DOCUMENT RESUME

ED 035 860 AL 002 263

AUTHOR RICHARDS, JACK C.
TITLE A PSYCHOLINGUISTIC MEASURE OF VOCABULARY SELECTION.
PUB DATE AUG 69
NOTE 25P.; REVISED VERSION OF PAPER READ AT THE ANNUAL MEETING OF THE CANADIAN LINGUISTIC ASSOCIATION, YORK UNIVERSITY, TORONTO, JUNE 1969

EDRS PRICE EDRS PRICE MF-$0.25 HC-$1.35
DESCRIPTORS ASSOCIATION (PSYCHOLOGICAL), CULTURAL CONTEXT, *ENGLISH (SECOND LANGUAGE), LANGUAGE INSTRUCTION, *NOMINALS, *PSYCHOLINGUISTICS, *VOCABULARY, WORD FREQUENCY, *WORD LISTS

ABSTRACT

SEVERAL BASIC PROBLEMS IN THE FIELD OF THE SELECTION OF VOCABULARY FOR TEACHING ENGLISH AS A FOREIGN LANGUAGE ARE DISCUSSED. THE NATURE OF WORD FREQUENCY AND WORD AVAILABILITY ARE CONSIDERED, ALONG WITH THEIR LIMITATIONS AS MEASURES OF THE USEFULNESS OF CONCRETE NOUNS. WORD FAMILIARITY IS PROPOSED AS A PSYCHOLINGUISTIC MEASURE FOR NOUN SELECTION, AND SOME EXPERIMENTAL EVIDENCE PRESENTED TO DEMONSTRATE ITS VALIDITY. THIS IS A PRELIMINARY REPORT OF A STUDY WHICH UPDATES THE "GENERAL SERVICE LIST" OF MICHAEL WEST THROUGH ESTABLISHING WORD FAMILIARITY FIGURES FOR SOME 5000 NOUNS AS WELL AS UPDATED FREQUENCY FIGURES FOR WRITTEN AND SPOKEN ENGLISH. (AUTHOR/DO)

A PSYCHOLINGUISTIC MEASURE OF VOCABULARY SELECTION

by Jack C. Richards
International Center for Research on Bilingualism
Laval University
Québec, *Canada*

August 1969

Abstract

Several basic problems in the field of the selection of vocabulary for teaching English as a foreign language are discussed. The nature of word frequency and word availability are considered, along with their limitations as measures of the usefulness of concrete nouns. Word familiarity is proposed as a psycholinguistic measure for noun selection, and some experimental evidence presented to demonstrate its validity.

CONTENTS

*INTRODUCTION**

"Since it is impossible to teach the whole of a language, all methods must in some way or another, whether intentionally or not, select the part of it they intend to teach"(1). Methods of selecting items for foreign language teaching may reflect differing views of language, and of language teaching. In choosing grammatical items, reference may be made to unobserved relations between items at the level of deep structure. Words, on the other hand, may be chosen according to their frequency of usage. Some 80 word-lists this century have been based on this principle. It seemed evident that if words were useful they would be used often; to establish a basic word-list for language teaching, it was necessary simply to count a wide sample of spoken or written discourse.

Yet many teachers have questioned the usefulness of word-frequency lists. To teach vocabulary, it is necessary to recreate the contexts in which words are used, choosing situations which the learner is likely to encounter. The vocabulary of many familiar situations however, does not rank highly in frequency lists. *Soap, bath, cushion, chalk,* and *stomach,* are not within the first 2, 000 words of Thorndike and Lorge's list(2). Teachers and course designers have often had to ignore the frequency lists and rely on their own discretion as to which nouns should be taught. This does not always make for consistency or reliability.

The research described here was designed to provide more reliable sources for the preparation of language teaching texts and materials, through accurate methods of vocabulary selection. To begin, it was necessary to analyse the advantages and limitations of existing criteria for selection.

* *This is a revised version of a paper read at the annual meeting of the Canadian Linguistic Association, York University, Toronto, June 1969.*

1. LIMITATIONS OF FREQUENCY LISTS

A word-frequency list is an arrangement of words in order of descending frequencies. In practice this means that words are arranged according to their degree of generality or grammaticality. For the language teacher, it means that some of the most teachable words -- the concrete nouns -- may be the least accessible, occurring in the third, fourth or fifth thousand range rather than within the first few hundred words.

Fries believed that this was the result of poor sampling. "We do not need...more general counts which include the function words, but counts limited to "things" and "qualities". There are many common necessary words that do not get into print, especially into the kind of publication that furnished the bulk of the material counted"(3). Michéa has shown however, that the instability of concrete nouns is rather a reflection of the nature of word frequency(4).

The effect of a frequency count is to reduce a corpus to a set of frequencies in which the value of a given frequency is necessarily relative to that of the other frequencies. Since the words at the top of the list occur with a very high frequency (the first 250 words representing as much as 80% of a text), there is only a small percentage left to be shared by the thousands of other words in the language. The most frequent words in the language are words with grammatical, abstract, or general meaning. The grammatical words are frequent because it is impossible to produce a sentence without them. Both grammatical and abstract words have a variety of meanings and hence several opportunities to occur in any given text. The Oxford English Dictionary

gives a large number of senses for the following words: (3:82) make (97), go (94), give (64), of (63), do (54), keep (58), put (57), up (67), with (45). These are words with high and stable frequencies in word-lists. Not being context bound, they can occur in many situations.

On the other hand infrequent words are generally words with fewer meanings. They may be confined to one or two particular contexts. Words like *thimble, blackbird, songbook,* have only one occurrence in Kučera and Brown's corpus of over a million running words[5]. Many common concrete nouns have very low frequencies in this and other word-frequency lists. Infrequent words are highly context bound, carry a great amount of information, and are liable to vary from one topic of discourse to another. The differences between these two classes of words are shown in Figure 1.

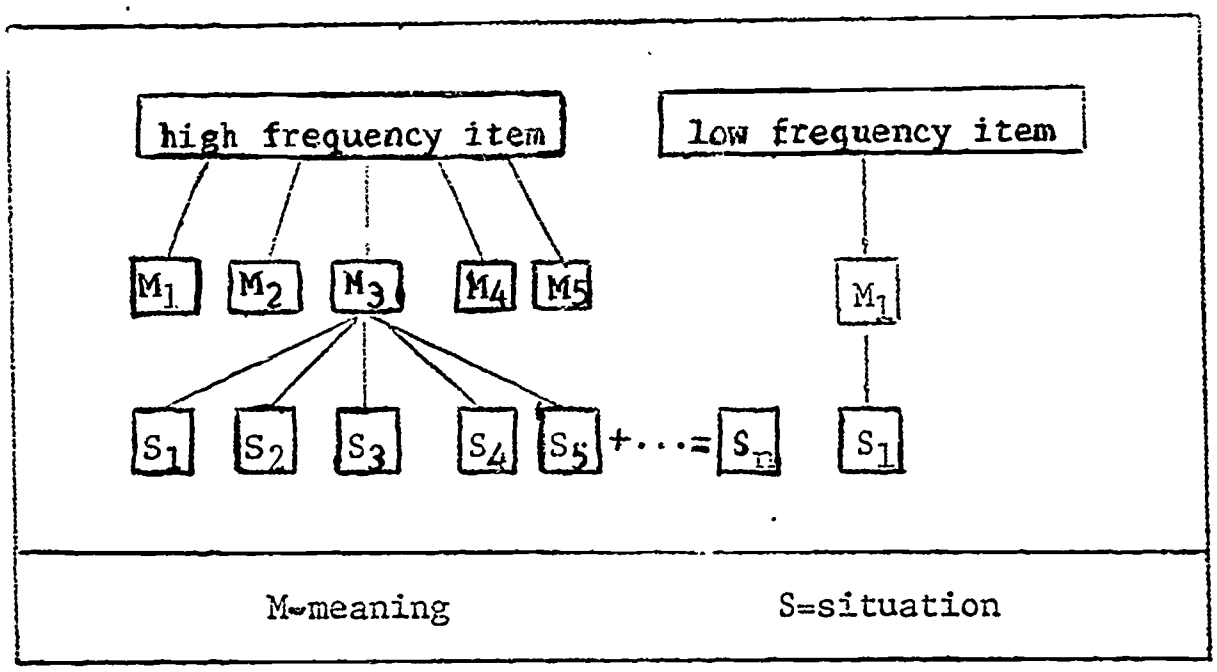

Figure 1. *High frequency items have multiple meanings and are context free. Low frequency items have fewer meanings and are context bound.*

Word frequency, as reflected in corpuses of spoken or written language, is a stable index for those members of the lexicon with grammatical, general, and context-free meaning, such as the structural words, and the most frequent verbs, adjectives, adverbs, and nouns, but is less reliable for context-bound words such as the majority of concrete nouns. For measures of the usefulness of such words we must look elsewhere.

2. OTHER TYPES OF WORD-LISTS

Most of the word-lists utilized for teaching, have been produced by language teachers such as Palmers, Bongers and West, or by educational psychologists such as Thorndike. In other disciplines vocabulary has been studied for quite different reasons.

In the field of association psychology, studies of word behaviour have supplied a great deal of information on the psychological grouping of words. Numerous word association studies have been carried out to provide support for learning theories, and theories of categorization and conceptualization(6). More recently, word association has been examined for evidence of the deep structural organization of language(7). Responses which subjects give to words in free-association tests, while seemingly random at the level of surface structure, are seen to be systematic at the level of deep structure. Various word-lists have been produced by psychologists and learning theorists to demonstrate the patterns of association between members of the lexicon.

2.1 *VERBAL CATEGORY TESTS*

The word association studies which are perhaps most relevant to the field of vocabulary selection, are those of W.A. Bousfield and his associates. Bousfield has studied the effects of the domination or organization imposed on words by particular mental concepts. In one set of experiments, students were given categories such as *weapons, the house, food,* and asked to list four words belonging to each category(8). There was a high degree of agreement between the words which were supplied. An example of the results may be seen in Table 1.

2.2 *AVAILABILITY*

A similar type of measurement was developed independently in France as part of the *français fondamental* project of 1954(9). It was found that the frequency lists established for this project gave insufficient importance to many common nouns. Michéa elaborated the concept of word availability in an attempt to rate nouns according to their degree of association within specific categories and situations. Situations, Michéa argued, call words to mind according to their degree of availability. "An available word is a word which though not necessarily frequent, is always ready for use, and comes to mind when it is needed. It is a word which, belonging to normal associations of ideas, emerges whenever such associations come into play...This is why it is possible to attribute to many concrete nouns a degree of availability within a particular associative group, whereas statistics based on the analysis of texts are unable to alot them any stable and well defined place in the order of frequencies."4:23

To locate the available words, lists of categories and centers of interest were drawn up. These were used to elicit the most available words for a number of basic situations. This information was then used to correct the figures obtained by frequency alone. Availability has also been used to establish vocabulary lists for teaching other varieties of French [10], and for Irish[11], and Spanish [12], but the concept of a center of interest or situation has never received precise definition. This concept presents a number of theoretical and practical difficulties.

3. CENTERS OF INTEREST

In language teaching, the concept of a center of interest has been used to refer to a procedure for the classroom presentation of words. Hornby writes, "The word family that is most useful in the classroom is that which provides a group of words that enables the learner to form associations, associations of the sort which help him to think in the new language. If I think, hear or see the name Egypt, I think of Cairo, the Nile, the Suez Canal, the Sphynx, the Pyramids, sand and desert, irrigated fields. Brazil suggests coffee and the Argentine calls to mind Tangos and frozen meat".[13] Jones describes a situation as "an internally cohesive semantic group".[14] There may be two quite different sorts of relation between the members of these semantic groups.[7:141-159]

The first is a relationship formed through contextual contiguity; that is, through the occurrence of items together in the real world. This is the case of the words which come to mind when we think of Egypt. We think of the things which we would expect to see and find in Egypt. A quite different type of grouping takes place when a concept such as "pointed

objects" is thought of. This might call to mind *pencil, rocket,* and *arrow.* These are examples of the category "pointed objects"; they are not necessarily found together in real life.

The words elicited in availability and verbal category tests can be classified into responses of these two types. One type is produced when subjects are asked to list their responses to categories or centers of interest which are class nouns. "Professions", "games", "animals" and other categories of this type were the basis of the *français fondamental* availability study of 1954(9). The responses to these categories are not words which would be associated together in the real world. Irish children produced words such as *lion, elephant* and *tiger* as responses to the category "animals", although these animals are presumably not a part of the Irish landscape. Similarly Irish children produced *carpenter, teacher, doctor, shopkeeper* and *builder* as the most frequent responses to the category "professions". In both cases subjects provide illustration of their ability to conceptualize, rather than a recollection of items seen together in experience.

A second type of center of interest is found in topics such as "Going on a voyage" or "Having a meal at the table", which elicit items which *are* associated together in experience. "Going on a voyage" elicits nouns like *bag, train ticket, hotel* and *camera.* "Having a meal at the table" elicits *knife, fork, plate* and *butter.* Many centers of interest produce responses of both types. "Illness" may be interpreted as Type 1; that is, as if the instructions were "think of the names of different types of illnesses". Or it may be interpreted as Type 2, as if the instruction was "imagine yourself sick in hospital and tell me all the things you would see there".

In the Irish availability study of 1966[11], most of the responses to the center of interest described as "Kind of illnesses, injuries, cures", can be divided into these two different types of response. Responses of Type 1 are such words as *measles, flu, headache, chickenpox, mumps, cancer, fractured skull, polio.* Responses of Type 2 are *medicine, doctor, tablet, hospital, nurse, bed, pill, ambulance, ointment.* It has not been made clear to date if availability tests are designed to measure words of the first type -- the vocabulary of semantic categories -- or words of the second type -- the vocabulary of particular situations. Additional factors make availability an unsatisfactory measure of the usefulness of concrete nouns.

The basic problem arises with the selection of centers of interest, situations and categories. No objective or precise way of choosing these seems possible. Nor is it easy to define centers of interest for many common nouns; *handle, glue,* and *cigarette* for example do not readily classify themselves situationally. Even if an objective method for selecting centers of interest were discovered, the question of their evaluation would remain. Are some situations more important than others, or are all words of a given rank within centers of equal importance? It would appear that each time a center of interest is chosen, an independent index is added to the criteria for vocabulary selection. What is needed is a method of evaluating nouns independently of centers of interest, giving a single scale rather than a multitude of independent indices. Word familiarity may provide such an index.

4. WORD FAMILIARITY

Important work on the establishment of word familiarity as a stable and meaningful psycholinguistic measure has been done by Henrion[15], Noble[16], and Fraisse[17]. It has

recently been used as a measure of bilingual dominance[18]. Word familiarity is an attempt to measure the degree of importance people attribute to words. This may be measured by asking subjects to rate words on a scale which indicates the degree to which they expect to hear, see or use words. Word familiarity may be interpreted differently according to the type of word which is rated. For members of the lexicon with stable frequencies, such as the common verbs, adjectives, adverbs and nouns, word familiarity confirms the psychological reality of word frequency. Subjects rate the more frequent members of these word classes as having more familiarity than the less frequent words[22]. For concrete nouns, word familiarity may reflect the familiarity of an item for person. Such a rating may be independent of frequency. *Tooth-paste* is not a frequent word, although it is a familiar word, since it designates something which is used by many people every day.

Since a culture may be defined partly in terms of the distribution of linguistic norms, word familiarity can be used as a measure of cultural homogeneity and diversity. Fraisse found that the familiarity of certain nouns differed according to the intellectual aptitude of the subjects, their instructional level, and their socio-professional level. Students rated *book* as the most familiar noun in experiment 1, below. Vocabulary learning, more especially nominal learning, is closely related with the domains in which a speaker is involved. A speech community is characterized by definite norms of language and behaviour. While at the individual level, a speaker's vocabulary defines his personal history, we can abstract from the level of ideolect to the common norms which define a speech community. According to Fishman[19], these social domains identify the major spheres of activity of a culture. The family, educational, and religious domains

for example, are easily recognized. These are institutionalized spheres of activity in which specific language behaviour occurs. Yet in word familiarity the general cultural significance of concrete nouns can be measured without reference to social domains or centers of interest.

5. THE INVESTIGATION OF WORD FAMILIARITY

To find out what word familiarity measures, and how it differs from other measures of vocabulary, a number of different experiments were performed. In the first place it was necessary to discover if subjects would give stable responses when asked to indicate the degree of familiarity of a random sample of nouns. Since a considerable amount of information about noun availability is already available, it was also necessary to devise tests which would enable us to compare word familiarity and word availability, and to measure the degree to which word familiarity ratings differed from one group of subjects to another.

5.1 *EXPERIMENTAL DESIGN*

Although the following experiments were carried out in an attempt to measure the familiarity of the objects designated by nouns, preliminary sampling showed that subjects gave the same results when asked to estimate the frequency of use or experience of a word, as they did when estimating the frequency with which they encountered the object designated by a word. Following Noble, subjects were asked to estimate their frequency of use or experience of a word. (See Appendix). Three experiments were performed; two familiarity tests, and one availability test.

The tests were given to native English speaking university students and high school pupils.

5.2 *RESULTS*

5.2.1 Experiment 1

192 concrete nouns selected at random from among the letters *a, b, c,* of the *Canadian Reader's Dictionary* (20) were arranged in random order in booklet form. The test was given to 38 subjects. Their responses produced the ranking shown in Table 4. The results seem to demonstrate that word familiarity may provide an evaluation scale for nouns which do not rank high on frequency lists. *Blackboard* is third in familiarity in Table 4 yet has only two occurrences in Kučera and Brown's corpus.

5.2.2 Experiment 2

To test the effects of sample size and word order on a familiarity ranking, 26 words selected from the letters *a, b, c,* of the dictionary, were mixed with a different group of randomly selected nouns from the same dictionary. 40 subjects completed the test. The 26 nouns were then mixed with a different group of randomly selected nouns and a different group of 40 students rated them. The coefficient of correlation between the two ratings came to .775. The correlation between the two ratings seems to demonstrate as Fraisse and others have found, that word familiarity is something real and measurable.

5.2.3 Experiment 3

The purpose of this experiment was to find out if word familiarity measures the same thing as word availability. 40 subjects were first given two centers of interest -- "parts of the body" and "clothing". They were instructed to write 15 words for each of the two centers. The results were totalled and ranked. The words produced were then distributed at random among a larger sample of nouns selected from the dictionary. Two weeks later, the same subjects were asked to rate the list according to familiarity. The coefficient of correlation between the two different rankings produced was calculated, giving .534 and .56 respectively. (See Table 3).

The correlation between the first 12 words on each list was also calculated. The correlation for "clothing" was .823, while that for "parts of the body" was .35. This latter figure may be the result either of sample size, or of minor affects in testing conditions. A larger group of subjects or a larger list of words might produce a higher correlation. The experiment does suggest however, that in general terms, word familiarity and word availability give similar information about concrete nouns, although word familiarity does not specify the domains with which they are associated.

6. COMPARISON OF FAMILIARITY AND AVAILABILITY

Table 2 compares words from the familiarity ranking of Table 4 with availability ratings for these words in the Irish and Acadian tests. The availability

ratings for these words are only meaningful with reference to particular centers of interest, since as with *book,* the degree of availability may change from one center to another. Sometimes a slight modification in the description of a center of interest means that a word is not elicited, as in the case of *bottle* in the Irish tests. Particular ratings may reflect merely the success with which a domain or center of interest is described, rather than the independent overall degree of availability of a word. Word familiarity on the other hand, appears to rate words to each other independently of how appropriate they are for particular contexts.

CONCLUSIONS

Concrete nouns may have unstable and insignificant ranks in a word-frequency list, but significant and stable positions in a familiarity list which indicates the degree to which people expect to hear, see, or use words. Word familiarity may differ according to the social, cultural, and intellectual level of the speaker, since each of these variables is likely to be reflected in differences in vocabulary knowledge and use. For a given social group, word familiarity can be measured without reference to social domains or centers of interest. In availability testing, domains and centers of interest have to be hypothesized. They raise questions of selection and evaluation which make them a difficult measure for vocabulary selection. In language teaching they may be more useful at the level of presentation, for once words have been selected, they can be arranged for teaching purposes into as many or as few centers of interest as are required. Although both word familiarity and

word availability appear to measure the cultural significance of nouns, the advantage of word familiarity is that it produces a single index for nouns, rather than a number of independent indices.

A test is now being prepared to obtain familiarity ratings on some 5,000 nouns. The figures given will be used in conjunction with other measures of vocabulary usage, to provide a basic word-list for the teaching of English as a second or foreign language.

Table 1

MOST FREQUENT RESPONSES TO "FURNITURE" CATEGORY

chair	392
table	346
bed	225
couch	122
desk	103
sofa	101
lamp	77

Seven most frequent responses to the category labelled "an article of Furniture" according to Bousfield et al. 1957[9].

Table 2

FAMILIARITY AND AVAILABILITY COMPARED

WORD	FAMILIARITY RANK	AVAILABILITY RANK	DESCRIPTION OF CENTER OF INTEREST
book (livre)	1	Irish - 7	school life, activities, furniture
		Irish - 31	recreations, hobbies, indoor pasttimes
		Acadian -2	l'école, ses meubles et son matériel scolaire
		Acadian - 30	les jeux et distractions
		Acadian - 113	la maison
bedroom (chambre à coucher)	2	Irish - 18	the home or house
		Acadian - 18	la maison
blackboard (tableau)	3	Irish - 2	school life, activities, and furniture
		Acadian - 3	l'école, ses meubles et son matériel scolaire
bottle (bouteille)	5	Irish - 0 (not elicited)	
		Acadian - 17	les objets placés sur la table
		Acadian - 138	la cuisine, ses meubles et les ustensiles qui s'y trouvent
bath (bain)	6	Irish - 32	the home or house
		Acadian - 137	la maison

Familiarity is shown as a single index and availability as a multiple index.

Table 3

VOCABULARY OF CENTERS OF INTEREST BY AVAILABILITY AND FAMILIARITY

CENTER OF INTEREST "Clothing"		CENTER OF INTEREST "Parts of the body"	
Availability	*Familiarity*	*Availability*	*Familiarity*
1 *shoes*	1 *shoes*	1 *eye*	1 *hair*
2 *socks*	2 *pants*	2 *leg*	2 *hand*
3 *sweater*	3 *socks*	3 *arms*	3 *mouth*
4 *shirt*	4 *sweater*	4 *foot*	4 *eye*
5 *pants*	5 *shirt*	5 *ear*	5 *face*
6 *hat*	6 *jeans*	6 *nose*	6 *finger*
7 *coat*	7 *jacket*	7 *finger*	7 *arm*
8 *tie*	8 *skirt*	8 *hand*	8 *head*
9 *jacket*	9 *stockings*	9 *head*	9 *foot*
10 *skirt*	10 *nylons*	10 *knee*	10 *leg*
11 *blouse*	11 *dress*	11 *toes*	11 *teeth*
12 *dress*	12 *pyjamas*	12 *mouth*	12 *neck*
13 *shorts*	13 *slacks*	13 *elbow*	13 *nose*
14 *underwear*	14 *coat*	14 *hair*	14 *lips*
15 *gloves*	15 *tie*	15 *neck*	15 *elbow*
16 *overcoat*	16 *boots*	16 *stomach*	16 *fingernails*
17 *rubbers*	17 *underwear*	17 *teeth*	17 *toes*
18 *vest*	18 *hat*	18 *shoulder*	18 *thumb*
19 *boots*	19 *overcoat*	19 *ankle*	19 *heart*
20 *scarf*	20 *suit*	20 *brain*	20 *stomach*
21 *slacks*	21 *blouse*	21 *face*	21 *shoulder*
22 *belt*	22 *belt*	22 *back*	22 *brain*
23 *suit*	23 *slip*	23 *fingernails*	23 *knee*
24 *stockings*	24 *trousers*	24 *heart*	24 *ear*
25 *ski-jacket*	25 *shorts*	25 *lips*	25 *cheek*
26 *jumper*	26 *T-shirt*	26 *cheek*	26 *ankle*
27 *jeans*	27 *scarf*	27 *thumb*	27 *back*
28 *nylons*	28 *ski-jacket*	28 *lungs*	28 *lungs*
29 *T-shirt*	29 *gloves*		
30 *pyjamas*	30 *rubbers*		
31 *ear-muffs*	31 *vest*		
32 *slip*	32 *jumper*		
33 *trousers*	33 *ear-muffs*		
Correlation = .534		Correlation = .56	

Table 4

WORD FAMILIARITY RATINGS FOR 192 NOUNS *

	WORD	FAMILIARITY SCORE	RANK		WORD	FAMILIARITY SCORE	RANK
1	*book*	1120	1	45	*bus driver*	680	45
2	*bedroom*	994	2	46	*cash register*	666	46
3	*blackboard*	984	3	47	*aspirin*	656	47
4	*breakfast*	980	4	48	*branch*	644	48
5	*bottle*	960	5	49	*automobile*	640	49
6	*bath*	956	6	50	*broom*	640	49
7	*car*	956	6	51	*booklet*	636	51
8	*bathroom*	952	8	52	*balcony*	632	52
9	*arm*	944	9	53	*bacon*	632	52
10	*butter*	944	9	54	*banana*	628	54
11	*bed*	940	11	55	*blizzard*	624	55
12	*apartment*	936	12	56	*basket*	622	56
13	*bill*	928	13	57	*bench*	616	57
14	*bread*	920	14	58	*brake*	614	58
15	*building*	888	15	59	*belt*	614	58
16	*bag*	876	16	60	*ball*	612	60
17	*blanket*	872	17	61	*alcohol*	604	61
18	*boy*	860	18	62	*armchair*	604	61
19	*brush*	836	19	63	*cat*	592	63
20	*brother*	826	20	64	*back*	590	64
21	*bank*	824	21	65	*camera*	586	65
22	*alarmclock*	824	21	66	*brochure*	580	66
23	*box*	824	21	67	*bandage*	570	67
24	*bird*	820	24	68	*bathrobe*	570	67
25	*beer*	806	25	69	*carrot*	568	69
26	*book case*	800	26	70	*carpet*	564	70
27	*air*	792	27	71	*ceiling*	564	70
28	*brain*	780	28	72	*ash*	558	72
29	*blouse*	768	29	73	*beach*	556	73
30	*boot*	768	29	74	*atlas*	552	74
31	*cake*	760	31	75	*actor*	542	75
32	*button*	756	32	76	*backyard*	538	76
33	*bus*	748	33	77	*aeroplane*	524	77
34	*apple*	744	34	78	*aunt*	522	78
35	*airport*	728	35	79	*bikini*	520	79
36	*bowl*	724	36	80	*bungalow*	518	80
37	*animal*	716	37	81	*boat*	512	81
38	*baby*	712	38	82	*account*	512	81
39	*bed-spread*	700	39	83	*cabinet*	508	83
40	*ankle*	700	39	84	*astronaut*	508	83
41	*beard*	688	41	85	*army*	494	85
42	*blood*	684	42	86	*album*	488	86
43	*basement*	684	42	87	*aluminium*	486	87
44	*candy*	684	42	88	*banknote*	482	88

* *Figures obtained through rating words marked as very often-32, often-16, sometimes-8, rarely-4, never-2.*

Table 4 (suite)

	WORD	FAMILIARITY SCORE	RANK		WORD	FAMILIARITY SCORE	RANK
89	*bell*	480	89	136	*belly*	326	136
90	*ambulance*	480	89	137	*carnation*	324	137
91	*bicycle*	476	91	138	*baywindow*	320	138
92	*bark*	474	92	139	*arrow*	320	138
93	*butcher*	472	93	140	*ambassador*	320	138
94	*cabin*	472	93	141	*beam*	318	141
95	*candle*	470	95	142	*boiler*	316	142
96	*bay*	470	95	143	*canary*	316	142
97	*auditorium*	470	95	144	*cane*	316	142
98	*cap*	460	98	145	*bull*	316	142
99	*bone*	456	99	146	*cabaret*	312	146
100	*barber*	456	99	147	*bear*	312	146
101	*board*	446	101	148	*cave*	310	148
102	*brassiere*	446	101	149	*badminton*	310	148
103	*basin*	446	101	150	*archbishop*	308	150
104	*bean*	444	104	151	*carport*	304	151
105	*ambulance*	444	104	152	*canal*	300	152
106	*cabage*	440	106	153	*ammonia*	298	153
107	*café*	440	106	154	*banner*	296	154
108	*bracelet*	424	108	155	*boarding house*	288	155
109	*bible*	422	109	156	*castle*	284	156
110	*apron*	420	110	157	*canyon*	284	156
111	*axe*	418	111	158	*biceps*	280	158
112	*bullet*	412	112	159	*blazer*	274	159
113	*brussel sprouts*	408	113	160	*basilica*	268	160
114	*architect*	406	114	161	*blueprint*	264	161
115	*bishop*	404	115	162	*anchor*	264	161
116	*accountant*	400	116	163	*almanac*	258	163
117	*ale*	398	117	164	*brewery*	252	164
118	*captain*	394	118	165	*alligator*	248	165
119	*canoe*	394	118	166	*camel*	246	166
120	*bolt*	390	120	167	*begonia*	244	167
121	*air hostess*	390	120	168	*aqueduct*	244	167
122	*cage*	386	122	169	*acetone*	236	169
123	*air matress*	380	123	170	*ape*	232	170
124	*baseball*	378	124	171	*barrister*	232	170
125	*bomb*	376	125	172	*Adam's apple*	230	172
126	*bee*	368	126	173	*bayonet*	224	173
127	*barman*	362	127	174	*brocade*	216	174
128	*beggar*	362	127	175	*arsenic*	206	175
129	*bat*	360	129	176	*bangle*	200	176
130	*ant*	360	129	177	*amethyst*	200	176
131	*apricot*	360	129	178	*cameo*	198	178
132	*barn*	356	132	179	*aster*	192	179
133	*attic*	354	133	180	*acetylene*	192	179
134	*bladder*	342	134	181	*airship*	180	181
135	*calf*	328	135	182	*arcade*	180	181

Table 4 (suite)

	WORD	FAMILIARITY SCORE	RANK		WORD	FAMILIARITY SCORE	RANK
183	*breeches*	172	183	188	*battlement*	140	188
184	*bowler hat*	168	184	189	*albatros*	134	189
185	*bison*	168	184	190	*anvil*	134	189
186	*antelope*	168	184	191	*bi-plane*	132	191
187	*boomerang*	142	187	192	*blow pipe*	106	192

APPENDIX

TEST INSTRUCTIONS FOR WORD FAMILIARITY TESTS

This is a test to find out how often you have come in contact with certain words. You will be given a list of words. You are to rate each word according to the number of times you have experienced it by placing a check mark (✓) in one of the five spaces provided for your rating. The five possible ratings are described by the words: VERY OFTEN, OFTEN, SOMETIMES, RARELY, NEVER. This means that you have seen, or heard or used the particular word in writing or speech either:

VERY OFTEN: (You have seen or heard or used the word nearly every day of your life).

OFTEN: (You have often seen or heard or used the word).

SOMETIMES: (You have sometimes seen or heard or used the word but not often).

RARELY: (You have seen, heard or used the word only rarely).

NEVER: (You have never used or heard or seen the word).

There may be some words which you have used or heard more often than you have seen. Or there may be some words which you have seen more often than you have used or heard. In such cases give the word the highest rating of the three.

REFERENCES

1. W.F. Mackey, *Language Teaching Analysis.* London:Longmans 1965.

2. E.W. Thorndike and I. Lorge, *The Teacher's Word Book of 30,000 Words.* New York:Columbia Teachers College 1944.

3. C.C. Fries and A.E. Traver, *English Word Lists.* Washington: American Council on Education 1940,92.

4. R. Michéa, Basic Vocabularies, *New Research and Techniques for the Benefit of Modern Language Teaching.* Strasbourg:Council for Cultural Cooperation 1964, 19-33.

5. H. Kučera and W. Francis, *Computational Analysis of Present-Day American English.* Providence:Brown University Press 1967.

6. W.A. Bousfield et al., *The Connecticut Free Associational Norms.* (Technical Report No. 35) Storrs:University of Connecticut 1961.

7. J. Deese, *The Structure of Association in Language and Thought.* Baltimore:The Johns Hopkins Press 1965.

8. W.A. Bousfield et al., *Cultural Norms for Verbal Items in 43 Categories.* (Technical Report No. 22) Storrs:University of Connecticut 1957.

9. G. Gougenheim et al., *L'élaboration du français fondamental (2e éd.).* Paris:Didier 1964.

10. W.F. Mackey et al., *Le Vocabulaire disponible en France et en Acadie.* Quebec:Centre de traitement de l'information (in the press).

11. Colman O Huallachain et al., *Buntus Gaeilge, Réamhthuarascail.* Dublin:(?) 1966.

12. W.F. Mackey, Trends and Research in Methods and Materials, *Languages and the Young School Child.* London:Oxford University Press 1968, 69-83.

13. A.S. Hornby, The Contextual Procedure-Word Families, *English Language Teaching* I:93-94.

14. R.M. Jones, Situat[illegible]nal Vocabulary, *International Review of Applied Linguistics.* 6:165-173.

15. P. Henrion, Statistique et Vocabulaire, *Les Langues Modernes.* 43:238-245.

16. C.E. Noble, The Meaning-Familiarity Relationship, *Psychological Review*. 60:89-98.

17. P. Fraisse, Fréquence et familiarité du vocabulaire, *Problèmes de Psycho-linguistique*. Paris:Presses universitaires de France 1963, 157-167.

18. R.L. Cooper and L. Greenfield, Word Frequency as a Measure of Bilingualism, *Bilingualism in the Barrio*, Fishman, Cooper, Ma et al. Bloomington:Indiana University Press (in the press).

19. J.A. Fishman, Sociolinguistic Perspective on the Study of Bilingualism, *Linguistics*. 39:21-49.

20. M. West and W.F. Mackey, *The Canadian Reader's Dictionary*. Don Mills:Longmans 1968.