiday. Do you prefer Toronto or Montreal?" The conversational tone of the question and social content belied the fact that it was still part of the data-gathering for predictions. Thus the teachers answered in real conversational language and dropped their *predictive or scholarly mode*. As every one of the teachers involved in the predictions answered this question in an identical fashion regardless of and contrary to their own previous predictions, the anecdotal results of this little scheme proved quite interesting, as will be evident in the chapter on results.

2.4 ANALYSES OF DATA

Analyses were conducted on four main aspects of the study:

1. the transcription, which items were selected from the corpus and why,
2. the frequency and utility of lexical items and syntactic structures in the corpus,
3. the broad syntactic classifications developed in order to more simply describe the data in the corpus, and
4. the lexical categories, keyword and frequent versus infrequent distinction with regard to percentages of use of formulaic speech.

THE TRANSCRIPTION

The corpus resulting from the pair-discussions was analyzed, transcribed, and inserted into a relational database. In order to determine how the native speakers stated their preferences in English, all utterances in the corpus which carried either positive or negative semantic load were selected for transcription. These utterances were then inserted into a database under a number of categories. The sentences chosen were those which ***clearly*** showed the speaker's sentiment regarding a particular item. Sentences in which the researcher could not ascertain the speaker's sentiment toward the item in question were not included.

Long descriptive passages or merely factual statements were not included; only those utterances which obviously, syntactically or semantically, showed the speaker's preference, his like or dislike or hate or disinterest, his positive or negative feeling, were selected for transcription.

As was stated, utterances which indicated any speaker's feeling toward an item were chosen for transcription. However, there were emotive utterances in the corpus consisting of only gagging sounds, grunts, and single word expletives as well as very long complex grammatical arrangements. As this research investigated both syntactic and semantic frequency of occurrence, utterances without verbal structure, hence syntax, were not included for analysis; such utterances as *No good* or *Grossly disgustingly barfers* were left out. (In any case, these utterances constituted less than one percent of the total so were insignificant in number.) In order to be included, the utterance had to contain some kind of verbal structure, so utterances such as *That be good*, even though they may not conform to prescriptive grammar rules, were included. Each utterance was inserted into the database as a separate unit without its context, i.e. without the utterances preceding and following it, so each entry is a complete utterance, a complete syntactic and semantic unit. Moreover, each entry

is classified in a number of ways: by speaker, by question category, by syntax, by lexical category, and by other features (see below).

In order to insert the information into the computer, it was necessary to develop a system by which the utterances could be categorized, described, and labeled in terms of structural, semantic, and socio-linguistic factors. Further information regarding the description and explanation of the data-fields and their purposes in the transcription system appears in appendix 4.

To summarize: each entry in the database had to meet certain conditions: it had to express semantic force clearly, it had to contain some verbal structure, and it had to be a complete syntactic unit; also each entry in the database carried certain information with it: speaker, question type, syntax, lexis, etc.

FREQUENCY AND UTILITY COUNTS

As was mentioned above, each utterance was marked for certain features: the main word carrying the semantic force was marked in one field, and the (simplified) syntactic structure was marked in another field. In this way, frequency counts could easily be done; every utterance of a particular syntactic form could be isolated and counted; every utterance using, for example, the words *good* or *prefer*

or *favourite* could be counted. In addition, items using particular syntactic forms in combination with particular words could be counted. Thus frequency of both lexis and syntax could be determined.

All utterances in the corpus were described and classified for structure and for specific adjectives, verbs, and nouns which carry the semantic load; further, the number of utterances employing a particular means of expressing preference were counted and classified, resulting in frequency charts. For the purpose of this research, any lexical item or structure which occurred in more than one percent of the utterances in the entire corpus, was termed frequent. The demarcation line of one percent was arbitrarily chosen, but, as will be evident in Chapter 3, the one-percent line was a clear natural break distinguishing the common from the rare items.

Manes and Wolfson (1983) also suggest that "the combination of a restricted semantic set and an even more highly restricted set of syntactic structures" (p. 123), may result in a formula; in reverse, in order to have a formulaic utterance, it may be necessary to have restricted sets of both structures and lexis. In the corpus, all the items which meet the prerequisite of Frequency make up a very restricted set, both for structure and lexis. But Frequency and set restriction may not be the only factors

involved in determining whether an item is formulaic. According to Manes and Wolfson (1983, p.123), "Compliments are ... highly structured formulas which can be adapted with minimal effort to a *wide variety of situations* in which a favourable comment is required or desired."

So, in order for an utterance to be formulaic, it must not only be *Frequent* (and therefore of a restricted set), it must also be used in a wide variety of situations by a variety of speakers. The only variety of situations which occurred in the corpus was the differing Categories (three or more, two, or one topic), so, for the purpose of the research, any utterance in the corpus which occured in all three Categories, and was used by speakers in all three socio-economic groups, was accepted as displaying *utility*. Utterances with both *frequency* and *utility*, according to the above definitions, will be called ***formulaic***.

SYNTACTIC TYPES

For the purpose of analysis, sentences indicating preference, choice, attitudes towards one object in relation to others were inserted into the data-base under a number of descriptions (see appendix 4). Once the data were analyzed under these simplified grammatical descriptions, the results were investigated to ascertain if larger patterns or trends could be found. A few very clear patterns began to emerge

suggesting that the entire corpus could be accounted for with the use of a small number (5) of utterance Types, with Subtypes. By reducing all utterances into only a few main Types, not only the language from the corpus, but also the teachers' predictions and the textbook language could be described and compared with facility. Commonalities and differences between actual speech and predicted speech could be noted.

LEXICAL CATEGORIES

All utterances in the corpus were classified according to their structure. This syntactic typing gives us information regarding frequency of structures used, but, generally, it provides no information regarding frequency of specific verbs, nouns, and adjectives used. We may know, for example, that forty-seven percent of the total corpus consists of Type 4 sentences, (Complete descriptions of Types follow in Results) indicating that adjectives determine the semantic force, but we do not know which adjectives fill that function. Consequently, it was necessary to categorize the words which carried the semantic load, *keywords*, by frequency as well as by structure. For simplicity's sake, the keywords were counted and categorized as either Frequent (F) or Infrequent (IF) according to the previously mentioned 1% demarcation point. With the use of

the keyword label and the Frequent/Infrequent distinction, it was possible to spot social-class variation in use of Frequency versus Infrequency for both lexis and structure: viz, higher use of IF items may indicate greater variation in the individual's speech, and less use of IF items may indicate less variation in speech. The IF speech (use of IF items in speech) can be expressed in terms of percentages of total speech; the higher the percentage of IF use, the greater the variation in the speaker's language. Greater percentages of IF use for any social group may indicate greater variation and thus greater range of structure and lexis available to or used by that group, and lower percentages may mean just the opposite.

The keyword and the F/IF labeling also aid in the search for formulas. By the earlier definition, in order to be formulaic, an item must be F and have utility. Combining findings from the social-class variation regarding use of F/IF items may lead to relationships between social class and use of formulaic speech.

CHAPTER III

RESULTS

3.1 RESULTS OF THE CORPUS ANALYSIS

The eighteen speakers produced approximately thirteen and one-half hours of taped conversation. From this taped discussion, 1292 individual utterances indicating preference were selected according to the criteria described earlier.

SYNTACTIC TYPES

Analysis of the utterances in the corpus into Syntactic Types reveals that they can be categorized into five different sentence Types according to the element carrying the semantic force. A description of the Types (plus the conventions used in the notation) will be followed by Table 1, showing the frequency of occurrence of each Type.

Type 1-A: NP + (NEG) + prefer + NP/VP

I prefer it, I prefer Montreal, I would prefer living in the city.

Type 1-B: NP + (NEG) + prefer + NP/VP + to/than/over + NP/VP

I prefer it to having bad teeth, I prefer to live in Montreal than in that hole.

Type 2-A: NP + (NEG) rather + VP (NP)

I'd rather not go, She'd rather lose weight.

Type 2-B: NP + (NEG) rather + VP (NP) + than + (VP)/(NP)
I'd rather eat salty stuff than sweet stuff. I'd sooner stay here than go to Toronto for a holiday.

Type 3-A: NP + (NEG) + (int) + like + NP/VP
I like movies, I don't like poutine, I really hate going to the dentist, I don't mind it.
Type 3-B: NP + (NEG) + (int) + like + NP/VP + comp + NP/VP
I like this one better than the first one we listened to, She loves chocolate less than chips.

Type 4-A: NP + (NEG) + be + (int) + adj + (NP)
It was beautiful, It wasn't so great, That sounds good, It tasted really awful, That was better.
Type 4-B: NP + (NEG) + be + (int) + comp + NP/VP
Shmarm is better than violence, The transportation system in Toronto is way more efficient than Montreal's.

Type 5-A and **Type 5-B:** All others unaccounted for by the above classes including:
-modal auxiliary verbs (e.g. *I'll eat cherries*)
-verbs of the type NP (it) + bother + NP (personal pronoun)
It bothers me, It doesn't interest me, it bugs me, it appeals to me.

- verbs which do not satisfy the affective condition of Type 3 such as trust, e.g. *I trust him completely where I don't trust Mulroney at all.* Other verbs of this type: suspect, disbelieve, etc. -sentences in which the semantic load is carried by a noun, e.g. *Bingo is my favourite, I'm a fan of chocolate.*

-Finally, all other sentences not covered by any of the above categories will be included in this Type.

CONVENTIONS:

A-Types = utterances in which only one element is mentioned.

B-Types = utterances in which two elements are mentioned, i.e. compared or contrasted "complete sentences".

NP = Noun Phrase, --noun, determiner + noun, Sentence or clause, etc-- e.g. NP + like + NP = *I like this book, Hilda likes books, He liked the book I gave him,* etc. NP's do not carry the semantic load except in the subType 5 sentences with nouns as critical elements.

VP = any Verb Phrase which does not carry the semantic load, e.g. *I like singing, I like to sing, I'd rather do it than watch it.*

(...) = optional elements. (NEG) = optional negative in verbal structure.

LIKE = all affective verbs --love, hate, adore, mind, etc---

int = intensifiers --very, too, really, etc--

comp = comparative joiners --better than, more than, less than, as much as, etc--

Be = linking or copular verbs (stative verbs) --be, taste, feel, look, seem, sound, appear, smell, etc--

adj = any semantically positive or negative adjective.

Notes = all verbs are cited in the base form. Adverbs (or phrases) and prepositions (or phrases) are not noted as they did not appear to carry the semantic load in any case. Modals which did not carry the semantic force of the utterance were not recorded e.g. *I would love to go, I would prefer it.* Embedded clauses which did not change the structure of the preference were not noted: I *think it was really good, The one we saw with Hilda was stupid.* NP + be + (int) + adj.

TABLE 1

STRUCTURAL TYPES BY SOCIAL CLASS

	M-class	W-class	T-Class	Total	(%)
Type 1-a (prefer)	3	6	3	12	
Type 1-b	0	0	2	+2	
				14	1%
	M-class	W-class	T-Class	Total	(%)
Type 2-a (rather)	2	7	13	22	
Type 2-b	2	2	3	+7	
				29	2%
	M-class	W-class	T-Class	Total	(%)
Type 3-a (like)	156	180	138	474	
Type 3-b	0	4	1	+5	
				479	37%
	M-class	W-class	T-Class	Total	(%)
Type 4-a (+ adj)	167	258	158	583	
Type 4-b	7	12	5	+24	
				607	47%
	M-class	W-class	T-Class	Total	(%)
Type 5-a (other)	55	64	39	158	
Type 5-b	1	1	3	+5	
				163	12%
				1292	100%

As is evident from Table 1, Types 1 and 2, i.e. *prefer* and *rather* constructions, are used infrequently in the corpus. *Prefer* forms account for one percent of the total; *rather* forms account for two percent.

The most common ways of expressing preference or dislike are with Types 3 and 4, *like* verbs and stative verbs with adjectives. They occurred in 84% of the cases.

COMPARISONS

Note that in Table 2, sentences involving (descriptions

of) only one item and (comparisons of) two items are included under each Type, distinguished by the -A and -B. Examples of sentences with B-Types, with both elements are :

People with a suntan look healthier than people with no suntan,
They have a larger intelligence than Australians.
There isn't another car as good as that one.

Examples of A-Types with only single elements are:

It's good stuff.
I hate it.
I like it.

Analysis of the data reveals that the A-Types are used more frequently than the B-Types. Table 2 shows the breakdown of the utterances according to the number of elements mentioned.

TABLE 2

FREQUENCY: TYPES A AND B

	M-class	W-class	T-class	Total	
Type A (1-5) (one element)	383	515	351	1249	97%
Type B (1-5) (two elements)	+10	+19	+14	+43	+3%
Totals	393	534	365	1292	100%

Combining all the B-Types results in only 43 utterances, 3% of the total. The remaining 97% use only one

element. Moreover, all the comparative sentences used in the corpus constitute a very small percentage of the total. Table 3 below shows the numbers of comparative and superlative structures used by both class and by Category.

TABLE 3

FREQUENCY: COMPARATIVES AND SUPERLATIVES

Types C:	M-class	W-class	T-class	Total
Comparatives:	12	14	12	38
	Cat. A	Cat. B	Cat. C	
	13	16	9	
Types S:				
Superlatives:	M-class	W-class	T-class	Total
	7	11	3	21
	Cat. A	Cat. B	Cat. C	
	8	7	6	

Comparative constructions with one element total 38, three percent of the total corpus; comparatives with two elements total 29, two percent of the total. Thus all comparative constructions total only five percent of the entire 1292 utterances. Superlatives total 21, just less than two percent. Other Types with two elements account for 14 additional sentences or 1%.

CATEGORIES

Tables 4 and 5 show the breakdown of the utterances according to the elicitation condition, i.e. the Category and the number of elements being discussed.

TABLE 4

	CATEGORIES			
	Cat. A (discuss 3 or more items)	Cat. B (compare 2 items)	Cat. C (describe 1 item)	Total
# of utterances:	500	492	300	1292
% of total:	39%	38%	23%	100%

As can be seen, Categories A and B, discussing more than a single item, generated more speech than did Cat. C. The ratio is approximately 5/5/3 for total number of utterances in each Category. (See appendix 9.)

TABLE 5

STRUCTURAL TYPES BY CATEGORY

	Cat. A	Cat. B	Cat. C	Total
Type 1	0	10	4	14
Type 2	10	12	7	29
Type 3	178	169	132	479
Type 4	286	237	132	655
Type 5: Nouns:	12	56	15	83
Other Verbs:	12	6	9	27
Other:	2	2	1	5
	500	492	300	1292

Great deviation from the standard 5/5/3 ratio occur in Types 1, 4, and 5. *Prefer* is used more with two elements than with either three or one, and adjectives with stative verbs occur most frequently with three or more elements.

SOCIAL CLASS VARIATION IN SPEECH

The main variations in speech between the three classes can be seen in the use of Frequent versus Infrequent items in structural Types 3, 4, and 5.

Type 3:

Like verbs comprise a total of 479 utterances, 37% of the total corpus. This group consists of fifty verbs: the five Frequent verbs, those used more than 1% of the time, are: like (45%), love (15%), want (11%), hate (7%), and mind (3%), adding up to 81% of this Type. 45 other verbs, none of which is used more than one percent of the time, make up the IF subgroup, Infrequent verbs. See appendix 8 for complete list of verbs and their frequency.

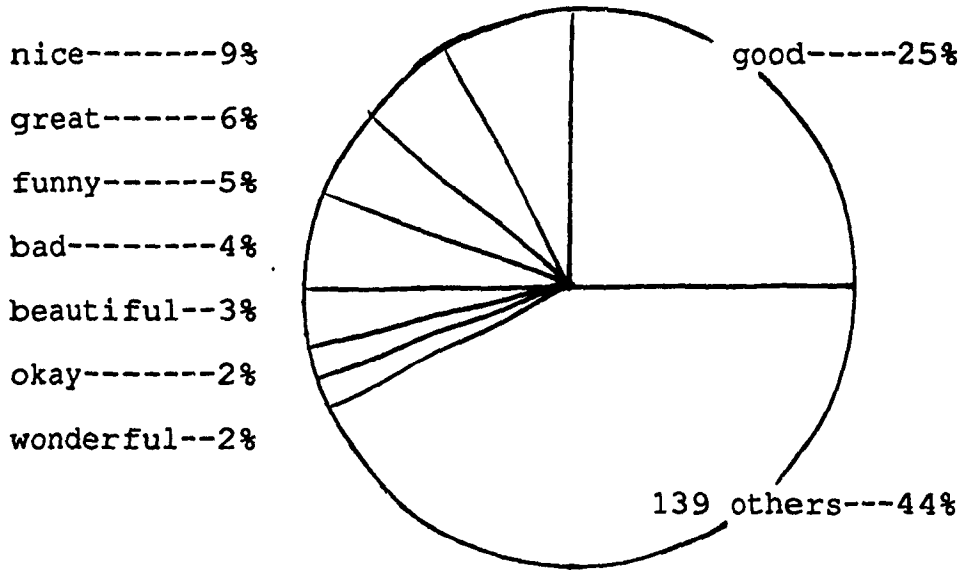
Type 4:

The Type 4 structures *be + adj* accounted for 47% of the total corpus, i.e. almost half of the preferences shown were made with adjectives expressing the semantic force. As with the Type 3 sentences, the adjectives in Type 4 can be broken down into two main classes: the Frequent and the Infrequent. Frequent adjectives include: good --including better and best-- (25% of adjective use), nice (9%), great (6%), funny (5%), bad --including worse and worst-- (4%), beautiful (3%), okay (2%), and wonderful (2%). Thus, the Frequent adjectives account for 56% of adjective use. The remaining 44%, or 273 utterances, use a total of 139

different adjectives, none of which are used more than one percent of the time. However, there were 48 other adjective occurrences in the corpus in which the main verb was not *be*. For the general structural Typing, as the structure was different, these uses were included in Type 5, but they will be incorporated into Type 4 here to more comprehensively describe adjective distribution. (Consequently all the figures total 655 rather than 607; frequency ratios do not alter with the inclusion of the non-be adjectives. The total numbers change but the ratios are not affected.) See figure 1 for graphic representation of adjective distribution.

FIGURE 1

DISTRIBUTION OF ADJECTIVES



Type 5:

Type 5 consisted of 163 sentences of various structures: eighty-three sentences in Type 5 utilized nouns for meaning, sixty-one of them with be and twenty-two with other verbs; forty-eight sentences used adjectives with verbs other than *be* (mentioned in the Type 4 data above); twenty-four sentences used thirteen different verbs of the type *It annoys me*. Four used modal auxilliary verbs for meaning, and five were isolates in that they fit no other group.

Sixty-one different nouns were used, but, unlike the other Types, none of the nouns in this group were Frequent: *favourite* was used only nine times (11% of total noun use, but less than 1% of the total corpus), *fan* (6%), *shit* (6%), *son-of-a-bitch* (4%), *charisma* (4%), *bullshit* (2%), and *no way* (2%). These seven nouns only account for 35% of the total noun use, resulting in 54 nouns used only once each, 65% of the total.

The M-class used twelve Type 5 verbs; the W-class used only six. However, with so few entries in this category, there is no meaningful Frequent/Infrequent distinction. Also, the isolates are not treated here, as, with only five entries, they account for less than 0.4 percent of the total and are thus insignificant.

Table 6 presents the breakdowns of Frequent vs. Infrequent item use by Type and class. No detailed statistical analyses were performed; the ratios are presented in terms of percentages of the class totals. For example, the M-class speakers used *like* verbs (Type 3) a total of 156 times. 116 of those uses were with Frequent verbs, and the remaining 40 were with the IF verbs. Thus, the IF verb use was 10% of the total, and the F verb use was 29% of the total. Using this notation system allows comparison between the various classes and their use of Types as well as of F and IF items within those Types.

TABLE 6

FREQUENT VS. INFREQUENT USE BY TYPE AND CLASS

	M-c	(% of M-c)	W-c	(% of W-c)	T-c	(% of T-c)
Total # S's :	393		534		365	
Average # S's:	66		76		73	
Type 3 (Like verbs):						
Frequent (# of S's.)	116	29%	160	30%	115	31%
Infrequent	40	10%	24	5%	24	8%
Total Type 3	156	39%	184	35%	139	39%
Type 4 (adjectives):						
Frequent	94	24%	193	36%	90	25%
Infrequent	98	25%	94	18%	86	24%
Total Type 4	192	49%	287	54%	176	48%
Type 5 (other verbs)	12	3%	6	1%	9	2.5%

From Table 6, it can be seen that the W-class speakers

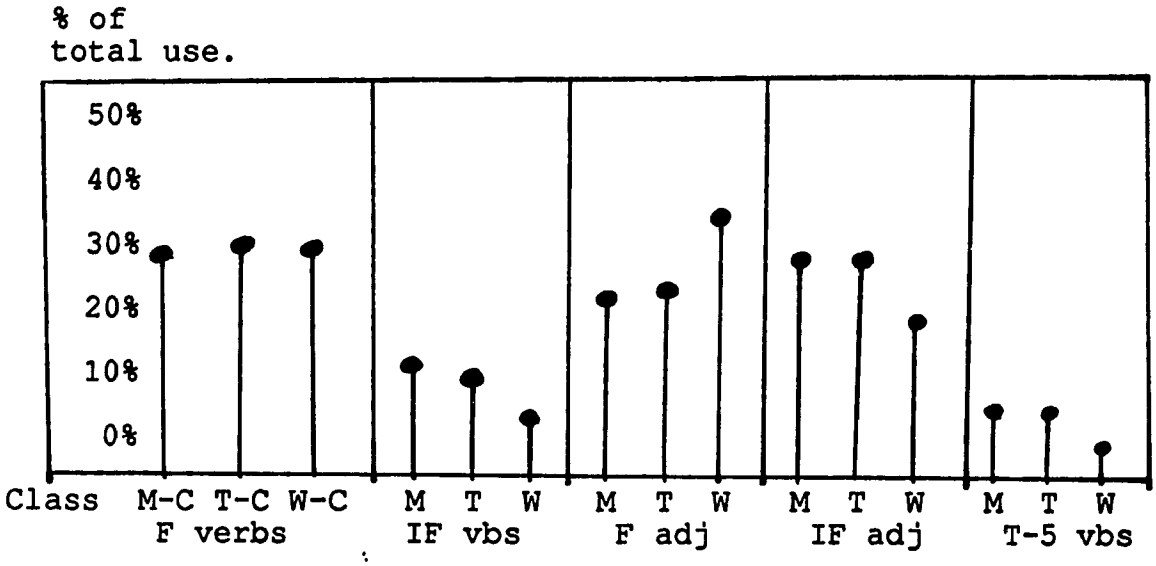
produced more utterances per capita (76) than did either of the other two groups (73, 66). They simply talked more.

In addition, the W-class speakers used a much higher percentage of Type 4 sentences with frequent adjectives, *It's good, It's nice;* 36% of W-class sentences were of this type, while only 24% of M-class and 25% of T-class utterances were of the *be + [frequent] adj* form. The W-class also used utilized a lower ratio of IF Type 3 verbs (5%) than did either of the other two groups (10%). The most common form for M- and T-class speakers was Type 3-A, 29% and 31%.

38% of M-class and 34.5% of T-class utterances utilized infrequent elements, whether verb or noun; but only 24% of W-class utterances used IF elements. This means that W-class speakers used fewer adjectives and verbs more often than did the others; W-class speakers controlled a narrower range of lexical items. There is no apparent pattern for occurrence of nouns, so they are not mentioned here. Figure 2 below shows the distribution of class speech in terms of the Frequent/Infrequent distinction.

FIGURE 2

DISTRIBUTION OF F/IF SPEECH BY CLASS



FORMULAIC EXPRESSIONS

The Frequent items in the corpus are represented in Table 7.

TABLE 7

LIST OF FREQUENT ITEMS

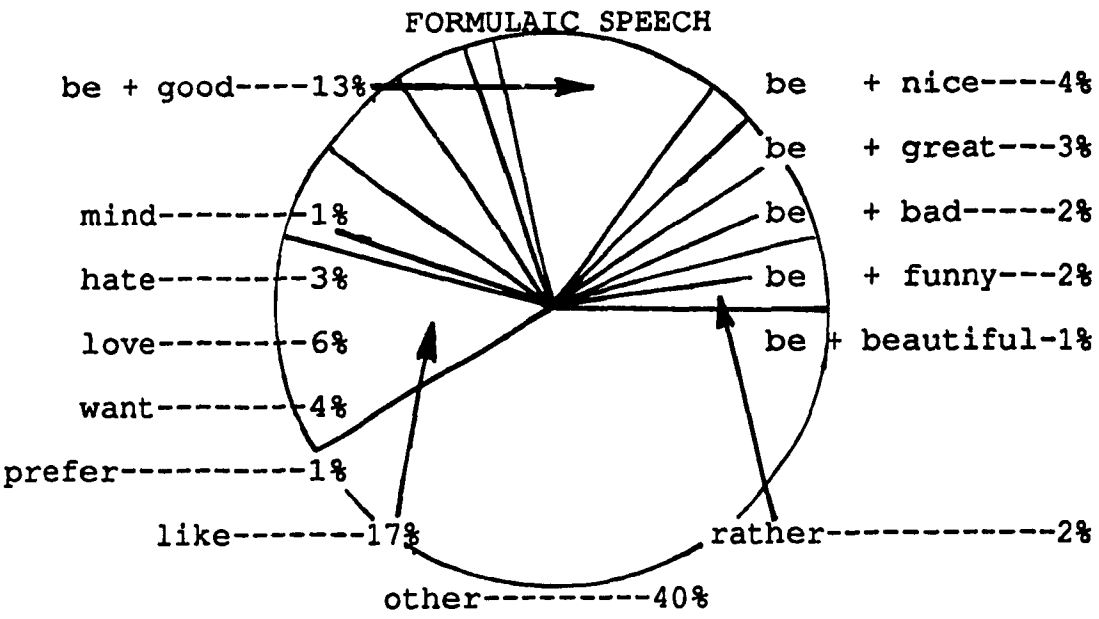
Type:		Percent of total:	Utility*:
1-A:	prefer + NP/VP	1%	N
2-A:	rather + VP	2%	Y
3-A:	like + NP/VP	17%	Y
	want + NP/VP	4%	Y
	love + NP/VP	6%	Y
	hate + NP/VP	3%	Y
	mind + NP/VP	1%	Y
4-A:	be + good	13%	Y
	nice	4%	Y
	great	3%	Y
	bad	2%	Y
	funny	2%	Y
	beautiful	1%	Y

*Utility: Y (yes) indicates that the item is used in all categories and by all speakers. N indicates that it is not used in all categories.

The above 13 items account for 60% of the corpus, and can be graphically represented in Figure 3. *Prefer* is the only item listed which does not have Utility (no examples in Category A), but as it has Frequency, it is included here in any case. A very few other items --Type 4-A: *awful, fine, lovely, neat, terrible;* Type 3-A: *can't stand, and enjoy*-- display utility but not frequency and are thus not listed as formulaic. There were no nouns or non-Type 3 verbs with either frequency or utility. Accepting that the conditions for formulas are Frequency and Utility, the above

thirteen items represent the Formulaic speech in the corpus.

FIGURE 3



Manes and Wolfson (1983) found similar results in their corpus for compliments.

-54% of the compliments used {NP + is/look + (int) adj} (similar to Type 4).

-16% used {I + (int) + like/love + NP} (similar to Type 3)

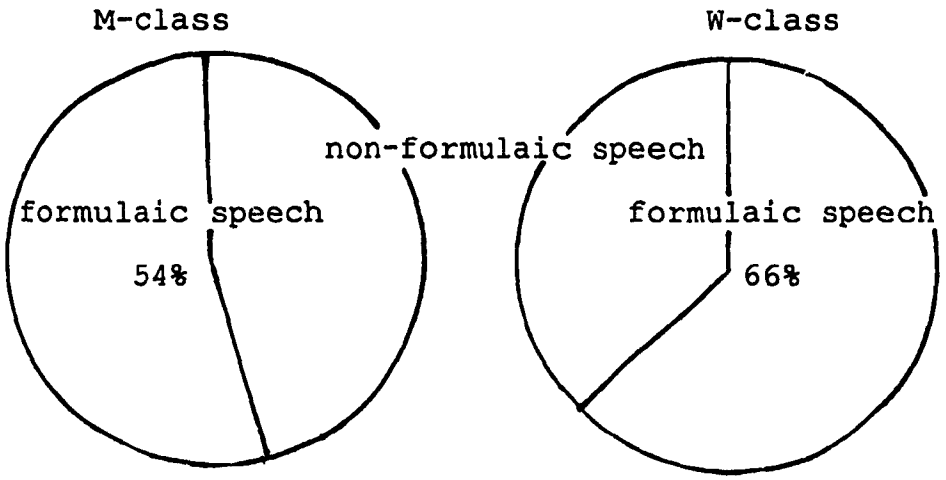
Although their notation system varies slightly from the one in the present study, it is obvious that similarities in the results exist. The two most common forms used to compliment are also the most common forms to express preference.

FORMULAIC VARIATION BY CLASS

According to our earlier definition of Formulae, the results shown in the above tables suggest that the W-class used a higher percentage of Formulaic expressions than did either the M-class or the T-class. The W-class speakers appeared to used a narrower range of both structures and lexical items to show preference; they used more frequent items and fewer infrequent ones, i.e. more repetition and less variety. The M-class and the T-class speakers were comparable in most regards; this stands to reason as, according to our use of Blishen's' occupation scale, teachers are Middle-class speakers.

FIGURE 4

FORMULAIC SPEECH BY CLASS



3.2 TEACHERS' PREDICTIONS

The Teachers' Predictions were gathered in the four-part questionnaire described earlier, the findings of which will be presented in order here.

Part I:

The first part of the prediction questionnaire elicited predictions regarding the language to be elicited in the corpus. The teachers predictions are presented here according to Category --A (three or more discussion items), B (two items), C (a single item). Ten teachers made predictions, but, as none of the predictions were made by all ten teachers, each Type is followed by a number less than ten which indicates how many teachers actually suggested that particular items would occur; in fact, the highest is six for Type 1 --*prefer*. This means that six of the ten predictors believed that with three or more items being discussed, the speakers would use the word *prefer*; five in Type 1-B indicates that half of the teachers believed that Cat. A would elicit NP + prefer + NP/VP + to/over/than + NP/VP, e.g. *I prefer swimming over suntanning*. All the predictions are expressed in terms of the Type system used earlier and throughout; Type-C and Type-S are introduced here only to show prediction of Comparatives and Superlatives.

TABLE 8

	PREDICTIONS		
Type	Cat. A:	Cat. B:	Cat. C:
1-A:	6	2	0
1-B:	5	1	0
2-A:	5	0	0
2-B:	0	0	0
3-A:	4	1	6
3-B:	0	0	0
4-A:	4	2	2
4-B:	3	4	1
4-C*:	2	6	0
4-S*:	5	1	0
5-A:	5 (modals)	2	0
5-B:	0	0	0

*1 = comparative constructions with Be verbs and adjectives.
*2 = superlative constructions with Be verbs and adjectives.

PREDICTIONS FOR THE CORPUS

1. The teachers predicted great variation between the language elicited in the three Categories--Cat. A cf. B, C. One predictor said, for example,

> *In this one (A), they'll say stuff like What do you think about? How come? I'd like to go...because...,* and use comparatives like *It's nicer here than there, I've never been there* and superlatives and descriptions like *if ... then... I will, I prefer this one to that one* and comparatives, for example, *if we go there we will do x and Y.*
>
> And in this one (B), I think they'll say *I think...if you ask me... You must be out of your mind,* gambits, invective, comparatives, *US is bigger*

> *than Canada, more money better films, Canada would do X. America would do Y,* lots of woulds. And here, oh, this one (C) would be completely different. *I like it but it gives me a headache, I remember when we made ice-cream. Eating ice-cream makes you fat,* but, however, opinions, like *Going to the dentist is for morons. It makes me sick, I can't stand ... I can't take ...*

All but one of the other teachers followed the above pattern of predicting that different language would be elicited in the three Categories. The actual corpus results showed no great variation between categories except for Type 1 -*prefer*- and Type 4 -adjectives; *prefer* was used in Cat. B but not in A, yet it was predicted more in Cat. A and less in B. Type 4 was predicted most for comparative adjectives, but as we have seen, comparatives were infrequent for this purpose.

2. The teachers predicted widespread and frequent use of Types 1, 2, 4-B, 4-C, and 4-S (*prefer, rather,* comparatives, comparatives with both elements, and superlatives). As was seen, none of these occurred with frequency. Half of the teachers predicted the use of *prefer X to Y,* but there is not one example of this in the entire corpus. There is one example of *prefer X over Y* and one of *prefer X than Y,* neither of which was predicted.

3. The teachers predicted frequent use of modals to express preference. *I wouldn't do that, I wouldn't see a violent movie like that, I couldn't go on rides.* Except for

modals as parts of other verbal elements, such as *would like, would rather, would prefer, can't stand, and can't take,* there were only four uses of modals to show preference, a very infrequent occurrence.

4. The most predicted Type was (*prefer*) Type 1 and the least predicted Type was (adj) Type 4 which runs directly contrary to corpus results.

Part II:

In this section of the prediction questionnaire, the teachers were presented with a list of possible preference structures. They were asked which they thought they would be most likely to use and told that they need not stick to the list. Thus, the list presented to them was meant to provide suggestions but not to limit their choices. So, what did the teachers believe that they would say?

Six teachers stated that they would most commonly say, Type 1-B, *I prefer X to Y.* There are no examples of this in the corpus, not even from the teachers themselves.

Six teachers believed that they would use the form, Type 3-B, I *like X better than Y,* and five thought they would say Type 4-B, *X is better than Y.* The majority of the teachers, then, predicted that they would use the B-Types most commonly in their own speech. As was seen, the B-Types accounted for only 3% of the corpus.

One teacher suggested that she would say *I adore X but I detest Y.* There were no examples of detest in the corpus, and only three uses of adore.

One predictor did not predict any of the above. On the contrary, he stated that all the forms in the list sounded stilted and predicted that he would use such structures as: *X is good, Y is bad, I like it, I don't like it.* In all his intuitions of his own speech production, and in the earlier predictions for the corpus, in Part I, only this one individual predicted accurately. (This individual has the longest teaching record, with over twenty years of ESL experience. Perhaps he has learned what is easisest to teach?)

Part III:

In this part of the predictions, the teachers were asked to describe what they taught in their ESL classes. Nine of the ten stated that they would teach all the forms listed, except for *adore, detest,* and *fond of,* depending on the level of the class. The correct predictor above stated that he would only teach what he thought people actually said, X is good, Y is bad, as above, but that he would not discourage use of the other forms if they should appear in the class. The other nine believed that all of the structures were required in their B-Types or their "complete" forms (term quoted from several predictors), such

as *I p~(· X to Y* and *I like X better than Y;* not one of the nine suggested alternate forms, nor did they question the frequency or utility of the B-Types. All nine maintained that they taught what they themselves produced in casual conversation; they maintained that they themselves said *I prefer X to Y* and other B-Types, so they taught them.

Part IV:

The final part of the questionnaire consisted of a single question asked outside the context of the formal interview. The researcher asked the predictors where they would prefer to spend a week for a holiday. The question for every teacher was exactly the same: *Do you prefer Toronto or Montreal?* In every case, the answer was *I like ....* Not one of the teachers said *I prefer...* This fourth part of the questionnaire was conducted after the interview had been completed and the predictor had been told that he/she was finished. In every case, the fourth question took place only minutes after the final predictions had been made, but, even so, the teachers did not answer with the preference forms that they had just predicted. Obviously, this anecdotal aspect of the Predictions Data does not prove that all intuitions are inaccurate, as it may still be entirely possible that *prefer* is used in other contexts. But, in conjunction with the results from the corpus in comparison with the first two parts of the Predictions, the

use of *I like...* in a case where *prefer* had just been predicted does tend to suggest that intuitions do not accurately describe reality.

3.3 TEXTBOOKS

As was expected, it appears that teachers' intuitions regarding language use may not be entirely accurate. The next question is *How accurate are the textbooks that the teachers use?* Because textbooks tend to be written by teachers, and because they are not generally based on corpus research, it stands to reason that the same inaccuracy will appear in the ESL texts. In order to check this hypothesis, several texts were examined. Since the corpus was produced entirely by Canadian speakers, and since the predictions were made by Canadian teachers, it was intended that the textbooks examined would be entirely Canadian as well. However, a sufficient number of Canadian texts which included information on preferences or likes/dislikes could not be found. Thus, the books chosen for the comparisons were two Canadian and two American conversation texts.

As was mentioned earlier, the only books investigated were those which specified appropriate language use according to register, which dealt with oral language, and which clearly labeled the function of expressing preference,

like/dislike, want/desire, or some comparable classification. The four texts were:

Take Part: Speaking Canadian English (Engkent and Bardy, 1986),

Functioning in English (Mendelsohn, Laufer, and Seskus, 1984),

Speaking Naturally (Tillit and Bruder, 1985),

Expressways English for Communication (Molinsky and Bliss, 1988).

Take Part: Speaking Canadian English:

In its introduction, this text mentions that:

> Register, the language appropriate to a particular context, is an important part of language use. It is, however, often neglected in language teaching... *Take Part* focuses on the conversational register of everyday English and explains what characterizes it and how it is different from more formal English. (p.vii).

Take Part clearly states that it is a conversation text, and it describes the language according to this level of formality. Some examples of preference language in the text are:

> *I would rather watch the movie than the hockey game.* (p.43).
> *Would you prefer an extended or a short holiday? Why?* (p.84).
> *Anything would be better than this dump.* (p.62)
> *I saw some ads for apartments that didn't sound half bad.* (p.62)

> *I'd rather walk than take that car to the wedding.* (p.63).

Note that this book presents B-Type utterances first and exclusively. All of the sentences in the text with *prefer, rather,* and *better* explicitly mention both compared elements regardless of the register. This agrees with the teachers' predictions for use but runs contrary to the production in the corpus. With regard to comparatives, the text includes B-Types in every case with no mention of the possibility of dropping the second element in casual conversation. There are no examples of A-Types. The only mention of register for comparatives is with regard to the object pronoun, *he* in formal language versus *him* in informal.

> *This book is more expensive than that one.*
> *She plays better than him.*
> *This story is not as interesting as that one.*
> (p.63).

In contrast, it was found in the corpus, that as the second element was generally dropped altogether, neither *he* nor *him* tended to be used in conversational language.

So, both for verbs -- *prefer* and *would rather* -- and for comparative adjectives -- *better, more expensive than,* and *as interesting as* -- the text presents B-Types, agreeing with the teachers' predictions in contrast to the language produced in the corpus.

Functioning In English:

In its introduction, this text also stresses the importance of sociolinguistic awareness.

> As such, language must be seen as being governed not only by linguistic rules (grammar, phonology, etc.) but also by sociolinguistic rules (setting, topic, relationship between speakers. etc.) which determine the effectiveness, appropriateness and perceived meaning of any chunk of language. (p.iii).

Some examples of their "key words and expressions" for expressing likes and dislikes were (p.21):

Strong Like	***Strong Dislike***
I love...	*I hate...*
I adore...	*I detest...*
I really like...	*I can't stand...*
Great!	*Forget it!*
Terrific!	*Not at all!*
Super!	*Not my type of thing!*
Marvelous!	*Awful!*
Wonderful!	*Terrible!*
	Dreadful!
	Yuk!
Weak Like	***Weak Dislike***
I like...	*I'm not mad/crazy about...*
It's all right.	*I'm not very fond of...*
It's okay.	*I don't like...*
Nice.	*I dislike...* etc.

And on the following page (p.22):

> *It's not bad.*
> *It's so-so. (so-so was not used).*
> *I prefer X to Y.*

And as examples of the above forms in an explanation of an exercise: (p.24)

> *The first time you are approached about food, you might answer: "I adore shrimp and I detest broiled fish." The second time, when asked about music, you*

> *should not use adore, detest, but say: "I'm extremely fond of rock music but I can't stand folk music."*

Functioning in English, then, presents as preference language a distribution of Frequent, Infrequent and unused structures and lexical items according to the corpus results. Some of the language clearly reflects that of the corpus, e.g. Type 4-A, *It's not bad;* but some lexical items and expressions such as Type 1-B, *I prefer X to Y* and Type 3, I *detest + NP/VP* are not at all represented in the corpus; there are also no examples of *fond of, mad about, dislike, not my type of thing,* or *dreadful.* On the other hand, this text lists a number of sample short exchanges using exactly the type of utterances elicited in the corpus, Type 4-A, [..be + adj..]: *She's a nice person, I think it's a terrific sport,* and *It's so cosmopolitan.* Examination reveals that some structures and lexis in *Functioning in English* agree with the findings in the preference corpus yet others agree with the teachers' predictions but not with the corpus.

Speaking Naturally:

This text specifically lists its common expressions from formal to informal. Although it has no special section on preferences or on likes and dislikes, it does include

many of the utterances in question when dealing with other functions. For example, in the chapter on expressing anger and resolving conflict, it lists: (p.62)

It annoys me when...
I don't like it when...
I can't stand it when...
It burns me up when...
I hate it when...

These are much the same forms which are used for expressing dislike. The text presents Type 3-A items only, two frequent --*like, hate*--, two infrequent --*can't stand, annoy*--, and one which does not appear in the corpus --*burn*.

In another chapter, the books lists, from formal to informal: (p.70)

I would like to compliment you on...
I just love your...
I really like your...
This (cheese) is super.
That's neat.
That's nice.
That's not bad. etc.

With the exception of the first entry, this list of expressions closely resembles the language elicited in the corpus. The authors have used the structure Type 4 which is the single most common structure elicited in the entire corpus and Type 3 which is the second most common structure in the corpus. However, concerning the formality of the utterances, the text indicates that as one reads down the list, the expressions become less and less formal. That means that *I just love your...* is more formal than *That's*

nice, but according to the preference corpus, this is simply not so; all speakers used both forms in all contexts. The discrepancy here may be a result of the difference in functions, expressing compliments versus expressing preference or like, but there is no obvious reason why this should influence the grammatical forms used.

As far as complimenting goes, the language does indeed seem quite "*Natural*"; but for other functions, such as expressing anger, the language seems to be quite at odds with what is actually produced. Again, it appears that the authors only partially agree with the corpus.

Expressways:

Expressways is a series of six texts intended "to provide dynamic communicative practice that involves students in lively interactions based on the content of real-life contexts and situations" (p.vii). The authors state that the books are conversational and functional, covering a broad variety of contexts. They also point out that (p.vii)

> While some texts make a point of giving students a range of ways of expressing a function, from extremely polite to very impolite, we have chosen to take the middle ground and concentrate on those expressions that would most frequently occur in normal polite conversation between people in various settings. *Expressways* does offer a variety of registers, from the formal language someone might use in a job interview, with a customer, or when speaking with an authority, to the informal language

someone would use when talking with family members, co-workers, or friends.

In addition, *Expressways* presents preferences, likes and dislikes, satisfaction/dissatisfaction, complimenting, and want/desire as functions and clearly marks in the *Summary of Functions* each instance when these items occur. Starting with 1-A, the first book in the series, and continuing to 3-B, the last book in the series:

Texts 1-A and 1-B:	
Expressing Likes:	I like to run. I like to run a lot. I like comedies.
Expressing dislikes:	I don't like to run. I don't like comedies very much.
Complimenting:	This cake is delicious. It's excellent. They're excellent.
Preference:	Would you prefer rice or a baked potato? I'd prefer a baked potato.
Texts 2-A and 2-B:	
Preference:	What would you prefer to do? What would you rather do? I'd prefer... I'd rather have... I'd like to.... I'd rather not...
Satisfaction/ Dissatisfaction:	The _____is/are very good/excellent wonderful/superb/perfect/out of this world/ very tasty/ delicious/ etc..
Texts 3-A and 3-B:	
Likes/dislikes:	Don't you like...? It's very nice.
Preference:	I'd rather not... I'd prefer not to...
Likes/dislikes:	I like....

	I love.... I don't like/enjoy ...very much. I don't (particularly) care for... I'm not (really) crazy about... I hate....
Complimenting:	That was a very good/quite a/some ... I thought it was excellent/ wonderful/ terrific/magnificent/fabulous/ superb. It's one of the ...er....s I've ever ...ed.

ETC.

In every case in which the authors believe preference should be expressed, they include the expressions which they have chosen for the functions; in every case, those structures are Types 1 and 2, *prefer* and *rather*. In presenting Type 1 for expressing preference, the text agrees with the teachers' predictions for Category A.

The authors suggest that in order to express preference, the only structures which can be used are prefer and rather; furthermore, to express like/dislike, one can only employ Type 3 items, *like/love/ enjoy/hate*. They do not consider the enormous overlap from one function to another. Under the function of complimenting, the text includes the Type 4, [...be + adj...] structure with good, wonderful, and other adjectives. Type 4 was found to be the most common structure in the corpus, yet not one of the utterances of this structure was a compliment; all were used

to indicate like/dislike and preference. The structural forms appear to resemble those found in the corpus, but their functions, the uses to which they are put, do not correspond. In the corpus, Types 1 and 2, *prefer* and *rather,* were used for the purpose of expressing preference but far less frequently than *Expressways* would suggest and less frequently than other Types which are not mentioned here.

Some of the forms suggested in the text do not appear in the corpus --*fabulous, superb, prefer not to*-- but as the book is American, while the speakers were all Canadian, and as the levels of formality may vary from the casual conversation of the corpus, these discrepancies may not be significant.

How similar, then, are the *Expressways* presentations to the findings in the corpus? For preference, *Expressways* agrees with the teachers' predictions on the use of Type 1 items. But with regard to other language functions, *Expressways* presents materials which resemble the findings of the corpus: use of Types 3-A and 4-A with Frequent lexical items. As with the other texts, Expressways appears to partly agree with the teachers' predictions and to partly agree with the corpus.

3.4 SUMMARY OF RESULTS

1. The most common forms used to express preference and likes were Types 3-A and 4-A, (*I like it,* and *It is good*), but the teachers and textbooks predicted use of Type 1-A/B, 4-C/S, and 5-A (*prefer X, prefer X to Y, X is better, X is the best,* modals). What was most frequently predicted is not what was most frequently used.

2. Teachers predicted great variation between Categories; for the most part, those variations did not occur in the corpus, and where variations did occur, the teachers did not predict them accurately.

3. The sociolinguistic variations which did occur in the corpus were not predicted or mentioned by either the teachers or the textbooks.

4. The teachers predicted consistently that the language elicited in the corpus would closely resemble their own speech, which in turn is what they would teach; i.e. they made the same predictions for Parts I, II, and III of the prediction questionnaire. The textbooks generally agreed with the teachers on preference forms predicted. That is, the teachers and the textbooks agreed, but the corpus did not.

CHAPTER IV

DISCUSSION

4.1 HYPOTHESES

There were two working hypotheses on which this research was based: the first hypothesis was that the language predicted and taught by the teachers and included in textbooks would differ from the actual language used by native speakers to express preference; i.e. it was expected that given a particular language activity in a particular situation teachers (and textbooks) would not accurately predict what people actually said. So, the first question was *Are students being introduced to what is used?* And the expected answer was *No, not entirely.*

The second hypothesis was that the language elicited in the sample would be reducible to a set of formulas of the Manes and Wolfson (1983) style. The question was *Can operational definitions for formulas be devised, and can the language in the corpus be reduced to a series of formulas according to those definitions?* The expected answer was *Yes.*

4.2 FINDINGS

The main findings from the data were with regard to predictions, formulas, and social class variations.

PREDICTIONS

There is a mismatch between teachers' predictions and textbook offerings and actual speech.

Contrary to prediction, there was infrequent use of prefer, would rather, like better, is better (comparatives), and is best (superlatives). In addition, the predicted complete forms --B-Types-- of comparatives, I like X better than Y, I prefer X to Y, X isn't as interesting as Y, etc. did not appear with frequency in any of the Categories. Moreover, Renouf's (1986) predictions for favourites --my favourite x is, I like x best, I prefer x, with full syntax-- did not occur. Also, modals were not Frequent in any Category.

Unpredicted by the teachers and the textbooks, what the speakers most commonly used to express preference was simple sentences with NP + like + NP (Type 3) and NP + be + adj (Type 4) with Frequent verbs and adjectives; in fact, using only thirteen items, the Frequent Type 3 verbs --like, love, want, hate, mind-- and the Frequent Type 4 adjectives -- good, nice, great, funny, bad, okay, wonderful, and beautiful-- accounts for 60% of the total corpus. The remaining 40% consists of 245 infrequent items. A typical

example of an interchange in the corpus is:

Speaker # 7-a: Where would you rather be, Vancouver or Calgary?
7-b: Vancouver.
7-a: Why?
7-b: Calgary is so bare and so bland outside the city. Downtown Calgary is okay. There is a lot happening there. It had a lot of outdoor restaurants. I wouldn't like to live there.

In the exchange, Type 2 (rather) is used to ask the initial question, but the answer consists of three Type 4 responses (be + adj) and one Type 3 (like). This style of passage is typical throughout the corpus with Types 1 or 2 (prefer or rather) being used to ask the initial question, but with the other Types being utilized in the answers. As the responses to the questions are generally longer than the questions themselves, and as prefer and rather are generally elided in the answers, the result is a far greater use of other Types than of the predicted prefer and rather. (However, the greater use of prefer and rather for questions may only be an artifact of the questionnaire, yet it might indicate that these two forms should be taught more for questions and less for statements.)

Except for two cases, there were no great effects caused by the Category of question being addressed/ discussed; viz, whether the subjects spoke about one, two, or three or more items did not, for the most part, noticeably alter their use of any of the preference forms.

The situations in which the Category did influence the forms used were: (1) Type 4, more adjectives with *be* were used to describe three or more items than either two items or one item. (2) *Prefer*, Type 1, was used most frequently in Cat.B. describing two items. The predicted use of *prefer* in Cat. A but not in Cat. C proved to be completely erroneous; there were no examples of *prefer* with three or more items but there were examples in Cat. C. with only one item. The predicted frequent use of superlatives in Category A --discussing three or more items-- did not materialize; neither did the predicted use of comparatives with Category B --discussing two items-- occur. Rather, use of both superlatives and comparatives appeared throughout the corpus unaffected by the number of items being discussed.

FORMULAS

Utterances showing preference are reducible to a few basic constructions --a restricted number of structures with a restricted number of lexical items. These basic constructions, which may be formulas, can account for a large percentage of the utterances used in the corpus. As is stated above, thirteen items account for 60% of the corpus, and by our earlier definitions of frequency, utility, and restriction, they may be formulaic.

SOCIAL CLASS VARIATIONS

W-class speakers spoke the most. However, they varied their speech the least of the three groups; they used fewer different adjectives, fewer different verbs, and fewer different structures than did either the M-class or the T-class. The W-class utilized the smallest range of language, both structures and lexis. The M-class and the T-class were comparable in most regards. By our earlier definitions, the W-class speech consisted of 66% formulaic utterances; the M-class and the T-class consisted of only 54% formulaic language. These results imply that the W-class speakers in the corpus relied more heavily on formulas in their speech; obviously, however, no conclusions can be drawn based on such a small corpus.

As the greater W-class use of formulas is an issue beyond the scope of the present study, it will not be dealt with further. But, it is apparent that the implications of greater formulaic use of language at W-class level speech should be considered and researched further. (Perhaps, if an ESL student plans to be a cook, a dishwasher, or a labourer and thus to join the working class, he only needs to concentrate on the language of that group, especially the formulaic utterances. Do we as teachers want our ESL students to speak like W-class natives? On the other hand, if a student intends to be upwardly mobile, he may want to

learn the formulas first then to concentrate on the language forms of the higher socio- economic groups. Or the opposite may be argued; if someone is working-class, perhaps he should be directed to learn non-formulaic middle-class language to aid in his future social and upward mobility. Or, for native speakers, if formulas vary from class to class, and if upward mobility depends on speech, perhaps people wanting to climb the social/sociological ladder could learn the formulaic speech of the next group. All this is pure conjecture at this point. This research was certainly not intended to provide any answers to language differences in social classes in Canada, but it may provide some minor insights and directions for further research.)

INCIDENTAL FINDINGS

In order to seek out other patterns or relationships between factors and speech, several other points were investigated for frequency in the corpus: *because* clauses, *when* clauses, compound sentences, other embedded or complex sentences, sentence fragments, questions using preference forms, teacher talk, profanity, male/female variation in speech, Western versus Eastern Canada speech variation, and age variation. These factors were all flagged in the data base, but no obvious patterns were found. The very fact that no clear distinctions presented themselves is thought-

provoking. For example, this might mean that social class differences influenced the speech of the corpus more than did region, age, or sex. It might prove very enlightening to investigate all the possible relationships between these factors: for example, an older BC male may use entirely different speech than a younger Quebec female; what is formulaic for one may not be for the other. None of the factors showed remarkable results in themselves, but in combination with the other factors, clear trends may appear. The investigation of personal and geographical factors which play a role on Canadian speech would be a very interesting topic for further research, but also beyond the scope of this present study.

4.3 IMPLICATIONS OF INACCURATE PREDICTIONS

The language produced in the corpus may accurately represent Canadian conversational English or it may just be an artifact of the data-gathering procedure. The speakers chosen may not be representative of the general English-speaking population of Canada -- there were only eighteen of them and they do not represent all social classes. Furthermore, the numbers of both predictors and texts were small. Therefore, care should be taken with regards to generalizing the results of the study. Other factors which may limit the generalizability of the results are:

--The discussion topics: Perhaps the language elicited may be used exclusively with the topics discussed. Other topics may elicit quite different language. This is doubtful, however, in light of the similarity of the Manes and Wolfson (1983) findings for compliments.

--The relationships between the speakers: All the speakers were well-known or intimate to each other. More formal relationships may elicit more formal language. More formal language may be more or less formulaic.

--Age: The average age of the speakers was: W-class, 50; M-class, 40; T-class, 36. It is possible that the W-class speech differed because the subjects were older, and that the M-class and the T-class speech closely resembled each other due to closeness in age. However, upon cursory analysis, no other obvious age-related variations were found.

--Class: As was mentioned, there are no speakers from the higher social levels, doctors, professors, or architects. Therefore, generalizations will not extend to these groups.

--Individual Analyses: Individual results were not compared or analyzed, so the findings are not known to be significant.

Despite the above limitations, the results of the study indicate interesting trends. The study underscores

the possibility of the discrepancy between what is actually used and what is considered important to teach. In the study, the teachers stated that their predictions were based on what they themselves would say. In addition, for the most part, it appears that the material presented in ESL textbooks closely resembles the predictions of the teachers. The teachers predicted the language which would be elicited by native speakers given certain tasks; but the native speakers did not conform to the teachers' expectations. Therefore, the disparity lies between the actual language produced and the intuitions of the teachers and the textbook writers.

The teachers doing the predicting stated that they taught the language that they believed they used. They taught their ESL students the language of their predictions. We have seen the discrepancies between the language predicted and taught and the language used. This means that the ESL students are learning or, at least, being taught language that is not used a majority of the time in casual native-speaker speech.

Moreover, the teachers believed that the situations and topics in the questionnaire would elicit certain forms. Supposedly then, the teachers might supply similar contexts --situations and topics-- to their students and expect them to use those predicted forms, as was done in *Expressways*.

The students would be practising language which may be used only very rarely by native speakers. Thus, given similar real-life situations and topics, the ESL students would spontaneously produce these non-native structures. The only ones who spoke according to the teachers' predictions would be the ESL students who had learned those structures in that context. This is not authentic language. The students are learning language forms which do not occur for situations in which that language would not be used.

Three of the teachers who made the predictions were also speakers for the corpus, and they predicted the *prefer X to Y* form, but did not use it. One of the corpus speakers who doubled as a predictor made her predictions ten minutes before her taping session; one made her predictions one day after her taping, and one made her predictions three weeks later. Regardless of when they made their predictions on their own speech production, they did not behave as they believed they would. Their intuitions for their own speech were no more accurate than for the speakers in general.

Presumably, given similar prediction tasks for other language functions, the teachers would predict with similar inaccuracy. Presumably also, teachers are teaching inauthentic language not only for preferences but for other functions as well. Assuming that these leaps of logic carry some validity, studies must be done to ascertain what

exactly native speakers *do* say; teachers might do well to rely less on intuition for their teaching materials and more on empirical studies to more closely approximate authentic native speaker language.

4.4 IMPLICATIONS OF FORMULAS

The findings on formulaic speech in the corpus are not surprising. According to research done on native speaker conversation and speech acts, the importance and frequency of formulas in L1 speech seems evident. For example, Sorhus (1975), from analysis of her own corpus linguistics study, suggests that twenty percent of spoken English may consist of "fixed expressions". Antal (1985) (from his own intuitions) believes that few (if any) utterances are novel or creative in the Chomskyan sense, and that spoken language merely consists of combinations of set formulas, old and new combinations of old and new elements. Pawley and Syder (1983) propose that "Memorized clauses and clause-sequences form a high proportion of the fluent stretches heard in everyday conversation" (p.208) and that "A minority of spoken clauses are entirely novel creations" (p.205). Unfortunately, little of the theory behind formulaic utterances has been based on data; most hypothesizing is intuitive. Thus the need for empirical data gathering and analysis.

If we accept that formulas constitute a significant percentage of spoken English language, the question remains of the role of formulas in English language acquisition. If formulas play any role in any aspect of language acquisition, we should find out what those formulas are. Such evidence does exist. For example, Ruth Clark in her seminal article *Performing Without Competence* (1973) suggested that formulas played an integral role in acquisition in that they developed from purely automatic memorized utterances without internal structural analysis (later called routines) into partly creative, partly memorized wholes or frames with open slots (later called patterns). Her *interface theory* spawned further hypothesizing which led to other viewpoints regarding the role of formulas. Various positions regarding formulas in both L1 and L2 acquisition now exist; evidence for the various theories can be found in the literature:

-the interface position. Routines evolve into patterns which evolve directly into creative language. (Ellis, 1984; Fillmore, 1976; Peters, 1977, 1983; Rescorla and Okuda, 1987). Both Peters and Ellis provide strong arguments for the importance of formulas in language acquisition.

-the non-interface position. Routines evolve into patterns, but creative construction develops independently.

This theory is supported by research with children in both L1 (Brown, 1973; Nelson, 1975; Dore, 1974) and L2 acquisition (Hatch, 1972; Hakuta, 1974; Wagner-Gough, 1975) and by research with adults in L2 acquisition (Hanania and Gradman, 1977). However, Ellis (1984) states:

> ...irrespective of whether 'familiar sentences' evolve into rules for performing 'novel sentences', formulaic speech is in itself an important aspect of communicative competence (p.80).

If formulas do, in fact, exist for various language functions, and if these formulas could be empirically discovered and gathered together, the resulting body of language might greatly facilitate both L2 learning and L2 teaching. How might knowledge of formulas assist in L2 acquisition/learning? Research suggests that effective use of formulas might lead to an increase in L2 input and thus to an increase in L2 acquisition. Krashen suggests that the ability to manipulate L2 formulas appropriately may result in more fluent speech and thus greater conversational skills. The native speaker, therefore, may respond better by feeling more comfortable and so talking more; the final result may be greater input and greater eventual learning (Krashen's acquisition) for the L2 learner. (Krashen and Scarcella, 1978; Kirkland, 1984). Other research supports the theory that use of formulas leads to creative speech (Clark, 1973). In addition, formulaic use of language would

resemble child L1 acquisition and thus be a more natural progression (Hakuta, 1974; Hatch, 1972) and would assist in information-holding capacity, processing time, and memory limitations (Peters, 1983).

How might knowledge of formulas assist in L2 teaching? Research suggests that providing simple common utterances - formulas - from the earliest stages of learning might increase motivation in learners (Gatbonton and Segalowitz, 1988). (Fries and Lado (1943) based their entire text, *English Pattern Practices,* on the use of patterns --slots and frames-- intended to lead to habit formation and thus automatization of new forms.) Moreover, knowing the language actually used by native speakers would aid materials writers and syllabus designers by streamlining materials necessary to be taught and would aid teachers by advising them in what to teach, what to focus on, and what to present when, in order to more closely approximate native speaker language. The fact that the Manes and Wolfson (1983) Compliment study found virtually the same result for formulas as the present study suggests that the formulas cross the boundaries of function in language. It could be that various language functions rely on a very small set of common formulaic structures --slots and frames or patterns-- with only the lexis function-dependent. If this is true, an enormous number of functions, contexts, situations, and

topics, an enormous amount of language, could be controlled with only a small set of formulas. Such cross-over would certainly raise the L2 learner's fluency and thus his conversational skills.

In addition, knowing the situations in which particular language would be elicited might provide the teacher with more authentic communicative tools and techniques for language teaching. Cathcart (1989) states that "simulated excerpts may serve to mislead students about the nature of everyday interactions" (p. 105). Teachers would not need to waste time dreaming up situations in which certain expressions might be used; they could have a bank of empirically validated and authentic situations/contexts on hand. Nor would teachers need to spend time concentrating on such utterances as *I would like to compliment you on your hat* or *I prefer living in the country to living in the city,* if patterns such as *I like your hat, it's nice* and *I like the city, it's nice* would suffice.

4.5 CONCLUSION

This study has attempted to answer two questions: *Do ESL teachers relying on their intuitions accurately predict and therefore teach authentic spoken language?* and *Can this spoken language be reduced to formulas?*

What was found first was that neither the teachers nor the textbooks accurately predicted the language which was used. By extension, it was suggested that teachers' and textbook writers' intuitions are probably faulty for other functions as well. Therefore, ESL students may be studying/learning language which is inauthentic in that it is either contextually inappropriate or simply unused by native speakers.

Second, formulas were investigated and isolated in the corpus. The formulaic utterances were shown to account for approximately 60% of the language used to express preference -likes and dislikes. However, none of the formulaic utterances were those predicted. Clearly then, what is being taught is not what is being used.

Because similarities were found between the formulas for Preferences and those for Compliments, it was hypothesized that a common core of formulas might exist which would extend to other functions as well.

Knowledge of formulaic language would facilitate teaching, textbook writing, and learning, so further research is necessary in order to isolate the patterns that make up the formulas of our language. More research is necessary to find the contexts in which those formulas are elicited from native speakers. We as teachers should not rely so heavily on our intuitions, which are necessarily contaminated by our training in prescriptive grammar and our familiarity with academic, often written language. We should teach most frequent and useful language, but first we require empirical research to discover what it is.

REFERENCES

Antal, L. (1985). Is It Really True That Nearly All Sentences Are Novel Ones? *Word*, 36, 3: 195-202.

Blishen, B. (1958). The Construction and Use of an Occupational Class Scale. *Canadian Journal of Economics and Political Science*, 24: 519-531.

Blishen, B. (1967). A socio-ecomonic index for occupations in Canada. *The Canadian Review of Sociology and Anthropolgy*, 4: 41-53.

Blishen, B. (1976). A revised socioeconomic index for occupations in Canada. *The Canadian Review of Sociology and Anthropology*, 13: 71-79.

Borkin, A. and Reinhart, S. (1978). "Excuse Me" and "I'm Sorry". *TESOL Quarterly*, 12, 1: 57-70.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge: Harvard University Press.

Cathcart, R. (1989). Authentic Discourse and the Survival English Curriculum. *TESOL Quarterly*, 23, 1: 105-126.

Clark, R. (1974). Performing Without Competence. *Journal of* Child Language, 1: 1-10.

Coulmas, F. (1981). *Conversational Routine*. The Hague: Mouton Publishers.

D'Amico-Reisner, L. (1983). An Analysis of the Surface Structure of Disapproval Exchanges. In Wolfson, N. and Judd, E. (Eds.) *Sociolinguistics and Language Acquisition*. Rowley, Mass.: Newbury House.

Dore, J. (1974). A Pragmatic Description of Early Language Development. *Journal of Psycholinguistic Research*, 4: 423-430.

Drazdauskiene, M-L. (1981). On Stereotypes in Conversation. Their Meaning and Significance. In Coulmas, F. (Ed.) *Conversational Routine*. The Hague: Mouton Publishers.

Dulay, H., Burt, M. and Krashen, S. (1982). *Language Two*. New York: Oxford University Press.

Ellis, R. (1984). Special Constructions in the Classroom. In Ellis, R. (Ed.) *Classroom Second Language Development*. Oxford: Pergamon Press.

Ervin-Tripp, S. (1969). Sociolinguistics. In Berkowitz, L. (Ed.) *Advances in Experimental Social Psychology*, Vol. 4. New York: Academic Press.

Ervin-Tripp, S. (1976). "Is Sybil There?" *The Structure of* American English Directives. Language in Society, 5: 27-76.

Ferguson, C. (1981). The Structure and Use of Politeness Formulas. In Coulmas, F. (Ed.) *Conversational Routine. The Hague: Mouton Publishers.*

Fillmore, L.W. (1976). The Second Time Around: Cognitive and Social Strategies in Language Acquisition. Ph.D. Dissertation, Stanford University.

Forcese, D. (1975). *The Canadian Class Structure.* Toronto: McGraw-Hill.

Fries, C.C. and Lado, R. (1943). *English Pattern Practices.* Ann Arbor: University of Michigan Press.

Gatbonton, E. and Segalowitz, N. (1988). Creative Automatization: Principles for Promoting Fluency Within a Communicative Framework. *TESOL Quarterly*, 22, 3:473-492.

Hakuta, K. (1974). Prefabricated Patterns and the Emergence of Structure in Second Language Acquisiton. *Language Learning*, 24, 2: 287-298.

Hanania, E. and Gradman, H. (1977). Acquisition of English Structures: A Case Study of an Adult Native Speaker of Arabic in an English Speaking Environment. *Language Learning*, 27, 1: 75-91.

Hatch, E. (1972). Some Studies in Language Learning. *UCLA* Workpapers in Teaching English as a Second Language, 6: 29-36.

Holmes, J. (1983). The Structure of Teacher's Directives. In Richards, J. and Schmidt, R. (Eds.) *Language and Communication.* London: Longman.

Keller, E. (1981). Gambits. Conversational Strategy Signals. In Coulmas, F. (Ed.) *Conversational Routine.* The Hague: Mouton Publishers.

Kirkland, B. (1984). Talking Faster Sooner. *TESL Talk,* 15 (1,2): 68-72.

Krashen, S. and Scarcella, R. (1978). On Routines and Patterns in Language Acquisition and Performance. *Language Learning,* 28, 2: 283-300.

Manes, J. and Wolfson, N. (1981). The Compliment Formula. In Coulmas, F. (Ed.) *Conversational Routine.* The Hague: Mouton Publishers.

Mendelsohn, D., Laufer, R. and Seskus, J. (1984). *Functioning in English.* Toronto: Dominie Press.

Nelson, K. (1981). Individual difference in language development: Implications for development and language. *Developmental Psychology,* 217: 170-187.

Ocke-Schwen Bohn (1986). Formulas, frame structures, and stereotypes in early syntactic development: some new evidence from L2 acquisition. *Linguistics,* 24: 185-202.

Olshtain, E. and Cohen, A. (1983). Apology: A Speech-Act Set. In Wolfson, N. and Judd, E. (Eds.) *Sociolinguistics and Language Acquisition.* Rowley, Mass.: Newbury House.

Pawley, A. and Syder, F. (1983). Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency. In Richards, J. and Schmidt, R. (Eds.) *Language and Communication.* London: Longman.

Peters, A. (1977). Language Learning Strategies. *Language,* 53, 3: 560-573.

Peters, A. (1983). *The Units of Language Acquisition.* New York: Cambridge University Press.

Pomerantz, A. (1978). Compliment Responses: notes on the cooperation of multiple constraints. In Schenkein, J. (Ed.) *Studies in the organization of conversational interaction.* New York: Academic Press.

Porter, J. (1965). *The Vertical Mosaic.* Toronto: University of Toronto Press.

Renouf, A. (1986). The Elicitation of Spoken English. In Tottie, G. and I. Backlund (Eds.) *English in Speech and Writing*. Uppsala: Almqvist and Wiksell.

Rescorla, L. and Okuda, S. (1987). Modular Patterns in Second Language Acquisition. *Applied Psycholinguistics*, 8: 281-308.

Schlegloff, E. (1968). Sequencing in Conversational Openings. *American Anthropologist*, 70: 1075-1095.

Sorhus, H. (1975). To Hear Ourselves--Implications for Teaching English as a Second Language. *English Language Teaching Journal*, 30, 1: 211-221.

Svartvik, J. and Quirk, R. (Eds.) (1979). *A Corpus of* English Conversation. Stockholm: Liberlaromedel Lund.

Tillit, B. and Bruder, M. (1985). *Speaking Naturally*. Cambridge: Cambridge University Press.

Verschueren, J. (1981). The Semantics of Forgotten Routines. In Coulmas, F. (Ed.) *Conversational Routine*. The Hague: Mouton Publishers.

Vihman, M. (1982). Formulas in first and second language acquisition. *Stanford Papers and Reports on Child Language Development*, 18: 75-92.

Wagner-Gough, J. (1975). Comparative Studies in Second Language Learning. MA Thesis. TESL Department UCLA.

Wolfson, N. (1983). An Empirically Based Analysis of Complimenting in American English. In Wolfson, N. and Judd, E. (Eds.) *Sociolinguistics and Language Acquisition*. Rowley, Mass.: Newbury House.

Wolfson, N., D'Amico-Reisner, L. and Huber, L. (1983). How to Arrange for Social Commitments in American English: The Invitation. In Wolfson, N. and Judd, E. (Eds.) *Sociolinguistics and Language Acquisition*. Rowley, Mass.: Newbury House.

Wolfson, N. and Manes, J. (1979). "Don't Dear Me". *Working* Papers in Sociolinguistics, 53.

Yorio, C. (1980). Conventionalized Language Forms and the Development of Communicative Competence. *TESOL Quarterly*, 19, 4: 432-443.

Appendix 1 **Corpus subjects: Personal Information.**

W-Class: Seven working-class native English speakers with the following occupations and educations:

1. -female, waitress, age 55?, retired, completed grade eight, born in Nova Scotia, monolingual.
2. -female, photographer (Blishen's class five), age 37, grade twelve, born in B.C., monolingual.
3. -male, travelling salesman, unemployed at the time of the sample, age 44, grade eleven, born in Ottawa, fluent French.
4. -female, nurse, age 48, grade twelve plus three years nurse's training, no property ownership, no university education, born in Saskatchewan.
5. -female, nurse, age 38, grade twelve plus two years nurse's training, working one day per week at the time of the sample, born in B.C. monolingual.
6. -male, bus driver, age 65, retired, grade eleven, born in Alberta, monolingual.
7. -female, housewife, age 60, wife of bus driver, grade twelve, no university, born in B.C., monolingual.

M-class: Six middle-class native-speakers with the following occupations and education:

1. -female, graphic artist, age 26, B.A. in Fine Arts, born in Montreal, speaks some French.
2. -male, writer, age 33, B.A. in Creative Writing, born in Montreal, speaks some French.
3. -male, publishing company owner, age 65, retired, two years university, born in Quebec, speaks fluent French.:
4. -female, homemaker, age 40, wife of businessman/ manager with a salary of $65,000.00 per year. Graduated from high-school, owned home plus summer home, born in Montreal, speaks a little French:
5. -female, Post Office clerk, age 38, eight years university, owned two homes, currently enrolled in M.F.A., born in Ottawa, speaks a little German and French.
6. -male, electrician, age 38, two years university, owned home, $40,000.00+ salary, born in B.C., monolingual.

T-class subjects participating in the corpus: The first two ESL teachers were only subjects for the corpus; the next three were both corpus subjects and predictors.

1. -female, age 30, M.A. APLI in progress, taught 5 years ESL, born in England, 14 years in Canada, (no discernible non-Canadian accent), speaks fluent French, Spanish, and Hebrew.
2. -female, age 35, M.A. APLI in progress, 14 years ESL, born in Calgary, fluent French.
3. -female, age 43, M.A. APLI, taught 7 years public school, taught two years ESL with adults, born in Montreal, speaks a little Fench and a little Spanish.
4. -female, age 38, M.A. APLI, taught 15 years ESL, born in Montreal, speaks fluent French and German.
5. -female, age 31, M.A. APLI in progress, taught 3 years ESL, born in Ontario, speaks a little French and Spanish.

Appendix 2

T-Class Subjects Participating in the Predictions. Personal Information.

6. female, age 42, M.A. APLI. in progress, born in Quebec, 15 years ESL.
7. female, age 40, M.A. APLI in progress, 11 years ESL, born in Montreal, some French
8. male, age 45, M.A. APLI in progress, 21 years ESL, born in Montreal, fluent in French.
9. female, age 45, M.A. APLI in progress, 9 years ESL, born in Ontario, some French.
10. female, age 38, M.A. APLI., 8 years ESL, born in Ontario, fluent French.
11. male, age 37, M.A. APLI in progress, born in B.C., monolingual,
12. male, age 37, M.A. APLI. in progress, born in New Brunswick, 11 years ESL, some French.

Appendix 3

Instructions for Corpus Speakers.

(Read by the researcher to the speakers before they began their recording.)

You will see a list of questions in front of you. Please read them, and if the questions are interesting and easy to talk about, do so. If you have nothing to say about a particular question, don't worry about it, just go on to the next question.

Don't turn off the machine even if you think you have gotten off the topic or if you think what you have said is not correct.

Just talk. Agree or disagree or whatever. If you have no comment, say so. If you have a lot to say, say a lot.

Appendix 4

Database Transcriptions:

Every utterance was categorized and labeled in seven ways and some were labeled in an extra eighth. Each label is known as a data field in the database.

1. Speaker. The interviews were numbered in order of occurrence, and each speaker was lettered either A or B. Therefore, the first field in the datafield is the interview number followed by the speaker number. As there were a total of eighteen speakers in the corpus, there are also eighteen possible entries in this field: numbers one to nine followed by letters A or B. Labelling utterances by speaker enables the researcher to discern not only exactly who said any particular entry in the datafield but also to collect all phrases spoken by any one speaker. Thus individual speakers can be compared to each other.

2. Category. Each utterance is labeled either Category A, B, or C depending upon which sheet of the questions was being discussed at the time. Questions with three or more items to discuss occur in category A, questions with two items occur in category B, and single items occur in C. Labeling categories allows the researcher to distinguish which types of utterances and which semantically positive or negative words were produced by which type of question. Lexical items and structures can then be checked for

frequency in general as well as for utility in a number of different situations.

3. Sex. Each utterance is labeled either m or f depending upon the sex of the speaker. Labelling sex facilitates gender discrimination in the utterance analysis. The information on gender is already included in the Speaker number (1-a is already known to be woman, 9-a is known to be a man) according to the personal history and records, but adding an separate field for gender greatly facilitates the data analysis in retrieving files thus saving computer time.

4. Socio-economic-linguistic level. Each utterance is labeled one of m, w. or t: m for middle-class speakers, w for working-class speakers, or t for teachers. Consequently, the language produced by the three groups can be compared in a number of contexts and for a number of points: syntactic, semantic, lexical.

5. Keyword. Each utterance is labeled by the word which most clearly determines its positive or negative connotation, what Manes and Wolfson have termed "the positive or negative semantic load" (1983, p.116). This may be a verb such as *like, love, adore, hate,* a noun such as *fan, favourite, lover, slime, hypocrite,* an adjective such as *good, bad, wonderful,* or others. In cases where utterances contain more than one keyword, the most significant word is chosen to be key and the entire utterance is marked in a later category, conjunction. In this way, common utterances with a single keyword can be recorded simply, but less common multiple keyword utterances can be recorded with the same system yet easily distinguished within the later category. Labeling the keyword allows access to the vocabulary used. The database can be commanded to list all utterances using a certain keyword, so analysis of utterances and their frequency is simplified. In addition, other field restraints may be invoked at the same time. For example, the database can be told to list all uses of prefer in category B by working-class speakers. This result may then be compared to the same data for middle-class speakers.

6. Structure. In order to facilitate retrieval of an otherwise unwieldy number of utterances in the database, each utterance is labeled by its syntactic structure. Items stored in this way can then be recalled by structural pattern as well as by other fields in the data system. The structure labeling system is as follows: (all examples are taken from the data; e.g. (1-a) indicates interview number one, speaker A.

[np] = all nouns or noun phrases (except it, that, this, these, those and they) whether in subject, object, or complement position. (3-b) *I hate it* is then represented by [np + hate + np] but (5-a) *That sounds good* is [it + be + adj].

[it] = all examples of subject pronouns it, that, these, those, and they as subject of the pertinent verb in the sentence. Otherwise the subject will be [np]. It is not distinguised as object of the verb, only as subject. Thus, sentences such as (3-b) *It was great* are marked [it + be + adj] but *I loved it* would be [np + love + np].

[vp] = all verbs or verb phrases which do not occur as the main verb in the sentence, eg., (4-a) *I would like to go and buy a nice book on St. Lawrence Main*....is [np + m like + vp {to go and buy} + adj + np + pp]

[be] = all occurrences of the actual verb be plus any occurrences of copular or linking verbs *look, feel, seem, appear, sound, taste, smell.* (1-b) *It looks disgusting,* (2-a) *It was enticing,* (1-b) *Sharks taste great,* and (2-b) *It sounds gory,* would all be classified as [it or np + be + adj].

[----] = all verbs other than *be* verbs (above) occurring as main verb in the sentence are labeled as such. Thus (2-a) *I hate dentists* would appear as [np + hate + np]. *It sucks* [it + suck]. All verbs are cited in the present tense.

[op] = sentence openers or gambits including opinion markers such as *I think..., I believe..., I'm sure....* Consequently, (3-a) *I think it's great* would be [op + it + be + adj].

[m] = all modal or auxiliary verbs . May only occur before the main verb or the negative marker [n].

[n] = all adverbial negative markers such as not. May only occur before the main verb and/or after the auxiliary marker.

[adj] = all instances of semantically positive or negative adjectives whether preceeding a [np] or in complement position following [be].

[adj2] = as with adjectives [adj] but all instances of comparative adjectives when semantically positive or negative, eg., *bigger, better, more.*

[adj3] = all instances of superlative adjectives, eg., *the best, the most, least.*

[adj2/]= all instances of comparative adjectives followed by *than* and a second noun or verb phrase. For example: *Montreal is more cosmopolitan than Toronto* would be transcribed as [np + be + adj2/ + np]

[int] = all instances of intensifiers of adjectives, nouns, or verbs. ([int] could be formally labeled adverbs) For example: *This coat is too heavy* or *That car really moves* or *The movie was totally stupid. Too, really,* and *totally* would be classified as intensifiers and labeled [int].

[S] = relative clauses, noun clauses, or embedded sentences of other types which do not contain the connotative/sematically loaded words. For example: (1-a) *The place that we went with Samantha for her birthday was an excellent place* would be transcribed [np {the place} + S {that we went with Samantha for her birthday} + be {was} + adj {an excellent} + np {place}]

[adv] = all instances of semantically positve or negative adverbs or adverbial phrases would have been labeled [adv], but in fact none occurred.

[pp] = prepositional phrase. Mentioned only where relevant.

7. Sentence. The actual sentence is written in its entirety in this field. In cases where the utterance is unfinished, the written words are followed by [...].

8. Conjunction. In this field, sentences are marked for additional specific grammatical, semantic, or lexical considerations.

-both: All of the utterances which contain both the x and the y element are marked both; called B-Types later.

-comp: all the utterances which use any form of comparative or superlative but not containing both the x and y elements are marked comp (comparative). As a result of marking the comparative, superlative, and both features, the file of sentences or utterances containing these features can instantly be retrieved and referred to. Rather than having to search through the entire file to locate and count the number of superlatives used, this file may simply be opened and examined.

-odd: In sentences with appositives or transformed subjects or objects, as in the following *Montreal, it's nice. Toronto, it's nice, too. Icecream, I like,* which are structurally described (in field 6) as [np, np + be + adj]

or [np, np + like] the conjunction field label employed is *odd*. Another use of the odd marking is in cases where more than one keyword could be chosen for an individual utterance. In, for example, (6-a) *I hate the asshole* either *hate* or *asshole* could be said to indicate dislike or non-preference. Rather than inserting the utterance into the databank several times and thus skewing the data, it is simply marked as odd and considered in a separate file.

-frag: Another marker in the conjunction field is *frag* which indicates a sentence fragment. Frag is only utilized when the utterance, if completed, would clearly have expressed preference. For example, (6-b) *I don't like the...I don't know.*

-question: The last use of the conjunction field was to indicate preference forms in questions, thus the marker *question*. Any question asking for the partner's preference was recorded and marked here.

Appendix 5

TEACHERS' PREDICTIONS:

Name: _____
Interview by phone: _____ in person: _____
Other: _____.
Age: _____
Sex: _____
Education: _____
Years of Teaching: _____
Place of Birth: _____
Length of time
in Montreal: _____

##########

PART I.

If the following questions were given to native speakers of Canadian English, what language - verbs, nouns, adjectives, structures, notions, functions, etc - do you think would most likely be elicited? What language would a native speaker use? What would he/she say?

Question A.

1. On the following page you will see a list of movies currently playing in Montreal. Choose one. Then suggest to your partner that you go to that particular movie and tell him/her why you chose it. If your partner chooses another movie, try to convince him/her to go to yours.

2. You may go to either Halifax or Vancouver for a short holiday, or you may stay home. Choose one. Why? Convince your partner.

Predictions for question A: _____

_____

_____

_____

Question B. Compare and contrast the following: How do you feel about them?

1. Bill 101----------Bill 178.

2. Canada----------The United States

Predictions for question B: _____

_____

_____

_____

Question C. How do you feel about the following?

1. Visits to the dentist.
2. Ice cream.

Predictions for question C:_____

_____

_____

##########

PART II.

If you wanted to tell someone that you believed one thing superior to another, what forms of English would you most likely use? Would you say:

I prefer x to y.
I like x better than y.
I would rather x than y.
I love x, but I hate y.
X is good. Y is bad.
X is good. Y is not as good as x.
X is better than y.
While (although) fond of x, I feel less than admiration for y.
I adore x. I detest y.

Or something else entirely?

_____

_____

_____

##########

PART III. What would you teach?

_____

_____

##########

PART IV. Trick question/ Which do you prefer?

_____

Appendix 6

Corpus Question Sheet: Montreal

I. Discuss the following.

1. On the following page you will see a list of movies currently playing in Montreal. Choose one. Then suggest to your partner that you go to that particular movie and tell him/her why you chose it. If your partner chooses another movie, try to convince him/her to go to yours.

2. Look at the sheet of restaurants. You are hungry; choose a restaurant and discuss your choice with your partner. You will go out for dinner with your partner.

3. Look at the listings for *Just For Laughs* and the film festival. Choose one. Discuss it with your partner. Choose time, date, etc. You will attend the show with your partner.

4. Look at the sheet of resorts. Choose one for a vacation. You will go for a long holiday with your partner. Discuss.

5. You have received a job-offer somewhere in the world. Choose a place and a job. Your partner and you will go to wherever the job is. Discuss with partner. You may go anywhere or do any job.

6. You may go to either Halifax or Vancouver for a short holiday, or you may stay home. Choose one.

7. You may have either lobster, meat-loaf, or something else for dinner. Choose one.

8. You may spend the day shopping, on the beach, or at home. Choose one.

9. You may watch a porno movie, a romance, or something else. Choose one.

10.You may do anything you want to do right now. Discuss anything you feel strongly about.

II. Comparison Questions.

Compare and contrast the following.
How do you feel about them?
Please discuss with your partner.

1. Bourassa----------Parizeau.
2. Bill 101----------Bill 178.
3. Toronto----------Montreal.
4. Westmount----------East Montreal.
5. Mulroney----------Trudeau.
6. poutine----------french fries.
7. Canada----------The United States
8. big cities----------small towns.
9. Italian food----------Chinese food.
10. English----------French.

III.

Opinion Questions
How do you feel about the following?
Discuss with your partner.

1. Visits to the dentist.
2. Chocolate cake.
3. Ice cream.
4. The new federal tax.
5. Montreal winter.
6. Montreal summer.
7. La Ronde.
8. Abortion.
9. Suntanning.
10. Claude Ryan.

Corpus Question Sheet: British Columbia

I. Discuss the following.

1. On the following page you will see a list of movies and videos. Choose one. Then suggest to your partner that you see that particular movie and tell him/her why you chose it. If your partner chooses a different movie, try to convince him/her to see yours.

2. Look at the sheet of restaurants. You are hungry; choose a restaurant and discuss your choice with your partner. You will go out for dinner with your partner.

3. Look at the sheet of resorts. Choose one for a vacation. You will go for a long holiday with your partner. Discuss.

4. You have received a job-offer somewhere in the world. Choose a place and a job. Your partner and you will go to wherever the job is. Discuss with partner. You may go anywhere or do any job.

5. Look at the list of cars. Choose one. Why did you chose it?

6. You may go to either Halifax or Vancouver for a short holiday, or you may stay home. Choose one.

7. You may have either lobster, meat-loaf or something else for dinner. Choose one.

8. You may spend the day shopping, on the beach, or at home. Choose one.

9. You may watch a porno movie, a romance, or something else. Choose one.

10.You may do anything you want to do right now. Discuss anything you feel strongly about.

II. Comparison Questions.

Compare and contrast the following.
How do you feel about them?
Please discuss with your partner.

1. Vander Zalm----------Barret.

2. _____----------_____

3. Vancouver----------Calgary.

4. Uplands----------Esquimalt (for Victoria)
   Shaugnessy----------East Van. (for Vancouver)
5. Mulroney----------Trudeau.

6. muffins----------croissants

7. Canada----------The United States

8. big cities----------small towns.

9. Italian food----------Japanese food.

10. English----------French.

III.

Opinion Questions.

How do you feel about the following?
Discuss with your partner.

1. Visits to the dentist.
2. Chocolate cake.
3. Ice cream.
4. The new federal tax.
5. Winter.
6. Summer.
7. The PNE.
8. Abortion.
9. Suntanning.
10. The Minister of Education.

Appendix 7

PERSONAL INFORMATION SHEET: CORPUS W-CLASS AND M-CLASS

1. Interview number: _____
2. Age: _____ (scale)
3. Sex: _____
4. Education: _____ (scale)
5. Type of employ: _____
6. Place of Birth: _____
7. Length of time in Montreal: _____ (scale)
8. Relationship with partner: _____
   (spouse, friend, acquaintance, etc.)
9. Other languages, language history: _____
10. Salary: _____ (scale)
11. Other: _____

PERSONAL INFORMATION SHEET: TEACHERS

1. Interview number: _____

2. Age: _____ (scale)

3. Sex: _____

4. Education: _____ (scale)

5. Employment history: _____

   Years of teaching ESL: _____ (scale)

6. Place of Birth: _____

7. Length of time in Montreal: _____ (scale)

8. Relationship with partner: _____
(professional, spouse, friend, acquaintance, stranger)

9. Other languages, language history: _____

10. Salary: _____

11. Other: _____

Appendix 8

Lists of Items Used and Their Frequencies

8-A ADJECTIVES: Infrequent

Adjective:	# of uses:	Adjective:	# of uses:	Adjective:	# of uses:
all right	3			amazing	2
anxious	1			asinine	1
awful	10			bare	1
big	2			bland	1
boring	5	bloody	1	cheap	4
civilized	1			clean	2
clever	1			colourful	1
courteous	1			crafty	1
cute	2			delicious	2
dirty	3			disgusting	4
dreadful	1			droopy	1
dull	4			enjoyable	1
enticing	1			evil	1
excellent	5			expedient	1
expensive	8			explicit	2
fantastic	4	fattening	1	filthy	1
fine	7			flaky	1
fluffy	1			friendly	1
				glad	2
gorgeous	4			gory	1
greasy	1			gross	7
gruesome	1			handsome	1
happening	1			hard	3
healthy	4			higher	1
hilarious	3			horrible	5
hot	2			homely	1
ignorant	1			impressed	1
impressive	1			inspiring	1
intelligent	2	intimate	1	interesting	6
keen	2			killer	1
lacking	1			larger	2
least	1			long	3
lousy	1			loveable	1
magnificent	1			marvellous	1
miserable	1			more	5
mushy	1			narrow-minded	1
neat	9			nervous	1
guilty	1			crazy	4
happy	1			right	3
smart	1	serious	1	stimulating	1
nuts	1			over-emphasized	1
over-rated	1			painless	1

painful	1	perfect	1
pretty	2	prudent	1
questionable	1	reasonable	4
refreshing	1	repulsive	1
respected	1	revolted	1
revolting	1	ridiculous	2
romantic	1	rotten	1
rough	1	rude	1
sad	1	scary	3
scungy	2	sell-out	2
sensuous	1	shitty	2
shmucky	1	sick	2
slimy	1	splendid	1
stuck-up	1	stuffy	1
stupid	5	sunny	1
super	4	superb	3
surreptitious	1	swelled	1
tasteless	2	terrible	7
terrific	2	thrilling	1
tight-ass	1	tired	3
tiresome	1	trapped	1
trivial	1	uncivilized	2
unfair	2	unneccessary	1
violent	1	weird	2

Adjectives: Frequent

Adjective:	# of uses:	Adjective:	# of uses:
good	129	bad	22
better	26	worse	5
best	10		
great	40	nice	57
funny	31	beautiful	16
okay	15	wonderful	11

8-B VERBS

Type 3: Infrequent

Verb:	# of uses:	Verb:	# of uses:
adore	3	barf	1
believe in	1	can't stand	8
can't take	3	enjoy	9
fed up	1	fuck	1
get	1	go for	5
going to	2	handle	1
ignore	1	interfere	1
leave	2	light up	3

miss	1	be into	1
care	2	feel like	1
kill	1	trust	3
be out of	1	resent	3
respect	2	screw	3
stink	3	stomach	1
suck	4	take	1
vote for	1		

Type 3: Frequent

Verb:	# of uses:	Verb:	# of uses:
hate	32	like	152
not like	64	love	71
not love	1	mind	15
prefer	14	rather	29
want	29	not want	26

Type 5: (It annoyed me. The movie interested me.)

Verb:	# of uses:	Verb:	# of uses:
annoy	1	bother	1
destroy	1	tickle/fancy	1
impress	1	interest	5
appeal	3	bother	1
bug	1	matter	2
turn on	3	satisfy	1
suit	1		

8-C NOUNS

Infrequent: all nouns in this category were used only once.

advantage	ass	asshole	belle	booboo
brain	bummer	champagne	character	charm
chocoholic	class	disadvantage	disaster	drag
dullard	enthusiast	evil	factory	
flapper	gentleman	heartburn	hell	hero
hypocrite	jerko	joke	junk	
lantern-jaw	mess	missiles	mole	
non-person	nuts	personality	pig	presence
price	prick	redneck	rowdyism	
sell-out	shmuck	shots	slut	snobs
something	sweetheart	the ones	thieves	traffic
void	warfare			

Frequent: (used more than once).

Noun:	# of uses:	Noun:	# of uses:
bullshit	2	charisma	3
fan	5	favourite	9
no way	2	s.o.b.	3
shit	5		

Appendix 9

Category Explanation

It seems logical to assume that A, the first category containing the questions with the greatest number of elements, would be the category that generated the largest number of utterances. Category B, with comparisons between two elements generated almost the same number; but category C, with only one element, generated only 300 utterances. Two reasons might account for the fewer number of responses in C: the speakers were getting tired of talking by the end of the sample, so they spoke less about each question, and each question held only one item to discuss so neither comparison nor contrast were invited.