WPLC Vol.1 No.2 (1981), 294-310

British and American Influences on Canadian English

H. J. Warkentyne and A. C. Brett
University of Victoria

1.0 INTRODUCTION

Canadian English is often characterized as a variety of English closely resembling American English, yet exhibiting many features typical of British English. This generalization presents certain problems in that there is a great diversity of dialects in both the United States and Britain. American English in this context presumably refers to the standard variety based on the regional speech of the North and West, sometimes referred to as General American (Thomas 1947:142-147). The model of British English for such a comparison usually is Southern Standard English.

The choice of Southern Standard British English as a basis for comparison is, in a sense, unfortunate in that many of the features of Canadian English introduced by the early settlers from the United Kingdom will thus be overlooked. The preponderance of the early immigration to Canada from Britain was not from southern England, but from northern areas of the British Isles; e.g., Yorkshire, Scotland, and Ireland. In 1871, for example, of over two million British immigrants residing in Canada 846,414 were Irish (Scargill 1977:10, figure quoted from Urquhart and Buckley 1965:18). It is interesting to note that a century or so earlier the same northern British dialects were the main contributors to what became the North and West dialect of the United States (Kurath 1971: 12-21). It is not inappropriate to reason, therefore, that the immigrants from northern Britain and the already established (American) Loyalist settlers found some compatability in each others speech ways; and,

rather than the new immigrants quickly abandoning their dialects, there ensued a blending or levelling of these closely related dialects which gave rise to a new dialect, Canadian English. Features of Canadian English such as the raised diphthongs [ʌy] and [ʌw], and neutralizations of the low back vowels, to name but a few, can be attributed to the influence of northern British English dialects. Although the close relationship of Canadian English to the northern dialect of the U.S. is undeniable, scholars in this field have largely neglected to examine the influences of British dialects other than Standard English of southern England.

The purpose of the study reported here was to determine the position of Canadian English in relation to American and British English using data obtained by means of the 1972 Survey of Canadian English (Scargill and Warkentyne 1972). Twenty-two items were selected on the basis that the choice of responses represented divided usage, one of which was typical of American English, and the other related more closely to British English. Data consisting of the responses to these items by informants who completed the Survey questionnaire were used in the analysis described below.

2.0 THEORETICAL CONSIDERATIONS

Implicit in the notion of determining the position of Canadian English with respect to American (U.S.) and British (U.K.) usage is the assumption that there exists some form of a continuum or line joining the extremes of U.S. and U.K. usage, and that we can measure the position of Canadian English along this line. Further, we are assuming that a set of items extracted from the questionnaire can represent this continuum in some way, and that the responses to these items may be used to make measurements along it. The validity of these assumptions must be tested before any attempt is made to determine the position of Canadian English, and certainly

before any conclusions are drawn concerning the meaning and importance of the position determined.

One model for the latent continuum along which we propose to make measurements, which contains methods of testing for its existence, is provided by Guttman scalogram analysis. The essential criterion that a set of questionnaire items must satisfy in order that it be considered to comprise a scale in the Guttman sense is cumulativity. Items are ordered according to 'difficulty' which, in our case, means that they are ordered according to the proportion of informants who have selected the U.K. usage response. The set of items so ordered would satisfy the cumulativity criterion perfectly if no informant selected a U.S. response for an item less 'difficult' than an item for which he had selected the U.K. response. Some latitude for variability is provided for in the method so that a number of errors or failures to fit the perfect scale pattern are tolerated and a set of items may still be regarded as constituting a scale.

Two properties or statistics frequently computed from responses to a set of items being tested as to whether or not it constitutes a scale are Coefficients of Reproducibility (REP) and Scalability (SCAL). REP is used to measure the fraction of responses which conform to the perfect scale response patterns, and SCAL measures, in effect, the fractional decrease in the number of errors observed over the maximal number of errors expected, given the numbers of informants selecting U.S. or U.K. responses for each item. By convention, if the fraction of errors is less than 0.10 or 10%, so that the value computed for REP exceeds 0.90, the set of items may be considered a scale. (In addition to applying the arbitrary value of 0.90 for REP as the critical value for cumulativity, we employed a Monte Carlo procedure, which we describe below, to test statistical significance of REP values.) Alternatively, if the ratio of the

number of errors observed to the maximal number expected is less than 0.40, so that the value of SCAL exceeds 0.60, the set may be regarded as a scale. If a set of items satisfies these criteria, one can infer the existence of a latent continuum underlying the items and can use scores computed from responses to items in the set to make measurements along this continuum.

3.0 THE DATA

3.1 The Sample

The original Survey of Canadian English data file containing the coded responses of 15,575 informants, consisting of grade nine students and their parents, was used to create a subfile containing the records only of informants born in the province in which they responded to the questionnaire. This subfile contained 11,561 records of which 710 were provided by informants in British Columbia, 969 from Alberta, 1326 from Saskatchewan, 950 from Manitoba, 1103 from Ontario, 1132 from Quebec, 1370 from New Brunswick, 1892 from Nova Scotia, 1391 from Prince Edward Island and 718 by informants from Newfoundland. A second subfile was created from the first by extracting 2000 records from it by a random process with the restriction imposed that the number of records drawn from each province be proportional to its English-speaking population. Thus a random sample was produced which could be construed as representative of the nation as a whole. The size of this sample was governed by the size of the original sample from Ontario, the province in which the largest number of English-speaking Canadians resides.

3.2 The Linguistic Items

Items were selected from the Survey questionnaire which offered a choice between a response that could be identified as representing preference for U.S. usage and a response that could be regarded

	U.S. usage		U.K. usage		*either*
Pronunciation					
lever	/lɛ́vər/	(16)	/lɪ́vər/	(71)	(13)
new	/nu/	(36)	/nyu/	(44)	(17)
student	/stʊ́dənt/	(65)	/styʊ́dənt/	(26)	(9)
schedule	/skɛ́ǰəl/	(78)	/šɛ́dyʊl/	(15)	(7)
(the letter) *z*	/zi/	(11)	/zɛd/	(78)	(11)
lieutenant	/lutɛ́nənt/	(69)	/lɛftɛ́nənt/	(22)	(9)
leisure	/lɪ́žər/	(67)	/lɛ́žər/	(24)	(9)
ate	/et/	(97)	/ɛt/	(3)	
aunt	/ænt/	(80)	/ant/	(20)	
progress	/prɒgrɪs/	(51)	/progrɪs/	(49)	
route	/rawt/	(22)	/rut/	(69)	(9)
Vocabulary					
	fall	(60)	*autumn*	(10)	(30)
(cloth)	*napkin*	(46)	*serviette*	(33)	(21)
(paper)	*napkin*	(37)	*serviette*	(43)	(20)
	toward	(29)	*towards*	(54)	(17)
(french)	*fries*	(69)	*chips*	(31)	
	mailman	(75)	*postman*	(8)	(17)
Spelling					
	color	(25)	*colour*	(52)	(23)
	center	(36)	*centre*	(40)	(24)
	traveled	(18)	*travelled*	(75)	(7)
	defense	(22)	*defence*	(54)	(24)
	gray	(27)	*grey*	(43)	(30)

Table 3.1 Questionnaire items showing U.S.-U.K. dichotomy with relative frequencies of responses (in percent) as observed in the random national sample.

as reflecting U.K. usage for purposes of this study. (Some items also allowed the ambivalent response of *either*.) A total of 22 such items were identified; eleven of these were classified as dealing with pronunciation, six with vocabulary, and five with spelling. The items selected are listed in Table 3.1 as representing U.S. (i.e, General American) or U.K. usage. Relative frequencies of the responses for each of the 22 items were determined for the random national sample of 2000 informants. The frequencies observed are recorded in Table 3.1 as percentages, appearing in parentheses following the responses to which they refer.

The reader will readily notice that the members of sets comprising the responses cannot all be exclusively assigned to U.S. or U.K. usage. Although by no means exclusively, the items assigned to U.S. usage are commonly used in General American, but not in British English. One exception to this criterion is *serviette*, the equivalent of which is U.S. *napkin* and U.K. *table-napkin*. According to Oxford English Dictionary older forms of *serviette* (*serviot, serviat, servit,* etc.) were used in Scotland exclusively. In the nineteenth century it was introduced into England with its French spelling, and presumably gained wide currency. However, *serviette* later came to be considered vulgar, and has disappeared from standard usage in England. From this account we may conclude, therefore, that *serviette* was introduced into Canadian English by British settlers.

4.0 ANALYSIS

4.1 Scaling Tests

A computer program was written to test the scalability in the Guttman sense, of the 22 items. The algorithm employed in this program followed that of Anderson (1966) used in the SPSS program Guttman Scale procedure (SPSS:528-539). This program had to be written

since the SPSS program permits scales consisting of, at most, twelve items. For this study it was desirable that large numbers of different sets of items be tested in an automatic way without the necessity of preparing different control cards for each set.

Scaling tests were performed using the national sample. For each subset of items tested, subjects who gave the ambivalent response *either* or who failed to respond to an item were excluded. The set of all 22 items was found not to comprise a scale, nor was any of the 22 subsets of 21 items found to comprise a scale. Since it was impractical to test all possible subsets consisting of from three to twenty items, we decided to follow a different procedure in the search for a subset which would comprise a scale.

From the set of all 22 items, the 22 subsets of 21 items each were generated by removing, in turn, each of the 22 items. Each of these subsets was tested for scalability, and the subset yielding the highest value of the Coefficient of Scalability (SCAL) was selected for the next step. On the basis of this subset 21 subsets were produced and tested, again retaining the subset yielding the highest SCAL. This procedure was repeated until a set of items was found which exhibited sufficiently high values of SCAL and REP that we could consider it to satisfy the conventional criteria for a Guttman Scale. The systematic elimination of items to produce successively smaller subsets was performed automatically by the computer program. The subset of five items comprising a scale derived from the 22 items is shown in Table 4.1 as the composite scale. Values of SCAL and REP obtained on the five-item composite scale were 0.5265 and 0.9070, respectively. It will be noticed that the SCAL value is somewhat below the 0.60 criterion value. Reducing the scale to four items by removing *lieutenant* would have yielded the values SCAL = 0.6516 and REP = 0.9360. However, we

chose to retain *lieutenant* and to deal with the five-item set on the grounds that the Coefficient of Reproducibility (REP) was sufficiently high that the five-item set could be regarded as a scale.

Composite scale	Pronunciation scale
center/centre (spelling)	*new*
z (pronunciation)	*student*
lieutenant (pronunciation)	*z* (the letter)
mailman/postman (vocabulary)	*ate*
ate (pronunciation)	

Vocabulary scale	Spelling scale
fall/autumn	*color/colour*
napkin/serviette (paper)	*center/centre*
mailman/postman	*traveled/travelled*

Table 4.1 Items comprising the scales

The pronunciation, vocabulary, and spelling subsets each were subjected separately to the stepwise scale derivation procedure. Scales obtained were a four-item pronunciation scale (SCAL = 0.6201, REP = 0.9121), a three-item vocabulary scale (SCAL = 0.6492, REP = 0.9163), and a three-item spelling scale (SCAL = 0.5394, REP = 0.8484, see further discussion below). The four subsets representing the four scales are shown in Table 4.1.

To decide if a set comprised a scale, we applied a test of scalability in addition to that provided by the REP and SCAL criterion values. We generated fifty independent sets of responses to

the 22 items which were random but subject to the condition that frequencies of responses to each item be the same as those observed in the national sample. For each of these fifty sets of randomly generated responses to the five items we took to comprise a scale, we determined a value for REP. The mean of these fifty values was computed, obtaining the value 0.8927, with a standard error, 0.003529. We tested the set of fifty REP values for skewness and kurtosis, and found the REP values to be distributed symmetrically about the mean, and observed that the distribution was mesokurtic. We then tested the distribution of REP values against a normal distribution with the same mean and variance by means of a Kolmogorov - Smirnov one-sample test, and found that the REP values did not differ significantly from a normal distribution at less than the 0.01 level. We therefore concluded that REP values computed from random responses were normally distributed; and that, if responses observed in the national sample were random, the statistic computed by dividing the difference between the mean REP from the random responses and the observed REP from the national sample by the standard error would be distributed standard normally. We tested the hypothesis that the responses observed in the national sample were random by determining the probability associated with the value of this statistic from a table of normal probabilities. We concluded that, if responses to the five items on the composite scale by the 2000 subjects in the national sample were random, the probability of a value of REP as high as the one observed would be less than one in a thousand. Therefore, we rejected the randomness hypothesis and concluded that responses to the five items were not random and hence, this set derived from the set of 22 items could be regarded as comprising a scale.

The sets of four pronunciation and three vocabulary items also

passed our randomness test and therefore satisfied our criterion for scalability, as well as the REP and SCAL criteria. Although the set of three spelling items failed to satisfy the REP and SCAL criteria, the value of REP obtained was sufficiently large that the chance of responses to these items being random was very small. We therefore concluded that this set could also be regarded as comprising a scale.

4.2 Statistical Measurements

Scores on the four subsets comprising scales were computed for the national sample as follows. For each item in the subset, a subject was assigned a score of -1 if he indicated a preference for the U.S. usage and a score of +1 if he selected the U.K. usage. In order to include the responses of subjects who chose the ambivalent response *either* to one or more items in a subset, a score of 0 (zero) was assigned to this response. Subjects who had failed to answer an item in a subset were excluded. Each subject's score for a subset of items was computed by summing his scores for individual items in the set. A mean score for the subset of items was computed by summing scores of subjects' given scores for the subset and dividing this sum by the number of subjects retained. This mean score was then divided by the number of items in the subset to produce a mean score per item.

The following mean scores per item were obtained: the composite scale, -0.279; the pronunciation scale, -0.154; the vocabulary scale, -0.371; and the spelling scale, +0.291. Mean scores per item and standard errors per item for all informants in the national sample on the four subsets of items are reported in Table 4.2. On the same four subsets, mean scores per item and standard errors per item for students and parents separately extracted from the national sample,

	Composite Scale	Pronunciation Scale	Vocabulary Scale	Spelling Scale
All informants	-0.279 (0.008)	-0.154 (0.008)	-0.371 (0.010)	+0.291 (0.012)
Students	-0.346 (0.008)	-0.222 (0.010)	-0.428 (0.012)	+0.280 (0.014)
All Parents	-0.154 (0.014)	-0.027 (0.016)	-0.266 (0.018)	+0.314 (0.022)
Parents by education				
Level 1	-0.196 (0.016)	-0.057 (0.019)	-0.282 (0.022)	+0.285 (0.025)
Level 2	-0.121 (0.056)	-0.030 (0.070)	-0.278 (0.059)	+0.278 (0.089)
Level 3	-0.048 (0.032)	+0.071 (0.026)	-0.221 (0.035)	+0.436 (0.046)

Table 4.2 Mean scores per item and standard errors per item (in parentheses) for students and parent subsamples of the national sample.

	Composite Scale	Pronunciation Scale	Vocabulary Scale	Spelling Scale
Students				
B.C.	-0.390 (0.013)	-0.190 (0.016)	-0.532 (0.018)	+0.191 (0.023)
Alta.	-0.401 (0.011)	-0.141 (0.015)	-0.556 (0.017)	-0.129 (0.021)
Sask.	-0.368 (0.011)	-0.168 (0.013)	-0.484 (0.016)	+0.050 (0.017)
Man.	-0.370 (0.012)	-0.132 (0.016)	-0.358 (0.016)	+0.107 (0.021)
Ont.	-0.290 (0.011)	-0.249 (0.013)	-0.352 (0.016)	+0.466 (0.016)
Que.	-0.363 (0.010)	-0.157 (0.013)	-0.503 (0.014)	+0.270 (0.016)
N.B.	-0.346 (0.010)	-0.220 (0.012)	-0.704 (0.013)	+0.143 (0.017)
N.S.	-0.351 (0.010)	-0.222 (0.012)	-0.605 (0.014)	+0.016 (0.016)
P.E.I.	-0.413 (0.011)	-0.235 (0.014)	-0.663 (0.015)	+0.016 (0.018)
Nfld.	-0.441 (0.018)	-0.269 (0.021)	-0.240 (0.023)	+0.112 (0.023)
Parents				
B.C.	-0.105 (0.026)	+0.119 (0.030)	-0.335 (0.035)	+0.364 (0.041)
Alta.	-0.241 (0.016)	+0.006 (0.020)	-0.434 (0.021)	-0.043 (0.028)
Sask.	-0.214 (0.014)	-0.053 (0.017)	-0.339 (0.018)	+0.041 (0.023)
Man.	-0.171 (0.018)	+0.123 (0.021)	-0.165 (0.023)	+0.088 (0.032)
Ont.	-0.089 (0.019)	-0.066 (0.023)	-0.181 (0.022)	+0.580 (0.025)
Que.	-0.146 (0.028)	+0.023 (0.032)	-0.283 (0.035)	+0.330 (0.044)
N.B.	-0.254 (0.014)	-0.115 (0.017)	-0.604 (0.019)	+0.097 (0.022)
N.S.	-0.220 (0.011)	-0.067 (0.013)	-0.409 (0.016)	+0.088 (0.017)
P.E.I.	-0.329 (0.014)	-0.136 (0.018)	-0.481 (0.020)	+0.102 (0.021)
Nfld.	-0.456 (0.020)	-0.215 (0.024)	-0.175 (0.026)	+0.147 (0.028)

Table 4.3 Mean scores per item and standard errors per item (in parentheses) of scale subsets for students and parents by province

and for parents on the basis of their educational level are also shown in Table 4.2. Mean scores and standard errors per item were computed for students and for parents by individual province, and are reported in Table 4.3.

5. RESULTS

5.1 Significance Levels

In the description of the results of the statistical tests, if we are able to reject the null hypothesis of the test at the 0.05 (5%) level of significance, than we refer to the difference between two means being tested for equality, or the difference between a mean and some constant as significant. At the 0.01 (1%) level we describe the difference as highly significant.

If, in a one-sample test, a particular score under examination is negative, we conclude that the informants for whom the mean was calculated show a preference for U.S. usage; a positive mean, on the other hand, indicates a preference for U.K. usage. If the mean is not significantly different from zero, we conclude that there is no discernible tendency towards either U.S. or U.K. usage. For the one-sample tests involving the equality of mean scores per item to -1 (U.S. usage extreme), to +1 (U.K. usage extreme), and to 0 (zero) for the various subsamples of informants taken on the four scales, since the sample sizes were very large, a z-statistic was computed and critical values were taken from a table of areas under the normal curve. For the two-sample tests of significance of a difference between two means, we employed a t-statistic, with different critical values being employed to compensate, when necessary, for unequal sample variances.

5.2 National Patterns

In all tests of mean scores per item against the U.S. and U.K. usage extreme values, all means were found to be highly significantly

different from both the U.S. and U.K. extremes. We concluded from the results of these tests that, although subsamples of informants might indicate a preference for U.S. or U.K. usage in the categories of items tested, none of the subsamples tested indicated strictly U.S. or U.K. usage.

After testing mean scores for equality to zero, we concluded that informants comprising the random national sample show a highly significant tendency towards U.S. usage on the pronunciation and vocabulary scales, and a highly significant tendency towards U.K. usage on the spelling scale. This observation also applies to the subsample of students.

Parents from the national sample, including all three levels of education in the one subsample, show a significant (although not highly significant) tendency towards U.S. usage on the pronunciation scale, a highly significant tendency towards U.S. usage on the vocabulary scale, and a highly significant tendency towards U.K. usage on the spelling scale. For parents at the first level of education (not beyond high school), the tendencies toward U.S. usage in pronunciation and vocabulary and towards U.K. usage in spelling are all highly significant. For parents at the second educational level (beyond high school, but not university), the tendencies towards U.K. usage in spelling and U.S. usage in vocabulary are highly significant. However, the preference for U.S. usage in pronunciation is not significant. Parents from the national sample at the third level of education (university training) show highly significant tendencies towards U.S. and U.K. usages in vocabulary and spelling, respectively; but, unlike all other subsamples, they show a significant (although not highly significant) tendency towards U.K. usage in pronunciation.

In addition to the one-sample tests described above, we also

conducted two-sample tests of the equality of mean scale scores per item (SPSS:267-275). When mean scores for students were compared with mean scores for parents, we found that students show a highly significantly greater tendency towards U.S. usage on pronunciation and vocabulary than do parents. In the case of spelling, however, the difference between student and parent mean scores is not significant. When mean scores for students were compared separately with those for parents at the three educational levels, we found that students' preference for U.S. usage in pronunciation is highly significantly different from the preferences indicated by parents at all of the three levels. In the case of vocabulary, students show a highly significantly greater tendency towards U.S. usage than parents at the first and third educational levels. The difference between the preference exhibited by students and by parents at the second educational level for U.S. usage in vocabulary is not significant. The mean scores for students and first and second level parents on the spelling scale do not differ significantly. However, the tendency towards U.K. usage in spelling by parents at the third level of education is highly significantly greater than that of the students.

Comparison of the mean scores for parents at the three educational levels showed that the only significant differences occur between the first and third levels. These differences appear in the cases of pronunciation and spelling, and they are highly significant. Differences between parents at the first and second levels, and between second and third levels in pronunciation and spelling are not significant. In the case of vocabulary none of the differences amongst the three levels is significant.

5.3 Provincial Patterns

When mean scores on pronunciation for parents in each province

were compared with zero, we found the tendency towards U.S. usage in Saskatchewan, Ontario, New Brunswick, Nova Scotia, Prince Edward Island and Newfoundland to be highly significant. The tendency towards U.K. pronunciation in British Columbia and Manitoba is highly significant. Preferences indicated by parents in Alberta and Quebec do not differ significantly from zero on this scale. In the case of vocabulary the tendency towards U.S. usage demonstrated by parents in all provinces is highly significant. The tendency towards U.K. usage in spelling by parents in all provinces, except Alberta and Saskatchewan, is highly significant. This tendency is significant in Saskatchewan. In Alberta there is a slight preference for U.S. usage in spelling, but the mean score is not significantly different from zero.

For students, there is a highly significant preference for U.S. usage on both pronunciation and vocabulary scales. On the spelling scale, there exists a highly significant tendency towards U.K. usage in all provinces except Alberta and Prince Edward Island. In Alberta the tendency towards U.S. usage in spelling is highly significant. The preference for U.K. spelling in Prince Edward Island is not significant and therefore no conclusion can be formed regarding a tendency in this direction for students of this province.

6. DISCUSSION

The step-wise scale derivation procedure used in this study (described in 4.1) is not exhaustive, and therefore there is no guarantee that the scale obtained will be that one, amongst all those of the same size, with the highest Coefficient of Scalability. The procedure does, however, permit subsets of items satisfying scaling criteria to be derived from larger sets which do not satisfy scaling criteria, at moderate expenditure of computational time and money compared with what would be required for an exhaustive search.

Since the set of spelling items was small we were able to test all possible subsets of three items for scalability in this category. Although this may be a special case, we found that the stepwise procedure did produce that set of three items with the highest value of the Coefficient of Scalability.

The reader is cautioned that there are limitations on the degree of validity of the results presented here. For example, the distribution of the responses elicited by the 1972 survey was not random in the technical sense; and, therefore, the claim that our results reflect the usage of the whole English-speaking population of Canada is somewhat of an over-statement. However, inasmuch as the survey did succeed in sampling a cross-section of the population, albeit informally, our results can be accepted at a reasonable level of confidence.

REFERENCES

Anderson, Ronald E. 1966. A Computer Program for Guttman Scaling with the Goodenough Technique. *Behavioral Science*, Vol. (1), No. 3:235.

Scargill, M. H. 1977. *A Short History of Canadian English.* Victoria: Sono Nis Press.

Scargill, M. H. and H. J. Warkentyne. 1972. The Survey of Canadian English. *Enlgish Quarterly*, 5.47-104.

SPSS. 1975. N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent. *Statistical Package for the Social Sciences* (Second Edition) McGraw-Hill.

Thomas, Charles K. 1947. *An Introduction to the Phonetics of American English.* New York: Ronald Press.

Urquhart, M. C. and K. A. H. Buckley. 1965. *Historical Statistics of Canada.* Toronto: MacMillan.