Gerard, Linda D.

INTERACTION OF INFORMATION IN BILINGUAL LEXICAL MEMORY

City University of New York PH.D. 1985

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ___√___.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ___√___
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received _____
16. Other _____

_____

_____

University
Microfilms
International

INTERACTION OF INFORMATION IN BILINGUAL LEXICAL MEMORY

by

LINDA D. GERARD

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1985

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

8/22/85
date

Chairman of Examining Committee

8/22/85
date

Executive Officer

Don L. Scarborough

Arthur Reber

Carl Zuckerman

Supervisory Committee

The City University of New York

Abstract

INTERACTION OF INFORMATION IN BILINGUAL LEXICAL MEMORY

by

Linda D. Gerard

Adviser: Professor Don L. Scarborough

Two major models purport to describe the relationship between a bilingual's two languages. The independence model holds that there are separate linguistic memories for each language. The interdependence model holds that there is a single integrated linguistic memory. The main hypothesis of this study is that lexical information is represented in language specific lexicons and word recognition requires searching the appropriate lexicon.

Because no objective Spanish frequency counts comparable to English norms are available, frequency ratings were obtained for 224 English words from English monolinguals and for these words and their Spanish translations from Spanish-English bilinguals in Experiments 1 and 2. These highly reliable ratings were used to select stimulus words for Experiment 3.

The separate lexicons hypothesis predicts that when a bilingual is focused on one language (the target language), word recognition laten-

cies will be determined by the language specific lexicon. Spanish-English bilinguals were tested in a two part lexical decision word recognition experiment. In Part 1, the target language was either English or Spanish. The word stimuli were: 1) Noncognates (words with differently spelled translations, e.g., 'dog' and 'perro'); 2) Cognates (words with identically spelled translations, e.g., 'actual'); and 3) Homographic-noncognates (words that occur in both languages but with different meanings, e.g., 'red'). Noncognates and cognates have similar frequencies of usage in each language, but homographic-noncognates generally differ in frequency. Part 1 latencies were determined by the frequency of the words in the target language. Part 2 was an unanticipated cross-language transfer block. Latencies were again primarily determined by the frequency of the words in the target language. No cross-language facilitation of noncognate translations was found. However, there was equivalent cross-language facilitation of cognates and homographic-noncognates, (i.e., repetitions of the same spelling pattern across languages). This transfer was independent of the target language and frequency of usage in the target languages. The results of this experiment are consistent with the predictions of the separate lexicons hypothesis.

ACKNOWLEDGEMENTS

I thank my adviser Don Scarborough for his invaluable help in refining ideas and editing the final manuscript. I thank Professors Art Reber, Eric Heinemann, Gay Snodgrass, and Carl Zuckerman for their helpful comments and suggestions. I am grateful to Michelle Fisher for her assistance in transcribing the data in Experiments 1 and 2.

Special thanks go to my husband Harvey Meiderbach and my parents Catherine and Edward Gerard.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

A central issue in bilingual memory research concerns the nature of the relationship between a bilingual's two languages. Two major competing models purport to describe this relationship. The independence or separate hypothesis holds that there are distinct memories for each language so that information processing in one language will not automatically affect processing in the other language (Kolers, 1963, 1966a; Tulving & Colotla, 1971). In contrast, the interdependence or shared hypothesis holds that there is a single integrated memory (McCormack, 1977). It is not obvious which of these two models provides the most adequate description of bilingual linguistic memory. Recent research has yielded some apparently conflicting experimental evidence.

A number of experimental findings are consistent with the hypothesis of a single integrated memory. The Stroop interference task has been used to show interaction between a bilingual's languages (Dalrymple-Alford, 1968; Dyer,1971; Fang, Tzeng & Alva, 1981; Preston & Lambert, 1969). In the usual monolingual Stroop task, color words are printed in different ink-colors (e.g., the word 'blue' is printed in red or green ink). When subjects are asked to name the ink-colors, considerable interference is found in this condition. Ink-color naming latencies are longer and error rates are higher than in control conditions (e.g., when the color word and ink-color match, or when the word is a non-color word). In the bilingual version of the Stroop task,

color words are shown in one language (e.g., French: 'rouge', 'bleu') and subjects name the ink-color in the other language (e.g., English). Considerable interference is found, though not as great as in the monolingual task, when the color word does not match the ink-color. The amount of interference appears to depend on the extent to which the color words are graphemically and morphemically similar to the ink-color names in the response language.

Some studies of free-recall of mixed-language word lists have also found evidence for language integration (Kolers, 1966b; Glanzer & Duarte, 1971). These studies demonstrated that presentation of a word's translation can improve recall of that word as much as repetition of the word itself. Lopez & Young (1974), in a free-recall task employing unilingual lists, also found evidence for language integration. Initially their subjects read aloud a unilingual 'familiarization' list of either English or Spanish words. Next, subjects learned a unilingual test list either in the same or different language. The 'familiarization' list for control subjects was composed of words unrelated to the words on the test list, whereas the 'familiarization' list for the experimental subjects was identical to the test list in the same-language condition or translations of the test list in the cross-language condition. Equivalent positive transfer (positive transfer defined as the increase in mean number of words correctly recalled in the experimental conditions relative to the controls) was found in both the same-language and cross-language conditions.

A similar conclusion of language integration also appears to follow from a number of word classification studies in which the depen-

dent variable is decision latency. McCormack (1977) reported an unpublished study conducted in his laboratory by Colletta employing a modification of the Posner (1969) 'same-different' decision task. Bilingual subjects were required to decide, as quickly as possible, whether the two words of a word-pair were synonyms. Word-pairs were presented in the same language (English-English; French-French) or different languages (English-French; French-English). Colletta reasoned that if there were a single integrated linguistic memory, decision latencies for the same- and mixed-language trials should be equivalent. However, if there were separate language memories, then mixed-language decision latencies should be longer than same-language latencies because an extra translation step between languages would have been necessary prior to the synonymity judgment. Mean decision latencies for same- and mixed-language trials were found not to differ. Carramazza & Brones (1980) reported a semantic classification study in which bilinguals were required to decide whether a given word was an exemplar of a particular category. Category names were presented in either English or Spanish and the words to be classified were presented in either the same or different language. Carramazza & Brones found that mean decision latencies for same- and mixed-language trials did not differ and, using the same line of reasoning as Colletta, concluded in favor of a single integrated linguistic memory.

In contrast to this research, evidence for a degree of language specificity in the way information has been stored in memory has been obtained in other tasks. Kolers (1963) obtained associations to words and their translations from fluent bilinguals. In many cases the asso-

ciations given to a particular word differed markedly from the associations given to the word's translation. Kolers suggested that a word and its translation may have somewhat different conceptual representations.

Considerable evidence that bilinguals can remember the language of stimulus presentation has been obtained in a number of studies. In free-recall studies using mixed-language word lists, it has been found that words tend to be recalled in the correct language (Lambert, Ignatow & Krauthamer, 1968; Rose & Carroll, 1974). Rose, Rose, King, & Perez (1975) required subjects to recall English and Spanish sentences. Given that a sentence was correctly recalled (i.e., the meaning was correct), it was typically recalled in the correct language.

Kintsch (1970) found evidence for language specificity in memory in a continuous recognition task using mixed-language word lists. In one condition, German-English bilinguals responded positively if a word had been previously presented in the same language. Discrimination was excellent: 94% of same-language repetitions were correctly recognized and 92% of new words and translations were correctly rejected. In another condition, subjects responded positively if a word had been previously presented in either the same language or as a translation in the other language. Discrimination between old and new items was excellent, however, recognition was better for same-language repetitions as compared to translations. Rose et al. (1975) and Saegert, Hamayan, & Ahmar (1975) using different recognition tasks also found language specificity. Rose et al. had subjects read English and Spanish sentences. Each sentence was shown in a 4AFC recognition task

along with its translation, a new sentence in the same language, and a new sentence in the other language. Given that the meaning of a sentence had been correctly recognized, subjects recognized the language of initial presentation 79% of the time. Saegert et al. had subjects study and recall a list of French and English words. The recognition list consisted of the Arabic translations of the words on the initial list plus distractors. Given that subjects recognized an Arabic word as the translation of a word in the initial list, subjects were able to identify the initial language of presentation 83% of the time.

A similar conclusion of language specificity in memory also follows from research using the release from proactive interference paradigm. In this immediate memory task, the probability of a word being correctly recalled drops substantially over several trials when the stimulus words have been sampled from a particular semantic category (e.g., the words are all animal names). A change in the category of the stimulus words (e.g., the word encountered on trial 4 is the name of a fruit) substantially increases recall probability on that trial. The usual interpretation of this phenomenon is that the new stimulus has been encoded differently and thus its memory representation is somewhat resistant to interference produced by previous learning. Goggin & Wickens (1971) and Dillon, McCormack, Petrusic, Cook, & Lafleur (1973), using bilingual subjects, found substantial release from proactive interference with a change in language of the stimulus words without a change in semantic category of the words. This result suggests that language specific information has been encoded and retained in memory.

It would seem that neither a completely separate nor a completely integrated model provides an adequate description of bilingual linguistic memory. A more sophisticated view of bilingual memory has been proposed based on the consideration that certain types of linguistic operations may depend on the language being processed. It has been suggested that there is an integrated semantic memory in bilinguals but there may be separate language specific processes involved in encoding and/or retrieval of a word in the bilingual's lexicon(s) (MacNamara & Kushnir, 1971; Meyer & Ruddy, 1974). Meyer & Ruddy (1974) reported an experiment on the bilingual semantic facilitation effect that gave evidence both for semantic integration and language specific processes in word recognition. A previous monolingual study by Meyer & Schvaneveldt (1971) found that words are recognized more quickly if they are presented in the context of a semantically related word as compared to an unrelated word. For example, the word 'doctor' is recognized more quickly if it occurs in the context 'nurse' as compared to 'house'. Meyer & Ruddy required German-English bilinguals to make lexical decisions to pairs of items. Subjects responded positively if both items were words regardless of language and negatively if one or both items were nonwords. Pairs of semantically associated words, whether from the same (e.g., 'horse-cow'; 'pferd-kuh') or different (e.g., 'horse-kuh'; 'pferd-cow') languages were recognized quicker than pairs of unrelated words. The amount of facilitation for same-language and mixed-language pairs was equivalent suggesting an integrated semantic memory. However, mixed-language word-pairs were recognized significantly slower than same-language pairs, a result that

seems at odds with the lack of overall effect of mixing the languages of word-pairs in Colletta's synonymity judgment experiment. Further analysis also showed that same-language German pairs were recognized slower than same-language English pairs. These additional results provide strong evidence for language specific encoding/retrieval operations in lexical access.

Similarly, Hines (1978) reported evidence for language integration and language specificity in a 'same-different' judgment task. Subjects judged, as quickly as possible, whether English word-pairs or French word-pairs belonged to the same (e.g., 'cat-dog'; 'chat-chien') or different (e.g., 'cat-foot'; 'chat-pied') semantic categories. Before a word-pair was presented either a neutral prime (XXXXX) or a superordinate category name in the same or different language that might be a valid cue to the category of the word-pair was shown. Hines found equivalent effects of priming both within and across languages supporting semantic integration. However, large differences in decision latencies to English and French word-pairs were also found leading Hines to conclude that orthographic and/or phonetic codes for words in the two languages are differentially accessible.

Kolers and Gonzales (1980) noted that bilingual memory research has often neglected to consider whether bilinguals can exercise strategic control over memory access. Task demands can control encoding and retrieval strategies so that bilinguals may show evidence for language integration in some circumstances but not in others. Thus bilingual memory research should include an investigation of what types of tasks encourage language integration or independence.

Recent research by Kirsner, Brown, Abrol, Chadha, & Sharma (1980) using Hindi-English bilinguals, and Scarborough, Gerard, & Cortese (1984) using Spanish-English bilinguals, considered whether subjects would show automatic language integration in a lexical decision word recognition task under circumstances where there was no reason to integrate the two languages. In these experiments subjects judged, as quickly as possible, whether a letter string was a word in a particular language. In the first block of trials, subjects made lexical decisions to words from one language. Then, in an unanticipated transfer block, subjects made lexical decisions to words in the other language. Words in the transfer block included translations of words encountered in the first block. Under these circumstances, where subjects had no reason to actively translate the words in the first block, there was no facilitation of decision latencies to translations in block two. Rather, decision latencies to translations were equivalent to latencies to words being seen for the first time in the experiment. In contrast, exact repetition of a word within a language block produced substantial facilitation of decision latency. These experiments clearly demonstrate that priming in a lexical decision task is one type of language interaction that does not occur when subjects are focused on one language at a time.

Kirsner, Smith, Lockhart, King, & Jain (1984) using French-English bilinguals, considered whether subjects would show language integration in a mixed-language lexical decision experiment. In a mixed-language lexical decision subjects are shown words from both languages (e.g., 'chien', 'tete'; 'foot', 'cat') and distractor strings that are

derived from words from both languages but are not words in either language. Subjects make a positive response if a letter string is a word in either language and a negative response if a letter string is a nonword. The experiment consisted of two blocks of mixed-language lexical decision trials. The second block was an unanticipated transfer block in which some words from the first block were repeated exactly (e.g., 'chien' - 'chien'; 'dog' - 'dog'), and some words from the first block were shown as translations (e.g., 'tete' - 'head'; 'cat' - 'chat'). In the transfer block, latencies for words repeated within language were facilitated whereas no facilitation of translations was obtained. However, Kirsner et al. (1984) also demonstrated that transfer across languages can occur in mixed-language lexical decision. In another experiment, the priming task required subjects to translate words and consider how the word would be spelled in the other language. Under these circumstances, decision latencies to translations in the lexical decision transfer block were facilitated.

Scarborough et al. (1984) reported another experiment that was a rather demanding test of the degree to which word recognition in one language could function independently of (i.e., without facilitation by or interference from) vocabulary knowledge in the other language. In this experiment the bilinguals responded positively to words from one language (the target language). For half the session (pure condition), negative response trials included nonwords derived from the target and nontarget languages. For the rest of the session (mixed condition), negative response trials included nonwords derived from the target language and both words and nonwords from the nontarget

language. If the bilinguals can selectively access information about words in the target language, then in the mixed condition the subjects may fail to recognize words from the nontarget language and reject these words as if they were nonwords. On the other hand, if each stimulus is evaluated with respect to linguistic knowledge in general, then there may be considerable interference when a nontarget word is encountered, producing an increase in errors and an increase in decision latency on these trials. The results of the experiment clearly showed that all subjects, including a control group of English monolinguals, quickly rejected nonwords derived from the nontarget language in the pure condition and words and nonwords from the nontarget language in the mixed condition. More importantly, for all subjects nontarget words were rejected as quickly as nontarget nonwords.

Taken together the experiments of Kirsner et al. (1980), Scarborough et al. (1984), and Kirsner et al. (1984) clearly demonstrate that bilinguals can excercise control over access of information in memory. Bilinguals can process words of a language in a highly language specific manner.

The main experiment (Experiment 3) of the present study is a further strong test of the degree to which word recognition in one language can function independently of linguistic knowledge in the other language. The main hypothesis tested is that lexical information is represented in separate language specific lexicons (mental dictionaries) and that recognizing a word requires searching the appropriate lexicon. In addition, there is a prior process involved in encoding stimulus information into a form usable for lexical access. This en-

coding process may also be language specific to the extent that different languages involve different orthographies (e.g., Chinese and English). The test of these ideas involves the careful selection of words that differ in systematic ways in terms of their familiarity (frequency of usage), and overlap in orthography and meaning in the two languages. The separate lexicons hypothesis predicts that word recognition latencies should reflect the language specific properties of the words. Furthermore, practice at recognizing a word in one language should transfer to word recognition in the other language only insofar as the word shares a common encoding in both languages. The experiment is a two part lexical decision task. In the first part Spanish-English bilinguals are focused on one of their languages. Some of the words in first part of the experiment are words only in the target language that the bilinguals are focused on (e.g., 'dog' or 'perro', noncognate translations that have the same meaning but different spellings in each language). However, some of the words encountered in the first part of the experiment are actually words in both languages (e.g., 'actual', a cognate word that has the same meaning in both languages; 'red', a homographic-noncognate word that has a different meaning in each language and thus a different frequency of usage in each language). The spelling, meaning, and frequency of usage of noncognate, cognate, and homographic-noncognate words vary systematically across the two languages. In the second part, which is an unanticipated transfer block of lexical decision trials in the other language, words include translations and repetitions of cognates and homographic-noncognates. Specific predictions about decision laten-

cies to the three classes of stimulus words are derived below from the separate lexicons hypothesis, and from an integrated lexicon model of bilingual lexical memory.

Experiments 1 and 2 are necessary preliminary experiments. There are no readily available Spanish word frequency norms comparable in scope to the extensive English word frequency norms such as the Kucera and Francis (1967) count. In order to select stimulus words for the main experiment having the desired spelling, meaning, and frequency of usage characteristics across languages it was necessary to get subjective estimates of Spanish word frequency from Spanish-English bilinguals (Experiment 2). However, it was first necessary to determine whether word frequency judgments can be made reliably and to what extent such subjective frequency judgments are related to objective frequency counts. In Experiment 1 English monolinguals made subjective frequency judgments reliably and these subjective judgments were correlated with the Kucera-Francis English frequency count.

Experiment 1

The main purpose of Experiment 1 was to assess the reliability and comparability of subjective word frequency ratings obtained from three different scaling methods that researchers have used to investigate word frequency and other attributes of English words.

The first method, called categorization (CAT), is based on the general method used by Paivio (1965) and Paivio, Yuille and Madigan (1968) to scale the attributes of imagery and concreteness for a sample of 925 English nouns. This procedure employs a 7-point rating scale for the dimension of interest. In the present experiment, each stimulus word is rated on a 7-point subjective word frequency scale where a 1 represents extremely low frequency (rare) words and a 7 represents extremely high frequency (common) words. While Paivio et al. (1968) did not evaluate word frequency, the simplicity of this technique and the high intergroup reliability of average ratings found for both the imagery task (n = 15, r = .94) and concreteness task (n = 14, r = .94) suggest that the categorization method is both practical and useful for scaling word attributes.

The second method employed in this study is the Steven's (1956, 1958) subjective magnitude estimation (MAG) procedure. Both Shapiro (1969) and Carroll (1971) have used this method in scaling relative word frequency for a set of 60 English words spanning the entire range

of objective word frequency. Shapiro (1969) reported a minimum reliability for average ratings (inter-rater reliability) of .943 in his subject groups (ns = 20). Carroll (1971) logarithmically transformed each subject's magnitude estimates and converted the log values to z-scores to ensure comparable distributions of the transformed estimates across subjects. The data thus transformed yielded a reliability of .974 for the average ratings in the sample of college educated adults (n = 13).

The third method derives from a procedure that Carroll and White (1973a) used to obtain direct estimates of word frequency in logarithmic form for a sample of 65 rare and recent origin English words that did not appear in all the available objective word frequency counts. Ratings in this method, herein called line scaling (LIN), are obtained by providing the subject with a vertical line subdivided into equal intervals. The intervals represent equal log frequency, and this scale is marked at equal log intervals with examples of words of an appropriate objective log frequency. For each new word, the subject marks the position on this scale where he estimates the stimulus word should fall. Despite the restricted frequency range of their stimulus set, Carroll and White (1973a) found a reliability of .92 (n = 20) for the average ratings using this method.

In the first experiment, the three rating methods were compared to each other and to objective frequency measures. Both Shapiro's (1969) and Carroll's (1971) studies using the MAG procedure suggest that college educated adults can accurately estimate objective frequency, where accuracy is defined as the correlation between the sub-

jects' judgments and an objective word frequency count, (e.g., Kucera and Francis, 1967). In Shapiro's and Carroll's studies the correlations between average subjective and objective frequency when both were in log form ranged from .920 to .958. To date, judgments obtained by the other subjective methods have not been related to objective measures of word frequency.

Method

Subjects. Eighteen English speaking monolinguals (10 female) recruited from the Brooklyn College student population each served in a single experimental session of approximately two hours duration. All subjects were students satisfying an introductory psychology requirement. The subjects' experience with Spanish was limited to, at most, a single high school course.

Design. Within each experimental session a subject rated 224 English words in terms of subjective frequency of occurrence by each of the three methods: 7-point scale categorization (CAT), subjective magnitude estimation (MAG), and graphic line scaling (LIN). The order in which subjects performed these tasks was counterbalanced across subjects. Carroll (1971) noted that in the MAG procedure the objective relative frequency of the first word encountered on a list tends to establish an anchor point relative to which the remaining words are judged, but this anchoring effect does not seem to influence a subject's accuracy. To control for the effects of the first word encountered, there were three versions of the stimulus list: Version one

started with a low frequency word, version two started with a moderately high frequency word and version three started with a high frequency word. The three list versions were also counterbalanced across subjects.

Stimuli. Because the words in Experiment 1 were also to be used with bilingual subjects in Experiments 2 and 3, the stimuli were chosen to satisfy several criteria. Some of the words were Spanish-English cognates, i.e., words that have the same spelling and meaning in Spanish and English. A second set of words were homographic-noncognates, i.e., words that are spelled identically in Spanish and English but have different meanings. Finally, a third set of words were standard noncognates, i.e., words that are spelled uniquely in each language. The details of the stimulus selection for this experiment (and Experiment 2) are found in Appendix A.

Two hundred twenty-four English words, two to ten letters in length, were selected from The New College Spanish and English Dictionary (Williams, 1968) so as to span the range of objective word frequency. Twenty-three rare or colloquial words were not listed in the Kucera and Francis (1967) printed English frequency norms and were assigned objective frequencies of occurrence of 0.1 per million following a procedure used by Carroll and White (1973b). Based on the Kucera and Francis (K-F) norms, the median frequency of the word set was 25 per million, range 0.1 to 9816. Approximately 57 percent of the words were low frequency i.e., frequency of occurrence less than 50 per million.

The words were quasi-randomized so that word frequency and word-type (cognates and noncognates) were approximately uniformly distributed over the list. The words were computer printed in upper case type, 32 words to the page, and assembled into 7-page booklets. Page order was manipulated so that three versions of the list were formed starting with either a high, medium or low frequency word.

Procedure. Subjects were tested either individually or in small groups. Subjects received both written and verbal instructions for each of the three rating tasks (see Appendix B).

For categorization, subjects were told to use numbers from 1 to 7 to tell how commonly (frequently) each word is used in English, with the number 1 representing extremely infrequent (rare) words and 7 representing extremely frequent words. Subjects were given a copy of the scale for reference during the task. After brief practice, subjects rated each word in the CAT test booklets. Subjects were encouraged to use all 7 scale values in their ratings. At any time during the task subjects were free to inspect their previous ratings and change any given answer.

For magnitude estimation, subjects were told to assign numbers to words to to indicate frequency of occurrence. The familiar word 'doctor' was given as a reference and the subjects were told to assign a value of 100 to it. Then they were told if another word (such as 'is') seemed to be 100 times as frequent, they should give this word a value 100 times as large (i.e., 10,000). If, on the other hand, a word (such as 'echo') seemed to be only a tenth as frequent, they should give this word a value only one tenth as large (i.e., 10). Subjects

were told to use any numbers they wished (whole numbers and fractions or decimals, if necessary) with the exception of negative numbers. After practice, subjects rated each of the words in the MAG test booklets. A brief copy of the instructions, including examples, was given to the subjects for reference during the task. Subjects were free to change or inspect their previous responses at any time during the task.

Figure 1 shows a page reproduced from a LIN test booklet. Each page of the LIN test booklets had several anchor words printed to the left of a vertical line at intervals proportional to the anchor words' log objective (K-F) frequencies. The words to be rated were printed to the right of the vertical line. For this graphic line scaling task, subjects indicated how frequently each to-be-rated word is used in English by drawing a line from the word to its relative position on the vertical scale. Again, subjects were free to inspect or change any of their previous responses.

Figure 1. Example of graphic scale used in LIN task. Anchor words are printed to the left of the vertical line. Log K-F values (in parenthesis) shown for anchor words in Figure 1 did not actually appear on the scale in the test booklets. The figure also shows how a typical subject may have rated the words ELEVEN, TIME and FLUSH.

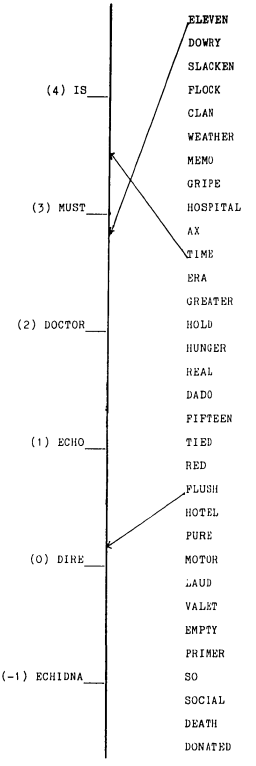

Figure 1

Results

The basic data used in analysis are as follows. The CAT data are the numbers (1 to 7) given to each word by each subject. The underlying metric of the LIN method scale itself is log(10) frequency, therefore the data are the logarithmic frequency values given to each word by each subject in this task. The MAG data are the numbers given to each word by each subject converted to base 10 logarithms. As previously noted, Carroll (1971) converted his subjects' log(10) magnitude estimates to z-scores prior to reliability analysis. In the present study no z-score conversions were made. In contrast to Shapiro (1969) and Carroll (1971), a reference value (i.e., 'doctor' = 100) was used in the present study to anchor the MAG scale. Furthermore, although neither Shapiro (1969) nor Carroll (1971) report reliabilities of log magnitude estimates, the high positive correlations between log magnitude estimates and log objective frequency found in both studies suggests that log magnitude estimates themselves can be highly reliable.

The following summary data were also obtained separately for each of the three rating tasks. For each of the 224 words the average rating and standard deviation (SD) of the ratings were calculated. The log K-F values, average ratings, and SD's of the ratings are given for each word as a table in Appendix D. For all three tasks, the lower the average rating of a given word, the lower the subjective frequency of that word. Table 1 gives, as an example, the average subjective ratings for a low, medium, and high objective frequency English word along with the mean and standard deviation of the ratings for the entire set of 224 words. As previously noted, for purpose of analysis

words that did not appear in the K-F count were assigned an objective frequency of occurrence of 0.1 per million (i.e., log K-F = -1) following Carroll and White (1973b). As is clearly illustrated in Table 1, subjects perceived the low frequency word 'dado' (assigned log K-F = -1) as an extremely low frequency word. The medium frequency word 'guilt' (log K-F = 1.52) was perceived as a higher frequency word than 'dado', and the high frequency word 'was' (log K-F = 3.99) was perceived as an extremely high frequency word.

The reliabilities of the subjective ratings obtained by each of the three scaling methods were determined by Ebel's (1951) analysis of variance method (Appendix C). A summary of the reliability analysis is given in Table 2. The primary interest is the reliability of average ratings (i.e., inter-rater reliability). The average ratings obtained by each method were highly reliable. The essentially identical reliability coefficients for the CAT, MAG, and LIN average ratings (.972, .961, and .970, respectively) indicate the three subjective scales are internally much alike. Based solely on reliability considerations there is no reason to prefer one subjective scaling method over another.

The standard errors of measurement for average ratings given in Table 2 are those obtained by Ebel's method, i.e.,

$$SEav = \sqrt{MSerror/k} \quad ,$$

where MSerror is the residual mean square from the analysis of variance and k is the number of raters. The three standard errors of measurement are more easily interpreted if they can be expressed on a common scale (e.g., log K-F units). To obtain corresponding standard

TABLE 1

Examples of average subjective frequency ratings obtained in Experiment 1. The grand mean and standard deviation of each of the complete sets of average ratings are also shown.

Word		Average Subjective Rating		
	LOG K-F	CAT	MAG	LIN
DADO	-1.000	1.222	0.023	-1.320
GUILT	1.519	4.778	2.267	1.606
WAS	3.992	6.889	3.548	3.429
Grand mean	1.226	4.306	2.018	1.370
SD	1.217	1.309	.757	1.270

errors of measurement in log K-F units the standard errors are multiplied by the following ratio:

$$SD_l/SD_m$$

where SD_l is the standard deviation of the set of 224 log K-F values and SD_m is the standard deviation of the set of 224 average ratings for a given rating task. Standard errors in log K-F units are also given in Table 2. These transformed standard errors imply that on the average the 18 monolingual raters are judging a word's frequency within approximately plus or minus a quarter of a log unit by any of the three rating methods.

It is perhaps more important, however, to consider the following empirical questions. Are the ratings obtained by any one of the scaling methods essentially equivalent to the ratings obtained by the other scaling methods? To what degree are the subjective ratings actually correlated with objective frequency (log K-F)? The intercorrelation matrix for average subjective ratings and log objective frequency is given in Table 3. Inspection of Table 3 not only shows that the average subjective ratings are highly correlated with objective frequency but the three sets of average subjective ratings themselves are highly intercorrelated, suggesting that the three scaling methods are redundant.

If indeed the average ratings obtained by one subjective method are merely linear transformations of average ratings obtained by the other subjective methods, then the methods are interchangeable. Any of the rating methods may be used to obtain subjective frequency ratings. This hypothesis of linear dependency (redundancy) among subjective

TABLE 2

Reliability of individual and average ratings of 224 English words. Standard errors of measurement are given in original scale units and on a common scale (log K-F units). Raters are 18 English monolinguals.

	CAT	MAG	LIN
Individual ratings			
Reliability (r_x)	.659	.578	.644
95% Confidence Limits for r_x	.618-.694	.533-.617	.625-.681
MSerror	.863	.405	.794
Average ratings			
Reliabilty (r_{Av})	.972	.961	.970
95% Confidence Limits for r_{Av}	.967-.976	.954-.967	.965-.975
SE of measurement	.219	.150	.210
SE of measurement(log K-F units)	.203	.241	.210

TABLE 3

Intercorrelation matrix of average ratings obtained by three subjective frequency scaling methods and objective word frequency. All correlations are based on N = 224 words. Subjects are English monolinguals.

	LOG K-F	CAT	MAG	LIN
LOG K-F	1.000	.815	.806	.803
CAT		1.000	.978	.977
MAG			1.000	.972
LIN				1.000

All correlations are significantly different from 0, $p < .001$.

rating methods can be reexpressed as follows: there is a single underlying common factor that accounts for the observed correlations among methods. The correlations among subjective rating methods were factor analyzed and the maximum likelihood solution for a single underlying factor was obtained (Morrison, 1976). The results of this analysis are summarized in Table 4. The statement that all residual off-diagonal correlations are zero indicates that the unique information in each of the three different rating scales does not account for any significant portion of the explained variance. A single underlying subjective frequency factor accounts for the observed correlations among rating methods, thus supporting the hypothesis of redundancy. This result further implies that the observed high correlations between subjective ratings and objective frequency are identical because one set of ratings is merely a linear transformation of any other set.

Given the exceptionally high degree of reliability and precision of the average subjective ratings, it is interesting to predict a word's objective log K-F value from its average subjective frequency value. This may be done by obtaining the three separate regression equations of average subjective frequency on log K-F. The standard errors of estimate for the prediction of log K-F from average ratings were .706, .721, and .726 log K-F units for the CAT, MAG, and LIN methods, respectively. The corresponding 95% confidence limits are ± 1.38, $\pm$ 1.41, and ± 1.42 log K-F units. Thus, a word's frequency as measured directly by its actual log K-F value could vary as much as 1.5 log units from its log K-F value predicted from the average subjective rating of the word. This result suggests a question which is

TABLE 4

Maximum likelihood factor analysis for CAT, MAG and LIN subjective ratings. Solution is shown for one underlying common factor. A factor loading is the correlation of the variable with the factor.

	Factor Loading	Unique Variance
CAT	.991	.017
MAG	.986	.028
LIN	.986	.028

All residual off-diagonal correlations are 0 (see text).

addressed in the discussion section: which measure of word frequency, subjective or objective, is the more valid measure of actual frequency of usage?

Discussion

The data clearly establish the CAT, MAG and LIN scaling procedures as equivalent methods for obtaining subjective frequency ratings. Not only are the average ratings obtained by each method merely linear transformations of each other and highly correlated with log objective frequency, but comparably high degrees of reliability and precision of ratings can be obtained with relatively small groups of raters.

Of the three scaling procedures, the subjects reported CAT scaling as the simplest and easiest task to do. Indeeed, CAT ratings of the list of 224 words were obtained in an average of 15 min, whereas MAG and LIN ratings required an average of 40 and 45 min, respectively. Based on these considerations, the CAT procedure is the most practical subjective scaling method.

To what extent are subjective frequency ratings merely duplicating information contained in objective frequency counts? Carroll (1971) suggests that despite the high correlations between subjective and objective frequency that were found for his groups of lexicographers and college-educated adults using the magnitude estimation procedure, subjective frequency estimates are not measuring exactly the same thing as objective frequency data. This seems a reasonable suggestion. If the log objective frequency variable is partialed out, the

three residual (partial) correlations among subjective ratings are very high. The partial correlation between CAT and MAG is .935, between CAT and LIN is .935, between MAG and LIN is .921 (df = 221, $ps <$ .001). This finding of high residual correlations among subjective ratings further raises the question of which type of measure, subjective or objective, is the more valid measure of actual frequency of usage.

As Carroll (1971) notes, it is well known that objective counts are subject to sampling bias. Objective counts are usually based on relatively small samples of text (e.g., 1 million words in the case of the K-F count) from a limited selection of literature types, whereas an adult has experienced a far greater number of word tokens in a wide variety of written and spoken contexts. Carroll notes that the frequencies of truly low probability words are often overestimated in objective counts. I would also suggest that the frequencies of colloquial or recent origin words that are used in everyday speech and informal writing but are rarely encountered in samples of formal text are underestimated in objective counts. Thus, words that are nominally low frequency according to objective counts may have actual frequencies of occurrence that are higher or lower. Carroll suggests that subjective frequency ratings may be inherently more valid than objective data because a person may be able to discount the various sources of sampling bias affecting objective counts. One aspect of the data supports this notion. If the 224 stimulus words are sorted into ranges of log K-F, it is found that the greatest variability of average ratings among words within a range occurs in the lowest objective fre-

quency range. Additionally, the results of Experiment 3, described below, indicate that subjective frequency ratings predict performance in a word recognition task better than objective frequency ratings.

Experiment 2

One problem facing researchers studying bilingualism is the limited availability of normative language data. One method of estimating the frequency of usage of a particular word in another language, such as Spanish, is to assign to that word the objective frequency value of its English dictionary translation equivalent. This procedure is not without difficulties as, for example, when multiple translation equivalents are given for a particular word, the researcher must depend on his judgment to select the translation equivalent that best matches the desired sense of the word. However, the results of the first experiment indicate that subjective rating procedures present an alternative method for obtaining frequency estimates in other languages.

In Experiment 2, the main questions asked are whether groups of fluent bilinguals can render reliable estimates of word frequency in English and Spanish, by each of three subjective rating methods, and to what extent these ratings correlate with objective measures of word frequency. Additionally, it will be determined whether the bilinguals' ratings of the English list are similar to the monolinguals' ratings (Experiment 1) in terms of reliability and correlation with objective frequency.

Method

Subjects. Eighteen Spanish-English bilinguals (11 female) were recruited from the Brooklyn College student population. Each served in a single experimental session of approximately two hours duration. Subjects were either paid volunteers or students satisfying an introductory psychology requirement.

All subjects completed a questionnaire on language background, current fluency and language usage. The native language of 15 subjects was Spanish (2 of these subjects reported learning English concurrently). Thirteen subjects rated their current English vocabulary as being equal to or greater than their Spanish vocabulary. All subjects could be considered fluent bilinguals on the basis of a 7-point self rating scale where a rating of 4 represents equal fluency in both languages. The median fluency rating was 4.50 (range 3 to 6). With the exception of 2 native English speakers who are Romance language majors, all subjects reported using Spanish primarily to converse with Spanish-speaking relatives and friends. English was generally used in academic and/or business settings. The bilinguals were generally more heterogeneous than the monolinguals in Experiment 1 with respect to age, academic background and general socio-cultural background.

Design. The basic design of this experiment is identical to that of Experiment 1. All subjects rated the stimulus list in terms of frequency of occurrence by each of three subjective scaling methods. Both task order and list version (low, medium or high frequency initial word) were counterbalanced across subjects. One group of 9 sub-

jects received the same English language stimulus list as the monolinguals in Experiment 1 and was asked to judge the English language frequency of each word to determine whether the bilinguals' perceptions of frequency of occurrence of English words was essentially identical to that of monlingual English speakers. The other half of the subjects received the corresponding Spanish language version of the list and rated the Spanish language frequency of occurrence of each word.

Stimuli. The Spanish language version of the stimulus list was derived by translating the English words into Spanish. Cognate words were unchanged. The details of this procedure are in Appendix A.

Procedure. The basic procedure in this experiment is identical to that followed in Experiment 1. All subjects received oral and written instructions in English about how to perform the rating tasks. Subjects assigned to the English condition were cautioned that the list was composed of English words and that the experimenter was interested only in the English language frequency of occurrence of each word. Practice was given using the same English words as in Experiment 1. Subjects in the Spanish condition were cautioned that the list was composed entirely of Spanish words and that the experimenter was interested only in the Spanish language frequency of occurrence of each word. Practice was given using the Spanish translations of the English language practice set.

Results and Discussion

The average rating and standard deviation of the ratings for each word, for each of the three rating methods, for each language, are shown in Appendix D along with the log K-F values of each word. The log K-F values for the Spanish words are the log K-F values for these words' English translations.

Separate reliability analyses of the three types of subjective ratings were done for the bilingual group judging English language frequency of occurrence and the group judging Spanish language frequency of the Spanish translation list. A summary of the reliability analysis is given in Table 5.

Uniformly high reliabilities of average ratings were obtained across methods within each language. Overall, the English raters provided more reliable estimates than Spanish raters. The difference between English and Spanish raters' reliability was significant for both MAG ($z = 2.86$, $p < .01$) and LIN ($z = 2.36$, $p < .02$) but nonsignificant for CAT ($z = .49$) by Fisher's test. The reliabilities of average ratings obtained in Experiment 2 are not directly comparable to those obtained for the monolinguals in Experiment 1 because the number of raters differs in the two studies. The Spearman-Brown prophecy formula (Edwards, 1967) can be used to adjust for the sample size difference. The adjusted reliabilities of average ratings for the bilinguals rating the English list are .950, .960, and .963 for CAT, MAG, and LIN, respectively. Overall, these reliabilities compare favorably with those obtained for the monolinguals being significantly lower than the monolinguals' only for the CAT task ($z = 2.39$, $p < .02$). The adjusted

TABLE 5

Reliability of individual and average ratings of 224 English words and corresponding Spanish translations. The standard errors of measurement for average ratings are expressed in original scale units and on a common scale (log K-F units). Nine different bilingual raters per list.

ENGLISH	CAT	MAG	LIN
Individual ratings			
Reliability (r_t)	.511	.570	.593
95% Confidence Limits for r_t	.462-.533	.522-.610	.547-.633
MSerror	1.435	0.356	0.978
Average ratings			
Reliabilty (r_{AV})	.904	.923	.929
95% Confidence Limits for r_{AV}	.885-.918	.908-.934	.916-.940
SE of measurement	.399	.155	.330
SE of measurement(log K-F units)	.378	.338	.324
SPANISH	CAT	MAG	LIN
Individual ratings			
Reliability (r_t)	.487	.429	.477
95% Confidence Limits for r_t	.438-.531	.380-.473	.428-.520
MSerror	1.426	0.355	1.175
Average ratings			
Reliabilty (r_{AV})	.895	.871	.891
95% Confidence Limits for r_{AV}	.875-.911	.847-.890	.871-.907
SE of measurement	.398	.199	.361
SE of measurement(log K-F units)	.394	.437	.401

reliabilities for the bilinguals rating the Spanish list are .945, .930, and .942 for CAT, MAG and LIN, respectively. These adjusted reliabilities are significantly lower than the monolinguals' (*ps* < .01). However, it is important to note that although the Spanish ratings were somewhat less reliable than the monolinguals' and bilinguals' English ratings, the Spanish ratings were still highly reliable. The language questionnaire indicated that the bilinguals use Spanish primarily as a conversational language and thus they currently do little reading in Spanish. Despite Spanish being the native language of most of the bilinguals, there was also considerable diversity among the bilinguals with respect to the extent of their formal training in Spanish and thus considerable variability among them with respect to reading experience and vocabulary exposure in general. The lower reliability of Spanish ratings may simply reflect the bilinguals' more variable and somewhat more limited experience with Spanish as compared to English.

As can be seen in Table 6, the three types of average subjective ratings are highly intercorrelated in both the English and Spanish conditions. Separate maximum likelihood factor analyses of the intercorrelations among subjective ratings formally confirmed the hypothesis of equivalence of the three types of ratings within each language. As in Experiment 1, we find that the CAT, MAG, and LIN scaling methods can be used interchangeably to obtain subjective frequency ratings.

Overall, English raters' subjective estimates were significantly more highly correlated with objective frequency than Spanish raters' estimates, (CAT, $\underline{z}$ = 3.79 MAG, $\underline{z}$ = 4.18 LIN, $\underline{z}$ = 3.55, *ps* < .001 by

TABLE 6

Intercorrelation matrices for the English and Spanish lists. Variables are average ratings obtained by the three subjective frequency scaling methods and log objective frequency. All correlations are based on N = 224 words. There are 9 Spanish-English bilingual raters per list.

ENGLISH

	LOG K-F	CAT	MAG	LIN
LOG K-F	1.000	.788	.780	.779
CAT		1.000	.947	.951
MAG			1.000	.948
LIN				1.000

SPANISH

	LOG K-F	CAT	MAG	LIN
LOG K-F	1.000	.608	.570	.607
CAT		1.000	.950	.945
MAG			1.000	.940
LIN				1.000

All correlations are significantly different from 0, $p < .001$.

Fisher's test). The bilinguals' English language ratings did not differ significantly from the monolinguals' ratings (Experiment 1) in terms of actual correlation with objective frequency ($ps > .05$). Furthermore, the bilinguals' English ratings were highly correlated with the monolinguals' ratings (minimum correlation .94). Thus, the data indicate that the bilinguals' subjective perceptions of English word frequency were essentially identical to that of the monolinguals.

However, the Spanish ratings do not correlate as highly with objective frequency estimates. The relatively lower correlation between subjective and objective frequency ratings for Spanish as compared to English can not simply be attributed to the somewhat lower reliability of the Spanish ratings. Rather, the bilinguals' subjective perceptions of Spanish word frequency may be quite valid (accurate). The lower correlation may be attributable to the inadequacy of the procedure used to estimate Spanish objective frequency (i.e., Spanish objective frequency estimates were based on K-F values of English translations). In Experiment 3 some evidence will be presented supporting this interpretation. Subjective frequency estimates of Spanish words predict performance in a speeded word recognition task better than estimates based on the K-F frequency of the translations.

Experiment 3

The basic purpose of this experiment is to assess independent and interdependent aspects of bilingual word perception. The experimental task is a simple word recognition task called lexical decision. On each lexical decision trial a subject sees a letter string and must

decide, as quickly as possible, whether the letter string is a real word in a particular language (the target language) or a nonsense item (nonword). Thus, the lexical decision task requires subjects to encode and retrieve from the mental lexicon information about individual words.

Word frequency is a powerful variable affecting word recognition in the lexical decision task (Scarborough, Cortese & Scarborough, 1977). Even though two words are well known to a subject, the more common of the two will be recognized more quickly. In general, this word frequency effect is used in the experiment to test several hypotheses about the organization and access of bilingual lexical memory.

The first question to be addressed is whether word recognition in bilinguals is based on language specific knowledge. A bilingual may have lexical information represented in separate lexicons, one for each language and recognizing a particular word requires searching the appropriate language lexicon. In contrast bilingual word recognition may depend on access to a common lexicon shared by both languages. Consider what might happen when a Spanish-English bilingual in a lexical decision in one language (the target language) encounters a letter string that is a word in both languages. In contrast to standard noncognate words that have unique spelling patterns in each language (e.g. the Spanish word for 'dog' is 'perro', and the Spanish word for 'frog' is 'rana'), there are two types of words that have common spelling patterns in both languages: cognates and homographic-noncognates. Cognates (e.g. 'central') have identical meanings and thus sim-

ilar relative frequencies of usage in each language. Homographic-noncognates have different meanings and thus can have different relative frequencies of usage in each language. For example, 'red' is a common color name in English but in Spanish it is a relatively uncommon word meaning 'net'; the word 'fin' is relatively uncommon in English but it is a very common Spanish word meaning 'end', (see Figure 2).

The separate lexicons model predicts that decision latencies will depend on the language specific characteristics of the words in the particular target language that the bilingual is focused on. Specifically, decision latencies should primarily be determined by the frequency of usage of the spelling patterns in the target language. Thus, if the homographic-noncognates 'red' and 'fin' should be encountered by a bilingual in an English-target lexical decision the bilingual should perform as an English monolingual. 'Red' should be recognized quickly with latency comparable to high frequency English noncognate words such as 'dog'; and 'fin' should be recognized more slowly with latency comparable to low frequency English noncognate words such as 'frog'. Should the homographic-noncognates 'red' and 'fin' be encountered in a Spanish-target lexical decision the opposite latency pattern should obtain. 'Red' is a low frequency Spanish word and its latency should be comparable to that of low frequency Spanish noncognate words such as 'rana'; and 'fin' is a high frequency Spanish word and its latency should be comparable to high frequency Spanish noncognate words such as 'perro'. In contrast, if recognition depends on lexical retrieval processes shared at least in part by both languages (i.e., access to a common lexicon) then homographic-noncognate latencies may

Figure 2. Example of homographic-noncognates that differ markedly in frequency of usage as English and Spanish words.

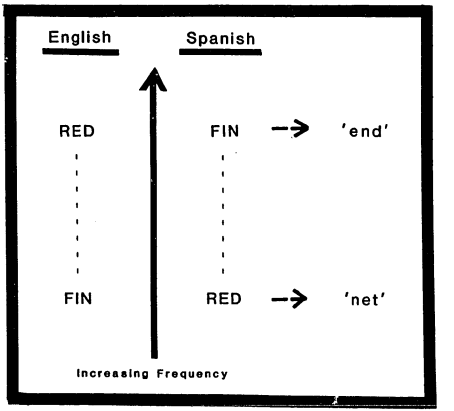

Figure 2

very well depend on the overall familiarity (overall frequency of usage) of these spelling patterns in both languages. Specifically, homographic-noncognates that are high frequency words in one language but low frequency words in the other language might have an overall familiarity (overall frequency of usage as a word regardless of language) similar to that of high frequency noncognate words in each language. Thus, homographic-nocognate latencies in either English- or Spanish-target lexical decisions should be similar to that for high frequency noncognate words. The two models thus make distinctly different predictions about homographic-noncognate latencies as compared to noncognate latencies.

In the first part of this experiment, the question of interest is whether Spanish-English bilinguals, making lexical decisions with one language as the target language, are influenced by the characteristics of the words in the other (nontarget) language, when the target items include single presentations of noncognates, cognates and homographic-noncognates.

Repetition of a word in lexical decision within a language produces a strong facilitation effect: decision times to subsequent presentations are generally faster and more accurate, with low frequency words showing the greatest benefit of prior exposure. Furthermore, this facilitation persists over many intervening trials (Scarborough et al., 1977). Other research has shown that this facilitation is not specific to the lexical decision task (Kolers, 1976; Dixon & Rothkopf, 1979; Scarborough, Gerard, & Cortese, 1979). This within-language repetition effect has two components. Repetition produces faster en-

coding and also faster lexical look-up (Forbach, Stanners & Hochaus, 1974; Kirsner & Smith, 1974; Scarborough et al., 1977; Scarborough et al., 1979). This repetition effect will be used to analyze further bilingual lexical access.

Although there might be separate lexicons, there may be either separate or common encoding processes in bilingual visual word recognition. For Hindi-English or Chinese-English bilinguals it may be obvious that there is language specific encoding for visual presentations of words in each language, however for Spanish-English bilinguals it is not so obvious. The second part of the experiment is an unanticipated block of cross-language transfer trials. These trials include second presentations of cognates and homographic-noncognates, and presentation of translations of noncognate words. If facilitative repetition effects are found in the cross-language transfer conditions an examination of which types of words are facilitated will determine whether such transfer effects are based on repetition of spelling, meaning or some interactive combination of spelling and meaning across languages. In the second part of the experiment, the independent lexicons model predicts that decision latencies for words should be determined by the frequency of usage in the target language. However, this model does not preclude some transfer from the Part 1 presentations based on visual similarity. Specifically, target words that are physically identical repetitions of Part 1 target words (i.e., words spelled identically in both languages: cognates and homographic-noncognates) might be recognized more quickly in Part 2 to the extent that encoding processes are the same for these words in both parts of

the experiment. If so, the degree of facilitation should be equivalent for both cognates and homographic-noncognates. There should be no facilitation of decision times for noncognate translations based on visual similarity.

The predictions for Part 2 derived from the integrated lexicon model are quite different. Again, this model predicts that overall familiarity of the spelling patterns should influence recognition latency. Cross-language transfer should depend in a complex way on the spelling and meaning relationships of the words across the two parts of the experiment. Because this model postulates that the same lexicon is searched for both languages, there may be transfer effects due to facilitation of lexical look-up and encoding facilitation. In general, the prediction is that cognates which share spelling and meaning across languages should show the greatest degree of facilitation, whereas noncognates and homographic-noncognates which share meaning and spelling, respectively, should show a lesser degree of facilitation.

Method

Subjects. Sixteen English monolinguals were recruited from the college student population and served in a single session of approximately 1 hour. One subject was eliminated because of his extensive Romance language background, leaving a total of 15 subjects (12 female). These monolinguals' experience with Spanish was limited to, at most, a single high school course. The monolinguals were paid volunteers or students satisfying an introductory psychology requirement. None of the monolinguals had been subjects in Experiment 1.

Thirty-three Spanish-English bilinguals were also recruited from the college student population and served in a single session of approximately 1 and 1/4 hours duration. One subject was lost due to computer failure and two were eliminated because of high error rates (greater than the a priori criterion of 25%), leaving a total of 30 subjects (18 female). The bilinguals were paid volunteers or students satisfying an introductory psychology requirement. None of the bilinguals had been subjects in Experiment 2.

The additional 15 minutes allotted to each bilingual's experimental session was used to collect background information on language acquisition, current fluency and language usage. The native language of 27 subjects was Spanish (10 of these subjects reported learning English concurrently). The native language of 2 subjects was English (both reported learning Spanish at about age 4). One subject's native language was Romanian but he learned Spanish at about age 4. While 25 subjects rated their current English vocabulary as being equal to or greater than their Spanish vocabulary, all can be considered fluent bilinguals on the basis of a 7-point self rating scale where a rating of 4 represents equal fluency in both languages. The median fluency rating was 4.25 (range 3 to 6). The bilinguals reported using Spanish mainly to converse with Spanish-speaking relatives and friends and English in academic and/or business settings. As a consequence of this differential language use, the bilinguals do little reading in Spanish. In general, the bilingual subjects' dominant language was English. In general, the Spanish-English bilinguals in the college student population are more heterogeneous with respect to age, academ-

ic background and general socio-cultural background than the typical monolingual students participating in research studies. In this experiment monolingual and bilingual subjects were reasonably well matched on these variables.

Stimuli. The critical stimuli for this experiment are cognates, homographic-noncognates differing in frequency of usage in English and Spanish, and noncognates the spelling patterns of which are unique to each language. These stimuli were selected from the word lists which had been subjectively rated for frequency of usage in English and Spanish in Experiments 1 and 2.

Because there was good agreement between average subjective ratings as English and Spanish words for both cognates and noncognates, the original sets of cognates and noncognates were used as stimuli. Thus there were 20 high frequency cognates (median frequency of occurrence 160 per million, range 61 to 2201), and 20 low frequency cognates (median frequency 3, range 0.1 to 12). There were 20 high frequency noncognates (median frequency 238 per million, range 64 to 1599), and 20 low frequency noncognates (median frequency 5, range 0.1 to 35). The frequencies reported here are Kucera-Francis frequencies of English words and dictionary translation equivalents. Some homographic-noncognates expected to differ across languages in frequency of usage on the basis of dictionary translation equivalence were found not to differ according to the average subjective ratings and thus were eliminated as stimulus candidates. Two homographic-noncognate sets were used. One set was composed of 15 words which are high frequency in English (median frequency 320, range 56 to 7250) but lower

frequency in Spanish (median frequency 3, range 0.1 to 169). The other set was composed of 22 words which are low frequency in English (median frequency 1.5, range 0.1 to 53) but higher frequency in Spanish (median frequency 153, range 56 to 9816), giving a total of 117 critical items.

In addition to the critical stimuli, a set of 30 high frequency (median frequency 260, range 60 to 1620) and a set of 30 low frequency (median frequency 7, range 0.1 to 36) noncognate words were selected for use as controls. These controls were divided into three subsets. One subset appeared in Part 1 and the remaining two subsets appeared in Part 2 of the experiment. The particular subset(s) used as controls in each part was counterbalanced over subjects.

The stimuli were presented on a CRT terminal driven by a Digital Equipment Corporation PDP8 computer. The stimuli appeared as uppercase black-on-white letters. A single letter subtended a horizontal visual angle of approximately 0.2^{0} and a vertical visual angle of approximately 0.4^{0} at the normal viewing distance of 61 cm. Thus a 5 letter word spanned a horizontal visual angle of approximately 1.25^{0}.

Design. Table 7 summarizes the assignment of stimuli to Parts 1 and 2 of the experiment for monolingual and bilingual subjects. In Part 1 of the experiment the entire test set of cognates, homographic-noncognates and noncognates was presented. A control set of noncognates seen only in this part was also presented. For monolinguals and one group of bilinguals, the noncognates appeared as English words. The second group of bilinguals saw the noncognates as Spanish words.

The nonwords appearing on negative trials were derived from non-stimulus cognates, homographic-noncognates, and English noncognates for the monolinguals and one group of bilinguals or Spanish noncognates for the second group of bilinguals. Thus for the monolinguals and one group of bilinguals the target language was English in the lexical decision task in Part 1, and the target language for the second group of bilinguals was Spanish.

In Part 2 of the experiment, the monolinguals continued to make an English target lexical decision in which the entire test set of cognates, homographic-noncognates and noncognate words was presented for a second time. A control set of noncognate words was presented for the first time in the experiment to provide an additional baseline for determining within-language repetition (facilitation) effects.

For bilinguals, the target language of the Part 2 lexical decision was the second language. As previously noted, subjects were not told about the Part 2 test involving the other language. For all bilinguals, the entire test set of cognates and homographic-noncognates was presented for a second time. The bilinguals with Spanish as the target language in Part 2 saw the Spanish translations of the noncognates which had appeared as English words in Part 1. Those whose target language was now English saw the English translations of the noncognates which had appeared as Spanish words in Part 1. An additional control set of noncognates (appearing as Spanish or English words as appropriate) was presented for the first time in the experiment to provide an additional baseline for assessing any cross-language repetition (facilitation) effects. The nonwords in this part included

TABLE 7

Sample of stimulus assignments in Experiment 3 for monolingual and bilingual subjects. Sample words shown are high frequency words in English. For nonwords the stimulus type designations refer to the words from which the nonwords were derived.

	Monolingual		Bilingual			
	English to English		English to Spanish		Spanish to English	
Stimulus Type	Part 1	Part 2	Part 1	Part 2	Part 1	Part 2
Words						
Cognates	Actual	Actual	Actual	Actual	Actual	Actual
Homographic-noncognates	Red	Red	Red	Red *	Red *	Red
Noncognates repeated within- or translated across-language	Dog	Dog	Dog	Perro	Perro	Dog
Noncognate controls- single presentations	Chair	Book	Chair	Libro	Silla	Book
Nonwords						
Cognates	Flaza	Flaza	Flaza	Flaza	Flaza	Flaza
Homographic-noncognates	Nemo	Nemo	Nemo	Nemo	Nemo	Nemo
Noncognates	Cardle	Cardle	Cardle	Veba	Veba	Cardle
Noncognate controls	Trenty	Blool	Trenty	Sandre	Liente	Blool

Note. * low frequency in Spanish

similar repetitions of Part 1 nonwords and a new set of nonwords derived from non-stimulus noncognate words.

While this design involves rather long lags between first and second presentations of items, it has previously been demonstrated that within-language repetition effects are robust for both monolingual and bilingual subjects (Kirsner et al., 1984; Scarborough et al., 1977; Scarborough et al., 1984).

Procedure. The experimental session consisted of approximately 600 lexical decision reaction time trials. Subjects were run individually in a single 1-hour session. The subject sat facing the CRT display in a dimly lit booth. A horizontal fixation line appeared at the beginning of each trial. The subject initiated each stimulus presentation by means of a foot switch. After a half second blank interval, the stimulus appeared on the screen. The subject indicated whether each stimulus was a word or nonword by pulling one of two levers. The right lever was used for word responses and the left for nonword responses by all subjects. The subject received immediate feedback as to the correctness of the response. The subject was given information about average response latency and accuracy at regular intervals during the course of the experiment. All subjects were instructed to respond as quickly as possible while trying not to make errors. For all subjects, Part 1 was preceded by a block of 50 practice trials and Part 2 was preceded by a 30 item practice block.

Results

The data for words (correct responses) were subjected to two analyses of variance (ANOVA), one using interactions with subjects and the other using interactions with stimulus words as the basis for error terms. The two ANOVA's were combined using the minF' statistic suggested by Clark (1973). Similar analyses were done separately for the nonword data. In both Parts 1 and 2, overall word error rates were low (6% in the English-target conditions and 11% in the Spanish-target conditions). Error rates covaried with latency showing no indication of a speed/accuracy trade-off. Because the analyses of error rates (arcsine transform) basically duplicated the results of the latency analyses, the results of the error analyses will generally not be reported.

First presentations - English-target conditions. Figure 3 shows the mean decision latencies and error rates for correct word responses in Part 1 for the English-target bilinguals and monolinguals.

The variables of interest in the analysis of the English-target conditions were subject groups (monolinguals and bilinguals), word frequency (high and low frequency), and word-type (cognates, homographic-noncognates, noncognates, and control noncognates). The test set of noncognates are noncognate words that are repeated for monolinguals or appear in translation for bilinguals in Part 2 of the experiment. The control noncognates are the set of noncognate words that are seen only in Part 1.

There was a significant main effect of word frequency (min$F'(1,147) = 69.90$, $p < .001$), low frequency words being recognized

Figure 3. Mean decision latencies and error percentages for correct responses to first presentations of words in the English-target conditions, Experiment 3.

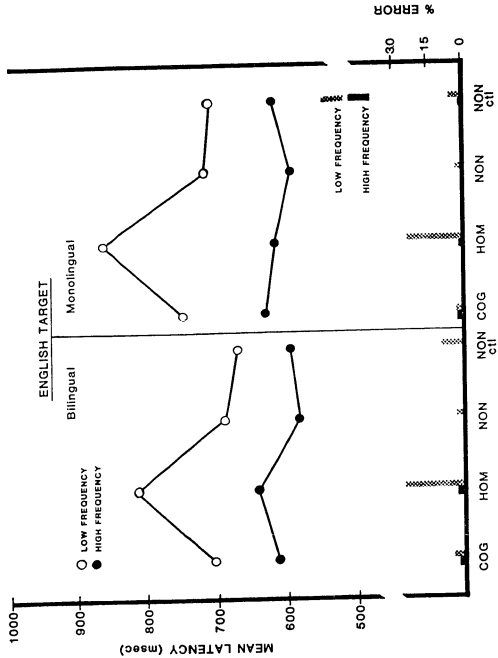

Figure 3

an average of 129 msec slower than high frequency words. Both a significant main effect of word-type ($minF'(3,177) = 8.56$, $p < .001$) and an interaction of frequency and word-type ($minF'(3,162) = 4.24$, $p < .01$) were obtained. Analysis of simple effects showed that the word-type effect and frequency by word-type interaction were attributable to the extremely slow responses to low frequency homographic-noncognates. Low frequency homographs were recognized significantly slower ($minF'(1,178) = 37.80$, $p < .01$) than low frequency cognates and noncognates. Low frequency cognate and noncognate latencies did not differ. Responses to the low frequency homographs are extremely slow simply because these words are very much lower in English frequency of usage (as measured by log K-F or subjective frequency) than words in the other low frequency groups. There were no significant differences among high frequency cognates, noncognates, and homographic-noncognates. The important point is that these results obtained both for the monolinguals who do not know Spanish and for the bilinguals. The English-target bilinguals showed the same latency pattern as monolinguals: there was no main effect of the monolingualism/bilingualism factor, $p > .10$, nor were there any interactions of the monolingualism/bilingualism factor with frequency or word-type, $ps > .25$. Similar analysis of the nonword data showed no significant differences between the English-target bilinguals and monolinguals. The mean latency to reject a nonword was 907 msec (14% error) for the bilinguals and 963 msec (14% error) for the monolinguals. Thus in the English-target lexical decision the bilinguals were effectively performing as English monolinguals consistent with the predictions from the separate lexicons model.

First presentations - Bilingual groups. Figure 4 shows the word data for the two bilingual groups. The right-hand panel of Figure 4 shows the data for the Spanish-target condition. It had been found that the English-target bilinguals showed the same latency pattern as monolinguals. The Spanish-target condition was analyzed separately to test the predictions from the two models.

Inspection of Figure 4 shows that homograph latencies are somewhat slower than noncognate latencies and cognate latencies are somewhat faster than noncognate latencies in the Spanish-target condition. However, planned contrasts among the cognates, homographic-noncognates and noncognates in the Spanish-target condition showed that the high frequency cognates, homographs and noncognates did not differ significantly, $ps > .10$. Low frequency homographs and noncognates did not differ significantly, $p > .10$. The homographs in the Spanish-target condition are somewhat less frequent than the nominally frequency matched Spanish noncognates. Both objective (log K-F) and subjective frequency accurately predict the homograph latencies. On the other hand, low frequency cognates were recognized 94 msec faster than low frequency noncognates, a difference that was marginally significant ($minF'(1,198) = 3.43$, $.05 < p < .10$).

Overall, the Spanish-target bilinguals responded 205 msec slower than the English-target group. The effect of word frequency was larger for the Spanish-target group, a mean of 210 msec as opposed to the mean of 114 msec for the English-target group. The overall slower decision latencies for the Spanish-target group perhaps reflects the fact that, although the bilinguals are proficient in Spanish, Spanish

Figure 4. Mean decision latencies and error percentages for correct responses to first presentations of words in the bilingual conditions, Experiment 3.

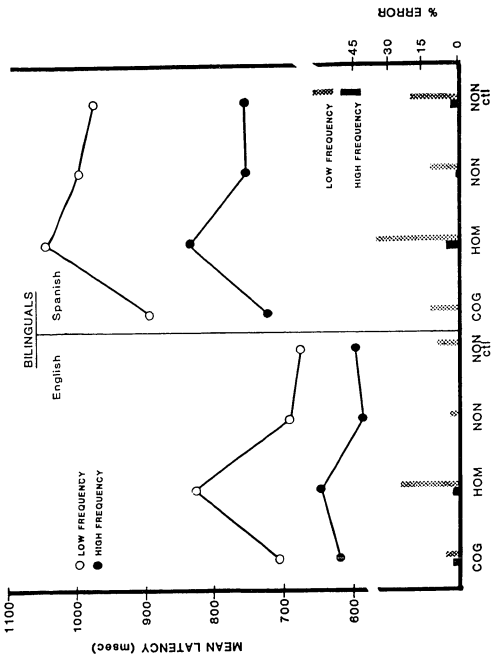

Figure 4

is primarily a conversational language for them and they do little reading in Spanish. Larger frequency effects for slower subjects are typically found in lexical decision. As in the English conditions, nonword latencies were longer than word latencies. The average latency to reject a nonword in the Spanish-target condition was 1279 msec (26% error).

Despite the overall difference in mean latency in the Spanish-target condition, the fact is that homographic-noncognate latencies in the Spanish-target condition were being determined by these words' frequency of usage as words in Spanish, consistent with the prediction from the separate lexicons model. A planned comparison of the two sets of homographic-noncognates in the English- and Spanish-target conditions, shown in Table 8, clearly indicates that decision latency for homographic-noncognates was primarily determined by frequency of usage of these words in the target language. English-target bilinguals were slow to recognize words such as 'fin' that are uncommon in English, but were 182 msec faster to recognize high frequency English words such as 'red'. When these same words appeared in the Spanish-target condition, the opposite pattern was found. Now Spanish-target bilinguals were quicker to recognize words such as 'fin' that are high frequency Spanish words, but were 211 msec slower to recognize words such as 'red' that are low frequency Spanish words. This interaction was highly significant ($\min F'(1,79) = 25.53$, $p < .01$). Error rates also showed a similar interaction.

In summary, English-target bilinguals showed the same latency and error pattern as English monolinguals; the frequency of usage of a ho-

mographic-noncognate as a word in the target language determined decision latencies. The integrated lexicon model is rejected. There are functionally separate lexicons, one for each language.

Subjective Frequency Ratings. In Experiments 1 and 2, it was shown that monolinguals' and bilinguals' subjective judgments of word frequency were highly reliable and well correlated with objective counts. On this basis it was proposed that subjective frequency ratings may be used as an alternative to objective frequency counts. It would be of great utility to bilingual researchers to use subjective frequency ratings because of the limited availability of normative frequency data in languages other than English. However, for subjective frequency ratings to be used as an alternative objective counts, it is also necessary to show that subjective ratings are consistent predictors of performance in a task in which objective frequency is known to be a predictor. As previously noted, objective frequency is a strong predictor of word recognition latency in lexical decision. Previous research has shown that decision latency is a linearly decreasing function of log frequency (e.g. log K-F) within a language (Scarborough et al., 1977; Scarborough et al., 1979; Kirsner et al., 1984).

Regression analyses of the mean frequency ratings and mean decision latencies for all 117 critical stimuli were done separately for the monolinguals, English-target and Spanish-target bilinguals. As is shown in Table 9, objective and subjective ratings are all good single predictors of decision latency. In the Spanish condition objective frequency is the poorest single predictor of all, a result that is not

TABLE 8

Mean response latencies, in msec, of correct decisions and error proportions (in parenthesis) for first presentation of homographic-noncognates in Part 1 of Experiment 3.

	Target Language		
	English		Spanish
	Monolingual	Bilingual	Bilingual
High Frequency in English but Low Frequency in Spanish e.g. 'red'	620 (.01)	646 (.02)	1048 (.36)
Low Frequency in English but High Frequency in Spanish e.g. 'fin'	866 (.24)	828 (.26)	837 (.04)

surprising because the objective frequencies of Spanish words were not obtained from Spanish norms but were estimated from log K-F values of English translations. In summary, on the basis of the traditional validation criteria of reliability, correlation with a standard index, and consistency of prediction, subjective frequency ratings can be used as an alternative to objective frequency counts.

Table 9 further shows that any of the three subjective indices is a better single predictor of decision latency than log K-F. Specific conditional tests based on stepwise analysis show that the objective frequency variable does not account for a significant increment in explained variance beyond the contribution of the single best subjective predictor in all three conditions. On the other hand, if the first variable entered in the regression is objective frequency, then any of the subjective frequency indices adds a significant increment to the explained variance. This additional finding suggests that even though subjective frequency ratings are highly correlated with objective counts, subjective ratings are not necessarily measuring exactly the same thing as the objective counts. Carroll (1971) argued that subjective ratings may be inherently more accurate indices of word frequency because subjects may be able to discount the sources of sampling bias that affect objective word counts. But subjective frequency ratings may not just be more accurate indices of frequency of usage, they might reflect a judgment of accessibility of the word that is influenced not only by frequency of usage, but by such factors as, e.g., recency of usage, typicality of orthography, part of speech, and inflection. If subjective frequency ratings reflect accessibility

TABLE 9

Summary of regression analyses of Part 1 decision latencies using objective and subjective ratings as predictors.

	English				Spanish	
	Monolingual		Bilingual		Bilingual	
Predictor	R	R^2 %	R	R^2 %	R	R^2 %
CAT alone	.775	60	.689	47	.747	56
MAG alone	.754	57	.759	58	.710	50
LIN alone	.761	58	.718	52	.692	48
Log K-F alone	.701	49	.659	43	.552	30
Best subjective and log K-F	.776	60	.760	58	.752	57
All subjective and log K-F	.776	60	.767	59	.759	58

then it is not surprising that subjective ratings are better predictors of performance in lexical decision, a task that clearly reflects accessibility. While it may not be possible to determine completely whether subjective frequency ratings are influenced by factors other than actual frequency of usage, certain aspects of the subjective rating and latency data suggest that the subjective ratings do not reflect general accessablity. If one assumes that cognates and homographic-noncognates are generally more familiar than noncognates for bilinguals because cognates and homographs are words in both their languages, then the frequency ratings for these words, obtained in Experiment 2, might well be influenced by this increased familiarity. If so, then subjective ratings of the cognates and homographs should be systematically higher than ratings of comparable frequency noncognates. Analysis of the CAT ratings of the stimulus words in Experiment 3 showed that neither cognates nor homographs were rated higher than comparable frequency noncognates in either language. Furthermore, bilinguals' ratings of the stimuli as English words were identical to the monolinguals' ratings. Bilingual's ratings of the Spanish stimuli tended to be somewhat higher than monolinguals' ratings of the English translations. But, this tendency obtained for cognates, homographs, and noncognates alike.

In the English-target lexical decision, monolinguals and bilinguals did not differ. In the Spanish-target lexical decision it had been found that cognates were recognized faster than noncognates, whereas homograph latencies did not differ from noncognates. Cognate latencies in this condition were predicted from the log K-F values and

from CAT ratings, using noncognate latency and frequency as baseline. If the subjective ratings reflect general accessibility, then one would expect the subjective ratings to yield a consistently better prediction of latency for cognates. Neither objective nor subjective frequency was a good predictor of cognate latency. Cognate latency deviated 98 msec from the prediction based on log K-F and 103 msec from the prediction based on CAT ratings. Some possible explanations of why cognate latencies in Spanish are faster than expected will be considered in the discussion section.

Transfer effects. Figure 5 shows the Part 1 (open circles) and Part 2 (closed circles) data for nonwords and words, averaged over frequency for words, for the monolinguals and the bilinguals. The data for bilinguals are averaged over language.

For monolinguals, the test sets of cognates, homographic-noncognates, and noncognates were repeated in Part 2. As noted above, repetition of a word within a language produces faster encoding and faster lexical search. Therefore, monolinguals should be faster in Part 2 for all repeated stimuli regardless of word-type: cognates and homographic-noncognates are known the monolinguals only as English words. Also, low frequency words are expected to show the largest effect of repetition as has been found in previous studies. This is because repetition affects a word's familiarity (i.e., relative frequency). The enhanced familiarity produced by repetition leads to faster lexical look-up within the mental lexicon. The enhanced familiarity is relatively greater for low frequency words. Figure 5 shows that repetition of a word within language substantially decreased response la-

tency and improved accuracy for all three word-types. The average facilitation effect of word repetition was 38 msec ($minF'(1,32) = 18.10$, $p < .001$). As expected, the effect of repetition was somewhat greater for low frequency words, 54 msec as opposed to 22 msec for high frequency words. However, this repetition by word frequency interaction was marginally significant ($minF'(1,37) = 3.73$, $.05 \quad p < .10$). No other interactions with the repetition factor were significant. Although responses to new noncognate control words in Part 2 were 18 msec slower than in Part 1, this difference is not significant.

In the Part 1 bilingual conditions, it was found that homographic-noncognate decision latencies were primarily determined by the frequency of usage of these words in the target language consistent with the separate lexicons model. Similarly, as is clearly shown in Table 10, homographic-noncognate latencies in Part 2 were still primarily determined by the frequency of usage of these words in the current target language. In Part 2, English-target bilinguals were slow to recognize words such as 'fin' that are uncommon in English, but were 182 msec faster to recognize high frequency English words such as 'red'. When these same words appeared in the Spanish-target condition, the opposite pattern was found. Now Spanish-target bilinguals were quicker to recognize words such as 'fin' that are high frequency Spanish words, but were 132 msec slower to recognize words such as 'red' that are low frequency Spanish words. This interaction was highly significant ($minF'(1,63) = 23.53$, $p < .01$). Error rates also showed a similar interaction. Thus, the Part 2 bilingual data reinforce the conclusion that lexical information is represented in functionally separate language specific lexicons.

Figure 5. Mean decision latencies, averaged over frequency for words, and error percentages for correct responses in Experiment 3, Parts 1 and 2. The left panel shows the effect of within-language repetition for monolinguals. The right panel shows the effect of cross-language repetition (cognates and homographs) or presentation of noncognate translations for bilinguals averaged over language.

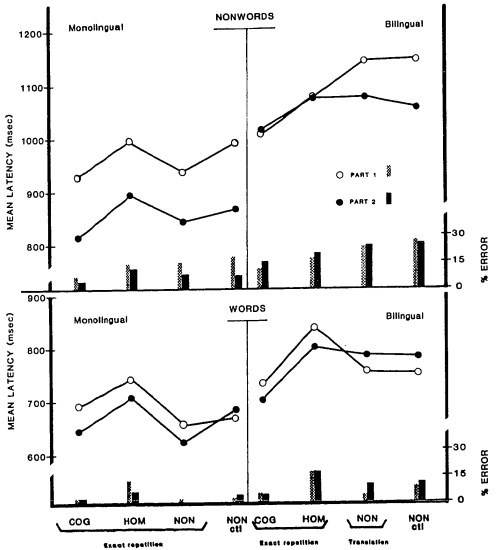

Figure 5

TABLE 10

Mean response latencies, in msec, of correct decisions and error proportions (in parenthesis) for re-presentation of homographic-noncognates in Part 2 of Experiment 3.

	Target Language		
	English		Spanish
	Monolingual	Bilingual	Bilingual
High Frequency in English but Low Frequency in Spanish e.g. 'red'	605 (.00)	702 (.04)	879 (.28)
Low Frequency in English but High Frequency in Spanish e.g. 'fin'	815 (.13)	884 (.32)	747 (.05)

For bilinguals, the test sets of cognates and homographic-noncognates were repeated in Part 2, and the translations of the test set of noncognates were shown. Little or no cross-language transfer is expected for bilinguals because the other language lexicon is being accessed in Part 2. However, if there is any cross-language transfer, such transfer should be restricted to cognates and homographic-noncognates because the bilinguals will have encoded the same visual pattern in both parts of the experiment.

The average effects of repeating the same orthographic pattern (for cognates and homographic-noncognates) or of presenting a translation (for noncognates) are shown in Figure 5. In contrast to the monolingual data, a significant repetition by word-type interaction was obtained ($\min F'(2,128) = 3.95$, $p < .05$). Repetition of cognates and homographic-noncognates decreased response latency by an average of 33 and 37 msec, respectively. There was no facilitation effect of prior presentations of noncognate translations. Rather in Part 2, noncognate translation latencies were 29 msec slower than Part 1 latencies. Separate analysis of the control noncognates showed that responses to new words in Part 2 were slower by 32 msec, similar to the effect seen for translations. The average repetition effect for cognates and homographs in the monolingual condition (40 msec) was similar to the repetition effect for these word sets in the bilingual conditions (35 msec). However, for the monolinguals the size of the repetition effect was somewhat greater for low frequency words, whereas for bilinguals the repetition by frequency interaction was not significant. The failure to find any cross-language transfer for noncognate trans-

lations is also consistent with the separate lexicons hypothesis. On the other hand, the cross-language transfer found for cognates and homographic-noncognates suggests that such transfer is an encoding facilitation based on the common spelling of these items in the two parts of the experiment. These results are examined in more detail below.

Language (English and Spanish) was included as a factor in the word-based ANOVA for transfer effects, whereas, transfer group (English-to-Spanish and Spanish-to-English) was included as a factor in the subject-based ANOVA. In both Parts 1 and 2, Spanish latencies were slower than English latencies. Because the language factor in the word-based ANOVA is not directly comparable to the group factor in the subject-based ANOVA, the foregoing minF' analysis of the bilingual conditions effectively required the data be averaged over target language. Thus any target-language specific repetition effects were not detectable in the analysis. Supplementary subject-based analyses using the control noncognate latencies as a baseline condition were undertaken to clarify the relationship between target language and repetition. The data for these analyses are difference scores generated by subtracting, for each subject, the high and low frequency control noncognate mean latencies from each of the corresponding high and low frequency cognate, homograph, and noncognate (translation) mean latencies. A comparison of the baseline-adjusted latencies for Parts 1 and 2 yields a measure of the repetition effect relatively uncontaminated by any individual subject's overall response latency and by general latency changes from Part 1 to Part 2.

In the monolingual analysis the 56 msec net effect of repetition was significant ($F(1,14) = 31.85$, $p < .001$). The repetition effect was greater for low frequency (94 msec) than for high frequency (17 msec) words ($F(1,14) = 12.39$, p .01). No other interactions with the repetition factor were significant. In the bilingual analysis the variables of interest were target language, repetition, word-type, and frequency. If indeed the cross-language repetition effect is independent of both the target language and the frequency in the target language then no interactions among the repetition, frequency, and target language factors should obtain. The 45 msec net effect of repetition was significant ($F(1,56) = 7.63$, $p < .01$). In contrast to the monolingual analysis the repetition by word-type interaction was significant ($F(2,112) = 5.73$, $p < .01$). Analysis of simple effects showed a clear cross-language repetition effect for both cognates (64 msec) and homographs (69 msec), that did not differ from each other. There was no facilitating effect of prior presentations of noncognate translations. Importantly, the repetition factor did not interact with the target language or frequency factors, $ps > .35$. Thus the cross-language repetition effect found for cognates and homographs was independent of target language and frequency within target language. This more extensive analysis of the cross-language repetition effect indicates that this effect is an encoding facilitation effect based on the visual identity of the cognates and homographs in both parts of the experiment.

The effects of nonword repetition are also shown in Figure 5. The monolinguals rejected repeated nonwords significantly faster in

Part 2 by an average of 105 msec (minF'(1,17) = 14.37, $p < .001$). The monolinguals also rejected the control (unprimed) nonwords 124 msec faster than the Part 1 controls (F(1,14) = 7.92, $p < .05$). Overall, the monolinguals became more proficient at rejecting nonwords in Part 2. The reason for this result is unclear but Scarborough et al. (1984, Experiment 1) report a similar finding for Spanish-English bilinguals in an English-target lexical decision. In their study, both repeated and new nonwords in Part 2 were 87 msec faster than nonwords in Part 1.

In contrast the bilinguals showed no overall effect of nonword repetition, $p > .10$, although translation derived nonwords were 65 msec faster in Part 2, similar to the effect seen for control nonwords. The bilinguals, like the monolinguals, rejected the control nonwords significantly faster (93 msec) in Part 2 (F(1,28) = 4.47, $p < .05$).

Discussion

The inclusion of orthographic patterns that are words in both languages provides for a strong test of the degree to which word recognition in one language can function independently of second language knowledge. Overall, the data are consistent with the predictions from the separate lexicons hypothesis, indicating that in lexical decision bilinguals can access memory in a highly language specific way.

In the first part of the experiment, English-target bilinguals showed no influence of their knowledge of the cognates and homographs as Spanish words. Indeed, the English-target bilinguals appeared to

be functioning as English monolinguals, showing the same latency and error pattern as the monolingual control group. Homograph latencies in the English-target condition were determined by the frequency of these words as English words. Similarly, the Spanish-target bilinguals showed no influence of their knowledge of the homographs as English words. Homograph latencies in the Spanish-target condition were determined by the frequency of these words as Spanish words. In the cross-language transfer block homograph latencies were primarily determined by the frequency of these words in the current target language.

There was no facilitation of noncognate translations in the cross-language transfer block, whereas repetition of the same orthographic pattern across languages produced equivalent facilitation of cognates and homographs. These results are consistent with the notion that transfer found for cognates and homographs is a general facilitation effect based on the visual identity of these words in both blocks of the experiment. Within-language repetition produced facilitation that was substantially greater for low frequency than for high frequency words.

Lexical decision requires three major processing stages: encoding, memory search (lexical look-up), and decision/response. Word frequency is generally considered to affect the search stage (Becker & Killion, 1977; Glanzer & Ehrenreich, 1979; Scarborough et al., 1977; Scarborough et al., 1979; Stanners, Jastrzembski & Westbrook, 1975). One explanation of the experimental results is that a bilingual's lexical memory is organized into separate lexicons, one for each lan-

guage, and that a bilingual can selectively access a particular lexicon. Presumably in the first part of the experiment, English-target bilinguals accessed the English lexicon while Spanish-target bilinguals accessed the Spanish lexicon and thus cognates and homographs are unaffected by any potential cross-language associations among these words. In part 2, presumably the bilinguals accessed the other language lexicon, so homograph latencies would be determined by these words' frequency characteristics in the current target language. Selective access further implies that noncognates should be unaffected by prior recognition of these words' translations in the other language.

Repetition can affect both encoding and memory search (Scarborough et al., 1977, Scarborough et al. 1979). Long lag repetition effects in lexical decision appear to be attributable to encoding and recognition processes that are specific to the surface form of the stimulus. Durso & Johnson (1979), Kolers (1975), Monsell & Banich (1982), and Scarborough et al. (1979) found no transfer and Kirsner & Smith (1974), and Monsell & Banich (1982) found markedly reduced transfer when the surface form of the second presentation differed substantially from that of the first presentation. The cross-language repetition effect found for cognates and homographs appears to be a general effect of encoding the same orthographic pattern in both blocks of the experiment. The size of the effect did not differ for cognates and homographs and the repetition effect was independent of target language and frequency within the target language. While the lack of transfer for noncognate translations can be understood in this

context, the lack of any transfer for translations strongly implies that the bilinguals are not responding on the basis of conceptual knowledge independent of the target language.

While I think this account of selective access is largely correct there is, however, one aspect of the data inconsistent with the predictions from the independence model. Cognates in the Spanish-target blocks (the subjects' non-dominant language) were faster than comparable frequency noncognates, regardless of whether word frequency is measured objectively or subjectively, suggesting some cross-language integration. In part 1, the advantage was 34 msec and 94 msec for high and low frequency cognates, respectively. Two recent studies have found a similar cognate advantage in lexical decision in bilinguals' non-dominant language. Caramazza & Brones (1979) found a 42 msec advantage for high frequency cognates and Cristoffanini, Milech, & Kirsner (in press) found a 100 msec advantage for low frequency cognates in English-target lexical decisions with Spanish dominant bilinguals. One explanation of this cognate advantage is that cognates are simply overrepresented in the weaker language lexicon. On the other hand, Caramazza & Brones (1979) interpret this cognate advantage as a semantic priming effect within the context of a model specifying separate lexicons for each language. Both graphemic and phonemic encoding are implicated in visual word recognition (Meyer, Schvaneveldt & Ruddy, 1974; Rubenstein, Lewis & Rubenstein, 1971). Cognates are orthographically identical but phonemically different in both languages. If access can proceed from a graphemic code and all lexical entries consistent with this encoding are automatically accessed (Clark & Clark,

1977; Tanenhaus, Leiman & Seidenberg, 1979) then it is possible that both the English and Spanish lexical entries are activated. Given that lexical access is faster in the bilingual's dominant language the lexical entry in the dominant language lexicon will be accessed first and will prime the corresponding lexical entry in the other language lexicon via a semantic (conceptual) link. I agree with this interpretation of the cognate advantage in Spanish as a semantic priming effect. However, it is important to note that cognates in the Spanish-target condition were not recognized as fast as cognates in the English-target condition. This result strongly implies that there are still language specific encoding/retrieval processes for cognates.

This account of automatic access of all lexical entries consistent with a given encoding presents some problem for the contention that homographic-noncognate latencies are completely determined by these words' characteristics in the target language. Like cognates, the homographs are orthographically identical in both languages. If in the Spanish-target condition the English lexical entry is accessed the subject might recognize the stimulus as an English word and then be forced to consider whether it is also a Spanish word before responding. This implies that there should be considerable interference in recognizing homographs in the Spanish-target condition: latency and error rates should be elevated. However, there is no evidence of any interference effects for homographs in the Spanish-target condition. Homograph latencies and error rates were not significantly different than that for noncognates. And, both objective and subjective frequency estimates predict homograph latencies extraordinarily well. CAT

ratings predicted homograph latencies within 7 msec and Log K-F predicted homograph latencies within 35 msec.

It would appear that the bilingual can restrict access to that portion of lexical/semantic memory which is germane to the task at hand. While it is not obvious how a bilingual can perform in this way we can not simply explain the results of the experiment by pointing to the fact that bilinguals consistently practice selectivity with respect to language perception and production. Monolinguals also show a considerable degree of selectivity in word recognition. Roydes & Osgood (1972) in a tachistoscopic threshold word identification task used homographs (e.g., 'box') that are more frequently used in a noun sense than a verb sense and homographs (e.g., 'want') that are more frequently used in a verb sense than a noun sense. When subjects were biased to expect nouns in the identification task thresholds were significantly lower for homographs with a more frequent noun sense than for homographs with a relatively less frequent noun sense. When subjects were biased to expect verbs in the identification task, the opposite pattern was found. Now thresholds were significantly lower for homographs with a more frequent verb sense than for homographs with a relatively less frequent verb sense.

CONCLUSIONS

The results of the main experiment are in agreement with the view that neither a completely separate nor a completely integrated model provides an adequate description of bilingual linguistic memory. We found strong evidence of language specific encoding/retrieval processes and some evidence of language integration at a conceptual level. The results also indicate that bilinguals can exercise a considerable degree of control over access to lexical information. When focusing on one of their languages they perform essentially as monolinguals, showing no interference from and highly restricted benefit from (in the case of cognates in Spanish) potential cross-language associations among words. Our proposal of selectively accessible separate lexicons is supported by the data. Other models are of course possible, but it would appear that any adequate model of bilingual linguistic memory must postulate language specific units (or processes) within an integrated semantic memory.

Appendix A

Stimulus Selection Procedures, Experiments 1 and 2

The New College Spanish and English Dictionary (Williams, 1968) was searched for instances of spelling patterns common to both languages and a pool of 556 two to ten letter words was obtained. The vast majority of the words in the pool were cognates that have the same meaning in both languages. Based on the Kucera and Francis (1967) printed English frequency norms, 20 (2 to 8 letter) high frequency cognates were selected. The median frequency of occurrence for this set is 160 per million, range 61 to 2201. Twenty (4 to 7 letter) low frequency cognates were also selected (median frequency 3 per million, range 0.1 to 12).

In the case of homographic-noncognates from the pool, a single spelling pattern has two distinct primary semantic meanings depending on language context. The frequency of occurrence of such a spelling pattern in Spanish, as measured by the K-F value for its respective English dictionary translation equivalent, may be similar to or markedly different from its English language frequency of occurrence.

Based on the above method of estimating Spanish language frequency of occurrence, a base set of 20 (2 to 6 letter) words that are low frequency words in both languages was selected. The median English language frequency of these words is 4.50 per million, range .01 to 30. The matching set of 20 (4 to 10 letter) English translations of

the Spanish meanings of these words has a median frequency of 2.00 per million, range 0.1 to 32. For example, the base set includes the items LAUD and CASTOR the translation set includes LUTE and BEAVER that are the Spanish meanings of the spelling patterns LAUD and CASTOR, respectively.

Instances of homographic-noncognates that are high frequency words in both languages were quite rare. A base set of 5 (2 to 6 letter) English words was selected (median frequency 166 per million, range 100 to 1984). The median frequency of occurrence for the 3 to 7 letter translation set is 204 per million, range 48 to 707. For example, the base set includes the items SON and SO, while the translation set includes the respective Spanish meanings SOUND and UNDER.

In like manner, common spelling patterns that differ markedly in frequency of usage between the two languages were identified. A base set of 27 (3 to 8 letter) words are low frequency in English usage and estimated to be relatively much higher in frequency with respect to Spanish usage. For this set the median frequency of occurrence in English is 1.88 per million, range 0.1 to 53. The corresponding translation set of 27 (2 to 9 letter) words has a median frequency of occurrence of 111.5 per million, range 53 to 9816. For example, SOY and FIN appear in the base set while AM and END, the respective Spanish meanings, appear in the translation set.

A base set of 20 (2 to 6 letter) words are high frequency in English usage (median frequency 239 per million, range 56 to 7250) and estimated to be relatively much lower in frequency with respect to Spanish usage. The corresponding translation set, 3 to 10 letters in

length, has a median frequency of occurrence of 3.25 per million, range 0.1 to 169. For example, the items ONCE and AS appear in the base set and the respective Spanish meanings ELEVEN and ACE appear in the translation set.

In addition, a control set of 40 English noncognate words, the spelling patterns of which are unique to English was included. The spelling patterns of the set of Spanish translations of these items are unique in Spanish. Half of these standard noncognate words are high frequency (median frequency 238 per million, range 64 to 1599) words, 4 to 6 letters in length. Half are low frequency words, 3 to 9 letters in length, (median frequency 5 per million, range 0.1 to 35).

This procedure yielded the English language stimulus list of 224 (2 to 10 letter) words that was rated by the English speaking monolinguals in Experiment 1 and by Spanish-English bilinguals in the English language rating task in Experiment 2.

An independent group of Spanish-English bilinguals rated the Spanish language frequency of occurrence of a Spanish language version of this stimulus list in Experiment 2 of this study. The equivalent Spanish language version of the stimulus list was derived as follows.

Cognate words (i.e. identical spelling pattern and primary semantic meaning in both languages) were unchanged. For the homographic-noncognates, the spelling patterns of which are common to both languages, the English language meaning of each item was rendered into Spanish to form the translation sets for this list version.

For example, the base set of items that are low frequency in English and Spanish includes the items LAUD and CASTOR. The translation

set is now 4 to 9 letters in length and includes the items ALABAR and RICINO that are the English meanings of the spelling patterns LAUD and CASTOR, respectively, as Spanish words. The base set of items high frequency in both languages includes the spelling patterns SON and SO. The translation set is now 3 to 9 letters in length and includes the items HIJO and ASI, the respective English meanings of the base items SON and SO. The base set of spelling patterns that are low frequency in English usage and relatively higher frequency in Spanish usage now has a corresponding translation set, 4 to 11 letters in length, that includes the items SOJA and ALETA that are the respective English meanings of the base items SOY and FIN. The base set of items high frequency in English usage and low frequency in Spanish usage now has a corresponding translation set, 3 to 11 letters in length, that includes the items VEZ and QUE, the respective English meanings of the base items ONCE and AS.

The noncognate control words were translated into Spanish. The high frequency set is now 4 to 7 letters in length and includes the words CASA and VIDA that correspond to the words HOUSE and LIFE on the English language list. The low frequency set is now 3 to 9 letters in length and includes CRANEO and CISNE that correspond to the words SKULL and SWAN on the English list. For ease of reference, a listing of the English and Spanish stimulus lists is included in Table A.1.

TABLE A.1

Homographic-noncognates, cognates and noncognate controls for generation of stimulus lists in Experiments 1 and 2. The English list is formed by combining columns labeled ENGLISH and BASE. The Spanish list is formed by combining columns labeled SPANISH and BASE.

WORD TYPE	ENGLISH	BASE	SPANISH
Homographic-Noncognates:			
English Low Frequency/ Spanish High Frequency	GIVEN	DADO	MOLDURA
	FIRST	PRIMER	CARTILLA
	KNOWLEDGE	SABER	SABLE
	WITHOUT	SIN	PECADO
	END	FIN	ALETA
	AM	SOY	SOJA
	TOWN	VILLA	QUINTA
	VALUE	VALOR	FIRMEZA
	TABLE	MESA	MESETA
	GREATER	MAYOR	ALCALDE
	WALL	PARED	MONDADO
	CHARGE	CARGO	CARGA
	SUN	SOL	SOLUCION
	POOL	BALSA	ALCORENOQUE
	DATE	CALENDAR	CALENDARIO
	OIL	OLEO	MARGARINA
	SHIP	NAVE	CUBO
	BIRTH	NATAL	PARTO
	CRISIS	TRANCE	RAPTO
	WEATHER	TEMPLE	TEMPLO
	SOFT	MOLE	TOPO
	VILLAGE	PUEBLO	ADOBES
	FOOT	PIE	PASTEL
	TASTE	GUSTO	DELEITE
	FIFTEEN	QUINCE	MEMBRILLO
	FOOL	MEMO	APUNTE
	WAS	ERA	SIGLO

WORD TYPE	ENGLISH	BASE	SPANISH
Homographic-Noncognates:			
English High Frequency/ Spanish Low Frequency	NET	RED	ROJO
	PUREE	PURE	PURO
	REMOVAL	QUITE	ENTERAMENTE
	ELEVEN	ONCE	VEZ
	PLOWED	ARE	ESTAN
	TIED	LIE	MENTIRA
	PUBLISHER	EDITOR	REVISOR
	PLEASES	PLACE	LUGAR
	EATS	COME	VENIR
	HOLD	TEN	DIEZ
	DONATED	DONE	HECHO
	BAA	BE	SER
	MARRIED	CASE	CASO
	TAVERN	BAR	BARRA
	DWELLED	MORE	MAS
	YEARNED	FINE	FINA
	ACE	AS	QUE
	SLACKEN	FILE	FILA
	SOCIALCLUB	CLUB	PORRA
	MINED	MINE	MIO

WORD TYPE	ENGLISH	BASE	SPANISH
Homographic-Noncognates:			
English Low Frequency/ Spanish Low Frequency	LUTE	LAUD	ALABAR
	KNAVE	VALET	AYUDA
	HEEL	TALON	GARRA
	WILLOW	SAUCE	SALSA
	VILE	RUIN	RUINA
	PLUNDER	PILLAR	PILAR
	SHEPHERD	PASTOR	CLERIGO
	STOCKING	MEDIA	MEDIOS
	DULL	MATE	COMPANERO
	BEAVER	CASTOR	RICINO
	TABLECLOTH	MANTEL	MESILLA
	DOWRY	DOTE	CHOCHEAR
	COAL	CARBON	CARBONO
	FLOCK	GREY	GRIS
	FLUSH	FLUX	FLUJO
	SNUFF	RAPE	ESTUPRO
	PIMPLE	PUPA	CAPULLO
	LOAFING	TUNA	ATUN
	GRIPPE	GRIPE	QUEJA
	OUCH	AX	HACHA
Homographic-Noncognates:			
English High Frequency/ Spanish High Frequency	UNDER	SO	ASI
	SOUND	SON	HIJO
	PRESENT	ACTUAL	EFECTIVO
	DOG	CAN	ENVASE
	ROYAL	REAL	VERDADERO

WORD TYPE	BASE
Cognates:	
High Frequency	HOTEL
	MOTOR
	SOCIAL
	AREA
	ANIMAL
	PLAN
	CENTRAL
	IDEA
	METAL
	COLOR
	CIVIL
	TOTAL
	LOCAL
	UNION
	NO
	NATURAL
	LABOR
	CONTROL
	GAS
	HOSPITAL
Cognates:	
Low Frequency	NASAL
	FRUGAL
	PERFUME
	DUAL
	JADE
	PEDESTAL
	CANAL
	ALTAR
	PLURAL
	HALO
	DRAGON
	MENU
	OVAL
	FERVOR
	DENTAL
	AROMA
	FACIAL
	SOFA
	SALON
	CLAN

WORD TYPE	ENGLISH	SPANISH
Noncognates:		
High Frequency	EMPTY	VACIO
	DEATH	MUERTE
	LOVE	AMOR
	BEFORE	ANTES
	WORD	PALABRA
	YOUTH	JOVEM
	HAND	MANO
	YEARS	ANOS
	GIRL	NINA
	EYES	OJOS
	HEAVY	PESADO
	BOOK	LIBRO
	DOUBT	DUDA
	STREET	CALLE
	CHAIR	SILLA
	HOUSE	CASA
	SUMMER	VERANO
	HEAT	CALOR
	LIFE	VIDA
	TIME	TIEMPO
Noncognates:		
Low Frequency	THIRST	SED
	SWAN	CISNE
	SKULL	CRANEO
	MIST	NIEBLA
	ACORN	BELLOTA
	BETRAYAL	TRAICION
	MYTH	MITO
	WALTZ	VALS
	YOLK	YEMA
	GUILT	CULPA
	OMEN	AGUERO
	HASTE	PRISA
	MELTED	DERRETIDO
	LEAP	SALTO
	WAX	CERA
	WAGER	APUESTA
	FEATHER	PLUMA
	EARTHWORM	LOMBRIZ
	THORN	ESPINA
	HUNGER	HAMBRE

Appendix B

Instructions for subjective frequency ratings, Experiments 1 and 2

Categorization - CAT

On each page of the booklet you will find a list of English (Spanish) words. We want you to tell how common (frequent) each word is by giving each word a number from 1 to 7. This is the scale you should use to give numbers to the words on the list.

7 extremely frequent word

6 very frequent word

5 frequent word

4 average frequency word

3 infrequent word

2 very infrequent word

1 extremely infrequent word

Here are some examples. The word IS (ESTA) is one of the ten most common English (Spanish) words. So the word IS (ESTA) is an extremely frequent word. If the word IS (ESTA) appears in the list, you would give it the number 7. The word ECHIDNA is an extremely infrequent (uncommon) word in English (Spanish). So if the word ECHIDNA appears in the list, you would give it the number 1. Remember to use all seven numbers to tell how frequent the words are. Write the number for each word on the line next to the word. There are no right or wrong answers. We are interested in how frequent you think each word is.

Magnitude Estimation - MAG

On each page of the booklet you will find a list of English (Spanish) words. We want you to tell how common (frequent) each word is by giving each word a number. The word DOCTOR is a fairly common English (Spanish) word. We will give the word DOCTOR a value of 100. You will give each word a number to tell how frequent the word is compared to DOCTOR. Here are some examples. Most people feel the word IS (ESTA) is 100 times more frequent than DOCTOR. So if IS (ESTA) appears on the list you would give IS (ESTA) a number 100 times larger than DOCTOR's value. Because DOCTOR's value is 100, 100 times 100 is 10,000. The number for IS (ESTA) would be 10,000. Many people feel that the word ZEBRA is 50 times less frequent than DOCTOR. So you might give ZEBRA a number 50 times smaller than DOCTOR's value. Because DOCTOR's value is 100, 100 divided by 50 is 2. The number for ZEBRA would be 2. If you think the word HOUSE (CASA) is 25 time more frequent than DOCTOR. You would give a HOUSE (CASA) a number of 2,500 because 100 times 25 is 2,500. Write the number for each word on the line next to the word. There are no right or wrong answers. We are interested in how frequent you think each word is as compared to DOCTOR. You may use any whole numbers, fractions or decimals that you wish. Just don't use negative numbers. Remember DOCTOR is 100.

Graphic Line Scaling - LIN

On each page of the booklet you will find a list of English (Spanish) words. You will also find a scale (vertical line). The scale is a frequency scale. The top of the scale represents extremely frequent

(common) English (Spanish) words and the bottom of the scale represents extremely infrequent (uncommon) English (Spanish) words. We want you to tell how common (frequent) each word is by drawing a line from the word to its place on the scale. We have already put five words on the scale for you. These same five words will appear on the scale in the same places on every page of the booklet. There are no right or wrong answers. We are interested in how frequent you think each word is. Before you start the booklet, here is a sample page for practice.

Appendix C

Reliability of Ratings via Ebel's (1951) Method

Hoyt (1941) was instrumental in developing reliability estimates of test scores based on analysis of variance (ANOVA) theory and procedures. Ebel (1951) extended this ANOVA methodology to the estimation of test reliabilities based on sets of ratings. A detailed explication of these methods is also given by Winer (1971). Ebel's method has most recently been applied by Shapiro (1969), Carroll (1971) and Carroll and White (1973a) to the reliability analysis of subjective word frequency ratings. The rating data collected by both Shapiro (1969) and Carroll (1971) were in the form of subjective magnitude estimates based on Steven's (1956, 1958) psychophysical scaling procedures. The rating data obtained by Carroll and White (1973) were derived in a different fashion. Their subjects assigned a log frequency value to test words by indicating where, on a graphic vertical scale, a test item would fall relative to reference words differing in objective log frequency.

The following is a brief summary of the ANOVA method of assessing reliabilty of ratings. Given that we have a sample of k estimates of some trait (or stimulus property) for each member of a sample of n items, each estimate is assumed to consist of a true trait component (T) and an error component (E). While the value of T varies among items, T is constant for a given item over all k estimates. The popu-

lation variance of T is designated σ^2_T. E varies among the k estimates for a given item but is assumed to be constant within each of the k sets of ratings. The population variance of E is designated σ^2_E. The total variance of the estimates is then: $\sigma^2_T + \sigma^2_E$.

Reliability of the individual ratings is defined as the proportion of total variance which is true trait variance.

$$r_I = \frac{\sigma^2_T}{\sigma^2_T + \sigma^2_E} \quad (1)$$

From the analysis of variance of a set of ratings we obtain a mean square for error (MS_E) and a between items mean square (MS_I). MS_E is a direct estimate of the population error variance.

$$MS_E = \sigma^2_E \quad (2)$$

However, MS_I is not a direct estimate of the true item variance. Specifically,

$$MS_I = k\sigma^2_T + \sigma^2_E \quad (3)$$

If we now solve equations 2 and 3 for σ^2_T and σ^2_E in terms of the sample statistics MS_E and MS_I and substitute into

equation 1, we arrive at the intraclass correlation formula for reliability of individual ratings:

$$r_I = \frac{MS_I - MS_E}{MS_I + (k-1)MS_E} \quad (4)$$

In many studies, and specifically in the present one, the experimental emphasis is on the degree of average inter-rater agreement rather than the reliability of the individual ratings. In other words, we are interested in assessing the reliabilty of the average ratings for the purpose of item selection or subsequent validity studies.

Ebel notes that reliability of average ratings is completely determined by the reliability of the individual ratings. The development of a direct formula for reliability of average ratings follows the same logic as the previous derivation.

For the case of average ratings, the definition of reliability as the ratio of the true item variance to the total variance is analogous to formula 1, i.e.:

$$r_{AV} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2} \quad (5)$$

where σ^2_E is the error variance of the item means (a standard error squared). As we previously noted, σ^2_E is directly estimated by MS_E from the one-way ANOVA. So,

$$\sigma_E^2 = \frac{MS_E}{k} \quad (6)$$

And as previously derived,

$$\sigma_T^2 = \frac{MS_I - MS_E}{k} \quad (7)$$

Substituting these values into equation 5 yields a direct formula for the reliability of average ratings:

$$r_{AV} = \frac{MS_I - MS_E}{MS_I} \quad (8)$$

It should be obvious at this point that the MS_E used in the foregoing reliability calculations includes a variance component attributable to between-raters mean differences in the use of the rating scale. For example, one rater may have used the numbers 1 to 10 and another the numbers 11 to 20, so that on the average these two sets of ratings will differ by a factor of 2 independent of the consistency of each rater's use of the numbers. In many circumstances it is desirable to remove this between-raters variance in estimating the reliability of a set of ratings. Specifically, Ebel states that between-raters variance should be removed when the ratings have been equated

from rater to rater: e.g. in the case of ranks or z-scores. Furthermore, between-raters variance should be removed if experimental decisions about the items are based on averages of complete sets of ratings. In the present study, the between-raters variance is removed in the calculation of reliability.

Computationally, the between-raters variance is isolated by performing a two-way ANOVA on the set of ratings using both ITEMS and RATERS as factors. The appropriate MS_E used in the reliablity calculations then becomes the $MS_{residual}$ (or equivalently, in the case of each item having been rated once by all raters, the ITEM by RATER interaction term from the two-way ANOVA).

Ebel's method also leads to estimates of precision (confidence limits) for the obtained reliabilities. Defining a variance ratio, F_s, as

$$F_s = \frac{MS_I}{MS_E} \tag{9}$$

then r_I may be expressed as

$$r_I = \frac{F_s - 1}{(F_s - 1) + k} \tag{10}$$

Under the assumption that the above sample values MS_I and MS_E have been drawn from independent populations, the true population variance ratio, F_p, may be greater or less than the observed

value. The maximum and minimum population ratio expected at a given confidence level may be read from a table for the central F distribution using the degrees of freedom for the sample values MS_I and MS_E. Let this value be called F_{crit}.

Specifically, the estimate of the upper limit of the variance ratio is obtained by multiplying the sample value F_s by the maximum ratio expected $F_{crit}(df_I, df_E)$ at the desired alpha level:

$$F_{pu} = F_s\left[F_{crit}(df_I, df_E)\right] \tag{11}$$

The estimate of the lower limit of the variance ratio obtained by multiplying the sample value F_s by the reciprocal value of the tabled critical ratio $1/F_{crit}(df_E, df_I)$. Note that at the lower limit the F table is entered with the degrees of freedom corresponding to MS_E as the degrees of freedom for the numerator.

$$F_{pl} = F_s\left[\frac{1}{F_{crit}(df_E, df_I)}\right] \tag{12}$$

Substituting these two limiting values into equation 10 yields the upper and lower confidence bounds for the reliability of individual ratings (r_I).

The formula for the reliability of average ratings may also be rewritten in terms of the variance ratio F_s:

$$r_{AV} = \frac{F_s - 1}{F_s} \tag{13}$$

Substituting the two limiting values previously derived into this equation yields the desired confidence bounds for the reliability of average ratings.

Appendix D

Summary of Frequency Ratings, Experiments 1 and 2

TABLE D.1

K-F values, log K-F values, means and standard deviations of subjective ratings of English words in Experiment 1. Eighteen English monolingual raters.

WORD	K-F	LOG K-F	CAT		MAG(LOG)		LIN	
			Mean	SD	Mean	SD	Mean	SD
DADO	0.1	-1.000	1.222	0.533	0.023	0.510	-1.320	0.215
LAUD	0.1	-1.000	1.833	0.764	0.682	0.626	-1.122	0.906
PRIMER	0.1	-1.000	2.833	0.764	1.498	0.482	0.299	0.860
OUCH	0.1	-1.000	5.167	1.384	2.405	0.667	1.913	1.549
FRUGAL	0.1	-1.000	2.278	1.044	1.075	0.573	-0.376	1.174
PUREE	0.1	-1.000	2.556	1.117	1.173	0.600	-0.286	1.204
TALON	0.1	-1.000	1.667	0.882	0.714	0.868	-1.258	0.811
ACORN	0.1	-1.000	3.167	1.167	1.444	0.782	0.670	1.181
TABLECLOTH	0.1	-1.000	4.500	1.424	2.246	0.742	2.026	0.775
SNUFF	0.1	-1.000	2.167	0.833	0.758	0.612	-0.861	1.035
PLURAL	0.1	-1.000	3.889	1.242	1.910	0.681	1.338	1.180
PARED	0.1	-1.000	2.444	0.762	1.234	0.699	-0.048	1.173
KNAVE	0.1	-1.000	1.889	0.936	0.586	0.812	-1.204	0.651
BAA	0.1	-1.000	2.278	1.660	0.799	0.934	-0.587	1.128
LOAFING	0.1	-1.000	3.111	1.048	1.620	0.588	0.793	1.147
BALSA	0.1	-1.000	1.556	0.685	0.434	1.147	-0.881	1.028
SOCIALCLUB	0.1	-1.000	3.833	1.167	1.657	0.853	0.983	0.637
DOTE	0.1	-1.000	2.111	0.809	1.031	0.583	-0.527	0.963
OLEO	0.1	-1.000	1.778	0.853	0.848	1.053	-0.745	1.180
NAVE	0.1	-1.000	1.833	1.067	0.552	1.264	-0.914	0.928
GRIPPE	0.1	-1.000	2.056	0.970	0.612	0.792	-0.499	1.087
PIMPLE	0.1	-1.000	4.889	1.242	2.287	0.718	1.494	0.736
PUEBLO	0.1	-1.000	2.056	0.705	0.803	0.686	-0.818	1.113
PUPA	0.1	-1.000	1.778	1.315	0.275	0.604	-1.116	0.756
TUNA	0.1	-1.000	5.056	1.268	2.278	0.542	1.970	0.765
SLACKEN	0.1	-1.000	2.944	0.911	1.098	0.814	-0.176	1.137
GRIPE	0.1	-1.000	3.611	0.891	1.557	0.691	0.566	1.210
SABER	1.	0.000	2.056	0.780	0.937	0.947	-0.509	1.131
JADE	1.	0.000	2.889	1.100	1.362	0.639	-0.099	1.247
SOY	1.	0.000	2.444	0.896	1.149	0.637	-0.469	1.118
VALOR	1.	0.000	2.778	0.975	1.112	0.830	-0.297	1.312
MESA	1.	0.000	1.833	0.833	0.456	0.906	-0.881	0.987
WALTZ	1.	0.000	2.722	0.870	1.089	0.932	0.037	1.189
YOLK	1.	0.000	3.611	0.951	1.606	0.665	0.927	1.127
DRAGON	1.	0.000	3.556	1.423	1.262	0.706	0.112	1.142
STOCKING	1.	0.000	4.667	1.155	2.267	0.843	1.589	0.870
LUTE	1.	0.000	1.944	0.970	0.719	0.824	-0.829	1.070
REMOVAL	1.	0.000	4.167	0.898	2.048	0.782	1.242	0.790
EARTHWORM	1.	0.000	3.167	1.302	1.297	0.653	0.048	1.257
SALON	1.	0.000	3.444	1.066	1.154	0.523	0.402	1.104
MEMO	1.	0.000	4.833	1.067	2.137	0.746	1.583	0.795
VALET	2.	0.301	2.722	1.044	1.108	0.603	-0.849	1.074

NASAL	2.	0.301	3.556	1.117	1.598	0.663	1.062	0.702
FIN	2.	0.301	3.167	1.014	1.393	0.726	0.276	0.841
PILLAR	2.	0.301	3.333	0.943	1.402	0.786	0.392	1.089
PLEASES	2.	0.301	4.444	1.301	2.282	0.627	1.848	0.729
HALO	2.	0.301	3.000	1.414	1.198	0.768	-0.372	1.307
OMEN	2.	0.301	2.500	0.833	1.409	0.621	-0.344	1.168
CASTOR	2.	0.301	2.167	0.833	1.058	0.671	-0.137	1.352
PLUNDER	2.	0.301	2.833	0.898	1.244	0.714	-0.176	1.315
MATAL	2.	0.301	2.833	1.302	1.138	0.800	-0.361	1.324
TAVERN	2.	0.301	3.444	0.762	1.553	0.576	0.379	1.084
FACIAL	2.	0.301	4.167	1.213	1.904	0.739	1.179	0.782
GUSTO	2.	0.301	2.778	1.133	1.255	0.628	-0.162	1.329
QUINCE	2.	0.301	1.611	0.756	0.361	0.769	-1.075	0.756
DOMRY	2.	0.301	2.278	0.803	0.818	0.847	-0.523	1.310
CLAN	2.	0.301	3.333	1.247	1.421	0.679	0.639	1.385
SWAN	3.	0.477	3.167	1.302	1.558	0.778	0.658	1.065
DWELLED	3.	0.477	3.167	1.118	1.536	0.649	0.687	1.027
SKULL	3.	0.477	3.389	0.826	1.471	0.753	0.825	1.321
MINED	3.	0.477	3.222	0.786	1.484	0.660	0.450	1.138
CANAL	3.	0.477	3.500	0.764	1.672	0.769	1.232	0.713
SOL	3.	0.477	1.889	0.936	0.470	1.361	-0.549	1.148
MANTEL	3.	0.477	3.000	1.000	1.095	0.742	0.028	1.218
EATS	3.	0.477	5.056	1.840	2.762	0.816	2.324	0.919
YEARNED	3.	0.477	2.889	0.809	1.264	0.639	-0.563	1.199
WAGER	3.	0.477	3.111	0.875	1.180	0.753	0.598	1.096
AROMA	3.	0.477	3.944	1.177	1.570	0.714	0.884	0.782
SHEPHERD	3.	0.477	3.333	0.816	1.376	0.865	0.115	1.370
BEAVER	3.	0.477	3.333	1.374	1.315	0.810	0.319	1.443
THORN	3.	0.477	4.111	1.149	1.626	0.961	0.888	0.826
THIRST	4.	0.602	4.667	1.202	2.345	0.719	2.039	0.725
TRANCE	4.	0.602	3.333	1.247	1.445	0.800	-0.331	1.265
FERVOR	4.	0.602	2.667	1.000	0.976	0.702	-0.686	1.023
MOLE	4.	0.602	3.556	1.212	1.275	0.704	0.636	1.050
VILE	5.	0.699	2.722	0.931	1.350	0.860	-0.156	1.166
PEDESTAL	5.	0.699	3.167	1.167	1.428	0.773	0.329	1.210
ALTAR	5.	0.699	3.611	1.458	1.545	0.658	0.865	0.582
PLOWED	5.	0.699	3.333	1.106	1.426	0.796	0.725	1.209
MENU	5.	0.699	5.056	1.026	2.236	0.695	1.815	0.598
RAPE	5.	0.699	4.611	1.253	2.154	0.614	1.856	0.810
BETRAYAL	6.	0.778	3.333	1.291	1.604	0.620	0.354	1.359
VILLA	6.	0.778	2.833	0.833	1.349	0.737	0.309	1.278
FEATHER	6.	0.778	4.111	1.449	1.767	0.942	1.403	0.567
SOFA	6.	0.778	4.889	1.197	2.206	0.765	1.917	0.714
AX	6.	0.778	4.056	1.129	1.657	0.875	1.195	1.029
DONATED	7.	0.845	4.111	1.329	1.867	0.613	1.278	0.588
CARGO	7.	0.845	3.333	1.202	1.479	0.671	0.229	1.179
OVAL	8.	0.903	4.167	1.167	1.571	0.553	1.172	0.897
WILLOW	9.	0.954	2.778	0.975	1.373	0.797	-0.156	1.234
DUAL	9.	0.954	3.167	0.898	1.478	0.770	0.749	0.990
HASTE	9.	0.954	3.333	0.816	1.650	0.790	0.699	0.808
MELTED	9.	0.954	4.556	0.956	2.078	0.622	1.561	0.486
HEEL	9.	0.954	4.389	1.671	1.658	0.521	1.237	0.723
PERFUME	10.	1.000	4.722	1.356	2.253	0.922	1.816	0.735
FLOCK	10.	1.000	3.222	0.786	1.322	0.760	0.440	1.505

FLUSH	11.	1.041	3.722	0.931	1.946	0.622	1.257	0.664
GREY	12.	1.079	4.444	1.257	2.043	0.471	1.656	0.836
DENTAL	12.	1.079	4.167	1.067	1.739	0.327	1.539	0.486
MEDIA	13.	1.114	4.222	1.083	1.849	0.609	1.387	0.832
MIST	14.	1.146	3.222	1.083	1.531	0.867	0.511	1.116
RUIN	14.	1.146	4.389	1.061	1.955	0.612	1.522	0.720
LEAP	14.	1.146	4.222	1.315	1.886	0.708	1.289	0.849
WAX	14.	1.146	4.667	1.414	1.873	0.987	1.387	0.999
PIE	14.	1.146	5.278	1.096	2.509	0.801	2.490	0.778
ACE	15.	1.176	3.500	1.167	1.701	0.897	1.029	0.658
PASTOR	17.	1.230	3.611	1.061	1.595	0.958	1.164	0.969
HUNGER	17.	1.230	4.889	1.370	2.617	0.670	1.828	0.840
SAUCE	20.	1.301	4.222	1.227	2.094	0.832	1.792	0.663
CALENDAR	20.	1.301	4.778	1.181	2.178	0.597	1.896	0.608
MATE	21.	1.322	3.889	0.936	1.874	0.788	1.383	0.502
PUBLISHER	23.	1.362	4.056	1.177	1.861	0.636	1.023	0.857
DULL	27.	1.431	4.611	1.208	2.167	0.715	1.867	0.987
CARBON	30.	1.477	3.778	1.133	1.713	0.589	1.040	0.999
FLUX	30.	1.477	2.222	0.853	1.028	0.759	-0.821	0.973
ERA	30.	1.477	3.833	1.167	1.587	0.722	0.549	1.456
COAL	32.	1.505	4.278	1.193	1.987	0.639	1.710	0.903
GUILT	33.	1.519	4.778	1.315	2.267	0.799	1.606	0.705
TIED	34.	1.531	4.778	1.083	2.197	0.564	2.008	0.776
NET	34.	1.531	3.778	1.083	1.911	0.698	1.211	1.031
MYTH	35.	1.544	3.611	1.253	1.464	0.544	0.321	1.090
MAYOR	38.	1.580	4.611	1.112	2.206	0.641	1.743	0.531
TEMPLE	38.	1.580	4.111	1.149	1.538	0.654	0.916	1.124
ELEVEN	40.	1.602	5.222	1.272	2.511	0.744	2.461	0.961
ROYAL	48.	1.681	3.667	1.000	1.646	0.816	0.828	0.826
SIN	53.	1.724	4.667	1.333	2.209	0.886	1.536	1.391
FOOL	53.	1.724	4.611	1.061	2.276	0.622	2.443	0.978
FIFTEEN	56.	1.748	4.889	0.994	2.470	0.796	2.356	0.761
PURE	56.	1.748	4.667	1.054	2.240	0.831	1.809	0.771
LIE	59.	1.771	5.556	1.066	2.576	0.851	2.756	0.727
MINE	59.	1.771	5.556	1.165	3.107	0.855	2.620	0.875
METAL	61.	1.785	4.667	1.155	2.198	0.639	1.865	0.839
TASTE	61.	1.785	5.611	0.756	2.744	0.522	2.622	0.662
MOTOR	62.	1.792	4.778	1.272	2.216	0.735	1.710	0.762
EMPTY	64.	1.806	5.444	1.117	2.464	0.525	2.138	0.622
BIRTH	66.	1.820	5.167	1.167	2.344	0.713	1.997	0.732
CHAIR	66.	1.820	5.833	1.014	3.084	0.817	2.638	0.694
ANIMAL	68.	1.833	5.667	1.054	2.894	0.761	2.484	0.755
FOOT	70.	1.845	5.667	1.247	2.546	0.694	2.318	0.740
VILLAGE	72.	1.857	4.667	1.106	1.908	0.909	1.312	0.936
DOG	75.	1.875	6.167	0.764	3.041	0.757	3.026	0.637
SOFT	77.	1.886	5.611	0.951	2.689	0.706	2.483	0.662
EDITOR	77.	1.886	4.222	0.916	1.966	0.908	1.313	1.054
FILE	81.	1.908	4.444	1.301	2.041	0.573	1.681	0.643
WEATHER	81.	1.908	5.500	1.258	2.596	0.727	2.316	0.630
YOUTH	82.	1.914	4.611	1.253	2.320	0.592	1.964	0.613
BAR	82.	1.914	5.000	1.202	2.438	0.871	1.937	0.703
CRISIS	82.	1.914	4.889	1.286	2.400	0.778	1.547	0.839
SHIP	89.	1.949	5.056	1.268	2.351	0.704	2.029	0.803
CIVIL	91.	1.959	4.111	1.197	1.882	0.510	1.184	0.565

OIL	93.	1.968	5.444	1.012	2.808	0.657	2.061	0.735
HEAT	97.	1.987	5.722	1.044	2.529	0.492	2.581	0.696
GAS	98.	1.991	6.056	0.848	2.894	0.685	2.641	0.835
ACTUAL	100.	2.000	4.444	0.831	2.166	0.538	1.887	0.666
DATE	103.	2.013	5.778	1.133	2.972	0.871	2.698	0.646
MARRIED	105.	2.021	5.667	0.943	2.968	0.749	2.359	0.684
HEAVY	110.	2.041	5.167	1.014	2.663	0.569	2.266	0.418
HOSPITAL	110.	2.041	4.833	1.118	2.237	0.431	2.052	0.569
POOL	111.	2.045	5.000	1.106	2.545	0.726	1.917	0.804
SUN	112.	2.049	6.056	1.079	3.069	0.847	2.891	0.659
DOUBT	114.	2.057	4.667	1.202	2.546	0.643	1.958	0.769
CHARGE	122.	2.086	4.889	1.197	2.377	0.681	1.975	0.880
HOTEL	126.	2.100	5.278	1.044	2.346	0.709	1.773	0.438
SUMMER	134.	2.127	5.889	0.994	2.586	0.668	2.324	0.962
COLOR	141.	2.149	5.333	1.106	2.640	0.677	2.423	0.704
KNOWLEDGE	145.	2.161	4.611	1.208	2.159	0.621	1.752	0.819
CLUB	145.	2.161	4.889	0.936	2.269	0.543	1.868	0.571
LABOR	149.	2.173	4.556	0.831	1.918	0.549	1.532	0.657
NATURAL	156.	2.193	5.278	0.989	2.597	0.841	2.008	0.801
WALL	160.	2.204	5.500	1.167	2.645	0.992	2.238	0.800
FINE	161.	2.207	5.833	1.067	3.027	0.718	2.778	0.612
CENTRAL	164.	2.215	4.333	1.202	1.869	0.748	1.476	0.754
TEN	165.	2.217	5.833	1.067	2.813	0.787	2.947	0.795
SON	166.	2.220	6.167	0.764	3.106	0.823	2.989	0.741
HOLD	169.	2.228	5.000	1.202	2.500	0.582	2.477	1.245
UNION	182.	2.260	4.444	1.461	2.064	0.712	1.314	1.050
GREATER	188.	2.274	4.778	1.181	2.281	0.635	1.935	0.736
BOOK	193.	2.286	5.667	0.882	2.986	0.661	2.974	0.831
IDEA	195.	2.290	5.500	1.118	2.588	0.712	2.211	0.676
RED	197.	2.294	5.444	1.012	2.698	0.858	2.670	0.597
TABLE	199.	2.299	5.889	0.994	2.805	0.598	2.721	0.837
VALUE	201.	2.303	4.889	1.048	2.270	0.595	1.971	0.725
SOUND	204.	2.310	5.389	1.061	2.508	0.702	2.111	0.857
PLAN	205.	2.312	5.056	1.026	2.342	0.463	2.226	0.744
TOTAL	211.	2.324	4.833	1.118	2.307	0.720	2.075	0.663
TOWN	218.	2.338	4.944	1.433	2.480	0.780	1.934	0.787
GIRL	220.	2.342	6.111	0.875	3.142	0.852	2.909	0.648
CONTROL	223.	2.348	4.778	1.083	2.270	0.530	1.598	1.186
AM	228.	2.358	6.889	0.314	3.379	0.764	3.590	0.582
LOVE	232.	2.365	6.389	0.678	3.378	0.804	3.125	0.521
STREET	244.	2.387	5.944	1.268	3.238	0.689	2.612	0.698
REAL	260.	2.415	5.389	1.297	2.776	0.623	2.630	0.790
WORD	274.	2.438	5.778	1.227	2.943	0.782	2.669	0.913
DEATH	277.	2.442	5.444	0.831	2.653	0.894	2.149	0.700
QUITE	281.	2.449	4.500	1.067	2.161	0.516	1.853	0.573
LOCAL	288.	2.459	5.000	1.000	2.143	0.679	1.881	0.481
DONE	320.	2.505	5.167	1.344	2.308	0.950	2.086	1.153
AREA	324.	2.511	4.667	0.882	2.083	0.641	1.729	0.727
CASE	362.	2.559	4.833	1.067	2.371	0.722	1.925	0.512
PRESENT	377.	2.576	5.222	1.227	2.764	0.758	2.063	0.683
GIVEN	377.	2.576	5.000	1.291	2.539	0.949	2.431	0.590
SOCIAL	380.	2.580	4.833	1.167	2.305	0.624	1.664	0.761
EYES	401.	2.603	5.500	1.167	3.007	0.597	2.538	0.620
END	410.	2.613	5.889	1.100	2.941	0.855	2.601	0.862

HAND	431.	2.634	5.889	0.936	3.050	0.778	2.877	0.559
ONCE	499.	2.698	5.778	1.083	2.472	0.808	2.402	0.743
PLACE	571.	2.757	5.389	1.112	2.856	0.777	2.557	0.646
WITHOUT	583.	2.766	5.333	1.155	2.660	0.828	2.563	0.523
HOUSE	591.	2.772	6.056	0.848	3.133	0.759	2.797	0.691
COME	630.	2.799	6.111	0.657	3.150	0.955	3.234	0.600
UNDER	707.	2.849	5.611	0.951	2.756	0.746	2.642	0.577
LIFE	715.	2.854	6.000	0.943	2.981	0.731	2.579	0.618
YEARS	949.	2.977	5.333	1.202	2.786	0.595	2.383	0.724
BEFORE	1016.	3.007	5.556	1.165	2.930	0.601	2.650	0.921
FIRST	1360.	3.134	5.611	1.008	2.658	0.694	2.383	0.680
TIME	1599.	3.204	6.000	0.816	3.238	0.806	3.101	0.764
CAN	1772.	3.248	6.611	0.591	3.388	0.749	3.481	0.410
SO	1985.	3.298	6.778	0.533	3.437	0.841	3.583	0.508
NO	2201.	3.343	7.000	0.000	3.770	0.853	3.862	0.285
MORE	2216.	3.346	6.111	0.737	3.004	0.834	3.080	0.707
ARE	4393.	3.643	6.889	0.458	3.499	0.822	3.711	0.407
BE	6377.	3.805	7.000	0.000	3.402	0.620	3.805	0.208
AS	7250.	3.860	6.889	0.314	3.612	0.747	3.788	0.207
WAS	9816.	3.992	6.889	0.458	3.548	0.859	3.429	0.942
Mean	265.	1.226	4.306		2.018		1.370	
SD	998.	1.217	1.309		0.757		1.215	

TABLE D.2

K-F values, log K-F values, means and standard deviations of subjective ratings of English words in Experiment 2. Nine Spanish-English bilingual raters.

WORD	K-F	LOG K-F	CAT		MAG(LOG)		LIN	
			Mean	SD	Mean	SD	Mean	SD
DADO	0.1	-1.000	2.000	2.000	0.300	0.561	-0.849	1.326
LAUD	0.1	-1.000	2.444	1.892	0.861	0.635	-0.296	1.506
PRIMER	0.1	-1.000	2.222	1.397	1.193	0.715	-0.083	1.102
OUCH	0.1	-1.000	5.222	1.315	2.150	0.777	2.128	1.058
FRUGAL	0.1	-1.000	1.778	1.030	0.810	0.775	-0.101	1.516
PUREE	0.1	-1.000	2.778	2.043	1.256	0.440	-0.400	1.184
TALON	0.1	-1.000	2.222	1.931	0.718	0.629	-0.904	0.877
ACORN	0.1	-1.000	2.556	1.499	0.959	0.461	-0.806	1.584
TABLECLOTH	0.1	-1.000	4.667	1.414	2.010	0.411	2.180	0.656
SNUFF	0.1	-1.000	2.667	1.414	1.034	0.555	0.210	1.066
PLURAL	0.1	-1.000	4.000	1.491	1.749	0.226	1.411	1.312
PARED	0.1	-1.000	2.556	1.641	1.057	0.761	-0.181	1.319
KNAVE	0.1	-1.000	1.667	1.054	0.717	0.583	-0.218	1.342
BAA	0.1	-1.000	1.889	1.286	1.035	0.817	-0.400	1.608
LOAFING	0.1	-1.000	3.222	0.916	1.559	0.401	0.564	0.905
BALSA	0.1	-1.000	2.000	1.414	0.497	0.671	-0.426	1.254
SOCIALCLUB	0.1	-1.000	3.667	1.414	1.837	0.605	1.356	0.869
DOTE	0.1	-1.000	1.889	0.737	0.766	0.652	-1.147	0.790
OLEO	0.1	-1.000	2.444	1.343	0.868	0.793	-0.558	1.157
NAVE	0.1	-1.000	2.222	1.315	0.957	0.680	-1.094	0.916
GRIPPE	0.1	-1.000	2.778	1.030	0.706	0.848	-0.503	1.546
PIMPLE	0.1	-1.000	5.000	1.247	2.275	0.608	2.232	1.046
PUEBLO	0.1	-1.000	2.778	1.315	1.045	0.650	-0.784	1.102
PUPA	0.1	-1.000	1.222	0.416	0.542	0.671	-1.241	0.509
TUNA	0.1	-1.000	4.556	1.707	1.885	0.448	1.868	1.189
SLACKEN	0.1	-1.000	2.222	1.030	0.911	0.563	-0.427	1.367
GRIPE	0.1	-1.000	2.333	1.333	0.766	0.809	-0.278	0.991
SABER	1.	0.000	1.667	1.054	0.489	0.493	-0.584	1.517
JADE	1.	0.000	3.000	0.943	1.573	0.424	0.117	1.041
SOY	1.	0.000	3.111	1.286	1.110	0.653	-0.636	1.288
VALOR	1.	0.000	2.889	1.912	1.223	0.586	-0.199	1.336
MESA	1.	0.000	2.111	1.912	0.586	0.679	-0.830	1.168
WALTZ	1.	0.000	2.444	1.257	1.119	0.558	-0.251	1.265
YOLK	1.	0.000	2.778	1.548	1.300	0.557	0.203	0.945
DRAGON	1.	0.000	3.000	1.700	1.242	0.586	0.879	1.141
STOCKING	1.	0.000	4.778	1.474	1.906	0.192	1.974	0.518
LUTE	1.	0.000	2.222	1.315	0.420	0.476	-0.127	1.349
REMOVAL	1.	0.000	4.556	1.343	1.935	0.648	2.086	0.724
EARTHWORM	1.	0.000	3.000	1.563	1.368	0.521	0.246	1.675
SALON	1.	0.000	3.222	1.315	1.322	0.892	1.029	1.151
MEMO	1.	0.000	4.333	1.491	1.763	0.414	1.036	1.211
VALET	2.	0.301	2.667	1.491	0.990	0.611	-0.266	1.085

NASAL	2.	0.301	3.667	1.764	1.633	0.314	0.759	1.150
FIN	2.	0.301	2.778	1.397	0.958	0.708	0.667	1.733
PILLAR	2.	0.301	3.111	1.197	1.566	0.338	0.507	0.990
PLEASES	2.	0.301	4.667	1.414	1.963	0.257	2.279	0.787
HALO	2.	0.301	2.778	1.397	1.149	0.928	-0.220	1.653
OMEN	2.	0.301	2.667	1.633	1.386	0.528	0.110	1.573
CASTOR	2.	0.301	2.667	1.054	0.883	0.604	-0.622	1.137
PLUNDER	2.	0.301	2.889	0.875	1.154	0.579	-0.064	1.118
NATAL	2.	0.301	2.222	1.315	1.145	0.721	-0.252	1.907
TAVERN	2.	0.301	3.000	1.333	1.460	0.833	-0.057	1.138
FACIAL	2.	0.301	4.778	1.315	1.770	0.363	1.167	1.583
GUSTO	2.	0.301	3.333	2.000	1.071	0.882	0.170	1.452
QUINCE	2.	0.301	2.333	2.055	0.620	0.653	-0.984	0.927
DOWRY	2.	0.301	2.444	1.165	0.740	0.704	-0.372	1.170
CLAN	2.	0.301	2.667	1.563	1.273	0.837	0.402	1.329
SWAN	3.	0.477	3.111	1.370	1.467	0.646	0.806	0.672
DWELLED	3.	0.477	2.778	1.474	1.378	0.583	0.448	0.971
SKULL	3.	0.477	3.778	1.474	1.730	0.215	1.417	0.675
MINED	3.	0.477	3.333	1.333	1.585	0.379	0.879	1.075
CANAL	3.	0.477	4.333	1.563	1.859	0.197	1.082	1.336
SOL	3.	0.477	2.444	1.771	0.873	0.701	-0.790	1.527
MANTEL	3.	0.477	3.667	1.700	1.096	0.535	0.320	1.457
EATS	3.	0.477	5.333	1.563	2.282	0.702	2.756	0.847
YEARNED	3.	0.477	2.556	1.165	0.946	0.661	-0.522	1.128
WAGER	3.	0.477	3.333	1.333	1.156	0.711	0.272	1.216
AROMA	3.	0.477	3.889	1.370	1.385	0.610	1.073	1.392
SHEPHERD	3.	0.477	2.778	1.548	1.278	0.537	0.170	0.932
BEAVER	3.	0.477	2.889	1.449	1.237	0.509	0.854	1.011
THORN	3.	0.477	3.667	1.247	1.440	0.541	1.237	0.817
THIRST	4.	0.602	4.556	1.343	2.118	0.715	1.869	0.901
TRANCE	4.	0.602	3.111	1.286	1.248	0.688	0.888	0.950
FERVOR	4.	0.602	2.333	1.247	1.028	0.715	-0.484	1.366
MOLE	4.	0.602	3.556	1.423	1.197	0.665	0.658	1.031
VILE	5.	0.699	2.000	0.816	0.918	0.523	-0.816	0.761
PEDESTAL	5.	0.699	3.556	1.165	1.600	0.323	0.703	1.086
ALTAR	5.	0.699	3.222	1.548	1.583	0.338	0.769	1.234
PLOWED	5.	0.699	2.222	1.030	1.058	0.638	-0.109	1.017
MENU	5.	0.699	5.222	1.227	2.191	0.671	2.031	0.747
RAPE	5.	0.699	4.556	1.499	1.903	0.588	1.640	1.264
BETRAYAL	6.	0.778	3.444	1.165	1.451	0.623	0.736	1.030
VILLA	6.	0.778	2.889	1.728	1.286	0.399	0.009	1.291
FEATHER	6.	0.778	3.667	1.563	1.710	0.645	0.978	0.982
SOFA	6.	0.778	5.222	1.474	1.928	0.515	2.084	1.581
AX	6.	0.778	3.444	1.641	1.481	0.419	1.226	1.324
DONATED	7.	0.845	4.000	0.816	1.749	0.214	1.157	1.231
CARGO	7.	0.845	3.444	1.165	1.630	0.431	0.586	1.251
OVAL	8.	0.903	3.111	1.523	1.522	0.568	1.746	0.920
WILLOW	9.	0.954	2.222	1.315	1.318	0.218	-0.740	1.014
DUAL	9.	0.954	3.222	1.397	1.604	0.319	0.073	1.302
HASTE	9.	0.954	2.778	1.133	1.095	0.739	0.639	1.182
MELTED	9.	0.954	5.111	1.286	1.787	0.345	1.602	0.750
HEEL	9.	0.954	4.444	1.707	1.795	0.549	1.533	1.217
PERFUME	10.	1.000	5.000	1.155	2.145	0.690	2.027	0.902
FLOCK	10.	1.000	2.444	1.066	1.090	0.538	-0.119	1.165

FLUSH	11.	1.041	4.333	1.247	1.882	0.642	1.389	0.700
GREY	12.	1.079	4.000	1.247	1.469	0.575	1.964	0.818
DENTAL	12.	1.079	4.556	1.343	1.912	0.747	1.767	0.527
MEDIA	13.	1.114	4.333	1.491	1.951	0.580	1.393	0.762
MIST	14.	1.146	2.889	0.994	1.468	0.375	-0.120	1.407
RUIN	14.	1.146	4.000	1.054	1.679	0.724	1.052	0.553
LEAP	14.	1.146	4.111	0.994	1.589	0.310	1.158	0.937
WAX	14.	1.146	4.667	1.414	1.946	0.438	1.633	0.773
PIE	14.	1.146	4.889	1.969	2.084	0.717	2.859	0.778
ACE	15.	1.176	3.444	0.497	1.627	0.448	1.297	0.985
PASTOR	17.	1.230	3.667	1.414	1.676	0.327	0.818	1.487
HUNGER	17.	1.230	4.667	1.764	2.145	0.580	2.330	1.005
SAUCE	20.	1.301	4.333	1.414	1.999	0.545	2.172	0.961
CALENDAR	20.	1.301	4.778	1.474	1.890	0.467	1.683	0.647
MATE	21.	1.322	4.111	1.449	2.105	0.456	2.223	0.987
PUBLISHER	23.	1.362	3.778	0.786	1.716	0.192	1.013	1.186
DULL	27.	1.431	4.444	0.956	1.933	0.460	1.726	1.506
CARBON	30.	1.477	4.333	1.563	1.737	0.228	0.782	1.199
FLUX	30.	1.477	2.000	0.943	0.819	0.559	-1.173	0.676
ERA	30.	1.477	3.222	1.474	1.535	0.462	1.454	0.946
COAL	32.	1.505	3.444	1.343	1.698	0.345	1.568	0.983
GUILT	33.	1.519	4.000	1.563	1.876	0.418	1.199	1.244
TIED	34.	1.531	4.889	1.595	2.014	0.443	1.994	0.964
NET	34.	1.531	4.222	1.397	1.823	0.300	1.564	0.865
MYTH	35.	1.544	2.889	0.994	1.334	0.560	0.051	1.332
MAYOR	38.	1.580	3.889	1.728	2.151	0.607	1.797	1.000
TEMPLE	38.	1.580	3.556	1.707	1.516	0.504	0.954	1.191
ELEVEN	40.	1.602	5.333	1.491	2.196	0.637	2.677	0.916
ROYAL	48.	1.681	3.889	1.595	1.477	0.637	1.231	1.196
SIN	53.	1.724	3.667	1.826	1.741	0.837	1.469	1.306
FOOL	53.	1.724	5.222	1.474	2.302	0.374	2.738	0.786
FIFTEEN	56.	1.748	5.111	1.449	2.091	0.671	2.383	0.818
PURE	56.	1.748	4.444	1.707	1.791	0.246	1.783	0.635
LIE	59.	1.771	5.111	1.100	2.315	0.618	2.652	0.916
MINE	59.	1.771	5.889	1.286	2.189	0.344	2.207	0.721
METAL	61.	1.785	4.556	1.707	1.921	0.234	1.743	0.765
TASTE	61.	1.785	5.556	1.165	2.320	0.505	2.368	0.501
MOTOR	62.	1.792	4.444	1.343	1.830	0.215	1.542	0.862
EMPTY	64.	1.806	4.667	1.333	2.197	0.628	2.277	0.841
BIRTH	66.	1.820	5.000	1.247	2.005	0.407	2.439	0.763
CHAIR	66.	1.820	5.889	1.197	2.456	0.693	3.051	0.645
ANIMAL	68.	1.833	5.111	1.449	2.342	0.604	2.443	0.698
FOOT	70.	1.845	5.444	1.257	2.139	0.492	2.531	0.631
VILLAGE	72.	1.857	4.444	1.165	1.781	0.414	1.542	0.677
DOG	75.	1.875	6.222	0.916	2.505	0.658	3.413	0.473
SOFT	77.	1.886	5.444	1.066	2.204	0.238	2.463	0.675
EDITOR	77.	1.886	4.222	0.916	1.838	0.225	1.407	0.753
FILE	81.	1.908	4.222	1.474	1.895	0.480	1.630	0.725
WEATHER	81.	1.908	5.667	1.247	2.250	0.475	2.649	0.950
YOUTH	82.	1.914	4.889	1.286	2.283	0.582	2.321	0.899
BAR	82.	1.914	4.444	1.771	1.881	0.301	1.899	0.998
CRISIS	82.	1.914	4.333	0.943	2.142	0.648	2.104	0.754
SHIP	89.	1.949	4.222	1.474	1.872	0.284	2.034	0.888
CIVIL	91.	1.959	4.778	1.315	1.805	0.373	1.157	0.747

OIL	93.	1.968	4.889	1.100	2.015	0.171	1.896	0.937
HEAT	97.	1.987	6.000	1.054	2.181	0.718	2.861	0.524
GAS	98.	1.991	5.111	1.728	2.375	0.534	2.631	0.946
ACTUAL	100.	2.000	4.556	1.641	2.035	0.252	2.191	0.925
DATE	103.	2.013	6.000	0.943	2.338	0.561	2.654	1.253
MARRIED	105.	2.021	5.111	1.370	2.219	0.553	2.206	0.769
HEAVY	110.	2.041	4.889	0.994	2.181	0.557	2.100	0.786
HOSPITAL	110.	2.041	5.000	1.491	2.173	0.336	2.078	0.865
POOL	111.	2.045	4.667	1.633	2.145	0.512	2.136	0.612
SUN	112.	2.049	6.111	0.875	2.075	0.469	2.680	0.926
DOUBT	114.	2.057	4.444	1.343	2.096	0.699	1.967	0.846
CHARGE	122.	2.086	5.444	1.423	2.244	0.651	2.368	0.947
HOTEL	126.	2.100	4.889	1.286	1.978	0.523	1.897	0.599
SUMMER	134.	2.127	5.444	1.257	2.151	0.625	2.729	0.736
COLOR	141.	2.149	5.778	1.133	2.326	0.627	2.651	0.869
KNOWLEDGE	145.	2.161	4.333	1.826	1.983	0.334	2.093	1.150
CLUB	145.	2.161	5.000	1.333	1.888	0.489	2.467	0.722
LABOR	149.	2.173	3.889	0.875	2.006	0.724	1.700	0.673
NATURAL	156.	2.193	5.000	1.491	2.145	0.628	1.723	0.696
WALL	160.	2.204	5.222	1.750	2.250	0.654	2.323	0.912
FINE	161.	2.207	5.556	1.066	2.438	0.659	2.586	0.797
CENTRAL	164.	2.215	4.778	1.618	1.925	0.198	1.381	0.856
TEN	165.	2.217	5.444	1.257	2.524	0.608	2.868	0.784
SON	166.	2.220	5.667	1.333	2.339	0.530	2.821	0.794
HOLD	169.	2.228	5.333	1.054	2.297	0.639	2.576	0.731
UNION	182.	2.260	4.667	1.333	2.126	0.535	2.371	0.700
GREATER	188.	2.274	5.000	0.943	2.133	0.646	2.040	0.700
BOOK	193.	2.286	5.778	1.618	2.596	0.600	2.864	0.909
IDEA	195.	2.290	5.000	1.247	2.052	0.202	1.868	1.398
RED	197.	2.294	5.667	1.333	2.142	0.339	2.803	0.482
TABLE	199.	2.299	5.444	1.165	2.431	0.583	2.566	0.821
VALUE	201.	2.303	4.556	1.343	1.890	0.436	1.512	1.085
SOUND	204.	2.310	5.333	1.414	2.076	0.291	2.380	1.025
PLAN	205.	2.312	5.000	1.054	2.110	0.506	1.672	1.020
TOTAL	211.	2.324	5.333	1.155	2.233	0.385	1.849	0.676
TOWN	218.	2.338	4.556	1.571	1.899	0.290	1.853	1.054
GIRL	220.	2.342	6.333	1.054	2.613	0.593	3.327	0.788
CONTROL	223.	2.348	4.444	1.571	1.945	0.697	1.821	0.990
AM	228.	2.358	6.111	1.728	2.707	0.582	3.642	0.517
LOVE	232.	2.365	6.333	0.816	2.659	0.538	3.190	0.516
STREET	244.	2.387	5.778	1.133	2.478	0.570	2.680	0.661
REAL	260.	2.415	5.000	1.563	2.204	0.488	3.021	0.546
WORD	274.	2.438	5.667	1.155	2.512	0.524	2.611	1.034
DEATH	277.	2.442	5.222	1.030	2.051	0.708	2.109	0.660
QUITE	281.	2.449	4.444	1.257	1.904	0.271	2.112	0.820
LOCAL	288.	2.459	4.889	1.286	1.966	0.400	1.871	0.646
DONE	320.	2.505	5.000	1.700	2.350	0.618	2.261	1.099
AREA	324.	2.511	4.889	1.449	1.848	0.378	1.859	0.953
CASE	362.	2.559	5.000	1.414	1.853	0.415	2.074	0.641
PRESENT	377.	2.576	5.333	1.491	2.209	0.574	2.154	0.930
GIVEN	377.	2.576	5.111	1.197	2.121	0.509	2.407	0.800
SOCIAL	380.	2.580	4.778	0.916	2.112	0.698	2.149	0.698
EYES	401.	2.603	5.667	1.155	1.829	0.452	2.622	0.413
END	410.	2.613	5.556	1.343	2.393	0.541	2.774	0.594

HAND	431.	2.634	5.889	1.100	2.412	0.570	3.062	0.278
ONCE	499.	2.698	5.444	1.343	2.049	0.380	2.232	1.399
PLACE	571.	2.757	5.111	1.370	2.288	0.625	2.469	0.734
WITHOUT	583.	2.766	5.444	0.956	2.324	0.536	3.023	0.902
HOUSE	591.	2.772	6.000	1.247	2.516	0.597	3.166	0.613
COME	630.	2.799	6.333	0.667	2.554	0.557	3.387	0.520
UNDER	707.	2.849	5.556	1.066	2.468	0.611	3.016	1.057
LIFE	715.	2.854	5.222	1.548	2.219	0.791	2.544	0.751
YEARS	949.	2.977	5.111	1.370	2.162	0.449	2.489	0.760
BEFORE	1016.	3.007	5.667	1.414	2.371	0.554	2.846	0.725
FIRST	1360.	3.134	6.111	0.994	2.434	0.488	3.158	0.500
TIME	1599.	3.204	6.000	1.491	2.448	0.578	3.042	0.995
CAN	1772.	3.248	6.333	1.247	2.674	0.548	3.543	0.486
SO	1985.	3.298	6.667	0.943	2.705	0.559	3.587	0.398
NO	2201.	3.343	6.889	0.314	2.742	0.555	3.821	0.373
MORE	2216.	3.346	5.556	1.343	2.409	0.589	3.284	0.579
ARE	4393.	3.643	6.889	0.314	2.727	0.551	3.277	1.563
BE	6377.	3.805	6.111	1.912	2.604	0.567	3.004	1.774
AS	7250.	3.860	6.556	0.956	2.560	0.645	3.320	0.987
WAS	9816.	3.992	6.889	0.314	2.686	0.493	3.757	0.296
Mean	265.	1.226	4.211		1.769		1.441	
SD	998.	1.217	1.287		0.556		1.239	

TABLE D.3

K-F values, log K-F values, means and standard deviations of subjective ratings of Spanish words in Experiment 2. Nine Spanish-English bilingual raters.

WORD	K-F	LOG-KF	CAT		MAG(LOG)		LIN	
			Mean	SD	Mean	SD	Mean	SD
MOLDURA	0.1	-1.000	3.444	2.061	1.386	0.701	1.320	0.594
ALABAR	0.1	-1.000	3.667	1.700	1.614	0.519	0.951	1.590
CARTILLA	0.1	-1.000	3.222	1.133	1.717	0.562	0.962	0.668
AX	0.1	-1.000	2.556	1.165	0.923	0.828	-1.152	1.204
FRUGAL	0.1	-1.000	2.444	1.423	1.256	0.617	-0.474	1.222
PURE	0.1	-1.000	4.000	1.247	1.574	0.860	0.717	1.548
GARRA	0.1	-1.000	3.333	1.886	1.272	0.867	0.362	1.618
BELLOTA	0.1	-1.000	2.556	0.956	0.982	0.778	0.151	1.341
MANTEL	0.1	-1.000	4.333	1.155	2.219	0.707	2.259	0.590
RAPE	0.1	-1.000	2.667	1.633	1.080	0.789	-0.036	1.210
PLURAL	0.1	-1.000	4.000	1.155	1.837	0.667	1.898	0.648
MONDADO	0.1	-1.000	3.333	1.826	1.068	0.866	-0.019	1.965
VALET	0.1	-1.000	3.333	1.764	1.178	0.564	0.241	0.996
BE	0.1	-1.000	3.556	1.571	1.808	1.118	1.373	1.495
TUNA	0.1	-1.000	4.333	1.333	2.076	0.523	1.702	0.880
ALCORENOQUE	0.1	-1.000	1.556	1.066	0.561	0.591	-0.828	0.861
CLUB	0.1	-1.000	5.000	0.943	2.055	0.431	1.439	0.773
CHOCHEAR	0.1	-1.000	2.444	1.641	1.233	0.996	-0.630	1.452
MARGARINA	0.1	-1.000	4.889	1.728	2.306	0.563	1.824	0.846
CUBO	0.1	-1.000	3.667	1.491	1.700	0.614	1.414	0.905
GRIPE	0.1	-1.000	4.556	1.571	2.198	0.747	1.210	1.364
PUPA	0.1	-1.000	1.556	0.685	0.686	0.577	-1.012	1.076
ADOBES	0.1	-1.000	2.778	1.133	1.399	0.551	0.433	1.092
CAPULLO	0.1	-1.000	2.889	0.994	1.586	0.491	0.492	1.341
ATUN	0.1	-1.000	2.889	1.100	1.221	0.793	0.199	1.559
FILE	0.1	-1.000	2.778	0.916	1.490	0.780	0.070	1.473
QUEJA	0.1	-1.000	4.444	1.343	2.326	0.748	1.942	0.998
SABLE	1.	0.000	2.556	1.832	1.463	0.704	-0.280	1.446
JADE	1.	0.000	2.222	0.916	0.946	0.698	-0.099	1.162
SOJA	1.	0.000	1.889	0.737	1.029	0.663	-0.487	1.526
FIRMEZA	1.	0.000	4.222	1.685	1.613	0.536	0.897	1.398
MESETA	1.	0.000	2.889	1.286	1.515	0.552	0.929	1.288
VALS	1.	0.000	2.444	1.257	1.030	0.693	-0.274	1.340
YEMA	1.	0.000	4.000	1.944	1.494	0.672	0.926	1.394
DRAGON	1.	0.000	3.444	1.707	1.584	0.485	0.621	1.808
MEDIA	1.	0.000	5.000	0.943	2.177	0.846	2.374	1.000
LAUD	1.	0.000	2.778	2.096	0.717	0.563	-0.722	0.992
QUITE	1.	0.000	4.556	1.571	2.224	0.543	1.949	1.176
LOMBRIZ	1.	0.000	3.556	1.832	1.655	0.427	1.102	1.072
SALON	1.	0.000	5.111	1.100	1.971	0.742	2.023	0.546
APUNTE	1.	0.000	4.556	1.499	1.894	0.981	1.937	0.795
AYUDA	2.	0.301	5.667	0.943	2.653	0.535	2.778	0.644

NASAL	2.	0.301	3.444	1.257	1.457	0.674	0.780	1.067
ALETA	2.	0.301	2.667	1.247	1.169	0.613	1.081	1.190
PILAR	2.	0.301	3.444	1.066	1.719	0.375	1.136	0.581
PLACE	2.	0.301	3.222	1.397	1.476	0.714	1.092	1.274
HALO	2.	0.301	3.111	1.792	1.841	0.533	0.496	1.849
AGUERO	2.	0.301	3.000	1.563	1.356	0.545	0.634	1.026
RICINO	2.	0.301	2.000	0.816	0.748	0.637	-0.426	1.722
PILLAR	2.	0.301	3.333	1.247	1.958	0.388	1.053	0.885
PARTO	2.	0.301	3.889	1.728	1.933	0.594	1.692	0.859
BAR	2.	0.301	4.556	1.423	2.031	0.529	1.574	0.967
FACIAL	2.	0.301	4.000	0.943	1.839	0.682	1.682	0.724
DELEITE	2.	0.301	3.111	1.370	1.373	0.784	0.401	1.652
MEMBRILLO	2.	0.301	2.556	1.641	1.195	0.604	-0.352	1.532
DOTE	2.	0.301	2.667	0.943	1.119	0.592	-0.402	1.141
CLAN	2.	0.301	2.556	1.066	1.192	0.750	-0.250	0.995
CISNE	3.	0.477	2.667	1.155	1.329	0.685	0.274	1.389
MORE	3.	0.477	2.889	1.728	1.299	0.716	0.256	1.796
CRANEO	3.	0.477	3.889	1.663	1.827	0.691	1.529	0.665
MINE	3.	0.477	2.444	1.641	1.004	0.811	0.431	1.293
CANAL	3.	0.477	4.778	1.315	1.984	0.388	2.168	0.689
SOLUCION	3.	0.477	4.667	1.155	2.121	0.463	2.259	0.472
MESILLA	3.	0.477	3.111	0.994	1.642	0.464	0.677	0.979
COME	3.	0.477	6.111	0.994	2.521	0.974	3.401	0.597
FINE	3.	0.477	3.667	1.886	1.796	0.616	0.648	1.823
APUESTA	3.	0.477	4.778	1.474	1.902	0.530	1.641	0.959
AROMA	3.	0.477	5.333	1.054	2.408	0.280	2.159	0.726
PASTOR	3.	0.477	4.333	1.700	1.958	0.313	1.642	1.379
CASTOR	3.	0.477	2.889	1.286	1.274	0.631	-0.413	1.511
ESPINA	3.	0.477	3.889	0.994	1.706	0.726	1.474	1.253
SED	4.	0.602	5.778	1.133	2.707	0.553	2.744	1.153
RAPTO	4.	0.602	2.333	1.563	1.163	0.862	1.018	1.435
FERVOR	4.	0.602	3.444	1.257	1.495	0.644	0.970	1.398
TOPO	4.	0.602	3.556	1.423	1.729	0.617	0.412	1.392
RUIN	5.	0.699	3.444	1.571	1.666	0.223	1.318	1.416
PEDESTAL	5.	0.699	4.111	1.728	1.647	0.521	0.837	0.836
ALTAR	5.	0.699	4.333	1.333	1.993	0.338	1.134	1.137
ARE	5.	0.699	4.111	1.969	1.524	0.920	1.252	1.351
MENU	5.	0.699	4.333	0.943	2.088	0.357	1.618	1.096
ESTUPRO	5.	0.699	2.000	1.333	1.118	0.764	0.089	1.487
TRAICION	6.	0.778	4.111	1.370	2.158	0.569	1.597	0.944
QUINTA	6.	0.778	4.444	1.771	2.181	0.491	1.790	0.902
PLUME	6.	0.778	2.444	0.956	0.870	0.688	-0.192	1.400
SOFA	6.	0.778	5.111	1.449	2.375	0.371	2.352	0.488
HACHA	6.	0.778	3.556	2.114	1.697	0.570	1.229	1.362
DONE	7.	0.845	2.889	1.370	1.343	0.901	0.841	1.652
CARGA	7.	0.845	4.444	1.423	2.252	0.713	1.460	1.472
OVAL	8.	0.903	3.444	1.257	1.522	0.427	0.729	0.462
SAUCE	9.	0.954	3.000	1.563	1.377	0.629	0.334	1.178
DUAL	9.	0.954	2.889	1.449	1.339	0.565	0.668	1.404
PRISA	9.	0.954	5.444	1.165	2.616	0.639	2.368	0.865
DERRETIDO	9.	0.954	3.333	2.055	1.723	0.896	1.019	1.027
TALON	9.	0.954	3.778	1.227	1.780	0.398	1.000	1.108
PERFUME	10.	1.000	5.111	1.197	2.363	0.353	2.132	0.547
GREY	10.	1.000	2.889	1.286	1.404	0.690	0.796	1.683

FLUX	11.	1.041	1.778	0.629	0.653	0.602	-0.863	1.053
GRIS	12.	1.079	4.222	1.474	1.800	0.797	1.573	0.700
DENTAL	12.	1.079	4.222	0.786	2.079	0.251	1.741	0.329
MEDIOS	13.	1.114	4.667	1.155	2.308	0.516	2.387	0.730
NIEBLA	14.	1.146	3.778	1.315	1.651	0.687	1.192	0.708
RUINA	14.	1.146	4.222	1.133	1.857	0.449	1.548	1.086
SALTO	14.	1.146	5.222	1.030	2.491	0.673	2.438	0.894
CERA	14.	1.146	3.778	0.786	2.064	0.702	1.726	0.737
PASTEL	14.	1.146	5.000	1.700	2.360	0.362	2.016	1.000
AS	15.	1.176	3.333	1.700	1.540	0.886	1.318	1.474
CLERIGO	17.	1.230	3.222	1.315	1.698	0.840	1.152	1.653
HAMBRE	17.	1.230	6.000	1.054	2.672	0.384	3.229	0.630
SALSA	20.	1.301	5.111	1.197	2.573	0.637	2.267	1.095
CALENDARIO	20.	1.301	5.111	1.370	2.401	0.531	2.541	0.634
COMPANERO	21.	1.322	5.222	1.315	2.033	0.917	2.309	0.914
EDITOR	23.	1.362	4.444	0.956	1.824	0.432	1.599	0.577
MATE	27.	1.431	4.111	1.595	1.784	0.485	0.612	1.539
CARBONO	30.	1.477	3.444	1.499	1.492	0.860	1.001	0.469
FLUJO	30.	1.477	3.000	1.764	1.101	0.748	-0.193	1.221
SIGLO	30.	1.477	4.667	1.155	1.905	0.471	1.672	0.836
CARBON	32.	1.505	4.111	1.595	1.992	0.575	1.404	0.636
CULPA	33.	1.519	4.778	1.548	2.287	0.674	1.861	1.376
LIE	34.	1.531	2.444	1.343	1.138	0.554	-0.062	2.004
RED	34.	1.531	3.222	1.618	1.547	0.744	1.024	1.267
MITO	35.	1.544	2.444	0.831	1.341	0.619	-0.167	1.244
ALCALDE	38.	1.580	4.222	1.685	1.738	0.387	1.652	0.771
TEMPLO	38.	1.580	3.222	0.786	1.512	0.598	0.814	0.534
ONCE	40.	1.602	5.000	1.491	2.255	0.392	2.504	0.927
REAL	48.	1.681	5.000	1.054	2.203	0.415	2.611	0.821
PECADO	53.	1.724	4.889	1.595	2.029	0.610	1.408	0.661
MEMO	53.	1.724	3.667	1.247	1.486	0.630	1.263	0.407
QUINCE	56.	1.748	4.778	1.315	2.047	0.312	2.311	0.776
PURO	56.	1.748	5.111	1.449	2.193	0.475	2.089	0.902
MENTIRA	59.	1.771	5.222	1.030	2.383	0.484	1.900	1.573
MIO	59.	1.771	6.111	0.567	2.821	0.536	3.142	0.601
METAL	61.	1.785	4.667	1.491	2.064	0.456	2.022	0.705
GUSTO	61.	1.785	6.000	1.155	2.445	0.453	2.879	0.825
MOTOR	62.	1.792	5.111	1.197	2.239	0.844	1.990	0.652
VACIO	64.	1.806	5.111	1.197	2.356	0.548	1.670	1.464
NATAL	66.	1.820	3.222	1.030	1.541	0.733	0.561	1.001
SILLA	66.	1.820	5.778	1.227	2.649	0.606	2.589	0.946
ANIMAL	68.	1.833	5.889	1.286	2.468	0.487	2.760	0.461
PIE	70.	1.845	4.667	1.491	2.368	0.543	1.988	0.932
PUEBLO	72.	1.857	5.444	1.343	2.242	0.388	2.362	0.559
CAN	75.	1.875	2.889	1.523	1.250	0.820	0.324	1.807
MOLE	77.	1.886	2.222	1.030	1.035	0.566	-0.887	1.102
REVISOR	77.	1.886	2.444	0.956	1.272	0.712	1.160	0.479
FILA	81.	1.908	4.556	1.499	1.968	1.118	1.459	0.769
TEMPLE	81.	1.908	3.111	1.197	1.357	0.753	0.456	1.352
JOVEN	82.	1.914	6.222	0.786	2.740	0.402	2.868	0.655
BARRA	82.	1.914	3.444	1.707	1.477	1.138	0.500	1.778
TRANCE	82.	1.914	3.667	1.414	1.500	0.746	0.502	0.955
NAVE	89.	1.949	3.444	0.956	1.534	0.452	1.191	1.111
CIVIL	91.	1.959	4.667	1.333	2.112	0.355	1.632	1.507

OLEO	93.	1.968	3.444	1.423	1.592	0.813	1.048	1.415
CALOR	97.	1.987	5.333	1.414	2.643	0.581	2.731	0.785
GAS	98.	1.991	5.556	1.066	2.279	0.760	2.370	0.891
EFECTIVO	100.	2.000	5.111	0.875	2.397	0.558	2.221	0.851
CALENDAR	103.	2.013	4.000	1.886	1.579	1.019	0.712	1.668
CASE	105.	2.021	3.333	1.700	1.563	0.769	1.064	1.403
PESADO	110.	2.041	5.111	1.523	2.376	0.475	2.141	0.916
HOSPITAL	110.	2.041	5.333	1.054	2.247	0.359	2.262	0.720
BALSA	111.	2.045	2.778	1.133	1.509	0.529	0.731	1.384
SOL	112.	2.049	5.444	1.066	2.805	0.654	2.387	0.812
DUDA	114.	2.057	4.778	1.227	2.440	0.689	2.500	0.541
CARGO	122.	2.086	4.222	0.916	1.864	0.430	1.653	0.651
HOTEL	126.	2.100	5.444	1.343	1.971	0.722	2.316	0.643
VERANO	134.	2.127	5.000	1.414	2.461	0.417	2.676	0.732
COLOR	141.	2.149	5.444	1.423	2.522	0.383	2.667	0.648
SABER	145.	2.161	5.778	0.916	2.624	0.529	2.673	0.638
PORRA	145.	2.161	2.556	1.771	1.080	0.748	0.083	1.300
LABOR	149.	2.173	4.556	1.165	2.007	0.789	1.691	0.670
NATURAL	156.	2.193	5.444	1.343	2.232	0.408	2.652	0.517
PARED	160.	2.204	4.889	1.595	2.170	0.431	1.740	1.461
FINA	161.	2.207	3.889	1.853	1.893	0.576	1.258	0.947
CENTRAL	164.	2.215	5.111	1.286	2.032	0.401	1.847	0.541
DIEZ	165.	2.217	5.222	1.474	2.121	0.335	2.788	0.763
HIJO	166.	2.220	6.222	0.916	2.803	0.521	3.436	0.647
TEN	169.	2.228	4.000	2.108	1.996	0.646	1.642	1.681
UNION	182.	2.260	5.000	1.054	2.136	0.402	1.640	1.355
MAYOR	188.	2.274	4.889	1.197	2.060	0.482	2.206	0.645
LIBRO	193.	2.286	6.111	0.737	2.804	0.686	3.132	0.900
IDEA	195.	2.290	5.778	1.315	2.455	0.440	2.809	0.554
ROJO	197.	2.294	5.000	1.333	2.323	0.354	2.297	0.534
MESA	199.	2.299	5.667	1.155	2.837	0.683	3.196	0.634
VALOR	201.	2.303	5.444	1.343	2.271	0.415	1.991	0.577
SON	204.	2.310	6.444	0.497	2.755	0.579	3.129	0.420
PLAN	205.	2.312	5.111	1.286	2.146	0.351	2.351	0.654
TOTAL	211.	2.324	5.333	1.155	2.322	0.425	2.294	0.878
VILLA	218.	2.338	4.222	1.133	1.900	0.312	1.744	0.614
NINA	220.	2.342	5.333	1.491	2.115	0.948	2.518	0.822
CONTROL	223.	2.348	5.000	1.414	2.213	0.389	2.119	0.303
SOY	228.	2.358	6.556	0.497	2.783	0.656	3.520	0.593
AMOR	232.	2.365	6.222	1.133	2.892	0.556	3.389	0.667
CALLE	244.	2.387	6.000	0.816	2.653	0.440	2.671	0.906
VERDADERO	260.	2.415	4.778	1.618	2.314	0.533	2.207	0.826
PALABRA	274.	2.438	6.000	0.816	2.613	0.421	2.999	0.681
MUERTE	277.	2.442	5.556	1.165	2.433	0.299	2.424	0.510
ENTERAMENTE	281.	2.449	4.222	1.685	2.258	0.688	1.980	1.067
LOCAL	288.	2.459	5.000	1.155	2.094	0.345	1.669	0.571
HECHO	320.	2.505	5.111	1.449	2.298	0.540	2.387	0.740
AREA	324.	2.511	5.000	1.247	2.133	0.357	2.091	0.800
CASO	362.	2.559	4.778	1.474	2.248	0.340	1.833	0.616
ACTUAL	377.	2.576	5.222	0.786	2.288	0.473	2.338	0.464
DADO	377.	2.576	4.556	1.641	1.654	0.935	1.872	0.845
SOCIAL	380.	2.580	5.000	1.333	2.301	0.503	2.274	0.553
OJOS	401.	2.603	5.889	0.994	2.794	0.657	2.672	0.625
FIN	410.	2.613	5.667	1.054	2.557	0.331	2.817	1.062

MANO	431.	2.634	6.000	1.054	2.535	0.399	3.021	0.610
VEZ	499.	2.698	5.667	1.333	2.596	0.511	3.116	0.538
LUGAR	571.	2.757	5.333	1.054	2.292	0.526	2.059	0.881
SIN	583.	2.766	5.889	0.737	2.697	0.354	2.878	0.535
CASA	591.	2.772	6.111	1.197	2.731	0.528	3.282	0.840
VENIR	630.	2.799	6.000	1.054	2.730	0.703	2.788	0.604
SO	707.	2.849	3.778	1.474	1.367	0.768	1.744	1.549
VIDA	715.	2.854	5.889	1.100	2.544	0.641	3.144	0.485
ANOS	949.	2.977	5.556	1.066	2.539	0.620	2.819	0.483
ANTES	1016.	3.007	5.556	1.066	2.696	0.546	2.698	0.688
PRIMER	1360.	3.134	5.889	0.875	2.430	0.529	2.660	0.759
TIEMPO	1599.	3.204	6.222	0.786	2.618	0.413	3.098	0.548
ENVASE	1772.	3.248	3.778	1.474	1.432	0.754	0.661	1.228
ASI	1985.	3.298	6.444	0.497	2.856	0.685	3.393	0.404
NO	2201.	3.343	7.000	0.000	3.000	0.588	3.811	0.455
MAS	2216.	3.346	6.556	0.497	2.877	0.448	3.543	0.384
ESTAN	4393.	3.643	6.778	0.629	2.834	0.436	3.810	0.354
SER	6377.	3.805	6.333	0.816	2.580	0.757	3.172	0.869
QUE	7250.	3.860	6.889	0.314	2.979	0.596	3.726	0.490
ERA	9816.	3.992	5.556	1.343	2.775	0.618	3.138	0.815
Mean	265.	1.226	4.303		1.932		1.562	
SD	998.	1.217	1.230		0.553		1.096	

References

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. Journal of Experimental Psychology: Human Perception and Performance, 2, 556-566.

Caramazza, A., & Brones, E. (1979). Lexical access in bilinguals. Bulletin of the Psychonomic Society, 13, 212-214.

Caramazza, A., & Brones, E. (1980). Semantic classification by bilinguals. Canadian Journal of Psychology, 34, 77-81.

Carroll, J. B. (1971). Measurement properties of subjective magnitude estimations. Journal of Verbal Learning and Verbal Behavior, 10, 722-729.

Carroll, J. B., & White, M. N. (1973a). Age-of-Acquisition norms for 220 picturable nouns. Journal of Verbal Learning and Verbal Behavior, 12, 563-576.

Carroll, J. B., & White, M. N. (1973b). Word frequency and age of acquisition as determiners of picture naming latency. Quarterly Journal of Experimental Psychology, 25, 85-95.

Clark, H. H. (1973). The language-as-a-fixed-effect fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behavior, 12, 335-359.

Clark, H. H., & Clark, E. V. (1977). Psychology and language. New York: Harcourt Brace Jovanovich.

Cristoffanini, P., Milech, D., & Kirsner K. (in press) Bilingual lexical representation: the status of cognates. Quarterly Journal of Experimental Psychology.

Dalrymple-Alford, E. C. (1968). Interlingual interference in a color-naming task. Psychonomic Science, 10, 215-216.

Dillon, R., McCormack, P. D., Petrusic, W., Cook, G., & Lafleur, L. (1973). Release from proactive interference in compound and coordinate bilinguals. Bulletin of the Psychonomic Society, 2, 293-294.

Dixon, P., & Rothkopf, E. (1979). Word repetition, lexical access, and the process of searching words and sentences. Journal of Verbal Learning and Verbal Behavior, 18, 629-644.

Dyer, F. (1971). Color-naming interference in monolinguals and bilinguals. Journal of Verbal Learning and Verbal Behavior, 10, 297-302.

Ebel, R. L. (1951). Estimation of the reliability of ratings.

Psychometrika, 16, 407-424.

Edwards, A. L. (1967). Statistical methods. New York: Holt, Rinehart and Winston.

Fang, S. P., Tzeng, O. J. L., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. Memory and Cognition, 6, 609-617.

Forbach, G., Stanners, R., & Hochaus, L. (1974). Repetition and practice effects in a lexical decision task. Memory and Cognition, 2, 337-339.

Glanzer, M., & Duarte, A. (1971). Repetition between and within languages in free recall. Journal of Verbal Learning and Verbal Behavior, 10, 625-630.

Glanzer, M., & Ehrenreich, S. L. (1979). Structure and search of the internal lexicon. Journal of Verbal Learning and Verbal Behavior, 18, 381-398.

Goggin, J., & Wickens, D. (1971). Proactive interference and language change in short-term memory. Journal of Verbal Learning and Verbal Behavior, 10, 453-458.

Hines, T. M. (1978). The independence of languages in bilingual memory. (Doctoral dissertation, University of Oregon). Dissertation Abstracts International, 1978-1979, 39, 1012-B. (University Microfilms No. 78-14, 314)

Hornby, P. A., ed. (1977). Bilingualism: psychological, social and educational implications. New York: Academic Press.

Hoyt, C. J. (1941). Test reliability estimated by analysis of variance. Psychometrika, 6, 153-160.

Kintsch, W. (1970). Recognition memory in bilingual subjects. Journal of Verbal Learning and Verbal Behavior, 9, 405-409.

Kirsner, K., Brown, J., Abrol, S., Chadha, N., & Sharma, N. (1980). Bilingualism and lexical representation. Quarterly Journal of Experimental Psychology, 32, 585-594.

Kirsner, K., & Smith, M. C. (1974). Modality effects in word identification. Memory and Cognition, 2, 637-640.

Kirsner, K., Smith, M. C., Lockhart, R. S., King, M. L., & Jain, M. (1984). The bilingual lexicon: Language specific units in an integrated network. Journal of Verbal Learning and Verbal Behavior, 233, 519-539.

Kolers, P. A. (1963). Interlingual word associations. Journal of Verbal Learning and Verbal Behavior, 2, 291-300.

Kolers, P. A. (1966a). Reading and talking bilingually. American Journal of Psychology, 79, 357-376.

Kolers, P. A. (1966b). Interlingual facilitation of short term memory. Journal of Verbal Learning and Verbal Behavior, 5, 311-319.

Kolers, P. A. (1975). Specificity of operations in sentence recognition. Cognitive Psychology, 7, 289-306.

Kolers, P. A., & Gonzalez, E. (1980). Memory for words, synonyms, and and translations. Journal of Experimental Psychology: Human Learning and Memory, 6, 53-65.

Kucera, H., & Francis, W. (1967). Computational analysis of present-day American English. Providence, R. I.: Brown University Press.

Lambert, W. E., Ignatow, M., & Krauthamer, M. (1968). Bilingual organization in free recall. Journal of Verbal Learning and Verbal Behavior, 7, 207-214.

Lopez, M., & Young, R. K. (1974). The linguistic interdependence of bilinguals. Journal of Experimental Psychology, 102, 981-983.

MacNamara, J., & Kushnir, S. L. (1971). Linguistic independence of bilinguals: The input switch. Journal of Verbal Learning and Verbal Behavior, 10, 480-487.

McCormack, P. D. (1977). Bilingual linguistic memory: The independence-interdependence issue revisited. In P. A. Hornby (Ed.), Bilingualism: psychological, social and educational implications. New York: Academic Press.

Meyer, D. E., & Ruddy, M. G. (1974). Bilingual word-recognition: Organization and retrieval of alternate lexical codes. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia.

Meyer, D. E., & Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. Journal of Experimental Psychology, 90, 227-234.

Meyer, D. E., Schvaneveldt, R., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word-recognition. Memory and Cognition, 2, 309-321.

Monsell, S., & Banich, M. T. (1982). Lexical priming: repetition effects across input and output modalities. Paper presented at the 23rd Annual Meeting of the Psychonomics Society, Minneapolis.

Morrison, D. F. (1976). Multivariate statistical methods. New York: McGraw-Hill.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired associate learning. Journal of Verbal Learning and Verbal Behavior,

4, 32-38.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology Monograph, 76(1, Pt. 2).

Posner, M. I. (1969). Abstraction and the process of recognition. In G. H. Bower, & J. T. Spence (Eds.), The psychology of learning and motivation (Vol. 3). New York: Academic Press.

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word test. Journal of Verbal Learning and Verbal Behavior, 8, 295-301.

Rose, R. G., & Carroll, J. F. (1974). Free recall of a mixed language list. Bulletin of the Psychonomic Society, 3, 267-268.

Rose, R. G., Rose, P. R., King, N., & Perez, A. (1975). Bilingual memory for related and unrelated sentences. Journal of Experimental Psychology: Human Learning and Memory, 1(5), 599-606.

Roydes, R. L., & Osgood, C. E. (1972). Effects of grammatical form-class set upon perception of grammatically ambiguous English words. Journal of Psycholinguistic Research, 1, 165- 174.

Rubenstein, H., Lewis, S., & Rubenstein, M. (1971). Evidence for phonemic coding in visual word recognition. Journal of Verbal Learning and Verbal Behavior, 10, 645-657.

Saegert, J., Hamayan, E., & Ahmar, H. (1975). Memory for language of of input in polyglots. Journal of Experimental Psychology: Human Learning and Memory, 1(5), 607-613.

Scarborough, D., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. Journal of Experimental Psychology: Human Perception and Performance, 3, 1-17.

Scarborough, D., Gerard, L., & Cortese, C. (1979). Accessing lexical memory: The transfer of word repetition effects across task and modality. Memory and Cognition, 7, 3-12.

Scarborough, D., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. Journal of Verbal Learning and Verbal Behavior, 23, 84-89.

Shapiro, B. J. (1969). The subjective estimation of relative word frequency. Journal of Verbal Learning and Verbal Behavior, 8, 248-251.

Stanners, R., Jastrzembski, J., & Westbrook, A. (1975). Frequency and visual quality in a word-nonword classification task. Journal of Verbal Learning and Verbal Behavior, 14, 259-264.

Stanners, R., Neiser, J. J., & Hall, R. (1979). Memory representation

for morphologically related words. Journal of Verbal Learning and Verbal Behavior, 18, 399-412.

Stevens, S. S. (1956). The direct estimation of sensory magnitudes-loudness. American Journal of Psychology, 69, 1-25.

Stevens, S. S. (1958). Problems and methods in psychophysics. Psychological Bulletin, 55, 177-196.

Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. Journal of Verbal Learning and Verbal Behavior, 18, 427-440.

Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. Cognitive Psychology, 1, 86-98.

Williams, E. B. (1968). The new college Spanish and English dictionary. New York: Amsco.

Winer, B. J. (1971). Statistical principles in experimental design. New York: McGraw-Hill.