

On the Structure of Mental Lexicon: Merging the Psycholinguistic and Construct Validation Studies

Hamideh Marefat
University of Tehran

Abstract

A major issue in testing is the investigation of the nature and structure of language proficiency. Psycholinguists, too, deal with the internal mental representation of linguistic components. This is exactly where the convergence of testing and psycholinguistics reveals itself. One of the key issues in psycholinguistics concerns the form in which words are represented in the mind. Results of the studies in this regard show that prior to lexical search, a morphological analysis of words is attempted. So the morpheme and not the word is a basic unit in the mental lexicon. In testing, too, specialists have been concerned with the structure of language proficiency and whether certain traits exist or not. There exists strong evidence signaling the existence of a number of traits like listening and speaking abilities. In this piece of research attempt has been made to see if the finding in psycholinguistics, that the morpheme and not the word is a basic unit in the mental lexicon, matches the evidence coming from factor analysis which is so commonly used in testing. In this study, vocabulary items in a TOEFL were divided into categories of words with various combinations of morphemes and administered on 132 subjects. Findings from construct validity studies, based on factor analytic procedures, were found to show that prefixed words make up a construct different from the words with derivational suffixes, implying that vocabulary knowledge is not stored in one place. This finding supports the decompositional model in psycholinguistics which suggests that lexicon is represented in a morphologically decomposed form. As far as the distinction between derived and inflected words is concerned, the results of factor analysis show that inflected words do not load on the same factor as the derivational words do. So, it is concluded that derivational and inflectional morphemes, too, are different constructs. The whole finding in this piece of research suggests that due to the problems in the correlational evidence for construct validity studies, and due to the consensus in the field that statistical validity alone can no longer qualify a test as a good language test, as the papers in *Language*

Testing (1985) put it, specialists in testing can get insights from psycholinguistic findings.

Introduction

The primary matter of concern in testing is measuring the testee's language knowledge. This assessment brings about complications as to the issues of validity and reliability. It is platitudinous to say that the most important quality of a test is that it should be valid and that the most important type of validity is construct validity (Davies, 1990: 36; Bachman, 1990, p.25; Henning, 1987, p.10). In the discussion of construct validity, Messick says that "a measure estimates how much of something an individual displays or possesses. The basic question [of construct validity] is, what is the nature of that something?" (Messick, 1975, p.957). In second language testing, it is common to devise tests having several parts. The rationale behind this inclusion is that any one of them tests one separate component which identifies the nature of the "thing" the test measured; this thing is referred to in the literature as construct. In essence, a construct is a hypothetical variable that is derived from a theory. Theories identifying the nature of such constructs have undergone many drastic changes ranging from the work of Lado and Fries to that of Canale & Swain. Following Fries and Lado, language ability was conceived of as including many aspects. The models they introduced distinguished skills from components of knowledge. So, the tests in line with this framework, which was influenced by the linguistic theory of structuralism, were based on grammar, vocabulary, phonology, graphology, etc. But as Bachman (1990) puts it their model has serious limitations: "It was not clear whether the skills were simply manifestations of the knowledge of components in different modalities and channels or whether they were qualitatively different in other ways" (p, 82). Moreover, as Bachman goes on to say such a model fails to account for the full context of language use "... the context of discourse and situation" (p.82).

Later on when Canale and Swain (1980) introduced their communicative competence model, attempt was made to construct test

tapping the constructs representing the components of this model. Canale and Swain claimed "pencil-and-paper tests now in use do not necessarily give a valid indication of second language learner's skills in performing in actual communicative situations" (p.34).

So, tests constructed in line with this theory do not measure linguistic knowledge per se; rather, they measure the productive use of language in situations. The different constructs this model of language ability contains include sociolinguistic competence, strategic competence, and grammatical competence.

At the time being, the radically different approaches to language test development, reflecting different views of language ability, are represented in the two language proficiency tests of TOEFL and IELTS currently administered worldwide. If we can have a clear picture of the constructs underlying language ability, being well aware of what we intend to test, it would be possible to construct tests tapping those constructs. But why is it that construct validity studies do not help us in achieving that clear picture?

How to Provide Evidence for Construct Validity?

According to Henning (1987), there are two major kinds of validity: empirical and non-empirical or theoretical. Underhill (1982, cited in Grotjahn, 1980) divides empirical validity into two types: concurrent validity and predictive validity; and non-empirical validity into three types: face, content, and construct validity. Underhill (1982) adds that in order to assess this latter type of validity, i.e., theoretical validity, "we have to rely primarily on intuition and introspection" (p. 161). Davies, too, asserting that empirical studies cannot improve the construct validity of a test says "what is most important is the preliminary thinking and the preliminary analysis as to the nature of the language learning we aim to capture" (Davies, 1990, p.37).

In line with this type of thinking, Messick (1989) says that the basic sources of validity evidence fall into a half dozen forms and mentions the first type of evidence as "engaging in judgmental and logical analysis as is done in documenting content relevance and representativeness" (p.49).

Thorndike and Hagen (1977), too, refer to the priority of introspection in construct validity when they say "evidence of validity is ... partly rational and partly empirical. Rational consideration of what is measured takes the center of the stage..." (p. 73).

Although construct validity is one of the types of non-empirical validity, as we see in the literature, it is treated quite like empirical validity, "not on the basis of intuition and introspection, but with the help of correlational analysis" (Grotjahn, 1986, p. 161). The methods for identifying the construct validity of tests are based on correlational studies or factor analytic procedures.

The basic notion underlying construct validity studies is that "persons high on the construct should score high on a variety of indicators of that construct" (Linn, 1989, p. 51). In this regard, Campbell and Fiske point out that in order to show the construct validity of a test, the multiple methods of measuring the same trait should converge, but the multiple traits should diverge, and what is needed is the correlations of the same trait assessed by various methods and correlations of different traits assessed by the same method (MTMM).

The other statistical technique pertinent to construct validation of a test is factor analysis. As Messick (1989) puts it "as the number of constructs and methods under study increases, it often becomes difficult to disentangle the threads of multiple influences by simply examining, however systematically, the surface patterns of correlation coefficients. Some method is needed for controlling certain variables while examining relationships among other variables. This can be accomplished by means of factor analysis" (p. 52). Factor analysis indicates the extent to which test scores measure certain basic mental skills, often described as factors that underlie performance on a variety of tests. This technique calls for the exercise of subjective judgement in identifying the factors and the use of experimental data for relating test scores and factors. As Thorndike and Hagen (1955) say "the factor analyst attempts to identify a small number of underlying factors that can account for the complete set of relationships among test variables. Each test has a loading on each of the factors... and the analyst tries to

arrive at a pattern of factor loadings that is simple and psychologically meaningful" (p.347-348).

Criticism of Correlational Evidence for Construct Validity

As far as the first technique is concerned, i.e., MTMM, it is all in all a correlational analysis and as Lindquist (1951) puts it "it is doubtful that any other statistical techniques have been so generally and widely misused and misinterpreted in educational research as have those of multiple correlations" (p.690). Grotjahn (1986), too, says that this approach is not a satisfactory one when he says "validation by correlational analysis tells us nothing about the individual mental processes going on in a learner when he or she takes a test" (p. 161). He is against relying on statistical findings because these procedures "will not provide us with a real understanding of what a specific language test measures" (p.160).

The second approach to construct validity, i.e., factor analytic procedures, is not without its shortcomings, either. As an instance, Grotjahn (1986) says that this approach is unsatisfactory because "the results of factor analysis depend heavily on the number and type of variables included in a study or on the specific factor analytic technique used" (p. 162). Thus, finding significant loadings on specific factors is not so much revealing and new strategies should be found. Helmut Vollmer, too, refers to this point when he says:

What we suggest is the application of alternative research strategies.... We will never really understand what correlations between tests of different skills mean, what they signify, and why some are higher than others - unless we better understand and are able to model more precisely the cognitive potentials and task specific operations on which performance in the various language tests depends. (Vollmer, 1981, p.115, cited in Grotjahn, 1986)

A Psycholinguistic Approach to Mental Representation of Language Knowledge

All the points regarding second language ability discussed above, i.e., the structure of language competence, the way knowledge of language is organized in the mind of the learner, and the more or less strong empirical evidence provided to find out what the structure of foreign language ability exactly is, are also central issues in psycholinguistics.

Much psycholinguistic research, too, has been concerned with the question of the nature of "the ability involved in human capacity called language" (Kess, 1976, p. xi) or with discovering the processes whereby the actual language behavior is acquired and maintained (Kess, 1976, p. 6).

Psycholinguists, in the desire of identifying the nature of language knowledge, started by examining the grammatical models offered by linguists to see if such models represent the language knowledge held by the user of language. So, for such a long time, psycholinguists like Miller and psycholinguists like Greene, were concerned with evaluating the claims made on the psychological reality of Chomsky's grammar. Of course, the reason why they took Chomsky's grammatical model as a starting point was Chomsky's own claims with regard to the psychological validity of his model: Obviously, every speaker of a language has mastered and internalized a generative grammar that expresses his knowledge of his language...

Any interesting generative grammar will be dealing, for the most part, with mental processes.... (Homsy, 1965, p. 8; cited in Steinberg, 1982, p. 14). Referring to this quotation from Chomsky, Steinberg (1982) says that it "indicates that Chomsky was arguing for a strict mentalistic interpretation of the rules of grammar which he was postulating" (p. 74). Moreover, when Chomsky (1967) says about deep structure that it "must be rediscovered by each child who learns a language" (p.399), he is claiming that deep structure has psychological reality, "otherwise a child couldn't be expected to rediscover it" (Steinberg, 1982, p.75). Being impressed by the claims Chomsky had made, psycholinguists started evaluating it from a psycholinguistic point of view. As far as the structures of a derivation are concerned, the question raised was whether the rules of the generative grammar are the actual rules used by speakers in the production of sentences. So many experiments were conducted and "it became obvious that the results could not be accounted for solely in terms of syntactic transformational operations" (Green, 1972, p. 107). Rather semantic factors play a crucial role and speakers of a language start with a

meaning in their mind rather than beginning with syntax as is the case in Chomsky's theory.

Throughout time, the concerns of psycholinguists changed. As Newmeyer (1988) says "during the 1960s, psycholinguists has concentrated primarily on syntactic variables in sentence processing. During the 1970s, they focused on higher level processes such as the comprehension and meaning of discourse or text, and on lower level processes such as the recognition of lexical and sub-lexical units" (p. 11).

As far as lexical processing is concerned, the questions to be dealt with include: what is the nature of the human word store or mental lexicon? How is the information related to the lexicon – including phonological, morphological, semantic and syntactic properties – organized and represented in the mind?

So, as can be seen from the previous paragraphs, there is a similarity between what psycholinguistic findings and construct validity studies imply concerning the structure of human language knowledge. The focus of what psycholinguistics has to offer regarding the representation of a certain linguistic component, i.e., vocabulary, in the mind and that which the construct validity approach offers.

The Structure of Vocabulary Knowledge from a Psycholinguistic Viewpoint

In psycholinguistics, many models have been offered regarding the architecture of mental lexicon which hypothesize that "a lexical entry consists of information (phonological, orthographic, semantic, etc.) stored in separate information-specific components. Lexical knowledge about a word is therefore not stored in one place, but is represented in different sub-lexicons" (Emmorey and Fromkin, 1988, p. 125). For instance, as far as morphologically complex words are concerned, it has been suggested that "affixed words ... are stored in their base form in the lexicon" (Taft and Forster, 1975, p. 638). Conducting various experiments, Taft and Forster concluded, "... a morphological analysis of words is attempted prior to lexical search" (p. 643) and thus, mental lexicon consists of both free and bound morphemes. But there have also been pieces of counter evidence and

the issue of the representation of morphologically complex words in the mind has become a controversial issue. As Laudanna et al (1994) put it there are two main contrasting suggestions in this regard in the literature. (1) The "fully non-decompositional models in which whole word representations are the only units of access in the lexicon" (Rubin, Becker, and Freeman, 1979). (2) The "completely decompositional models in which lexical access takes place through procedures of morphological decomposition of the input word and lexical entries are morphologically decomposed" (Taft and Forster, 1975). So, the evidence is not conclusive. Perhaps the reason why the picture is not so clear is that, as Laudanna et al (1994) put it, these theories "have not adequately taken into account a number of relevant dimensions..." (p. 297). The point that should be taken into account is the distinction between derivational and inflectional morphemes. As Aitchison (1987) puts it "inflectional suffixes are commonly added on as needed in the course of speech, but that derivational prefixes and suffixes are already attached to their stems" (p.117). Thus, it is possible that the alternative views on the representation of morphologically complex words are due to the different treatments towards these relevant dimensions. Laudanna et al (1994), too, refer to this point when they say "... there are both theoretical and empirical reasons for not considering prefixed words as a homogeneous class. The type of prefix included is one of the factors that may contribute to the dishomogeneity [of the finding in the literature]" (p. 299). Thus both linguistically and empirically, it is plausible to acknowledge the difference between various types of morphemes.

The Structure of Vocabulary Knowledge according to the Findings of a Construct Validity Approach

In testing, the existence or non-existence of the morphemes as separate factors or constructs must be shown by factor analytic studies.

As far as listening is concerned, evidence has been provided showing that listening does exist as a separate trait in second language ability (Buck, 1992, p. 352). As far as reading is concerned, Alderson

(1990) has tried to see if reading consists of distinct separately identifiable skills or it should be considered as a single aptitude. So many studies have also been conducted on the construct validation of tests of communicative competence and also on the construct validation of their components. The question of concern in this piece of research concerns the structure of words in the mind. The method of answering this question in testing is through statistical procedures of MTMM or factor analysis. It might be found out through these procedures that vocabulary is a construct. A further step in this process would be to see if the components of words also have construct validity; again through factor analysis.

The Study

Aims

The focus of research to be presented here is the comparability of the nature of the representation of morphologically complex words as psycholinguists suggest with the findings of construct validation studies in testing regarding the components of lexicon. Specifically speaking, this study aims at finding out whether the finding in psycholinguistics, that the morpheme and not the word is a basic unit in the mental lexicon, matches the evidence coming from factor analysis that is so commonly used in the field of testing. The prediction is that the results of factor analysis should show different loadings for the variously complicated words. If this turns out to be the case, then it will be shown that this single aspect of psycholinguistic findings maps the findings gained through construct validity studies.

A further aim of this study is to see if the much-debated issue in psycholinguistics regarding the different representations of derivational and inflectional morphemes can be resolved in any way by construct validity studies. More specifically, attempt will be made to see if derived and inflected words show comparable factor loadings or not. If it is found throughout this study that both inflected and derived words show comparable loadings, or if it is found that only the inflected words have a considerable loading, it will help to resolve the issue mentioned above in psycholinguistics. Because as Laudanna

et al (1994) put it: "the empirical evidence is not completely unambiguous when derived words (both prefixed and suffixed) are taken into consideration" (p. 296).

Research Questions

1. *Does the finding in psycholinguistics, that the morpheme and not the word is a basic unit in the mental lexicon, match the results of factor analysis of the different categories of the words in the vocabulary section of a TOEFL test?*
2. *Do the findings of the factor analytic procedures applied on the derived and inflected words provide a clear-cut picture so that the ambiguity in this regard in psycholinguistics can be resolved?*

Subjects

The subjects in this study were 132 Persian-speaking learners of English as a foreign language at Allameh Tabataba'i University, taking their Advanced Writing and Essay Writing courses. The subjects took the test as part of their course requirements. Of these, 110 were selected. The rest were excluded because they had wrongly answered more than half of the vocabulary items. The criterion for selection thus was for the subjects to have accurately replied more than half of the 30 items.

Material

The material in this study included the TOEFL test, and the statistical analysis was conducted on the data from the vocabulary section of this test.

Procedure

The vocabulary items of the TOEFL administered on the subjects were divided into various categories reflecting their morphological structure. For example, words like *dusk* and *feat* were included in the category of stems. Out of the 30 items in the vocabulary section of the TOEFL test administered, 5 were stems. Since the other categories also included stems, these 5 items were excluded from the analysis.

The other category included words having stems + derivational suffixes. This category included words having either one or two derivational morphemes, or words having an inflectional morpheme at the end, too. Examples include *markedly* and *decorated*. This category, including 9 items, is referred to with the initials **SD**, **SDD**, **SDI**, throughout this study. The next category included words having stems + inflectional morphemes. Examples include *needs*, *vats*, and *limits*. This category, including 7 items, is referred to with the initials **SI** throughout this study. The last category included items with prefixes + stems having either derivational or inflectional morphemes at the end. Examples from this category include words like *predetermined*, *advocates*, *discontented*, etc. This category, including 9 items, is referred to with the initials **PS**, **PSI**, **PSD**, and **PSDI**.

Following is a list showing the morphological structure of words with examples for each:

	SD	stem + derivational morpheme	
		<i>powerful</i>	
	SDD	stem + derivational morpheme + derivational	
morpheme		<i>markedly</i>	
	SDI	stem + derivational morpheme + inflectional	
morpheme		<i>decorated</i>	
	SI	stem + inflectional morpheme	
		<i>needs</i>	
	PSI	prefix + stem + inflectional morpheme	
		<i>discontented</i>	
	PSD	prefix + stem + derivational morpheme	
		<i>substantial</i>	
	PS	prefix + stem	
		<i>benign</i>	

For the analysis of the data, distributions, correlations, and factor analyses were computed using SPSS. As for the factor analyses, both a confirmatory ("factor analysis in which specific expectations concerning the number of factors and their loading are tested on sampled data", Kim and Mueller, 1978, p. 76) and exploratory ("factor

analysis which is mainly used as a means of exploring the underlying factor structure without prior specification of number of factors and their loadings, Kim and Mueller, 1978, p. 77) approaches were used. The aim was to find an evidence of the reality of separate constructs for the various parts of vocabulary items. First an exploratory factor analytic procedure was used to classify the subcategories. Initially, 12 factors were extracted as default, but since there were factors on which no variable loaded at .3, gradually, the number of factors was reduced. In the long run, the 5-factor solution was found to be the most interpretable one. Principal factor analysis with iterations and varimax rotation was run on the data in this study.

The data were analyzed three times:

1. All the items were included in the analysis but the subjects were divided into 2 groups based on their proficiency. The advanced group included those subjects whose score was .25 SD above the mean; and the weak group included those subjects whose score was .75 SD below the mean. The one standard deviation difference between the two groups was to make sure that the groups were quite distinct from each other. But factor analysis couldn't be applied on this data because as the SPSS warned there was no variance to factor. In other words the members in each group were so homogeneous that no variance was gained.

2. All the items were included (except for the stems, for the reason mentioned above); and all the subjects were included in the analysis and no distinction was made between the different levels of proficiency. The results related to this part are reported below.

3. All the subjects were included in the analysis but the items to which more than half of the subjects had answered wrongly were excluded from the study. In other words, items whose item difficulty was high were excluded. In this way, items 16, 19, 22, 23, 25, 27, 28, and 30 to which only, respectively, 53, 33, 54, 31, 50, 35, 36 and 19 ones (out of a total of 132) had replied accurately, were excluded. The results of the data gathered in this way are also reported below.

Results and Discussion

The results for the first type of analysis will not be reported here because of the problem of lack of variance as mentioned above.

The results for the second type of analysis are as follows. Table 1, provides the factor loadings for all the items in the test excluding those that have only stems without any affixation. After a step-by-step exclusion of un-interpretable factors, a 4-factor solution was chosen which has many variables loading at .30 or better.

Table 1. Varimax Rotated Factor Matrix for all the items with affixes

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
SDD9	.32667	-.12211	.26339	-.05005
SDD19	.36382	.02270	.01245	-.07033
SDI25	.24711	-.18890	-.07780	.47779
SD2	.18841	.07607	-.05885	-.29791
SD3	-.07239	.25599	.42129	.07252
SD11	.04398	.06710	.41188	-.03555
SD16	.38226	.31864	.13155	-.03328
SD26	.51892	.10367	.32453	.22426
SD28	.52702	.15133	.16587	-.03233
SI4	.05008	-.07358	.37182	-.08087
SI5	.31362	.01832	-.03112	.00612
SI6	-.13393	.01898	.20115	.13592
SI7	.00993	.41052	.07149	-.09834
SI17	.12635	.06475	.13051	.04779
SI22	-.07148	.11132	-.05796	.50314
SI24	.14637	.14410	.46364	.37413
PSI1	-.02689	.37584	.04309	.00945
PSI18	.13856	.26868	.12158	-.01483
PSI20	.17188	.52311	-.17432	.12391
PSI21	.10514	.30928	.31387	.20072
PSD12	.15211	.34004	.08248	.08875
PSD29	.02344	.25670	-.02908	-.02530
PSD23	.38650	.06541	-.20172	.05248
PS14	-.07836	.32372	.10249	.21310
PSDI15	.10363	.21326	.15367	.26736
Eigenvalue	2.32397	1.04149	.90505	.79237
Pct of Var	.9.3	4.2	3.6	3.2

Factor 1 includes 5 variables from the category with stems and derivational suffixes, all loading at between .30 and .53. This factor is labeled Derivation on the basis of the combination of the variables loading on it.

Factor 2 in this analysis appears to be a prefix factor since out of the 9 items having prefixes as part of their structure, with inflectional or derivational suffixes at the end, 7 items load on it. These variables have loadings on this factor ranging from .30 to .50.

Factors 3 and 4 seem to be weak factors, each accounting for only 3.6 and 3.2 percent of the variance respectively. But as can be seen from Table 1, variables including inflectional morphemes, i.e., items 25, 22, 24, and 15, load on factor 4. Perhaps the reason why the picture is not so clear is that the number of items with inflectional morphemes is less than the number of items in the other two categories.

In summary, the results of the analysis of scores for the different categories reflect much the same structure as the psycholinguistic findings reveal concerning the representation of the mental lexicon; in other words, prefixes, derivational morphemes, and inflectional morphemes have separate entries. This pattern with a factor loading on words having derivational suffixes as a factor loading on prefixes continues to show itself throughout all the subsequent analyses regardless of the inclusion of all items or exclusion of some.

The results for the next type of analysis are as follows.

Table 2. Varimax Rotated Factor Matrix for Items to which more than 50% of the Students Accurately Replied

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
SDD9	.39160	-.19151	-.00891	.07512
SD2	-.11886	.00376	-.00555	.72513
SD3	.27955	.31118	-.03390	.03839
SD11	.32554	.06275	-.03941	.14173
SD26	.54497	-.00009	.17353	.01451
SI4	.30214	.13216	-.43914	.11630
SI5	.12089	-.03974	.01422	.22793
SI6	.09962	.20143	-.17018	-.16351
SI7	.05937	.29303	.12917	.25378
SI17	.15192	.13194	-.03540	.00971
SI24	.58030	.14082	.05335	-.13065
PSI1	-.01039	.43079	.07742	.01113
PSI18	.24989	.06718	.22914	.25275
PSI20	.02824	.31038	.37066	.01009
PSI21	.38292	.35080	.08805	.14684
PSD12	.22272	.10161	.50364	.03773
PSD29	.03843	-.01090	.37842	.07365
PS14	.00864	.60187	-.01592	-.13424
PSDI15	.33233	.16589	.22467	-.17668
Eigenvalue	1.91211	.90307	.85244	.75730
Pct of Var	10.1	4.8	4.5	4.0

Table 2 provides the factor loadings on items to which more than half of the subjects replied accurately. Again a 4-factor solution was found to yield the most interpretable solution.

Factor 1 has four variables, out of 5, from the category of items with derivational suffixes. This factor is labeled derivational morpheme.

Factor 2 is easily identifiable. The four variables from the category of words having prefixes load on this factor at between .31 and .60. This factor is labeled prefix.

The third and fourth factors again are weak and each accounts for only 4.5 and 4.0 percent of the variance.

Again the results of this analysis, too, like the previous one, show that prefixes and derivational suffixes are quite distinct constructs.

Another factor analysis was run on the total score for each of the categories of the scores. For each subject, his score for each of the categories of words was calculated separately. Starting with an exploratory factor analysis and finding factors having little loadings, the researcher ended up with the confirmatory factor analysis and decided on a 3-factor solution.

Table 3. Varimax Rotated Factor Matrix for Total Scores on each Category of Words (to which more than 50% of the Subjects Accurately Replied)

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
SD	.22380	.21769	.48400	.00357
SI	.53932	.00128	.29930	-.02281
PSI	.52529	.39068	.10891	.04634
PSD	.06058	.47753	.12798	.00079
Eigenvalue	1.04765	.22976	.12469	-.00126
Pct of Var	26.2	5.7	3.1	-.0

As can be seen from Table 3, each of the three variables loads on one specific factor. Factor 1 is labeled inflection morpheme because the variable **SI**, i.e., words with stems + inflectional morphemes, loads on it at .46. Factor 2 is labeled derivational morpheme because the variable **SD**, i.e., items with stem + derivational morphemes, loads on it at .46. And Factor 3 has items with prefixes + inflectional morphemes + derivational morphemes loading at .46. This factor is labeled prefix.

Conclusion

In this study, the morphology of mental lexicon was viewed both from a psycholinguistic perspective and a construct validity perspective.

Findings from both approaches confirm the suggestion that lexical entry for derivationally suffixed and prefixed words is morphologically based. Findings from construct validity studies, based on factor analytic procedures, have been found to show that prefixed words make up a construct different from the words with derivational suffixes, implying that vocabulary knowledge is not stored in one place. This finding supports the decompositional model in psycholinguistics which suggests that lexicon is represented in a morphologically decomposed form. As far as the distinction between derived and inflected words is concerned, the results of factor analysis show that inflected words do not load on the same factor as the derivational words do. Thus, it is concluded that derivational and inflectional morphemes, too, are different constructs.

The whole finding in this piece of research suggests that the construct validity studies in the field of language testing and psycholinguistic investigations regarding the role of morphological structure in the representation of lexicon have much to share and broaden the insights regarding the representation of language knowledge.

Limitations of the Study and Suggestions for Further Research

It should be mentioned here that more research is needed in this area and more factors must be taken into account. Some of the limitations of this study will be mentioned below; some suggestions will also be made:

1. The number of words belonging to each category had better match with that of other categories. In this study, certain categories of morphologically complex words included only two words while other categories had seven words.
2. The frequency of the words was not comparable. That is why certain items, difficult even for advanced students were excluded in part of the analysis in this study.
3. Since the vocabulary test used in this study was a multiple choice one, the four choices in each item also included words which had their own morphological structure. So each of these choices had their own effects that had not been taken care of. To eliminate this

effect, the researcher suggests choosing a type of test in which such amalgamation would not occur.

4. This study was conducted on subjects who knew English as their second language. The question that raises here is whether the structure of mental lexicon in L1 is totally different from this one or it has a comparable structure.

5. This study presented items in a written form. The question is whether the organization of spoken words in the mental lexicon differs from that of written words. This study can be replicated with data from an aural representation of items.

6. Mental lexicon, besides its morphological structure, has semantic, grammatical, and phonological representation as well. Attempt should be made to take care of their loadings on factors so that they can also be separated and identified.

Received 3 September 2002

Accepted 31 December 2002

References

- Aitchison, J. (1987). *Words in the mind: An introduction to the mental lexicon*. Oxford: Blackwell.
- Alderson, J. C. (1990). "Testing reading comprehension skills." *Reading in a foreign Language*, Vol. 6: 2, 425-438.
- Buck, G. (1992). "Listening comprehension: Construct validity and trait characteristics." *Language Learning*, Vol. 42: 3, 313-357.
- Davies, A. (1990). *Principles of language testing*. Cambridge: Basil Blackwell, Inc.
- Emmorey, K. D. and V. A. Fromkin. (1988). "The mental lexicon." In F. Newmeyer (Ed.) *Linguistics: The Cambridge Survey*. Volume III: *Language, psychology and biological aspects*. Cambridge: CUP.
- Foss, D. J. and D. T. Hakes. (1978). *Psycholinguistics: An introduction to the psychology of language*. N. J.: Prentice Hall.
- Garnham, A. (1985). *Psycholinguistics: Central topics*. London: Methuen.

- Green, J. (1972). *Psycholinguistics: Chomsky and psychology*. England: Penguin Education.
- Henning, D. D. (1982). *Psycholinguistics: language, mind, and world*. New York: Longman.
- Kess, J. F. (1976). *Psycholinguistics: Introductory perspectives*. New York: Academic Press, Inc.
- Kim, J. O. and C. W. Mueller. (1978). Introduction to factor analysis: What it is and how to do it. Sage. University paper series on quantitative applications in the social sciences, series no. 07-013. London: Sage publications.
- Laudanna, A., B. Cristinaa and A. Cermele. (1994). "Prefixes as processing units." *Language and Cognitive Processes*, Vol. 9, 3: 295-316.
- Murrell, G. A. and J. Morton. (1974). "Word recognition and morphemic structure." *Journal of Experimental Psychology*, Vol. 102: 963-968.
- Palmer, A. S., D. J. M. Groot and G. A. Trostler. (1981). *The construct validation of tests of communicative competence*. Washington, TESOL.
- Steinberg, D. D. (1982) *Psycholinguistics, language, mind, and world*. New York: Longman.
- Taft, M. and K. L. Forster. (1975) "Lexical storage and retrieval of prefixed words." *Journal of Verbal Learning and Verbal Behavior*, Vol. 14, 634-647.