T.C. ANADOLU ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

CORRELATION AMONG EXACT WORD, CLOZENTROPY AND PSEUDO-CLOZENTROPY AS THREE SCORING METHODS IN EFL CLOZE TESTS

(Yüksek Lisans Tezi)

Emel ARIDURU

Eskişehir, 1990

ABSTRACT

In this study, which consists of five chapters, the cloze test is used as a classroom test and scored by three objective methods, exact word, clozentropy and pseudo-clozentropy. The aim of this study is to find out if there is a significant correlation among these three scoring methods in cloze testing and which one is more reliable, thus preferable to the others.

In the first chapter the background to the problem, the problem, the research criteria, its scope and limitations and the definitions of the basic terms used in the study are introduced. The second chapter reviews the literature relevant to the study. The research method is explained in the third chapter. The fourth chapter analyzes the results. In the fifth chapter discussions and suggestions for further studies are introduced.

ÖZET

Beş bölümden oluşan bu çalışmada "cloze test" diye bilinen boşluk doldurma testi belirli bir bilgiyi ölçmek için değil sadece notlama metodları arasındaki ilişkiyi test etmek amacıyla kullanılmıştır. Çalışmanın amacı cloze testi notlamada kullanılan "exact, word, 'clozentropy" ve bu çalışmada 'clozentropy' metodundan uyarlanan 'pseudo-clozentropy' arasında anlamlı bir ilişkinin var olup olmadığını saptamak, ayrıca bu üç metodla notlanan 'cloze test'in güvenirlik katsayısını hesaplayarak, en güvenilir, dolayısıyla en tercih edilebilir olanını saptamaktır.

Birinci bölümde problemin kaynağı, problem, çalışmanın alanı, sınırlılıkları ve tanımları yer almaktadır. İkinci bölümde çalışma ile ilgili kaynak taranmıştır. Üçüncü bölümde araştırma yöntemi açıklanmıştır. Dördüncü bölüm bulgularla ilgilidir. Sınav yoluyla elde edilmiş bulgular çözümlenmiş ve yorumlanmıştır. Son bölümde bu çalışmayla ulaşılan sonuçlar ve ileriye dönük araştırmalar için öneriler yer almaktadır.

CONTENTS

CHAPTER II

CHAPTER V

TABLES

CHAPTER I

INTRODUCTION

1.1. Background to the problem

Foreign language testing is an important part of every teaching and learning experience. Teachers can use tests to evaluate their efforts as well as those of the students. As Ingram (1977: II ed. Davies) claims, "Educational assessment is as old as education itself. As long as there have been teachers they have wanted to know how much their students have learned."

Madsen (1983) points out that during the last century and the early decades of this one testing was basically subjective and dependent on the personal impressions of teachers. After the subjective stage, testing entered a scientific stage in which objective evaluation of language specialists became important. During this stage each point of language, such as grammar, vocabulary, pronunciation points, was tested separately. According to Madsen, today testing is in the communicative

stage where the evaluation of "language use" is more important than the evaluation of "language form." Today's communicative tests tend to measure more than isolated language features, which were given importance in the scientific stage. These tests aim to show how well a person can function in the foreign language.

As mentioned above, there was a movement from rather subjective to highly objective testing. Madsen (1983) states that communicative tests seem to contain the best features of these two extremes. A relatively new test type, called the "cloze" test, has been widely used in the communicative stage. A cloze test is constructed by deleting words from a passage of prose. The deleted words are replaced by a standard blank space. Generally, the test consists of 50 blanks. The students read the passage and write the appropriate words in the blank spaces.

There are two deletion methods in cloze testing. One of them is the fixed ratio deletion method. In this method, words are removed at regular intervals (eg. every seventh word) from a story or essay. The other is the variable ratio deletion method. In this method, the testers delete some pre-determined parts of the text, such as articles, auxiliaries, modals, or prepositions.

In scoring the cloze tests there are different methods. The most widely used one is the "exact word" method. In this method, the scorer of the test gives credit only for the restoration of the exact words which occur in the original text. Each correct answer is credited with +1 in a 50-blank cloze test. At the end, the student's score is doubled and the mark out of 100 is obtained.

The exact word method is regarded as objective because it has a ready answer key which consists of the words removed from

the original text. Thus, the scorer does not refer to his or any other person's knowledge in scoring the test. This answer key also offers an easy way of scoring.

The exact word method expects the examinee to fill in the blanks with the very words used by the writer of the text. Although there may be other appropriate answers to complete the blanks, the scorer gives credit only to the exact answer.

Because the exact word method seems stringent, the appropriate word scoring method is also used (cf Oller 1979: 367). Although there is some variation in this kind of scoring method, in general terms, any word that fully fits the total surrounding context is counted as correct. One of the most difficult aspects of this kind of scoring method is deciding on the degree of appropriateness of the answer. In this respect, the scorer has to have a native-like control over the language in which the text is written. There should also be a simple or complicated scale that will be used to determine the degree of appropriateness. (cf sec.2.6.4.3.)

Darnell (1968-1970) introduced a new method of weighting degrees of appropriateness. He named this method "clozentropy". In this method the cloze test is firstly given to native speakers of the language in which the text is written. In a 50-blank test their answers to each blank are credited with a score between +1 and +0.1 relative to the frequency of occurrence (cf sec.3.5). These calculations are performed by a computer. Estimating the weighted points for each answer given by the native speakers to each blank, a "weighted-degree answer key" is constructed. Then each student's answers are entered into the computer and these answers are scored according to the weighted - degree answer key.

At the end the score is doubled and a mark out of 100 is obtained.

1.2. Problem

The exact word method has a limited answer key which gives credit only to the words which correspond exactly to those of the original text. The appropriate word methods have larger answer keys but the subjectivity of the tester who has to decide the apropriateness as well as the degree of appropriateness has to be taken into consideration. In this study, the clozentropy method, which uses a mathematical way to weight the degree of appropriateness, was correlated with the exact word method. The first question this study attempts to answer is whether there is a significant correlation between the exact word and the clozentropy scoring methods in cloze testing.

The words in the exact word method are preferred by the writer of the text but another writer might have preferred other words while writing the text. The Clozentropy answer key shows the answers given by different native speakers. Thus, these answers can be credited with +1 as in the exact word method. Then the testees can be scored according to this answer key. This variation of the clozentropy can be called "pseudo-clozentropy". The second question this study attempts to answer is whether there is a significant correlation between the exact word and the pseudo-clozentropy.

The third question this study attempts to answer is whether there is a significant correlation between the clozentropy method and the pseudo-clozentropy method.

The fourth question thisstudy attempts to answer is which of these scoring methods is more reliable, thus preferable to the others.

1.3. Research Criteria

The cloze test in this study was not used in order to test a particular language ability. It was only used as a classroom test.

The cloze test consisted of a passage of 380 words taken from pages 5-6 of "The Story of the International Civil Aviation Organization". There were 50 blanks and the first 32 words of the passage were left intact to provide a lead in context.

The cloze test was administered to 45 second year students from the Ground Services, Cargo and Catering department in the Aviation School at Anadolu University.

The cloze test was scored in three different ways by exact word, clozentropy and "pseudo-clozentropy" scoring methods. In the first method, only the responses corresponding exactly to the original passage were counted as correct. In the second method, the test was first given to 25 native speakers of English. Then, their answers to each blank were credited relative to the frequency of occurrence by a computer program (cf App. C 2). An answer key was constructed and the Turkish testees were scored according to this answer key. In the third method, each answer in the clozentropy answer key was credited with +1 and the students were scored on this basis.

Incorrect spellings were not penalized in any of the scoring methods.

1.4. Scope and Limitations

This study attempts to correlate three scoring methods the exact word, the clozentropy and the pseudo-clozentropy. Since appropriate word methods require the judgement of a tester who has native-like control of that language, they can be called subjective. However, the exact word, clozentropy and pseudo-clozentropy methods can be called more objective because they don't need the subjective evaluation of a tester.

In the clozentropy scoring method the answers given by the native speakers are not checked in terms of acceptability. They are unquestionably accepted as correct and are not retested for appropriateness.

Different reading passages with different difficulty levels were not used, only one text was selected on the assumption that it could be filled in by both the second-year students who followed the one-year intensive English program in the preparatory class of the Aviation School at Anadolu University, and by native speakers of English who don't have any specific background in aviation.

Fixed-ratio deletion procedure was used and "every seventh word" was deleted in the text. Deleting every seventh word procedure was selected randomly. In this study the variable ratio method was not used because the aim was not to test the students' ability to use certain linguistic features such as conjunctions or prepositions.

1.5. Definitions

The basic terms used in this study can be defined as follows.

Cloze test: A cloze test is constructed by deleting words from a passage of prose. Testees are requested to restore the missing items by filling in the blanks. (cf Oler and Conrad, 1971: 183)

Fixed Ratio Deletion Method: In this method the words are removed from a passage at regular intervals. In other words, every nth word is deleted (cf Cohen 1980: 91)

Variable Ratio Deletion Method: In this method the words to be deleted are determined by a rational selection procedure. e.g. delete only content words or function words, etc (cf Cohen, 1980: 92)

Exact Word Scoring Method: This involves giving credit only for restoration of the exact word which has been removed from the original text (cf Oller, 1979: 367)

Appropriate Word Scoring Method: There is some variation in what this type of scoring implies. For example, in one approach called the "any contextually acceptable answer" method, the word must be grammatically correct, but incorrect spellings are not penalized (Aitken 1977). Another approach distinguishes "acceptable restorations" (grammatical or correctly spelled) from "sensible restorations" (not necessarily grammatical or spelled correctly, but making sense in the context of the passage) (Cooper and Fishman, 1977) (cf. Cohen, 1980: 99).

Clozentropy Scoring Method: In this method the test is given to native speakers of the language which the text is written in. After obtaining the answers written in the blanks, a computer program is prepared to give weighted credit to each answer given for each blank in the test. In other words, the more frequent answers given by the native speakers receive higher credit than less frequent answers. The non-native speaking testees' answers are then scored according to the answer key constructed by the answers of the native speakers (cf Oller, 1979: 372)

Pseudo-Clozentropy Scoring Method: This method was derived from the clozentropy method for the benefit of this study. The given answers in the clozentropy were not credited considering their frequency of occurrence. Rather, they were credited with +1 as in the exact word method. Thus, the answer key does not reflect any level of appropriateness but considers all the answers equally acceptable.

CHAPTER II

REVIEW OF LITERATURE

2.1. Language Testing

Language testing involves both linguistics and psychology because it is concerned with learning. It is also experimental since it sets up learning tasks to study behaviour. Finally, it is evaluative because it uses statistical procedures to study that behaviour. Therefore, there are three strands in language testing: learning, experiment and evaluation (cf Davies, 1968).

2.2. Characteristics of a good test

Heaton (1975) describes four characteristics of a good test as 'validity', 'reliability', 'discrimination' and 'washback'. Finocchiaro and Sako (1983) make another classification

of this issue, i.e. 'validity', 'reliability', 'comprehensiveness' and 'practicality'.

2.2.1. Validity

A good test should measure what it is intended to measure and nothing else. The validity of a language test is determined by comparing the results it gives with some outside or independent criterion. There are different types of validity. Content validity is how throughly a language test samples the instructional objectives. Concurrent validity is obtained by correlating the language test scores with a criterion behaviour or an established language performance. Predictive validity is how well the test scorers predict language learning. Construct validity is obtained by examining correspondance between the content of the test and various components of the skill (writing, for example) being measured.

2.2.2. Reliability

This is the accuracy of a test in measuring consistenly what it supposed to measure. It is only concerned with the dependability of the test. In order to be considered reliable, a language test must produce consistent results upon repeated administrations to the same individual, or to give consistent information about the value of learning point being measured. A language test which is higly reliable is not necessarily valid. A test may have maximum consistency, but may not be measuring what it is specifically intended to measure.

2.2.3. Discrimination

A test should show the differences in performance of individual testees.

2.2.4. Washback

A good test has a positive influence on the learning and and teaching that takes place before the test.

2.2.6. Practicality

A good test is practical in terms of economy, scorability, administrability.

2.3. Types of Tests

Finocchiaro and Sako (1983) classify the types of tests as following:

2.3.1. Achievement Test

This test measures at a specific time how much of the foreign language taught during a given period has been learned by the student.

2.3.2. Proficiency Test

This test measures a person's knowledge and ability in a foreign language without regard to formal study or text used.

It generally defines the students' achievement with relation to a particular job, or school program or the real-life use of the language.

2.3.3. Aptitude Test

This test measures students' probable performance in learning a foreign language.

2.3.4. Diagnostic Tests

These tests are used to identify students' weaknesses and difficulties related to a particular area in order to plan the most remedial program before a course of study is begun.

2.4. Recent Historical Trends in Language Testing

At the 1975 AILA (Association Internationale de Linguistique Appliquée) World Congress Bernard Spolsky, in a keynote speech named "Language Testing: Art or Science?", identified three trends in the history of modern language testing: the prescientific trend, the psychometric - structuralist trend and the psycho-linguistic-socio-linguistic trend. (cf Valette, 1977).

2.4.1. The Prescientific Trend

The prescientific approach stemmed from the old grammar translation method. Students were given a passage and were simply

required to translate it into or from the target language. In other cases easy questions were given and scored by one or two teachers purely subjectively.

This trend was popular in the USA before 1920s and is still the approach of some teachers. In terms of objectivity and test reliability there is a lack of concern of statistical analysis.

Teachers were untrained in testing and as a result students were only required to label parts of a sentence and memorize lists of language patterns (I am, we are,you are, he/she/it is etc.). The person who knew how to teach was the best person to evaluate the students. Written exercises, e.g. translation, composition or isolated sentences, were commonly used.

2.4.2. The Psychometric Structuralist Trend

Under the influence of the audio-lingual approach and structuralist linguists, the theory of language testing developed by Lado and his followers assumed that language is a system of habits. It was thought that measurement can and should be precise and scientific. Objective test techniques were developed. Multiple choice items were begun to be widely used. Statistical methods were improved to measure the reliability and validity of tests as well as their correlation with other types of measurement instruments. The language experts used contrastive analysis to identify the problems faced by the second language learner. The resulting language tests were 'discrete-point' tests. The basic principle of the discrete-point approach involved each point of language (grammar, vocabulary, pronunciation or other

linguistic properties) being tested separately. The best examples of discretepoint are the standardized language tests which were developed and administered by the Educational Testing Service, the College Entrance Examination Board Achievement Tests, the MLA Coop Tests, the Graduate Record Tests and the Test of English as a Foreign Language (TOEFL) (cf. Valette, 1977).

2.4.3. The Psycholinguistic-Sociolinguistic Trend

Farhady (1979) reports that the shortcomings of discrete-point tests have been discussed in Oller 1975, Carroll 1961 and Jakobovitz 1970. Opponents of discrete point tests claim that in spite of their practicality, since the contribution of any item to assessing total language communication may be neither significant nor identifiable, answering individual items correctly may not mean mastery in communication (cf Spolsky, 1973). In other words, knowing a language is more than just knowing several thousand separate, discrete elements. The sociolinguists place the emphasis on communicative competence and introduce the concept of a situatinoal approach to testing.

Carroll (1961) used types of tests which focus on the total communicative effect of an utterance rather than its discrete linguistic components. He referred to this type of test as integrative, (cf Farhady, 1979).

Integrative tests are intended to assess the total communicative abilities of second language learners. Oller (1973) notes that integrative tests assess the communicative skills. Some well known integrative tests are the cloze test,

dictation, listening, reading comprehension tests and oral interviews (cf Farhady, 1979).

2.5. Dicrete Point vs. Integrative Testing

Farhady (1979) states that John Carroll (1961) described the distinction between discrete point and integrative tests. Although they are not completely different for practical purposes, their theoreatical bases contrast considerably, Traditionally, a discrete point test focuses attention on one point of grammar at a time (e.g. phonology, syntax or vocabulary). It also attempts to assess only one skill at a time (e.g. listening, speaking, reading or writing) and only one aspect of a skill (e.g. productive vs. receptive, oral vs. visual). Whereas discrete items take language skills apart, integrative tests put them back together. Discrete items assess knowledge of language one bit at a time; integrative tests attempt to test a learner's capacity to use many bits all at the same time.

If a test requires the testee to use more than one of the four language skills and more than one aspect of language, it can be considered integrative. Oller (1979: 38) states that "pragmatic task is any procedure or task that causes the learner to process sequences of elements in a language that conform to the normal contextual constraints of that language."

Lots of studies made on the development of integrative or pragmatic tests have widely used statistically well-established discrete point tests as the validation criteria, (cf. Oller and Conrad, 1971; Darnell, 1970; Oller 1972, Stubbs and Tucker 1974, Swain, Lapkin and Barik 1976).

2.6. Cloze as an Integrative Task

Soudek and Soudek (1983) reports that Wilson Taylor(1953 and follow-up article 1956) was the first person who designed a procedure in which the words are systematically deleted from an ongoing text and replaced by the blanks of uniform length. He called attention to the fact that when a testee fills in the blanks in the ongoing text, he is doing something similar to what Gestalt psychologists call "closure".

The founders of Gestalt psychology are Max Wertheimer, Wolfgang Köhler and Kurt Koffka. In Gestalt theory the law of 'closure' is crucial. It reflects the natural tendency of human beings to fill in the gaps in broken patterns, or to complete, to 'cloze' incomplete figures or acts.

Similar to Gestalt psychology, Taylor's cloze technique requires the testee to fill in blanks which have been deleted intentionally from an ongoing text.

Aitken (1977: 64) states that according to Carroll (1964) redundancy is a linguistic property which allows us to predict missing symbols from the context. Where there is interference in the communication channel, redundancy is very helpful since it reduces the possibility of errors and misunderstandings. Messages can be comprehended even though some parts of them are distorted. A cloze test is such a distorted message. When a student fills in the blanks in a cloze test he formulates hypotheses by sampling the information he has available. By sampling following sequences the student accepts or rejects his hypotheses.

Aitken (1977) reports that if the processes in receptive and productive language are not different but manifestations of the same underlying competence, as Halle and Stevens (1962), Miller (1964), Goodman (1971) and Oller (1973) show it would follow that a good general comprehension test of reading would be used as an overall language proficiency test.

Cohen (1980) states that the basic cloze has been tried out in English on learners of English who are studying in countries where English is the first language (Oller 1973), and in countries where English is not the first language. (Stubbs and Tucker 1974, Moy 1975, Anderson 1976, Cooper and Fishman 1977). Cloze tests also have been used in other languages such as Amharic, Thai, Vietnamese, German, Japanese, Russian, Spanish, French and Hebrew (Bowen 1969, Oller et al 1972, Briere et al 1978, Hanzeli 1979, Nir and Cohen 1981).

Cohen (1980) reports that three types of knowledge are necessary to complete a cloze test: linguistic knowledge, textual knowledge and world knowledge.Linguistic knowledge, that is to select a semantically appropriate item and to use it in a grammatically correct form, may sometimes be sufficient to complete the cloze items correctly. Textual knowledge, that is to perceive the cohesive relationships between that item and the rest of the sentence, may sometimes be necessary. Finally, world knowledge may sometimes be required when the deleted word conveys specific information i.e. when it is a proper name, a date or a numeral. Cohen(1980: 97) concludes that "... the cloze test can be termed a measure of general reading comprehension".

Aitken (1977) reports that cloze procedure has been used for different purposes. Taylor (1953) proposed that cloze can be used as a basis for measuring the readability of prose. Anderson (1971) used cloze to determine the appropriateness of the passages for an ESL class or individual. Jongsma (1971) stated that cloze is a valid and reliable measure of both specific and general reading comprehension. Oller (1979) states that Klare, Sinaiko and Stolurow (1972) pointed out that the average score of a group of subjects on a cloze test is an actual measure of readability, of the texts.

Aitken (1977) states that; cloze has also been used as an overall language proficiency test. Darnell (1970) used a very complex scoring method - called clozentropy- and obtained a high correlation .83 with the test of English as a Foreign Language (TOEFL) total test score. Oller (1972) used a cloze test and scored it by two different methods. He obtained very high correlations of .75 for exact word method and .83 for acceptable word method with the UCLA ESA Placement Examination. Stubbs and Tucker (1974), using a cloze test reported a high correlation of .76 with the English Entrance Examination of the American University of Beirut.

Cohen (1980) reports that Carroll, Carton and Wilds (1959) found their cloze instruments unreliable and questioned the validity of such a test in measuring foreign language proficieny. Oller (1979: 354) claims that this conclusion was a result of some problems related to their study "First, the criterion measure against which the cloze test were correlated was not another language test but was in fact the Modern Language

Aptitude Test - which itself has surprisingly weak validity (Carroll, 1967). A second difficulty was that the tests developed were not of the standard every nth word deletion ratio ... A third problem was that the generalizations were based on quite small sample sizes with subject who may have been atypical in the first place".

Cohen (1980) states that Carroll (1972) suggested that the ability to supply missing elements needed a special kind of ability, which is independent of verbal ability, to use redundancy in a passage. He emphasized that cloze was working with the 'local redundancy' of a passage, i.e. that the linguistic cues in the immediate environment of a missing word tended to supply it. Oller and others (Oller 1974; Chihara, Oller, Weaver, and Chàvez-Oller 1977) attempted to show that the ability to fill in cloze involves an awareness of the flow of discourse across sentences and paragraphs. Alderson (1978), on the other hand, did not find performance on the cloze as a whole to be based on awareness of the larger context.

2.6.1. Selecting the Material for Cloze

Oller (1979: 364) states that "In spite of the natural feeling that just any old text will not do, or that a suitable text must be a carefully chosen and possibly edited text, research has shown that the cloze procedure is probably appropriate to just about any text. True some may be more difficult than others, but it has been demonstrated that for some purposes (e.g. testing ESL proficiency of university level foreign students) the level of

difficulty of the task does not greatly affect the spread of scores that will be produced. This does not mean that a teacher or other educator would ever want to use a cloze task without some sense of the level of skill of the students for whom the test was invented, but it does mean that a fairly wide latitude can be tolerated by the technique."

Related to the question of whether the cloze is valid whatever text is used, some research shows that a difficult text will result in higher correlations with proficiency and criterion measures (Carroll et al 1959, Darnell 1968, Oller 1972) (cf Alderson, 1983).

2.6.2. Deletion Procedure

There are two main types of deletion methods. The most commonly used one is deleting every nth word in an ongoing passage. This procedure is called the fixed-ratio method. Another type is called variable-ratio method. Instead of deleting every nth word, the words to be deleted are selected on some other basis, for example only the function words, such as prepositions, conjunctions, articles and the like, are left out.

According to Oller (1979) if the purpose of the testing does not involve a need to assess student performance on some other particular grammatical form, type of content or the like, every nth deletion procedure will probably work best. He reports that (Oller, 1979: 365) 'though some researches have left the first sentence intact and have begun their deletions on the second sentence of a text, Klare, Sinaiko and Stolurow (1972)

contend that this is not necessary - though there is no harm in it either".

Oller (1972) reports that in the studies of Oller and Inal (1971), Oller, et al. (1972) and Oller and Nagato (1974) every nth word deletion ratio was used and reliabilities between .80 and .98 were achieved by half-score and Kuder richardson formulas.

Cohen (1980) states that, Oller (1973) found the cloze easier when there are more words between deletions because increased context improves comprehension. Alderson (1978) showed that changes in deletion frequency sometimes resulted in significant differences between tests. Less frequent deletion sometimes actually resulted in more difficult tests. He concluded that increasing the amount of context on either side of a cloze gap beyond five words had no effect on the ease with which that gap would be clozed.

2.6.3. Administering the Cloze Test

Aitken (1977: 61) gives an example of administering the test: "In the exercise below every seventh word has been taken out and blanks put in their place. Your job will be to guess which word has been left out of each space and write that word in the blank.

Remember these things:

(1) Write only one word in each blank

(2)Try to fill in every blank.Guess if you do not know

(3) You may leave difficult blanks and come back to them when you have finished

(4) Spelling mistakes will not be marked wrong

(5) Write neatly please

(6) Take as much time as you need. It usually takes about half an hour to do the exercise."

Oller (1979) reports that a 50 minute class period usually long enough to fill in a 50-blank-cloze-test.

2.6.4. Scoring Procedures in Cloze Test

According to Oller (1979), research literature shows that all of the scoring methods that have ever been investigated produce highly correlated measures. A more lenient scoring method produces a change in the mean score. Usually there is little change in the relative rank order of scores - that high scorers will remain high scorers and low scorers will remain low relative to all of the subjects taking the test.

2.6.4.1. The Exact Word Method

Scoring of cloze test can be carried out in various ways. One of the ways is called "exact word method". This method allows credit only for the restoration of the exact, original word deleted, minor misspellings apart (cf. Alderson, 1983).

Oller (1979: 367) reports that "The rationale behind the exact word scoring method goes back to Taylor (1953). He reasored that the ability of a reader to fill in the very word

used by a writer would be a suitable index of the degree of correspondance between the language system employed by the writer and that employed by the reader. However it can be argued that requiringthe reader to replace the very word used by the writer is too stringent a requirement. It certainly places a severe limit on the range of possible responses. In fact it would keep the creative reader from supplying an ever better word in some cases."

Oller (1979) states that exact word method correlates very highly with all of the other proposed methods. He concludes that this method probably loses verylittle information if the total score is the main object of interest; however, if one is interested in specific performances of examinees on particular items, much information may be lost.

Swain, Lapkin and Barik (1976) stated that when very fine discrimination is desired, or for instructional purposes, it may make more sense to use a more lenient scoring method. (cf. Oller, 1979).

2.6.4.2.'Appropriate' or 'Acceptable' Word Methods

Another approach is referred to as 'scoring for contextual appropriateness' by Oller (1979). Briefly it is to count any word that fully fits the total surrounding context as correct. There is some variation in this type of scoring. In 'any contextually acceptable answer method' the word must be grammatically correct but incorrect spellings are not penalized (Aitken, 1977). Another

one is 'acceptable restorations', here the restorations should be grammatical and correctly spelled. In 'sensible restorations' the words do not have to be grammatical or spelled correctly but should make sense in the context of the passage (Cooper and Fishman, 1977). Antoher way is to allow credit for a synonym of the deleted word. There is also the clozentropy method, which gives weighted credit for responses which are the same as responses given by a criterion group (usually native speakers of that language)

2.6.4.3. Weighting Degrees of Appropriateness

Some responses may seem to be more appropriate than others. Different weighting scales of appropriateness of the given responses have been suggested and used. In one of the ways of accomplishing this aim, distinctions are made between types of errors in relation to some linguistic analysis of the text. Categories of errors are distinguished along the following lines. The most serious is the one that violates the most obvious and strongest contextual constraints. The least serious is the one that violates the least obvious and the weakest contextual constraints.

Cohen 1980 reports that detailed, objective categories in order to assess the acceptability of answers have been developed and used (Clarke and Burdell 1977, Nir and Blum-Kulka, 1980) There had been more subjective categories (as suggested by Bowen 1969, Oller et al 1972). In Clarke and Burdell's (1977) study the 'code sheet' included a scale for 'syntactic

acceptability', 'semantic acceptability' and 'semantic change'.

Cohen (1980: 108) reports that Nir and Blum Kulka (1980) designed another taxonomy for interpreting deviant cloze scores. Their taxonomy includes (1) the type of deviation from the word in the original (e.g. lexical/semantic, grammatical or orthographic), (2) the level at which the response is unacceptable (micro-or macro-level), and (3) the likelihood that the respondent understood this portion of the text despite the error. It should be understood that the micro-context refers to the immediate context of the blank (i.e. the word, words or phrase that the deleted word is immediately associated with) and the macro-context refers to the sentence, paragraph, section or total passage."

Considering the issue of weighting degrees of appropriateness, Oller (1972) reports that for scoring purposes differentially weighting degrees of appropriateness had no significant effect on the overall varience produced. Complex weighting systems do not afford much improvement over the simpler scoring methods.

2.6.4.4. Clozentropy as a way of weighting degrees of appropriateness

Cohen (1980: 99) states that "the acceptable word approach to scoring has both semantic and grammatical components. The semantic component is not easy to deal with. By leaving out blanks, the teacher is in essence giving the student an opportunity to produce a correct collocation (juxtaposition of words) for the given environment and sometimes it is difficult

for a native reader (scorer) to judge whether a given collocation produced by a student is acceptable or not. It is useful to give the cloze passage to natives of the target language to fill out. Usually there is general agreement among them as to acceptable answers. These answers can serve as the basis for an answer key".

Alderson (Ed. Candlin 1986: 105) states that if a record of group performance were maintained by a computer, a student could receive credit for his response, relative to the frequency (difficulty, etc) weighting of responses given by criterion groups. This would obviate the need for test constructors to specify in advance what the correct answers should be, but rather allow the computer to grade according to what proved to be popular answers for criterion group. Indeed, such a procedure has already been developed for use with cloze tests, and is known as the clozentropy scoring procedure. Cousin (1983) has developed a clozentropy program using interactive terminals and a mainframe computer (IBM 4331) for use on self-assessment programmes within self-directed learning programmes. Overseas students taking the cloze tests are scored according to the frequency of responses given by peer non-native groups, at present. The aim, however, is eventually to produce entropy scores based on appropriate native-speaker criterion groups."

2.6.4.5. Darnell's Clozentropy Method

Darnell (1968) administered a cloze test to native speakers of English (200 natives participated in the test) on the assumption that in judging the adequancy of responses given by non-native

foreign students to the cloze items, the responses given by a native-speakers-group would be a reasonable criterion.

Darnell (1970: 36) reports that 'The new test employs a variation of cloze procedure (used previously to measure readability, comprehension and language aptitude) and an entropy measure derived from information theory which indexes the compatibility of an individual's responses with those of a selected criterion group. The term CLOZENTROPY was coined to name this combination of procedures."

He thought that some objective measures (e.g. multiple choice exams) tend to be highly prescriptive and use a simple "right-wrong" scoring system, but some measures which avoid this rigid standard (e.g. essay exams or interviews) use a very subjective evaluation. He states that Clozentropy uses a very simple method of constructing test items but uses a rather complex, mathematically precise scoring method to avoid the right-wrong judgement on items.

Darnell used some complex mathematical calculations in order to find the "compatibility score" of non-native speakers of English against the criterion group-native speakers. He found to what extent the non-native speaker tended to give responses that were unusual in the context of a native speaker group responses. When a native speaker was scored against the group which he was a member, the calculated score was named the "abnormality score".

In Darnell's study both native and non-native students were composed of four sub- categories: graduate engineers, graduate non-engineers, undergraduate engineers and undergraduate

non-engineers. He used four samples of prose representing different kinds of content and different levels of difficulty. Two of the samples represented engineering content. Two represented "liberal arts" content. According to Flesch readability formula, one sample of each kind was rather easy and one was rather difficult.

The 200 native students were scored against themselves so that each student's scores reflect his standing (abnormality) in the total group. Darnell (1970: 41) states that, "This factorial design for repeated measures can be described as a 2x2x50x2x2-two levels, graduate, and undergraduate; by two majors, engineering and "other", by subjects (nested within levels and majors); by two kinds of content, engineering and "liberal arts" by two levels of difficulty, easy and difficult".

He scored non-native students against themselves and against each sub-group of the native students. Darnell (1970: 42) reports that "the seven-way analysis of variance (repeated measures, nested design) can be described as a 2x2x10x2x2x2x2x-level by major by subjects by content by difficulty by level of the criterion group by major of the criterion group."

Some of the conclusions he arrived at can be summarized as follows (cf Darnell, 1970: 42-43): There is not an apparent difference between the graduate and undergraduate foreign students in their proficiency in the English language; therefore, the graduate foreign students are probably functioning at a greater disadvantage than undergraduates in English speaking universities. There are specialized languages, because it makes

a difference which criterion group foreign students are scored against. Foreign students seem to be most similar to undergraduates and non-engineers. According to him, this finding strongly suggests that the need for specialized language programs for foreign students. The optimal language test for a foreign student would take into account the kinds of material that he actually needs to use and the kinds of people with whom he needs to communicate. He claims that demanding "general proficency" in English of a foreign student is probably unreasonable and unnecessary.

2.6.5. Correlation Studies Between Exact Word Method and Other Scoring Methods in Cloze

Oller (1979: 367) reports that "Correlations as high as .99 have been observed between exact word scores and scores that allowed synonyms for the exact words as correct answers (Miller and Coleman, 1967). Usually the correlations are in the .9 and up range (Klare, Sinaiko, and Stolurow, 1972)and this also holds for non-native speakers of the language in question (Stubbs and Tucker, 1974, Irvine, Atai, and Oller, 1974, Oller, Baca and Vigil 1977, and their references). Thus the exact word technique, which is generally the easiest to apply, is usually (but not always) to be preferred."

Wainman (1979: 128) states that Taylor (1953) and Anderson (1972) have produced data which indicates that allowing half scores for synonyms and good enough answers only gives slightly higher scores and can be discounted.

Pike (1973) reports that there was a very high correlation between Darnell's (1970) scoring technique and exact word scoring technique. Both scoring methods produced strong correlations with the TOEFL. This latter result was also obtained by Irvine et al (1974). Rand (1978) found that a close test using acceptable word scoring actually proved more reliable than a cloze scored on an exact word basis.

Alderson (1983) reports that in a study he used five different scoring procedure of the cloze test exact word, any semantically acceptable word (SEMAC), identical form class (IDFC), any word from an acceptable form class which fulfilled the same grammatical function as the deleted word (ACFC) and any grammatically correct word regardless of form, class, function or meaning (GRCO). He (1983: 208) reports that "Consistently, scoring for any semantically acceptable word (SEMAC) produces among the highest correlations with the ELBA total. In particular it almost always correlates higher than the exact word scoring procedure."

(The ELBA test is used by several English and Scottish universities to screen their foreign students). (cf. Alderson, 1983: 207).

2.6.6. Interpreting Cloze Scores

According to Aitken (1977), Harris (1969) stated that several abilities which are necessary for reading comprehension are mutually dependent. Thonis (1970) stated that reading

comprehension depends on a knowledge of meanings of words, phrases and sentences and of arrangements of these according to the conventions of written English. Spolsky (1968) claimed that language proficiency means the facility with which an individual can cope with the communication needs of a given task or in a given situation.

Aitken (1977: 62) reports that "...since overall language proficiency and reading comprehension both depend on the individual's mastery of meaning on the various levels, as well as on the individual's experience bank, a good reading comprehension test is also a good test of overall language proficiency."

Anderson (1971)suggests that to determine what a given cloze score means in respect to a non-native student's understanding of a particular passage one can use the following criteria: A close score of 53 or above (by the exact word method) corresponds to 'independent level of reading'. At this level the passage is easy for the students. They can read it without the help of a teacher. A score between 44 and 53 corresponds to the traditionally called 'instructional level.' At this level the passage could be read with the help of a teacher. A score less than 44 corresponds to the 'frustrational level'. According to Anderson this level is not appropriate for instructional use. This level means that the passage is too difficult even with the help of a teacher. Research suggests that the actual percentages vary somewhat, the instructional level may range from 38 to 50 percent to 47 to 61 percent. Bormuth (1967) reports research on young native speakers of

English suggests that a score of 43 percent on a cloze test is comparable to a score of 75 percent on a standart multiple-choice test of reading comprehension.

CHAPTER III

METHOD

3.1. The Research Method:

This study was designed in order to correlate three different scoring methods, the exact word, clozentropy and pseudo-clozentropy methods in cloze testing. Cloze test and the scoring methods have been discussed in Chapter II (cf sec. s between 2.6-2.6.6.). This study took place in the Aviation School of Anadolu University. A cloze test was administered to 45 second-year students who enrolled in Ground Services, Cargo and Catering Departments. No teaching related to the text of the cloze test took place before the cloze test.

In order to construct the clozentropy-method answer key, 25 native speakers of English were given the cloze test. This answer

key was also used for the pseudo-clozentropy scoring procedure.

3.2. The Construction of the Cloze Test

In order to prepare the cloze a passage was chosen from the Introduction of a book, "Memorandum on ICAO, "The Story of the International Civil Aviation Organization". This text was selected since it is a very general introduction to aviation, so the three different departments (Ground Services, Cargo and Catering) could cope with it equally well. Besides, since the cloze had to be completed by the native speakers of English regardless of their background, it needed to be a very general passage in the field of aviation. (cf. App. A.1)

3.3. Deletion Procedure

In the passage, taken from the introduction of the book, every seventh word was deleted. Fixed ratio method in deleting (i.e. deleting every nth word) was preferred to variable ratio method (i.e. deleting some predetermined words) because the purpose of the study was not to search for any specific knowledge of any discrete part of the language, but only to correlate three scoring methods of cloze testing. The first sentence of the passage was left intact in order to give a general idea about the text (cf. sec. 2.6.2). Including the first sentence, there were 380 words. (cf. App. A.2)

3.4. The Administration of the Cloze Test

Fourty five students took the cloze test. They were given fifty minutes to complete the test (cf.sec. 2.6.3).

The cloze test was also administered to 25 native speakers of English. Some were English Instructors at different universities (15 native speakers) and some were working in the Airbase and TUSAŞ in Eskişehir (10 native speakers). Since the native participants in the cloze test were not working at the same place, they did not take the cloze test at a particular hour, they were all given the papers and they were required to fill in the blanks with as many appropriate words as they could in order to increase the number of given possible answers. They were given 50 minutes to fill it in but there is no clear record for all of them as to how many minutes it took them to complete it.

Aitken's (1977) instructions for the administration of the cloze tests were followed (cf. sec. 2.6.3. and App. A.2).

3.5. Scoring the Cloze Test

The papers filled in by the Turkish students were scored in three ways.

First, they were scored by the exact word method. Each answer given by a student was credited with +1 when it was the exact word in the original text. The scores were doubled and the marks out of 100 were obtained (App.B and App.D).

Second, they were scored by the clozentropy method by a computer program in the manner explained below:

The cloze test given to the native speakers of English was scored on the basis of the rationale behind Darnell's (1970)

clozentropy method. (cf sec. 2.6.4.5.) He assumed that the responses given by the native speakers would serve as a basis for judging the adequancy of responses given by the non-native speakers. He reasoned that the responses given more frequently by native speakers should be assigned a heavier weighting then less frequent responses. However, this method was not completely used in this study; namely, neither abnormality scores of native speakers of English, nor compatibility scores of non-native speakers of English namely Turkish students were obtained in this study (cf.sec. 2.6.4.5.) Since the aim of this study was only to correlate three scoring systems in the cloze test, the rationale behind the clozentropy method was sufficient for the design of the research. Therefore neither the abnormality nor the compatibility scores were calculated.

The cloze test was administered to 25 native speakers of English. Each answer given to each blank was noted down and the frequency of occurrence for each answer was counted. A list containing all the answers given in the 50 blanks was constructed. The data obtained was entered into the computer (cf. App.C.1). A computer program was designed considering the rationale behind the clozentropy method (cf App. C.2.) Every different answer given by the native speakers to any blank was credited relative to its frequency of occurrence. For example, for the 5^{th} blank two different answers were given, "of" and "before". The word "of" was given 24 times and "before" was given once. The procedure of calculating the weighted score of each answer is as follows: 24 is divided by 25 (24+1, total occurrence of both responses) and the result is 0.96 for "of", then 1 is divided by

25 and the result is 0.04 for "before". The total of all the weighted scores in each blank is always equal to +1, which is the score given by the exact word method when the answer is correct.

However, as was mentioned before, in order to increase the number of possible appropriate answers given by the native speakers to each blank, the native speakers were told to fill in the blanks with as many appropriate words as they could (cf. sec. 3.4.). In this case, the total number of given appropriate answers to each blank was not always 25; therefore, the computer program was designed in such a way as first to calculate the total number of given responses then to divide the frequency number of each appropriate answer by the total number, for example, in the first blank two answers were given: "of" occurred in 21 native speakers' papers and "by"occured in 6 native speakers' papers. The total number of answers is 21+6=27. The weighted score of "of" is found by dividing its frequency (21) by the total frequency (27). The result is that "of" is credited with 0.78 and "by" is credited with 0.22. This process was performed by the computer for all the answers given to the 50 blanks in the cloze test (cf. App. C.3).

Next, each student's answers to the 50 blanks were entered and another program was used to compute the score of each student (cf. App. C.4 and C.5.) After that, all the students' scores, which were out of 50, were listed (cf. App. C.6). Finally the scores were doubled and the marks out of 100 were calculated (cf. App. D).

Finally, the students were scored according to pseudo clozentorpy answer key, in which every answer in the clozentropy answer key was credited with +1, as in the exact word method (cf.

App. C.1). The scores were then doubled and the marks out of 100 were calculated.

3.6. Data Analysis

In the evaluation of the results Kendall's tau statistic was employed in order to calculate correlation coefficients (cf. Appendices) Robson defines Kendall's tau as an index or measure of the tendency of two rank orders to be similar. He states that, "It deals, not with the scores themselves, but with the order when they have been ranked in size, and it then measures the concordance or agreement between these rank orders" (cf Robson 1973: 55). If two or more subjects get the same score from the two different tests 'ties within conditions' procedure is followed. In this study "no ties within conditions" procedure was applied. (cf. Robson 1973: 57-59)

Kuder Richardson (21) reliability formula and method was employed for the reliability calculations (cf. App. E). This method gives a measure of equivalence homogeneity. The reliability is estimated from a single administration of the test and the application of a formula that takes into account the proportion of students who passed and failed each item (cf. Lado, 1961: 333)

CHAPTER IV

ANALYSIS OF RESULTS

In this study three cloze testing scoring methods, exact word, clozentropy and pseudo clozentropy, were correlated and the reliability coefficients of the cloze test which was scored by these three methods were calculated. In the following sections the questions asked in this study and the answers are stated.

4.1. Correlation between the Exact word and the Clozentropy

The first question this study attempts to answer is, whether there is a significant agreement between the exact word and the clozentropy methods in cloze testing.

In order to find out if there is a significant agreement between these two scoring methods in cloze testing Kendall's tau statistic was used. The calculations are given in Appendix. G. The result obtained from this statistic is τ:0.57. The table value

for 45 students is τ: 0.25. Because 0.57 is more than 0.25 there is a significant agreement between the two orderings of the data at the p: 0.05 level.

4.2. Correlation between the Exact Word and the Pseudo-Clozentropy

The second question this study attempts to answer whether there is a significant agreement between the exact word and the pseudo-clozentropy methods in cloze testing.

In order to find out if there is a significant agreement between these two scoring methods in cloze testing Kendall's tau statistic was used. The calculations are given in Appendix I. The result obtained from this statistic is 0.55. The table value for 45 student is τ :0.25. Because 0.55 is more than 0.25 there is a significant agreement between the two orderings of the data at the p: 0.05 level.

4.3. Correlation between the Clozentropy and the Pseudo-Clozentropy

The third question this study attempts to answer whether there is a significant agreement between the clozentropy and pseudo-clozentropy methods in cloze testing.

In order to find out if there is a significant agreement between these two scoring methods in cloze testing Kendall's tau statistic was used. The calculations are given in Appendix K. The result obtained from this statistic is τ : 0.68. The table

value for 45 students is τ: 0.25. Because 0.68 is more than 0.25 there is a significant agreement between the two orderings of the data at the p: 0.05 level.

4.4. Reliability of the cloze test scored by the three methods

The fourth question this study attempts to answer is which of these scoring methods is more reliable, thus preferable to the others.

In order to estimate the reliability coefficients of the cloze test scored by three different scoring methods Kuder Richardson (21) reliability formula and method was applied. The calculations are given in Appendix E.. The reliability coefficient obtained by using the exact word scoring method is $r_{e.w.}$: 0.946. The reliability coefficient obtained by using the clozentropy scoring method is r_{cl}: 0.869. The reliability coefficient obtained by using the pseudo-clozentropy scoring method is $r_{ps.cl.}$: 0.464. These coefficients show that exact word is more reliable than the other scoring methods, thus, preferable to them. The results also show that the clozentropy scoring method produced a higher reliability coefficient than the pseudo clozentropy scoring method.

CHAPTER V

DISCUSSIONS AND SUGGESTIONS FOR FURTHER STUDIES

5.1. Interpretation of the Results

In this study three cloze testing scoring methods, exact word clozentropy, and pseudo-clozentropy were correlated. The general purposes of this study were to find out if there was a significant agreement among these three scoring methods in cloze testing and to find out which one was more reliable thus preferable to the others.

The exact word method offers an easy way in scoring the cloze tests. The scorer has a ready list of the correct answers that are the original words which have been removed from the original text. Therefore the scorer does not have to decide which answers are acceptable, which answers are not. If minor

misspellings are overlooked, little knowledge of the target language would be sufficient for any scorer who has got the ready answer key of the exact word method.

This study questioned if there is a significant agreement between the exact word scoring method and other scoring methods which give also credit to the acceptable answers.

Acceptability is a problem for the scorer in cloze tests since it is the scorer who has to decide on the degree of acceptability of the given answers. He also has to have a considerable control over the language in which the text is written, in other words he must be a native or native like speaker. Moreover; he has to have either a simple or a complicated scale to weight the degree of acceptability or appropriateness of the answer given to the blank. At this point the clozentropy offers an objective evaluation in cloze tests because it uses a mathematical way to weight the degree of appropriateness. This method may be considered adventageous for the scorer because he does not have to specify in advance what the correct answers should be; but rather the computer evaluates the test according to what proved to be popular answers for the native speakers.

In this study the 'pseudo-clozentropy' method was derived from the clozentropy scoring method. Each answer given by the native speakers to each blank was credited with +1. This method was also correlated with the exact word and the clozentropy scoring methods.

The results of analyses show that there is a significant agreement between the exact word and the clozentropy, between

the exact word and the pseudo-clozentropy, between the clozentropy and the pseudo-clozentropy scoring methods in cloze testing. Besides, the exact word method is more reliable thus, preferable to the other two scoring methods.

In the light of these findings, it can be concluded that if a scorer wants to use a scoring method which allows acceptable answers, both the clozentropy and the pseudo-clozentropy can be employed due to their agreement with the exact word method as well as the agreement between them. However, according to the reliability findings the exact word method is preferable. Furthermore, the exact word method also offers the teacher/scorer an ease that does not exist in the other two, a ready made answer key which enables the scorer who may be a non-native speaker to obtain reliable scores. Therefore, it seems, from the findings ofthis study that especially in EFL situations the exact word scoring method is easier and more accurate.

Not all the Turkish subjects who took the cloze test completed all the blanks. (cf. Table 1). This may be because the text may have been difficult for them or they may not have taken it seriously since it was not evaluated as part of their course work, or simply because of the fact that they became bored. Since only the given answers were taken into consideration the blanks were not evaluated at all. This could effect the result reached by this study.

5.2. Suggestions for Further Studies

The following can be taken into consideration for future studies.

1-The cloze test could be given to a much larger group of native speakers to check if an increase in number of criterion group makes a difference in results.

2-An acceptable-word-method-weighting scale could be used. Then, the four methods could be correlated with respect to producing significant agreements.

3-Different texts with different difficulty levels could be used, and the study could be redone to see if difficulty level is an important factor in producing significant agreements among the exact word and the clozentropy and the pseudo-clozentropy.

4-Different deletion types could be used in either one text or in different texts of the same difficulty level to see if deletion procedure plays a role in producing significant agreements among the three scoring methods.

5-The same cloze test can be administered to students who are at different language levels such as elemantary, intermediate and advanced etc. in order to test if differences in language level cause a significant agreement among the three scoring methods.

REFERENCES

AITKEN K.G. (1977) "Using cloze procedure as an overall language proficiency test" TESOL Quarterly, Vol. 11, No:1, 59-67.

ALDERSON C. (1983) "The cloze procedure and proficiency in English as a foreign language" in J.W. Oller (ed) Language Testing Research Newbury House Publishers, 205-217.

ALDERSON, C. (1986) "Computers in language testing" in G.Leech and C.N. Candlin (eds): Computers in English Language Teaching and Research. Selected Papers from the 1984 Lancester Symposium, 98-110.

ALDERSON C. and Urquart A.H. (1983) "Background Discipline: Effect on Comprehension" in A. Hughes (ed). Current Developments in Language Testing. Academic Press,London, 123-127.

ALLEN J.P.B. and Davies A. (1977) (Eds). Testing and Experimental Methods. Oxford University Press.

BACHMAN L.F. (1985) "Performance on cloze tests with Fixed-Ratio and Rational Deletions" TESOL Quarterly. Vol.19, No:3, 535-556.

COHEN A.D. (1980) Testing Language Ability in the Classroom. Newbury House Publishers Cambridge.

DARNELL, D.K. (1970) "Clozentropy: A procedure for testing English language proficiency of foreign students". Speech Monographs, Vol.:37, 34-46.

DAVIES, A. (1968) (Ed.). Language Testing Symposium, Oxford University Press.

FARHADY, H. (1979) "The disjunctive fallacy between Discre-Point and Integrative tests"TESOL Quarterly, Vol.: 13, No:3, 347-357.

FİNOCCHIARO, M. and Sako S. (1983). Foreign Language Testing. Regents Publishing Company Inc.

GEFEN, R. (1979) "An experiment with cloze testing", ELT Journal, Vol.: 23, No: 2, Mac Millan Press, 122-126.

HEATON J.B. (1975). Writing English Language Tests, Longman Handbooks for Language Teachers.

IRVINE, P., Atai P., Oller J.W. (1974) "Cloze, dictation and the Test of English as a Foreign Language", Language Learning, Vol.: 24, No:2, 245-252.

KLARE, G.R., Sinaiko H.W., and Stolurow L.M., 1972 "The cloze procedure: a convenient readability test for training materials and translations", International Review of Applied Psychology, 21, 77-106.

KLEIN, Braley C. (1983) "A cloze is a cloze is a question" in J.W.Oller(ed) Language Testing Research, Newbury House Publishers 218-228.

LAESCH, K.B. and Kleeck A (1987). "The cloze test as an alternative measure of language proficiency of children considered for exit from bilingual education programs" Language Learning, Vol: 37, No:2.

LADO, R. (1961). Language Testing, New York: Mc Graw. Hill Book Campany.

MADSEN, H.S. (1983). Techniques in Testing, Oxford University Press.

MILLER, G.R. and Coleman E.B. (1967). "A set of thirty six prose passage calibrated for complexity", Journal of Verbal Learning and Verbal Behavior, Vol.: 6, 851-854.

OLLER, J.W. and Conrad C., (1971). "The cloze procedure and ESL proficiency", Language Learning, Vol.: 21, 183-196.

OLLER, J.W. and Inal N. (1971). "A cloze test of English prepositions" TESOL Quarterly, Vol.: 5, 315-326.

OLLER, J.W. Baca L., and Vigil F., (1977). "Attitudes and attained proficiency in ESL: a sociolinguistic study of Mexican-Americans in the Southwest". TESOL Quarterly, Vol.: 11, 173-183.

OLLER, J.W., (1979). Language Tests at School, Longman.

OLLER, J.W., (1972). "Scoring Methods and difficulty levels for cloze tests of ESL proficiency", Modern Language Journal, Vol.: 56, 151-158.

PORTER, D. (1983). "The effect of quantity of context on the ability to make linguistic predictions: a flaw of in measure of general proficiency" in A. Hughes (ed.). Current Developments in Language Testing. Academic Press, Loncon,63-74.

ROBSON, C. (1973) Experiment, Design and Statistics in Psychology, Penguin Books.

SHOHAMY, E. (1983). "Interrater and intrarater reliability of the oral interview and concurrent validity with Cloze procedure in Hebrew" in J.W. Oller (ed). "Issues in Language Testing Research", Newbury House Publishers, 229-236.

SOUDEK, M. and Soudek L.I. (1983). "Cloze after thirty years" ELT Journal, Vol.: 37/4, Oxford University Press,335-340.

SPOLSKY, B. (1968). "What does it mean to know a language? Or how do you get someone to perform his competence?" Paper read at the Second Conference on Problems in Foreign Language Testing, University of Southern California. Reprinted in Oller and Richards (1973), 164-176.

STUBBS, J.B. and Tucker G.R., (1974). "The cloze test as a measure of English proficiency", Modern Language Journal, Vol.: 58, 239-241.

SWAIN, M., Lapkin S. and Barik H.C. (1976). "The cloze test as a measure of second language proficiency for young children", Workingpapers in Bilingualism, Vol.: 11, 32-44.

VALETTE, R.M. (1977). Modern Language Testing, Harcourt Brace. Jovanovich, Inc.

WAINMAN, H. (1979). "Cloze testing of second language learners" ELT Journal, Vol.: 23, No: 2, Mac Millan Press, 126-132.

Memorandum on ICAO "The Story of the International Civil Aviation Organization" Published in Montreal, Canada by the Public Information Office of the International Civil Aviation Organization Thirteenth Edition, 1987, 5-6.

Appendix A-1

The text used in the construction

of the cloze test

INTRODUCTION

On 17 December 1903, in North Carolina, a frail structure of metal, wood and fabric struggled into the air and carried a single passenger 260 metres. This was the first recorded flight by a heavier-than-air powered machine, but it was also the culmination of experiments made by men of many nations during the previous century. For even at the moment of its birth, the aeroplane was a creation of no one nation or of no one technology. Today, little more than 80 years later, the international character of air transport is self-evident. The scheduled airlines of the world now carry more than 900 million passengers and fly around 8 000 million kilometres. The world is enveloped by a network of air routes. The air has become a highway for world commerce.

This development of the aeroplane into a major instrument of transport has brought with it international problems — the co-ordination of techniques and laws, the dissemination of technical and economic information — far beyond the ability of individual governments to solve. The need for safety and regularity in air transport involves the necessity of building aerodromes, of setting up navigation aids and of establishing weather reporting systems. The standardization of operational practices for international services is of fundamental importance, so that there may be no error caused by misunderstanding or inexperience. The establishment of such standards, standards for rules of the air, for air traffic control, for personnel licensing, for the design of aerodromes and for so many details of prime importance to air safety, all require more than national action.

The Second World War had a major effect upon the technical development of the aeroplane, telescoping a quarter-century of normal peace-time development into six years. A vast network of passenger and freight carriage was set up but there were many problems, both political and technical, to which solutions had to be found to benefit and support a world at peace. There was the question of commercial rights — what arrangements would be made for airlines of one country to fly into and through the territories of another. There were other concerns with regard to the legal and economic conflicts that might come with peace-time flying across national borders such as how to maintain existing air navigation facilities, many of which were located in sparsely settled areas. For these reasons the Government of the United States conducted exploratory discussions with other allied nations during the early months of 1944. On the basis of the talks invitations were sent to 55 allied and neutral States to meet in Chicago in November 1944. Of these 55 States, 52 attended.

AppendipX A-2

The cloze test used in this study

In the exercise below every seventh word has been taken out and blanks put in in their place. Your job will be to guess which word has been left out of each space and write that word in the blank.

Remember these things:

(1) Write only ONE word in each blank.
(2) Try to fill in every blank. Guess if you do not know.
(3) You may leave difficult blanks and come back to them when you have finished.
(4) Spelling mistakes will not be marked wrong.
(5) Write neatly please.

Duration: 50 min.s

On 17 December 1903, in North Carolina, a frail structure of metal,wood and fabric struggled into the air and carried a single passenger 260 metres.This was the first recorded flight __1__ a heavierthan-air powered machine, but it __2__ also the culmination of experiments made __3__ men of many nations during the __4__ century. For even at the moment __5__ its birth, the aeroplane was a __6__ of no one nation or of __7__ one technology. Little more than 80 __8__ later, the international character of air __9__ is self-evident. The scheduled airlines of __10__ world carry more than 900 million __11__ now and fly around 8000 million __12__. The world is enveloped by a __13__ of air routes, The air has __14__ a highway for world commerce. This __15__ of the aeroplane into a major __16__ of transport has brought with it __17__ problems-the coordination oftechniques and __18__, the dissemination of technical and economic __19__-far beyond the ability of individual __20__ to solve.

The need for safety __21__ regularity in air transport involves the __22__ of building aerodromes, of setting up __23__ aids and of establishing weather reporting __24__. The standardization of operational practices for __25__ services is of fundamental importance, so __26__ there may be no error caused __27__ misunderstanding or inexperience. The establishment of __28__ standarts, standards for rules of the __29__ for air traffic control, for personnel __30__, for the design of aerodromes and __31__ so many details ofprime importance __32__ air safety, all require more than __33__ action.

The Second World War had __34__ major effect upon the technical development __35__ the aeroplane, telescoping a quarter-century of __36__ peace-time development into six years. A __37__ network of passenger and freight carriage __38__ set up but therewere many __39__, both political and technical, to be which __40__ had to be found to benefit __41__ support a world at peace. There __42__ the question of commercial rights-what __43__ would be made for airlines of __44__ country to fly into and through __45__ territories of another, Therewere other __46__ with regard to the legal and __47__ conflicts that might come with peace-time __48__ across national borders such as how __49__ maintain existing air navigation facilities, many __50__ whichwere located in sparsely settled areas.

Appendix-B

THE EXACT WORD METHOD CHECKLIST

1	by	26	that
2	was	27	by
3	by	28	such
4	previous	29	air
5	of	30	licensing
6	creation	31	for
7	no	32	to
8	years	33	national
9	transport	34	a
10	the	35	of
11	passengers	36	normal
12	kilometres	37	vast
13	network	38	was
14	become	39	problems
15	development	40	solutions
16	instrument	41	and
17	international	42	was
18	laws	43	arrangements
19	information	44	one
20	governments	45	the
21	and	46	concerns
22	necessity	47	economic
23	navigation	48	flying
24	systems	49	to
25	international	50	of

Appendix C-1

Frequencies of answers in the clozentropy battery

1 OF = 21
1 BY = 6
2 WAS = 25
3 BY = 25
4 19TH = 6
4 20TH = 9
4 PREVIOUS = 5
4 LAST = 2
4 18TH = 1
4 THE = 1
4 PAST = 2
5 OF = 24
5 BEFORE = 1
6 PRODUCT = 17
6 RESULT = 1
6 INVENTION = 3
6 CREATION = 2
6 CONCEPT = 1
6 THOUGHT = 1
6 WORK = 1
7 ANY = 15
7 NO = 5
7 ONLY = 3
7 EVEN = 1
8 YEARS = 25
9 POWER = 3
9 TRAVEL = 14
9 PLANES = 1
9 IT = 1
9 TECHNOLOGY= 1
9 FLIGHT = 2
9 TRANSPORTATION = 1
9 TRANSPORT = 3
10 THE = 24
10 TODAY'S = 1
11 PASSENGERS = 20
11 PEOPLE = 8
12 MILES = 23
12 KILOMETRES = 6
12 HOURS = 1
12 PER-YEAR = 1
13 MASS = 7
13 NETWORK = 9
13 NET = 1
13 SYSTEM = 2
13 WEB = 4
13 PATTERN = 1
13 MYRIAD = 2
13 LOT = 1
13 MULTITUDE = 2
13 SERIES = 1
13 NUMBER = 1
13 PLETHORA = 1
13 MAZE = 1
14 BECOME = 22
28 SAFETY = 7
28 SAFE = 1
28 CLEAR = 1
28 RIGID = 1
28 INTERNATIONAL = 7
28 COUNTLESS = 1
28 UNIVERSAL = 1
28 FLIGHTS = 1
28 HIGH = 1
28 THESE = 2
28 SET = 1
28 MANY = 1
28 AERONAUTICAL = 1
29 PROCEDURES = 2
29 SKY = 3
29 OPERATION = 1
29 NECESSITY = 1
29 OPERATIONS = 1
29 AIRPORTS = 1
29 HIGHEST = 1
29 AIR = 4
29 RUNWAY = 1
29 WORLD = 1
29 ROAD = 1
29 OFFICES = 1
29 METHODS = 1
29 PRACTICE = 2
29 REGULATION = 1
29 ADMINISTRATION = 1
30 CONDUCT = 1
30 HIRING = 1
30 MANAGEMENT = 4
30 REQUIREMENTS = 3
30 TRAINING = 5
30 SELECTION = 1
30 SERVICES = 1
30 RECRUITMENT = 2
30 ALSO = 1
30 STAFF = 2
30 SAFETY = 1
30 FLYING = 1
30 CONVENIENCE = 1
30 DEVELOPMENT = 1
31 OFTEN = 2
31 FOR = 12
31 ROUTES = 1
31 TRAINING = 1
31 FACILITIES = 1
31 RUNWAYS = 1
31 CREATE = 1
31 WITH = 1
31 ALSO = 2
31 LIKEWISE = 1
32 FOR = 5
32 TO = 18

FILE: ANSWER FREQUENC A1

14 PROVIDED = 1
14 PROVEN = 1
14 BEEN = 1
14 CREATED = 1
15 PROJECTION = 1
15 DEVELOPMENT = 13
15 TRANSFORMATION = 2
15 EMERGENCE = 2
15 LEAP = 1
15 ADVANCEMENT = 1
15 ESTABLISHMENT = 1
15 INTRODUCTION = 1
15 AGE = 1
15 ERA = 1
15 EVOLUTION = 1
15 ADDITION = 1
16 SOURCE = 3
16 MODE = 10
16 FORM = 3
16 MEANS = 7
16 TYPE = 1
16 KIND = 1
16 FORCE = 1
16 WAY = 1
16 ROLE = 1
16 AREA = 1
16 METHOD = 1
17 VARIOUS = 1
17 CERTAIN = 1
17 SOME = 4
17 MANY = 17
17 SEVERAL = 1
17 TWO = 2
17 NUMEROUS = 1
18 TECHNOLOGY = 5
18 MANUFACTURING = 1
18 OPERATIONS = 1
18 RESEARCH = 1
18 APPLICATIONS = 1
18 PLANNING = 2
18 KNOW-HOW = 1
18 COMMUNICATION = 1
18 ROUTES = 1
18 SAFETY-MEASURES = 1
18 ORGANIZATION = 1
18 SAFETY = 2
18 PROCEDURES = 2
18 METHODS = 1
18 EQIUPMENT = 1
18 DESIGNS = 1
18 ADMINISTRATION = 1
18 TRAINING = 1
18 REGULATIONS = 1
19 LOGISTICS = 1
19 ASSISTANCE = 1
32 IN = 2
32 AND = 1
32 OF = 1
33 PRIMARY = 2
33 TEMPORARY = 1
33 BRIEF = 1
33 INDIVIDUAL =4
33 IMMEDIATE = 2
33 ONE = 3
33 SINGLE =1
33 SEPERATE = 1
33 REMEDIAL = 1
33 CASUAL = 1
33 UNILATERAL = 1
33 TO =1
33 SIMPLE =1
33 NATIONAL =2
33 JUST = 2
33 PASSING =1
33 NORMAL = 1
33 MINIMAL =1
33 LIMITED =1
34 A = 25
35 OF =25
36 THEN =1
36 THE = 1
36 RELATIVE =1
36 ONGOING = 1
36 POSSIBLE =1
36 NORMAL = 5
36 PREVIOUS = 2
36 GRADUAL = 2
36 EARLIER = 1
36 ONCE = 1
36 COMMERCIAL = 1
36 HAND =1
36 CONTINUAL =1
36 FORMERLY = 1
36 STEADY = 1
36 SLOW = 1
37 VAST = 4
37 NEW = 5
37 EXTENSIVE =1
37 IMMENSE = 1
37 WORLD-WIDE =3
37 MODERN = 1
37 NECESSARY =1
37 HUGE = 1
37 SIMPLE = 1
37 BASIC = 1
37 WIDE = 1
37 IMPROVED =1
37 LARGE = 3
37 MAJOR = 2
37 SMALL = 2

FILE: ANSWER FREQUENC A1

19 INFORMATION = 10
19 AID = 1
19 ADVANCES = 1
19 RESOURCES = 1
19 KNOW-HOW = 1
19 KNOWLEDGE = 2
19 FACTORS = 1
19 TIES = 1
19 DATA = 2
19 DICIPLINES = 1
19 BARRIERS = 1
19 MEANS = 1
19 PROBLEMS = 1
19 EXPERTISE = 1
20 COMPANIES = 2
20 COUNTRIES = 2
20 NATIONS = 9
20 AIRLINES = 2
20 PEOPLE = 7
20 SCIENTISTS = 1
20 EXPERTS = 1
20 CAPABILITY = 5
20 GOVERNMENTS = 1
20 MINDS = 1
20 MEN = 1
21 AND = 19
21 OF = 1
21 HAS = 1
21 STANDARDS = 2
21 RULES = 1
21 CONTROL = 1
22 PLANNING = 1
22 PROCESS = 3
22 NECESSITY = 6
22 PROBLEMS = 1
22 TASK = 4
22 PROBLEM = 3
22 WORK = 1
22 CONSTANT = 2
22 CHALLENGE = 1
22 SCIENCE = 1
22 EXPERTISE = 1
22 ACTION = 1
22 ADMINISTRATION = 1
23 NAVIGATIONAL = 8
23 TRAINING = 3
23 MORE = 1
23 NAVIGATION = 3
23 COMMUNICATION = 2
23 MEDICAL = 1
23 RADIO = 1
23 OF = 1
23 TRAFFIC = 1
23 SUPPORT = 1
23 SAFETY = 1
37 ROUGH = 1
37 COMPLICATED = 1
37 COMPLEX = 1
38 WAS = 17
38 WERE =1
38 AIRLINES = 2
38 HAD-BEEN = 2
38 HAS-BEEN = 1
38 HAD = 1
38 TO = 1
39 OBSTACLES = 1
39 PROBLEMS = 24
39 QUESTIONS =1
39 FLAWS = 1
39 HITCHES =1
39 COMPLICATIONS = 1
40 THERE = 1
40 SOLUTIONS = 17
40 ANSWERS = 7
40 ANSWER = 1
40 WAYS = 1
41 WAY = 1
41 AND =22
41 THE =1
41 THOSE-WHO =1
42 IS = 3
42 WAS =21
42 WERE =1
43 ROUTE =2
43 IT = 1
43 PROVISIONS = 7
43 RULES = 5
43 REGULATIONS = 4
43 REQUIREMENTS =1
43 AGREEMENTS = 1
43 RIGHTS = 3
43 LAWS = 1
43 ALLOWENCES = 3
43 EXCEPTIONS = 1
43 AIRSPACE = 1
43 PERMISSION =1
44 ANOTHER = 2
44 ONE = 14
44 A = 4
44 ANY-GIVEN = 1
44 EACH =1
44 ANY = 1
44 FOREIGN = 2
44 DIFFERENT =1
45 THE =19
45 NATIONAL =1
45 DOMESTIC =1
45 SOVEREIGN =3
45 WHICH = 1
45 OTHER = 1

FILE: ANSWER FREQUENC A1

23 PASSENGER = 1
23 TRAVEL = 1
24 TECHNIQUES = 1
24 PROCEDURES = 2
24 STATIONS = 8
24 SERVICES = 4
24 FACILITIES = 1
24 SYSTEMS = 4
24 CENTRES = 2
24 TOO = 1
24 FORECASTS = 1
24 CHANNELLS = 1
25 AIRPORT = 1
25 ALL = 4
25 SUPPORT = 1
25 SCHEDULED = 1
25 TECHNICAL = 1
25 MECHANICAL = 1
25 AIR = 3
25 THESE = 5
25 COMMUNICATIVE = 1
25 INFORMATIVE = 1
25 ITS = 1
25 SUCH = 1
25 MILITARY = 1
25 AIRLINE = 4
25 BASIC = 1
25 CONTROLLED = 1
26 THAT = 20
26 IMPORTANT = 1
26 AS = 1
26 THERE = 1
27 NOR = 1
27 BY = 22
27 THROUGH = 2
28 OPERATIONAL = 3
45 AIR = 1
46 PROBLEMS =18
46 QUESTIONS =2
46 CONSTITUENTS =1
46 FACTORS = 1
46 CONSIDERATIONS = 4
46 CONCERNS = 1
47 TERRITORIAL =2
47 NATIONAL = 2
47 POLITICAL =10
47 ECONOMIC = 2
47 TRADE = 1
47 DOMESTIC = 1
47 OTHER = 1
47 BUSINESS =1
47 REGIONAL =1
47 LOCAL = 1
47 MORAL = 3
47 FINANCIAL =1
47 ILLEGAL = 1
47 PERSONAL =2
47 TECHNICAL =1
48 FLIGHT = 11
48 TRAVEL = 9
48 SECURITY =1
48 WEALTH = 1
48 OPTIMISM =1
48 OPERATIONS =1
48 PASSAGE = 1
48 USE = 1
48 NAVIGATION = 1
48 MOVEMENTS = 1
49 TO = 24
49 GOVERMENTS =1
50 OF = 24
50 RESIDENTS = 1

Appendix C-2

The computer program which computes point for each answer in the clozentropy battery

FILE: POINT EXEC A

```
/* THIS PROGRAM GETS THE CORRECT ANSWER COUNTS FOR A QUESTION AND */
/* COMPUTES POINT FOR EACH WORD IN THE QUESTION                   */
   SAY 'ENTER QUESTION COUNT '
   PULL QUESTION_COUNT
   TEMPORARY = 0
   'ERASE ANSWER POINTS A'
   DO QUESTION_INDEX = 1 TO QUESTION_COUNT
     VMFCLEAR
     SAY 'ENTER THE ANSWERS FOR THE 'QUESTION_INDEX'. QUESTION'
     SAY '  ENTER BLANK TO EXIT'
     TOTAL_FREQUENCY = 0
     DO WORD_INDEX = 1 TO 100
       SAY  '  ENTER 'WORD_INDEX'. WORD AND ITS FREQUENCY'
       PULL WORD FREQUENCY
       IF WORD = ' ' THEN LEAVE
       WORD_ARRAY.WORD_INDEX = WORD
       WORD_FREQUENCY.WORD_INDEX = FREQUENCY
       TOTAL_FREQUENCY = TOTAL_FREQUENCY + FREQUENCY
     END
     WORD_COUNT = WORD_INDEX - 1
     DO WORD_INDEX = 1 TO WORD_COUNT
       WORD_POINT = WORD_FREQUENCY.WORD_INDEX / TOTAL_FREQUENCY
       SAY WORD_ARRAY.WORD_INDEX ' = ' WORD_POINT
       RESULT_LINE = QUESTION_INDEX WORD_ARRAY.WORD_INDEX '=' WORD_POINT
       EXECIO 1 DISKW ANSWER   POINTS  A ' (VAR RESULT_LINE'
    END
  END
```

Appendix C-3

The points of answers in the clozentropy battery

1 OF = 0.777777778
1 BY = 0.222222222
2 WAS = 1
3 BY = 1
4 19TH = 0.230769231
4 20TH = 0.346153846
4 PREVIOUS = 0.192307692
4 LAST = 0.0769230769
4 18TH = 0.0384615385
4 THE = 0.0384615385
4 PAST = 0.0769230769
5 OF = 0.96
5 BEFORE = 0.04
6 PRODUCT = 0.653846154
6 RESULT = 0.0384615385
6 INVENTION = 0.115384615
6 CREATION = 0.0769230769
6 CONCEPT = 0.0384615385
6 THOUGHT = 0.0384615385
6 WORK = 0.0384615385
7 ANY = 0.625
7 NO = 0.208333333
7 ONLY = 0.125
7 EVEN = 0.0416666667
8 YEARS = 1
9 POWER = 0.115384615
9 TRAVEL = 0.538461538
9 PLANES = 0.0384615385
9 IT = 0.0384615385
9 TECHNOLOGY = 0.0384615385
9 FLIGHT = 0.0769230769
9 TRANSPORTATION = 0.0384615385
9 TRANSPORT = 0.115384615
10 THE = 0.96
10 TODAY'S = 0.04
11 PASSENGERS = 0.714285714
11 PEOPLE = 0.285714286
12 MILES = 0.741935484
12 KILOMETRES = 0.193548387
12 HOURS = 0.0322580645
12 PER-YEAR = 0.0322580645
13 MASS = 0.212121212
13 NETWORK = 0.272727273
13 NET = 0.0303030303
13 SYSTEM = 0.0606060606
13 WEB = 0.121212121
13 PATTERN = 0.0303030303
13 MYRIAD = 0.0606060606
13 LOT = 0.0303030303
13 MULTITUDE = 0.0606060606
13 SERIES = 0.0303030303
13 NUMBER = 0.0303030303
13 PLETHORA = 0.0303030303
13 MAZE = 0.0303030303
14 BECOME = 0.846153846
28 SAFETY = 0.24137931
28 SAFE = 0.0344827586
28 CLEAR = 0.0344827586
28 RIGID = 0.0344827586
28 INTERNATIONAL = 0.24137931
28 COUNTLESS = 0.0344827586
28 UNIVERSAL = 0.0344827586
28 FLIGHTS = 0.0344827586
28 HIGH = 0.0344827586
28 THESE = 0.0689655172
28 SET = 0.0344827586
28 MANY = 0.0344827586
28 AERONAUTICAL = 0.0344827586
29 PROCEDURES = 0.0869565217
29 SKY = 0.130434783
29 OPERATION = 0.0434782609
29 NECESSITY = 0.0434782609
29 OPERATIONS = 0.0434782609
29 AIRPORTS = 0.0434782609
29 HIGHEST = 0.0434782609
29 AIR = 0.173913043
29 RUNWAY = 0.0434782609
29 WORLD = 0.0434782609
29 ROAD = 0.0434782609
29 OFFICES = 0.0434782609
29 METHODS = 0.0434782609
29 PRACTICE = 0.0869565217
29 REGULATION = 0.0434782609
29 ADMINISTRATION = 0.0434782609
30 CONDUCT = 0.04
30 HIRING = 0.04
30 MANAGEMENT = 0.16
30 REQUIREMENTS = 0.12
30 TRAINING = 0.2
30 SELECTION = 0.04
30 SERVICES = 0.04
30 RECRUITMENT = 0.08
30 ALSO = 0.04
30 STAFF = 0.08
30 SAFETY = 0.04
30 FLYING = 0.04
30 CONVENIENCE = 0.04
30 DEVELOPMENT = 0.04
31 OFTEN = 0.0952380952
31 FOR = 0.523809524
31 ROUTES = 0.0476190476
31 TRAINING = 0.0476190476
31 FACILITIES = 0.0476190476
31 RUNWAYS = 0.0476190476
31 CREATE = 0.0476190476
31 WITH = 0.0476190476
31 ALSO = 0.0476190476
31 LIKEWISE = 0.0476190476
32 FOR = 0.185185185
32 TO = 0.666666667

FILE: ANSWER POINTS A

14 PROVIDED = 0.0384615385
14 PROVEN = 0.0384615385
14 BEEN = 0.0384615385
14 CREATED = 0.0384615385
15 PROJECTION = 0.0384615385
15 DEVELOPMENT = 0.5
15 TRANSFORMATION = 0.0769230769
15 EMERGENCE = 0.0769230769
15 LEAP = 0.0384615385
15 ADVANCEMENT = 0.0384615385
15 ESTABLISHMENT = 0.0384615385
15 INTRODUCTION = 0.0384615385
15 AGE = 0.0384615385
15 ERA = 0.0384615385
15 EVOLUTION = 0.0384615385
15 ADDITION = 0.0384615385
16 SOURCE = 0.1
16 MODE = 0.333333333
16 FORM = 0.1
16 MEANS = 0.233333333
16 TYPE = 0.0333333333
16 KIND = 0.0333333333
16 FORCE = 0.0333333333
16 WAY = 0.0333333333
16 ROLE = 0.0333333333
16 AREA = 0.0333333333
16 METHOD = 0.0333333333
17 VARIOUS = 0.037037037
17 CERTAIN = 0.037037037
17 SOME = 0.148148148
17 MANY = 0.62962963
17 SEVERAL = 0.037037037
17 TWO = 0.0740740741
17 NUMEROUS = 0.037037037
18 TECHNOLOGY = 0.192307692
18 MANUFACTURING = 0.0384615385
18 OPERATIONS = 0.0384615385
18 RESEARCH = 0.0384615385
18 APPLICATIONS = 0.0384615385
18 PLANNING = 0.0769230769
18 KNOW-HOW = 0.0384615385
18 COMMUNICATION = 0.0384615385
18 ROUTES = 0.0384615385
18 SAFETY-MEASURES = 0.0384615385
18 ORGANIZATION = 0.0384615385
18 SAFETY = 0.0769230769
18 PROCEDURES = 0.0769230769
18 METHODS = 0.0384615385
18 EQIUPMENT = 0.0384615385
18 DESIGNS = 0.0384615385
18 ADMINISTRATION = 0.0384615385
18 TRAINING = 0.0384615385
18 REGULATIONS = 0.0384615385
19 LOGISTICS = 0.037037037
19 ASSISTANCE = 0.037037037
32 IN = 0.0740740741
32 AND = 0.037037037
32 OF = 0.037037037
33 PRIMARY = 0.0714285714
33 TEMPORARY = 0.0357142857
33 BRIEF = 0.0357142857
33 INDIVIDUAL = 0.142857143
33 IMMEDIATE = 0.0714285714
33 ONE = 0.107142857
33 SINGLE = 0.0357142857
33 SEPERATE = 0.0357142857
33 REMEDIAL = 0.0357142857
33 CASUAL = 0.0357142857
33 UNILATERAL = 0.0357142857
33 TO = 0.0357142857
33 SIMPLE = 0.0357142857
33 NATIONAL = 0.0714285714
33 JUST = 0.0714285714
33 PASSING = 0.0357142857
33 NORMAL = 0.0357142857
33 MINIMAL = 0.0357142857
33 LIMITED = 0.0357142857
34 A = 1
35 OF = 1
36 THEN = 0.0454545455
36 THE = 0.0454545455
36 RELATIVE = 0.0454545455
36 ONGOING = 0.0454545455
36 POSSIBLE = 0.0454545455
36 NORMAL = 0.227272727
36 PREVIOUS = 0.0909090909
36 GRADUAL = 0.0909090909
36 EARLIER = 0.0454545455
36 ONCE = 0.0454545455
36 COMMERCIAL = 0.0454545455
36 HAND = 0.0454545455
36 CONTINUAL = 0.0454545455
36 FORMERLY = 0.0454545455
36 STEADY = 0.0454545455
36 SLOW = 0.0454545455
37 VAST = 0.129032258
37 NEW = 0.161290323
37 EXTENSIVE = 0.0322580645
37 IMMENSE = 0.0322580645
37 WORLD-WIDE = 0.0967741935
37 MODERN = 0.0322580645
37 NECESSARY = 0.0322580645
37 HUGE = 0.0322580645
37 SIMPLE = 0.0322580645
37 BASIC = 0.0322580645
37 WIDE = 0.0322580645
37 IMPROVED = 0.0322580645
37 LARGE = 0.0967741935
37 MAJOR = 0.064516129
37 SMALL = 0.064516129

FILE: ANSWER POINTS A

19 INFORMATION = 0.37037037
19 AID = 0.037037037
19 ADVANCES = 0.037037037
19 RESOURCES = 0.037037037
19 KNOW-HOW = 0.037037037
19 KNOWLEDGE = 0.0740740741
19 FACTORS = 0.037037037
19 TIES = 0.037037037
19 DATA = 0.0740740741
19 DICIPLINES = 0.037037037
19 BARRIERS = 0.037037037
19 MEANS = 0.037037037
19 PROBLEMS = 0.037037037
19 EXPERTISE = 0.037037037
20 COMPANIES = 0.0625
20 COUNTRIES = 0.0625
20 NATIONS = 0.28125
20 AIRLINES = 0.0625
20 PEOPLE = 0.21875
20 SCIENTISTS = 0.03125
20 EXPERTS = 0.03125
20 CAPABILITY = 0.15625
20 GOVERNMENTS = 0.03125
20 MINDS = 0.03125
20 MEN = 0.03125
21 AND = 0.76
21 OF = 0.04
21 HAS = 0.04
21 STANDARDS = 0.08
21 RULES = 0.04
21 CONTROL = 0.04
22 PLANNING = 0.0384615385
22 PROCESS = 0.115384615
22 NECESSITY = 0.230769231
22 PROBLEMS = 0.0384615385
22 TASK = 0.153846154
22 PROBLEM = 0.115384615
22 WORK = 0.0384615385
22 CONSTANT = 0.0769230769
22 CHALLENGE = 0.0384615385
22 SCIENCE = 0.0384615385
22 EXPERTISE = 0.0384615385
22 ACTION = 0.0384615385
22 ADMINISTRATION = 0.0384615385
23 NAVIGATIONAL = 0.32
23 TRAINING = 0.12
23 MORE = 0.04
23 NAVIGATION = 0.12
23 COMMUNICATION = 0.08
23 MEDICAL = 0.04
23 RADIO = 0.04
23 OF = 0.04
23 TRAFFIC = 0.04
23 SUPPORT = 0.04
23 SAFETY = 0.04
37 ROUGH = 0.0322580645
37 COMPLICATED = 0.0322580645
37 COMPLEX = 0.0322580645
38 WAS = 0.68
38 WERE = 0.04
38 AIRLINES = 0.08
38 HAD-BEEN = 0.08
38 HAS-BEEN = 0.04
38 HAD = 0.04
38 TO = 0.04
39 OBSTACLES = 0.0344827586
39 PROBLEMS = 0.827586207
39 QUESTIONS = 0.0344827586
39 FLAWS = 0.0344827586
39 HITCHES = 0.0344827586
39 COMPLICATIONS = 0.0344827586
40 THERE = 0.037037037
40 SOLUTIONS = 0.62962963
40 ANSWERS = 0.259259259
40 ANSWER = 0.037037037
40 WAYS = 0.037037037
41 WAY = 0.04
41 AND = 0.88
41 THE = 0.04
41 THOSE-WHO = 0.04
42 IS = 0.12
42 WAS = 0.84
42 WERE = 0.04
43 ROUTE = 0.064516129
43 IT = 0.0322580645
43 PROVISIONS = 0.225806452
43 RULES = 0.161290323
43 REGULATIONS = 0.129032258
43 REQUIREMENTS = 0.0322580645
43 AGREEMENTS = 0.0322580645
43 RIGHTS = 0.0967741935
43 LAWS = 0.0322580645
43 ALLOWENCES = 0.0967741935
43 EXCEPTIONS = 0.0322580645
43 AIRSPACE = 0.0322580645
43 PERMISSION = 0.0322580645
44 ANOTHER = 0.0769230769
44 ONE = 0.538461538
44 A = 0.153846154
44 ANY-GIVEN = 0.0384615385
44 EACH = 0.0384615385
44 ANY = 0.0384615385
44 FOREIGN = 0.0769230769
44 DIFFERENT = 0.0384615385
45 THE = 0.703703704
45 NATIONAL = 0.037037037
45 DOMESTIC = 0.037037037
45 SOVEREIGN = 0.111111111
45 WHICH = 0.037037037
45 OTHER = 0.037037037

FILE: ANSWER POINTS A

23 PASSENGER = 0.04
23 TRAVEL = 0.04
24 TECHNIQUES = 0.04
24 PROCEDURES = 0.08
24 STATIONS = 0.32
24 SERVICES = 0.16
24 FACILITIES = 0.04
24 SYSTEMS = 0.16
24 CENTRES = 0.08
24 TOC = 0.04
24 FORECASTS = 0.04
24 CHANNELLS = 0.04
25 AIRPORT = 0.0357142857
25 ALL = 0.142857143
25 SUPPORT = 0.0357142857
25 SCHEDULED = 0.0357142857
25 TECHNICAL = 0.0357142857
25 MECHANICAL = 0.0357142857
25 AIR = 0.107142857
25 THESE = 0.178571429
25 COMMUNICATIVE = 0.0357142857
25 INFORMATIVE = 0.0357142857
25 ITS = 0.0357142857
25 SUCH = 0.0357142857
25 MILITARY = 0.0357142857
25 AIRLINE = 0.142857143
25 BASIC = 0.0357142857
25 CONTROLLED = 0.0357142857
26 THAT = 0.869565217
26 IMPORTANT = 0.0434782609
26 AS = 0.0434782609
26 THERE = 0.0434782609
27 NOR = 0.04
27 BY = 0.88
27 THROUGH = 0.08
28 OPERATIONAL = 0.103448276
45 AIR = 0.037037037
46 PROBLEMS = 0.666666667
46 QUESTIONS = 0.0740740741
46 CONSTITUENTS = 0.037037037
46 FACTORS = 0.037037037
46 CONSIDERATIONS = 0.148148148
46 CONCERNS = 0.037037037
47 TERRITORIAL = 0.0666666667
47 NATIONAL = 0.0666666667
47 POLITICAL = 0.333333333
47 ECONOMIC = 0.0666666667
47 TRADE = 0.0333333333
47 DOMESTIC = 0.0333333333
47 OTHER = 0.0333333333
47 BUSINESS = 0.0333333333
47 REGIONAL = 0.0333333333
47 LOCAL = 0.0333333333
47 MORAL = 0.1
47 FINANCIAL = 0.0333333333
47 ILLEGAL = 0.0333333333
47 PERSONAL = 0.0666666667
47 TECHNICAL = 0.0333333333
48 FLIGHT = 0.392857143
48 TRAVEL = 0.321428571
48 SECURITY = 0.0357142857
48 WEALTH = 0.0357142857
48 OPTIMISM = 0.0357142857
48 OPERATIONS = 0.0357142857
48 PASSAGE = 0.0357142857
48 USE = 0.0357142857
48 NAVIGATION = 0.0357142857
48 MOVEMENTS = 0.0357142857
49 TO = 0.96
49 GOVERNMENTS = 0.04
50 OF = 0.96
50 RESIDENTS = 0.04

Appendix C-4

The computer program which computes the total score of a student whose answers are entered from keyboard

```
FILE: COMPUTE  EXEC     A                    VM/SP CONVERSATIONAL MONITOR SYSTEM

/*
   THIS PROGRAM COMPUTES SCORE FOR A STUDENT WHOSE ANSWERS
    ARE ENTERED FROM KEYBOARD, AND WRITES THE RESULT ONTO
    FILE '<STUDENT NUMBER> RESULT A'
*/
/* GET POINTS FOR EACH QUESTION FROM FILE ANSWER POINTS A */
     EXECIO '*' DISKR ANSWER POINTS A
     POINT. = 0
     DO QUEUED()
       PULL QUESTION1 WORD1 . POINT.QUESTION1.WORD1 ,
            QUESTION2 WORD2 . POINT.QUESTION2.WORD2
     END
     SAY 'TYPE QUESTION COUNT AND PRESS ENTER KEY '
     PULL QUESTION_COUNT
     DO FOREVER
       SAY 'TYPE STUDENT NUMBER AND PRESS ENTER KEY '
       PULL STUDENT_NUMBER
       IF STUDENT_NUMBER = '' THEN LEAVE
       STUDENT_SCORE = 0
       VMFCLEAR
       SAY 'PROCESSING FOR 'STUDENT_NUMBER'. STUDENT '
       SAY ' '
       DO QUESTION_INDEX = 1 TO QUESTION_COUNT
         SAY 'TYPE ANSWER FOR 'QUESTION_INDEX'. QUESTION CR'
         SAY '  TYPE BLANK IF THERE IS NO ANSWER,AND PRESS ENTER KEY'
         PULL ANSWER
         ANSWER=STRIP(ANSWER,'BOTH',' ')
         RESULT_LINE = QUESTION_INDEX ANSWER POINT.QUESTION_INDEX.ANSWER
         EXECIO 1 DISKW STUDENT_NUMBER RESULT A ' (VAR RESULT_LINE'
         IF POINT.QUESTION_INDEX.ANSWER ¬= 0
           THEN STUDENT_SCORE = STUDENT_SCORE + ,
                    POINT.QUESTION_INDEX.ANSWER
           ELSE SAY '*****  ANSWER IS INCORRECT *****'
      END
      SAY 'TOTAL SCORE = 'STUDENT_SCORE
      RESULT_LINE = 'TOTAL SCORE FOR 'STUDENT_NUMBER'. STUDENT = ' ,
                              STUDENT_SCORE
      EXECIO 1 DISKW STUDENT SCORES A ' (VAR RESULT_LINE'
    END
```

Appendix C-5

A sample of a student's output calculated by the clozentropy

FILE: 17 RESULT A

```
1 WAS 0
2 WAS 1
3 BY 1
4 20TH 0.346153846
5  0
6 PRODUCT 0.653846154
7 NO 0.208333333
8 YEARS 1
9  0
10 THE 0.96
11 PASSENGERS 0.714285714
12 M LES 0
13  0
14  0
15  0
16  0
17 MANY 0.62962963
18  0
19  0
20 HX 0
21  0
22  0
23  0
24  0
25  0
26  0
27  0
28  0
29  0
30  0
31  0
32  0
33  0
34  0
35  0
36  0
37  0
38  0
39  0
40  0
41  0
42  0
43  0
44  0
45  0
46  0
47  0
48  0
49 TO 0.96
50  0
```

Appendix C-6

The raw scores of the students in the cloze test scored by the clozentropy method. The scores are out of 50.

FILE: OGRENCI CIKIS A1

OGRENCININLERIN PUANLARI:

1 .OGRENCININ PUANI = 6.41828791
2 .OGRENCININ PUANI = 7.69509322
3 .OGRENCININ PUANI = 8.42223486
4 .OGRENCININ PUANI = 3.72015603
5 .OGRENCININ PUANI = 6.76365262
6 .OGRENCININ PUANI = 10.7155869
7 .OGRENCININ PUANI = 7.08225044
8 .OGRENCININ PUANI = 6.06364604
9 .OGRENCININ PUANI = 9.28754447
10.OGRENCININ PUANI = 6.43917314
11.OGRENCININ PUANI = 10.4925405
12.OGRENCININ PUANI = 10.5374221
13.OGRENCININ PUANI = 22.9698074
14.OGRENCININ PUANI = 9.24024801
15.OGRENCININ PUANI = 10.62511148
16.OGRENCININ PUANI = 10.9809334
17.OGRENCININ PUANI = 15.5315472
18.OGRENCININ PUANI = 8.53636300
19.OGRENCININ PUANI = 4.80808933
20.OGRENCININ PUANI = 7.24094080
21.OGRENCININ PUANI = 4.94545878
22.OGRENCININ PUANI = 4.06043956
23.OGRENCININ PUANI = 5.68712076
24.OGRENCININ PUANI = 7.61155111
25.OGRENCININ PUANI = 4.78541653

Appendix C-6

Continued

26.CGRENCININ PUANI = 7.37471163

27.CGRENCININ PUANI = 5.70751415

28.CGRENCININ PUANI = 5.02946415

29.CGRENCININ PUANI = 6.37631755

30.CGRENCININ PUANI = 4.87349685

31.CGRENCININ PUANI = 5.84982925

32.CGRENCININ PUANI = 3.16833333

33.CGRENCININ PUANI = 6.80494886

34.CGRENCININ PUANI = 3.68454350

35.CGRENCININ PUANI = 4.95166007

36.CGRENCININ PUANI = 5.70232128

37.CGRENCININ PUANI = 8.68022069

38.CGRENCININ PUANI = 6.34221617

39.CGRENCININ PUANI = 2.80237504

40.CGRENCININ PUANI = 6.36367531

41.CGRENCININ PUANI = 5.81937866

42.CGRENCININ PUANI = 7.17354922

43.CGRENCININ PUANI = 6.22996763

44.CGRENCININ PUANI = 10.6790875

45.CGRENCININ PUANI = 2.82245357

Appendix D

The scores of the students in the cloze test scored by the pseudo clozentropy, exact word and clozentropy methods. All scores are out of 100.

Subjects	Pseudo clozentropy	Exact word	Clozentropy
1	26	12	12.8
2	34	18	15.3
3	36	16	16.8
4	18	12	7.4
5	26	16	13.5
6	34	20	21.4
7	32	16	14.1
8	22	14	12.1
9	26	18	18.5
10	22	14	12.8
11	36	22	20.9
12	38	22	21
13	70	50	45.9
14	34	26	18.4
15	42	22	21.2
16	32	22	21.9
17	46	30	31
18	36	18	17
19	14	10	9.6
20	24	14	14.4
21	16	10	9.8
22	10	8	8.1
23	2o	14	11.3
24	30	12	15.2
25	12	10	9.56
26	20	16	14.7
27	24	8	11.4
28	16	12	10
29	18	14	12.7
30	10	8	9.7
31	16	10	11.6
32	8	8	6.3
33	22	12	13.6
34	16	12	7.3
35	18	14	9.9
36	20	14	11.4
37	24	12	17.3
38	18	16	12.6
39	8	4	5.6
40	22	14	15.7
41	22	18	11.6
42	24	16	14.3
43	20	14	12.4
44	38	22	21.3
45	12	6	5.6

Appendix E

Formula and method for estimating equivalance reliability Kuder-Richardson (21)

Kuder-Richardson (21),

$$r = \frac{m}{m-1} \times \frac{S^2t - (\Sigma p - \Sigma p^2)}{S^2t}$$

m= number of items

S^2t= the variance of the test

p= the proportion of passes for each item

$$p = \frac{\text{number of passes}}{\text{number of people}}$$

Exact word

$$r_{ex} = \frac{50}{49} \times \frac{54.29-(7.59-3.69)}{54.29}$$

$$r_{ex} = 1.02 \times \frac{50.39}{54.29}$$

$$r_{ex} = 0.946$$

Clozentropy

$$r_{cl} = \frac{50}{49} \times \frac{48.06-(12.42-6.15)}{48.06}$$

$$r_{cl} = 1.02 \times \frac{41.79}{48.06}$$

$$r_{cl} = 0.869$$

Pseudo clozentropy

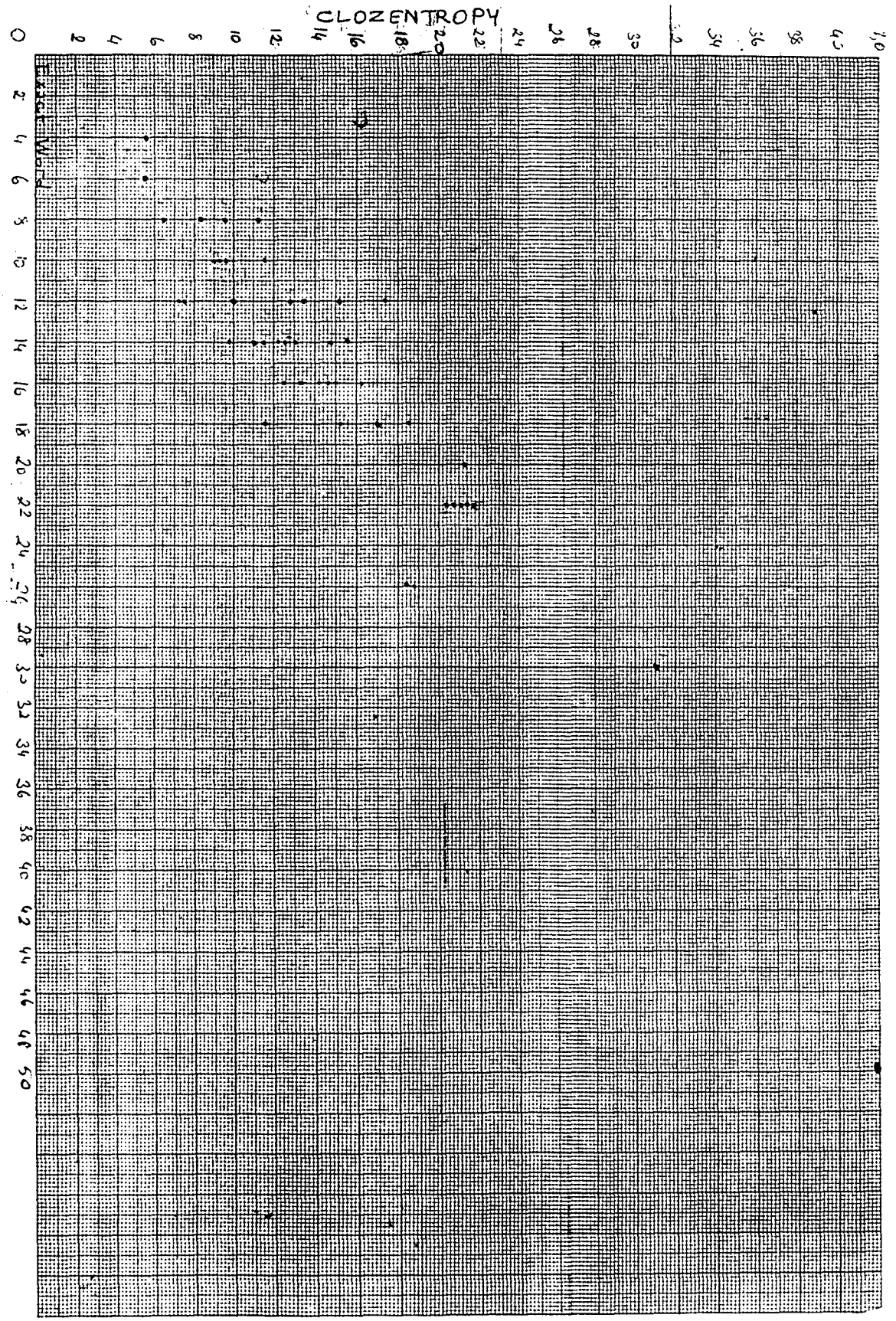
$$r_{pc} = \frac{50}{49} \times \frac{11.51-(12.42-6.15)}{11.51}$$

$$r_{pc} = 1.02 \times \frac{5.24}{11.51}$$

$$r_{pc} = 1.02 \times 0.455 = 0.464$$

Appendix F

The scattergram related to the correlation between exact word and clozentropy



Appendix G

Correlation between the Exact word and the clozentropy

Kendall's Tau- no ties within conditions

Subject

1 - 19
2 - 8
3 - 11
4 - 28
5 - 12
6 - 3
7 - 12
8 - 18
9 - 8
10 - 16
11 - 2
12 - 2
13 - 0
14 - 2
15 - 3
16 - 2
17 - 1
18 - 8
19 - 35
20 - 15
21 - 35
22 - 37
23 - 19
24 - 14
25 - 35
26 - 12
27 - 29
28 - 27
29 - 16
30 - 33
31 - 35
32 - 39
33 - 18
34 - 28
35 - 19
36 - 19
37 - 10
38 - 12
39 - 43
40 - 12
41 - 9
42 - 12
43 - 16
44 - 2
45 - 43

$S^+ = 779$

N (number of pairs) N=45

$$\frac{N \times (N-1)}{2} = \frac{45 \times 44}{2} = 990$$

$$S = 2\,S^+ - \frac{N \times (N-1)}{2}$$

$S = (2 \times 779) - 990$

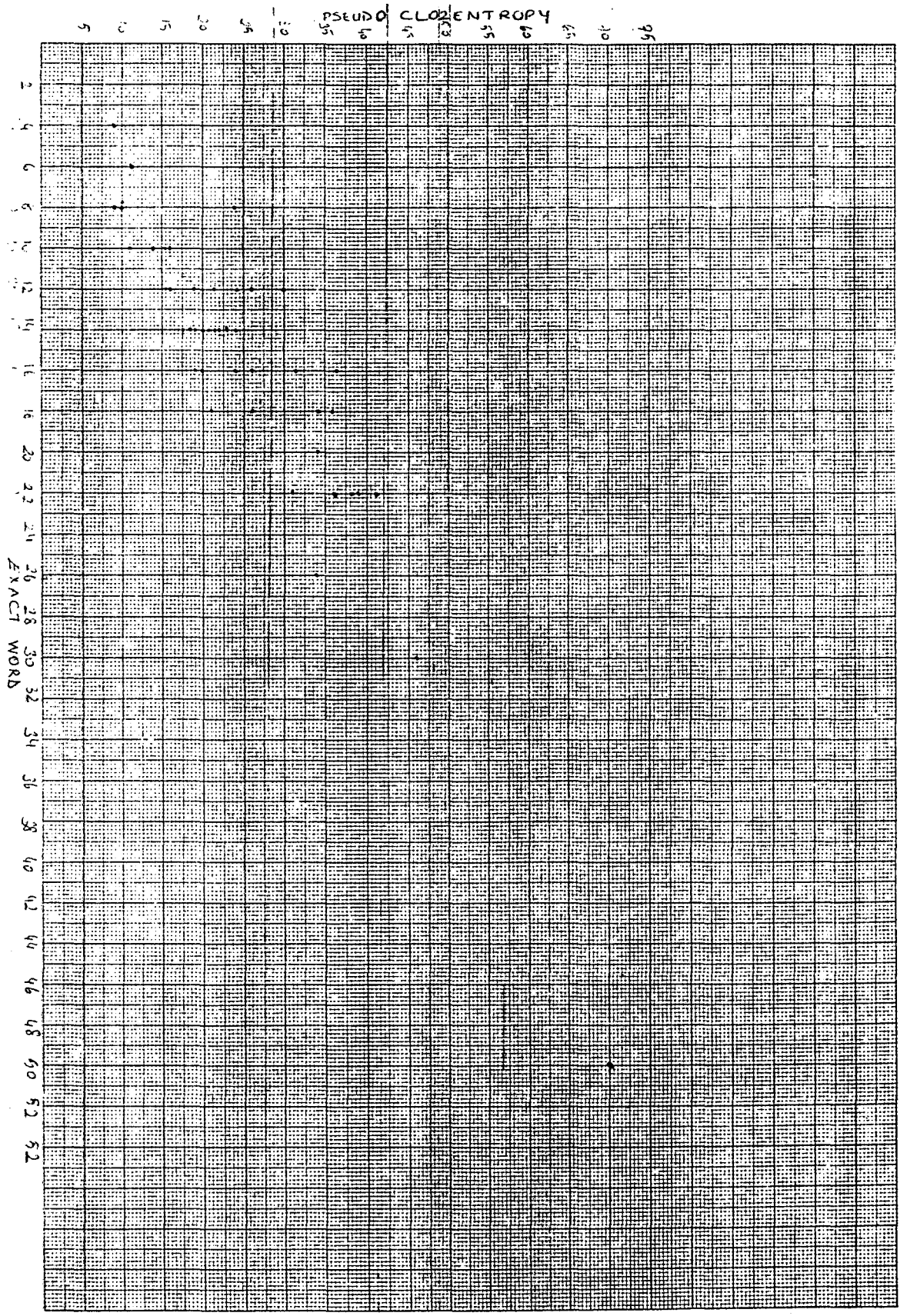
$S = 1558 - 990$

$S = 568$

$$\tau = \frac{S}{N(N-1)/2} = \frac{790}{990} = 0.57$$

Appendix H

The scattergram related to the correlation between exact word and pseudo clozentropy



Appendix I

Correlation between the Exact word and the Pseudo-clozentropy
Kendall's tau- no ties within conditions

Subject

1 - 13
2 - 6
3 - 13
4 - 24
5 - 13
6 - 6
7 - 13
8 - 16
9 - 11
10 - 16
11 - 2
12 - 2
13 - 0
14 - 2
15 - 2
16 - 3
17 - 1
18 - 5
19 - 35
20 - 15
21 - 35
22 - 39
23 - 16
24 - 13
25 - 35
26 - 13
27 - 17
28 - 28
29 - 17
30 - 39
31 - 35
32 - 39
33 - 17
34 - 28
35 - 18
36 - 17
37 - 15
38 - 13
39 - 43
40 - 16
41 - 9
42 - 12
43 - 17
44 - 2
45 - 39

$S^+ = 770$

N (number of pairs) N=45

$$\frac{N \times (N-1)}{2} = \frac{45 \times 44}{2} = 990$$

$$S = 2S^+ - \frac{N \times (N-1)}{2}$$

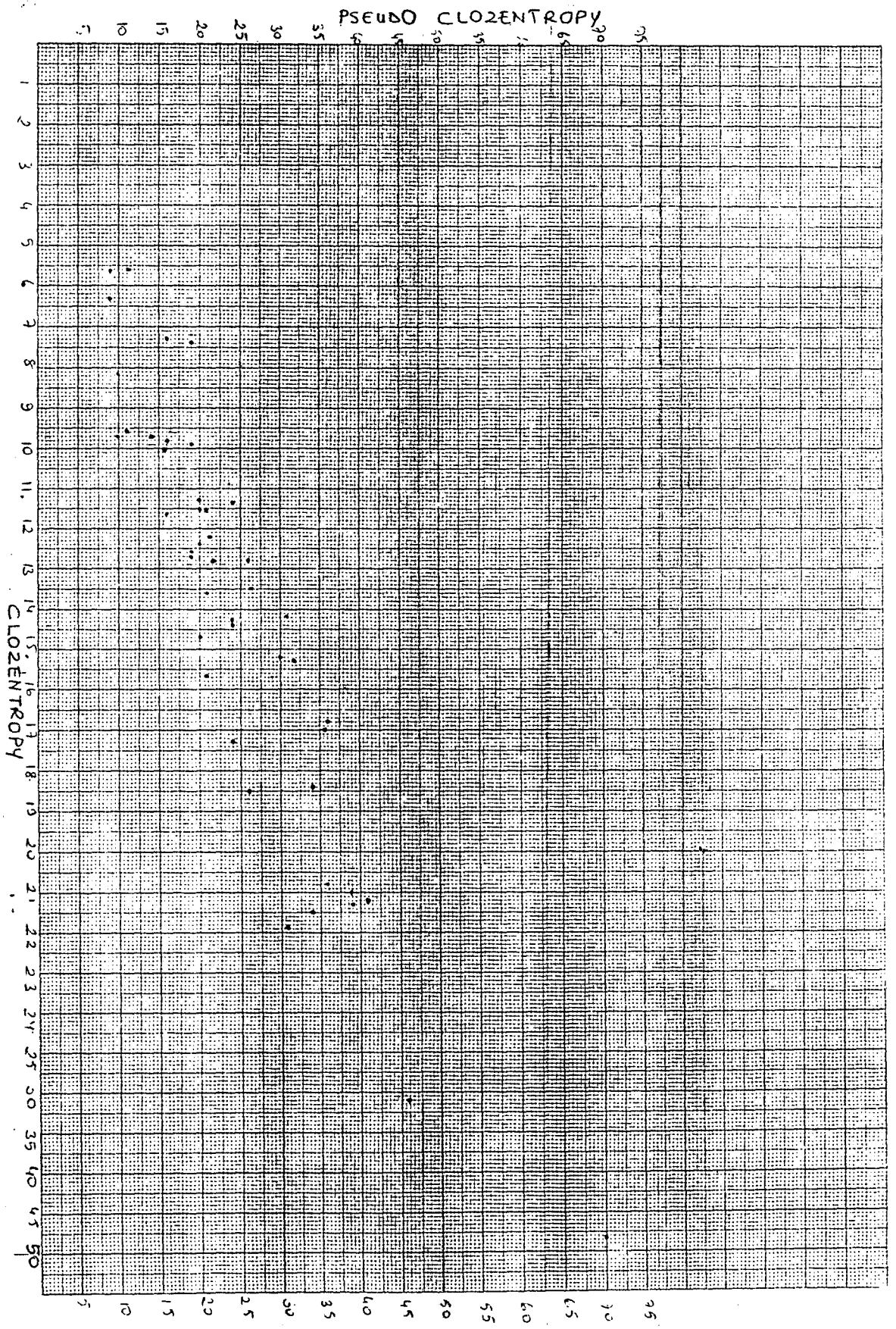
$$S = (2 \times 770) - 990$$

$$S = 550$$

$$\tau = \frac{S}{N(N-1)/2} = \frac{550}{990} = 0.55$$

Appendix J

The scattergram related to the correlation between clozentropy and pseudo clozentropy



Appendix K

Correlation between the Clozentropy and the Pseudo-clozentropy
Kendall's tau- no ties within conditions

Subject

1 - 15
2 - 10
3 - 5
4 - 30
5 - 14
6 - 2
7 - 11
8 - 21
9 - 8
10 - 19
11 - 5
12 - 4
13 - 0
14 - 6
15 - 2
16 - 2
17 - 1
18 - 6
19 - 33
20 - 14
21 - 32
22 - 39
23 - 25
24 - 11
25 - 33
26 - 16
27 - 17
28 - 32
29 - 24
30 - 33
31 - 28
32 - 42
33 - 18
34 - 33
35 - 30
36 - 25
37 - 10
38 - 24
39 - 42
40 - 13
41 - 22
42 - 14
43 - 23
44 - 2
45 - 39

$S^+ = 835$

N (number of pairs) N=45

$$\frac{N \times (N-1)}{2} = \frac{45 \times 44}{2} = 990$$

$$S = 2S^+ - \frac{N \times (N-1)}{2}$$

$S = (2 \times 835) - 990$

$S = 1670 - 990$

$S = 680$

$$\tau = \frac{S}{N(N-1)/2} = \frac{680}{990} = 0.68$$

Table I

The answers given by the 45 Turkish students and 25 native speakers of English to the cloze test. The exact word answer key is also included.

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
1	by	x by	6	22.2	x by	8	17.7
		x of	21	77.7	x of	6	13.3
					with	8	
					that	5	
					is	3	
					which	2	
					vehicle	2	
					aeroplane	1	
					height	1	
					mechanics	1	
					tool	1	
					on	1	
					aircraft	1	
					way	1	
					no answer	4	
2	was	x was	25	100	x was	37	82.2
					is	3	
					has	1	
					had	1	
					would	1	
					could	1	
					no answer	1	
3	by	xby	25	100	x by	21	46.6
					flight	3	
					the	3	
					from	2	
					science	1	
					many	1	
					busy	1	
					some	1	
					bold	1	
					research	1	
					for	11	
					no answer	9	
4	previous	previous	5	19.2			
		x 20th	9	34.6	x 20th	31	68.8
		x 19th	6	23	x 19th	7	15.5
		x last	2	7.6	x last	1	2.2
		x past	2	7.6	x past	1	2.2
		x 18th	1	3.8	x 18th	1	2.2
		the	1	3.8	this	2	4.4
					no answer	2	
5	of	x of	24	96	x of	5	11.1
		x before	1	4	x before	1	2.2
					was	6	
					that	4	
					since	1	
					after	1	
					had	1	
					fly	1	
					benefit	1	
					when	1	
					with	1	
					no answer	22	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
6	creation	creation	2	7.6			
		x product	17	65.3	x product	3	6.6
		x invention	3	11.5	x invention	1	2.2
		result	1	3.8	productor	2	
		concept	1	3.8	produced	1	
		thought	1	3.8	machine	3	
		work	1	3.8	vehicle	3	
					one	2	
					belonged	1	
					men	1	
					great	1	
					birth	1	
					flight	1	
					fly	1	
					made	1	
					either	1	
					first	1	
					industry	1	
					no answer	20	
7	no	x no	5	20.8	x no	27	60
		x any	15	62.5	x any	1	2.2
		x only	3	12.5	x only	1	2.2
		even	1	4.1	the	2	
					great	1	
					with	1	
					aviation	1	
					no answer	11	
8	years	x years	25	100	x years	38	84.4
					days	1	
					passengers	1	
					no answer	5	
9	transport	x transport	3	11.5	x transport	2	4.4
		travel	14	53.8	travelling	1	
		power	3	11.5	craft	4	
		flight	2	7.6	x flight	6	13.3
		planes	1	3.8	plane	4	
		it	1	3.8	lines	4	
		technology	1	3.8	line	4	
		transportation	1	3.8	which	2	
					that	1	
					company	1	
					control	1	
					traffic	1	
					carried	1	
					area	1	
					laws	1	
					no answer	11	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
10	the	x the	24	96	x the	31	68.8
		today's	1	4	all	2	
					new	2	
					transport	2	
					aviation	1	
					passengers	1	
					our	1	
					no answer	5	
11	passengers	x passengers	20	71.4	x passengers	23	51.1
		x people	8	28.5	x people	14	31.1
					miles	1	
					kilometres	1	
					metres	1	
					since	1	
					no answer	4	
12	kilometres	x kilometres	6	19.3	x kilometres	17	37.7
		x miles	23	74.1	x miles	19	42.2
		x hours	1	3.2	x hours	2	4.4
		per-year	1	3.2	metres	2	
					passengers	1	
					place	1	
					no answer	3	
13	network	x network	9	27.2	x network	1	2.2
		mass	7	21.2	using	5	
		web	4	12.1	way	2	
		system	2	6	plane	2	
		myriad	2	6	airplane	3	
		multitude	2	6	map	2	
		net	1	3	such	1	
		pattern	1	3	line	1	
		x lot	1	3	x lot	6	13.3
		series	1	3	machine	1	
		number	1	3	rule	1	
		plethora	1	3	traffic	1	
		maze	1	3	radar	1	
					technology	1	
					plan	1	
					planning	1	
					no answer	15	
14	become	x become	22	84.6	x become	1	2.2
		provided	1	3.8	got	14	
		proven	1	3.8	importance	2	
		been	1	3.8	had	2	
		created	1	3.8	not	3	
					line	1	
					built	1	
					best	1	
					route	1	
					flight	1	
					center	1	
					no answer	17	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
15	development	x development	13	50	x development	1	2.2
		transformation	2	7.6	kind	5	
		emergence	2	7.6	way	3	
		projection	1	3.8	one	2	
		leap	1	3.8	route	2	
		advancement	1	3.8	flight	2	
		establishment	1	3.8	line	2	
		introduction	1	3.8	commerce	2	
		age	1	3.8	joining	1	
		era	1	3.8	some	1	
		evolution	1	3.8	is	1	
		addition	1	3.8	centre	1	
					no answer	22	
16	instrument	mode	10	33.3	effect	4	
		means	7	23.3	flight	2	
		form	3	10	point	1	
		source	3	10	thing	1	
		type	1	3.3	character	1	
		kind	1	3.3	development	1	
		force	1	3.3	route	1	
		way	1	3.3	field	1	
		role	1	3.3	country	1	
		area	1	3.3	object	1	
		method	1	3.3	problem	1	
					place	1	
					service	1	
					line	1	
					airport	1	
					vehicle	1	
					use	1	
					no answer	24	
17	international	x many	17	62.9	x many	4	
		x some	4	14.8	x some	8	
		x two	2	7.4	x two	1	
		various	1	3.7	has	8	
		certain	1	3.7	own	3	
		several	1	3.7	these	1	
		numerous	1	3.7	solved	1	
					big	1	
					more	1	
					no answer	17	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
18	laws	technology	5	19.2	economic	8	
		planning	2	7.6	others	3	
		safety	2	7.6	difficulty	1	
		procedures	2	7.6	organize	1	
		manufacturing	1	3.8	money	1	
		operations	1	3.8	flight	1	
		research	1	3.8	transport	1	
		applications	1	3.8	tools	1	
		know-how	1	3.8	no answer	28	
		communication	1	3.8			
		routes	1	3.8			
		safety-measures	1	3.8			
		organization	1	3.8			
		methods	1	3.8			
		equipment	1	3.8			
		designs	1	3.8			
		administration	1	3.8			
		training	1	3.8			
		regulations	1	3.8			
19	information	information	10	37	difficulty	1	
		knowledge	2	7.4	ways	1	
		data	2	7.4	systems	2	
		x factors	1	3.7	x factors	1	2.2
		x problems	1	3.7	x problems	11	24.4
		aid	1	3.7	so	3	
		advances	1	3.7	concepts	1	
		resources	1	3.7	no answer	25	
		know-how	1	3.7			
		logistics	1	3.7			
		ties	1	3.7			
		disciplines	1	3.7			
		barriers	1	3.7			
		means	1	3.7			
		assistance	1	3.7			
		expertise	1	3.7			
20	governments	governments	1	3.1	problems	15	
		nations	9	28.1	has	1	
		people	7	21.8	make	1	
		capability	5	15.6	motion	1	
		countries	2	6.2	methods	1	
		airlines	2	6.2	questions	1	
		companies	2	6.2	ways	1	
		experts	1	3.1	them	1	
		scientists	1	3.1	find	1	
		minds	1	3.1	attempts	1	
		men	1	3.1	no answer	11	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
21	and	x and	19	76	x and	11	24.4
		standards	2	8	flight	5	
		x of	1	4	x of	5	11.1
		x has	1	4	x has	1	2.2
		rules	1	4	routes	1	
		control	1	4	was	1	
					is	1	
					aviation	1	
					fly	1	
					no answer	18	
22	necessity	necessity	6	23	aircraft	3	
		task	4	15.3	plane	2	
		process	3	11.5	design	2	
		problem	3	11.5	same	2	
		constant	2	7.6	next	1	
		planning	1	3.8	one	1	
		problems	1	3.8	make	1	
		work	1	3.8	safety	1	
		challange	1	3.8	flight	1	
		science	1	3.8	many	1	
		expertise	1	3.8	methods	1	
		action	1	3.8	need	1	
		administration	1	3.8	airplane	1	
					no answer	27	
23	navigation	navigation	3		the	4	
		navigational	8		this	2	
		training	3		people	1	
		communication	2		control	1	
		x more	1		x more	1	2.2
		medical	1		it	1	
		radio	1		first	1	
		x of	1		x of	1	2.2
		traffic	1		some	1	
		support	1		to	1	
		safety	1		for	1	
		passenger	1		no answer	30	
		travel	1				
24	systems	x systems	4	16	x systems	1	2.2
		x services	4	16	x services	2	4.4
		x stations	8	32	x stations	4	8.8
		x centres	2	8	x centres	1	2.2
		procedures	2	8	part	1	
		techniques	1	4	result	1	
		facilities	1	4	tower	1	
		x too	1	4	x too	2	4.4
		forecasts	1	4	them	1	
		channels	1	4	area	1	
					forecast	1	
					office	1	
					no answer	28	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
25	international	these	5	17.8	aeroplane	1	
		all	4	14.2	good	2	
		x airline	4	14.2	x airline	1	2.2
		x air	3	10.7	x air	6	13.3
		airport	1	3.5	flight	4	
		support	1	3.5	the	5	
		scheduled	1	3.5	passenger	1	
		x technical	1	3.5	x technical	2	4.4
		mechanical	1	3.5	many	1	
		communicative	1	3.5	ground	2	
		informative	1	3.5	crew	1	
		its	1	3.5			
		x such	1	3.5	x such	1	2.2
		military	1	3.5	no answer	18	
		bazic	1	3.5			
		controlled	1	3.5			
26	that	x that	20	86.9	x that	20	44.4
		important	1	4.3	and	2	
		as	1	4.3	flight	2	
		there	1	4.3	many	1	
					no answer	20	
27	by	x by	22	88	x by	10	
		through	2	8	of	4	
		nor	1	4	for	3	
					to	3	
					about	2	
					and	1	
					plot	1	
					the	1	
					any	1	
					accident	1	
					no answer	18	
28	such	safety	7	24.1	the	10	
		international	7	24.1	world	2	
		operational	3	10.3	air	2	
		x these	2	6.8	x these	1	
		safe	1	3.4	airplane	1	
		clear	1	3.4	aviation	1	
		rigid	1	3.4	flying	1	
		countless	1	3.4	airlines	1	
		universal	1	3.4	many	1	
		flights	1	3.4	flight	1	
		high	1	3.4	some	1	
		set	1	3.4	this	1	
		many	1	3.4	ICAO	1	
		aeronautical	1	3.4	general	1	
					standardization	1	
					no answer	17	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
29	air	air	4	17.3	aviation	4	
		sky	3	13	flight	4	
		procedures	2	8.6	aeroplane	3	
		practise	2	8.6	establishment	2	
		xnecessity	1	4.3	xnecessity	1	2.2
		operation	1	4.3	organization	2	
		operations	1	4.3	weather	1	
		airports	1	4.3	lines	1	
		highest	1	4.3	men	1	
		runway	1	4.3	we	1	
		world	1	4.3	traffic	1	
		road	1	4.3	laws	1	
		offices	1	4.3	no answer	23	
		methods	1	4.3			
		regulation	1	4.3			
		administration	1	4.3			
30	licensing	x training	5	20	x training	2	4.4
		management	4	16	handing	1	
		requirements	3	12	health	1	
		recruitment	2	8	and	1	
		staff	2	8	who	1	
		conduct	1	4	abilities	1	
		hiring	1	4	right	2	
		selection	1	4	on	1	
		services	1	4	service	1	
		also	1	4	buildings	1	
		safety	1	4	education	3	
		flying	1	4	condition	1	
		convenience	1	4	problem	1	
		development	1	4	happen	1	
					no answer	27	
31	for	x for	12	52.3	x for	4	8.8
		x also	1	4.7	x also	2	4.4
		often	2	9.5	them	2	
		routes	1	4.7	rules	2	
		training	1	4.7	that	1	
		facilities	1	4.7	others	1	
		runways	1	4.7	gives	1	
		create	1	4.7	the	1	
		with	1	4.7	aircraft	1	
		likewise	1	4.7	no answer	30	
32	to	to	18	66.6	for	14	
		x in	2	7.4	x in	1	
		x of	1	3.7	x of	4	
		and	1	3.7	on	1	
					about	1	
					flights	1	
					provide	1	
					no answer	22	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
32	to	to	18	66.6	for	14	
		x in	2	7.4	x in	1	2.2
		x of	1	3.7	x of	4	8.8
		and	1	3.7	on	1	
					about	1	
					flights	1	
					provide	1	
					no answer	22	
33	national	national	2	7.1	an	2	
		individual	4	14.2	the	2	
		x one	3	10.7	x one	3	6.6
		primary	2	14.2	all	2	
		immediate	2	14.2	this	2	
		just	2	14.2	its	2	
		temporary	1	3.5	these	1	
		brief	1	3.5	real	1	
		single	1	3.5	important	1	
		separate	1	3.5	on	1	
		remedial	1	3.5	that	1	
		casual	1	3.5	fire	1	
		unilateral	1	3.5	aviation	1	
		to	1	3.5	those	1	
		simple	1	3.5	no answer	24	
		passing	1	3.5			
		normal	1	3.5			
		minimal	1	3.5			
		limited	1	3.5			
34	a	x a	25	100	x a	7	15.5
					made	4	
					some	3	
					been	3	
					the	2	
					had	2	
					causes	1	
					many	1	
					started	1	
					caused	1	
					no answer	20	
35	of	x of	25	100	x of	22	48.8
					on	5	
					in	2	
					to	2	
					at	1	
					about	1	
					was	1	
					for	1	
					no answer	10	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
36	normal	normal	5	22.7	twenty	4	
		previous	2	9	20th	4	
		gradual	2	9	a	2	
		then	1	4.5	world	1	
		x the	1	4.5	x the	8	18.7
		relative	1	4.5	this	1	
		ongoing	1	4.5	nineteen	1	
		possible	1	4.5	no answer	24	
		earlier	1	4.5			
		once	1	4.5			
		commercial	1	4.5			
		hand	1	4.5			
		continual	1	4.5			
		formerly	1	4.5			
		steady	1	4.5			
		slow	1	4.5			
37	vast	vast	4	12.9	airplane	2	
		x new	5	16.1	x new	1	2.2
		world-wide	3	9.6	main	1	
		x large	3	9.6	x large	1	2.2
		x major	2	6.4	x major	1	2.2
		small	2	6.4	flight	1	
		extensive	1	3.2	plane	1	
		immense	1	3.2	certain	1	
		modern	1	3.2	doing	1	
		necessary	1	3.2	development	1	
		huge	1	3.2	no answer	34	
		simple	1	3.2			
		basic	1	3.2			
		wide	1	3.2			
		improved	1	3.2			
		rough	1	3.2			
		complicated	1	3.2			
		complex	1	3.2			
38	was	x was	17	68	x was	6	
		airlines	2	8	could	2	
		had-been	2	8	which	1	
		were	1	4	has	1	
		has-been	1	4	system	1	
		had	1	4	is	1	
		x to	1	4	x to	1	
					cpacity	1	
					it	1	
					no answer	30	
39	problems	x problems	24	82.7	x problems	15	33.3
		obstacles	1	3.4	other	1	
		questions	1	3.4	airlines	1	
		flaws	1	3.4	inventions	1	
		hitches	1	3.4	development	1	
		complications	1	3.4	reasons	1	
					no answer	25	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
40	solutions	x solutions	17	62.9	x solutions	1	2.2
		x answers	7	25.9	x answers	1	2.2
		there	1	3.7	they	2	
		answer	1	3.7	rules	1	
		ways	1	3.7	countries	1	
					it	2	
					its	1	
					were	1	
					for	1	
					have	1	
					no answer	33	
41	and	x and	22	88	x and	5	11.1
		way	1	4	for	4	
		the	1	4	to	4	
		those-who	1	4	it	2	
					by	1	
					no answer	29	
42	was	x was	21	84	x was	12	26.6
		x is	3	12	x is	7	15.5
		x were	1	4	x were	7	15.5
					are	1	
					no answer	18	
43	arrangements	provisions	7	22.5	they	2	
		rules	5	16.1	rule	1	
		regulations	4	12.9	cause	2	
		rights	3	9.6	solutions	3	
		allowences	3	9.6	were	1	
		route	2	6.4	law	1	
		x it	1	3.2	x it	3	6.6
		requirements	1	3.2	which	1	
		agreements	1	3.2	no answer	31	
		laws	1	3.2			
		exceptions	1	3.2			
		airspace	1	3.2			
		permission	1	3.2			
44	one	x one	14	53.8	x one	1	2.2
		a	4	15.3	the	7	
		another	2	7.6	20th	4	
		foreign	2	7.6	twenty	1	
		any-given	1	3.8	nineteen	1	
		each	1	3.8	every	1	
		any	1	3.8	many	1	
		different	1	3.8	this	2	
					developed	1	
					no answer	26	

Blank No.	Exact word method checklist	Answers given by the native speakers	Number	%	Answers given by the Turkish students	Number	%
45	the	x the	19	70.3	x the	8	17.7
		sovereigh	3	11.1	for	1	
		national	1	3.7	one	1	
		domestic	1	3.7	to	1	
		which	1	3.7	no answer	28	
		other	1	3.7			
		x air	1	3.7	x air	6	13.3
46	concerns	concerns	1	3.7	things	2	
		x problems	18	66.6	x problems	6	13.3
		considerations	4	14.8	rules	1	
		questions	2	7.4	conditions	1	
		constituents	1	3.7	airlines	1	
		factors	1	3.7	details	1	
					actions	1	
					no answer	32	
47	economic	economic	2	6.6	law	1	
		political	10	33.3	the	1	
		moral	3	10	its	1	
		territorial	2	6.6			
		national	2	6.6			
		personal	2	6.6			
		trade	1	3.3			
		domestic	1	3.3			
		other	1	3.3			
		business	1	3.3			
		regional	1	3.3			
		local	1	3.3			
		financial	1	3.3			
		x illegal	1	3.3	x illegal	8	17.7
		technical	1	3.3	no answer	34	
48	flying	flight	11	39.2	to	2	
		travel	9	32.1	for	1	
		security	1	3.5	among	1	
		wealth	1	3.5	of	1	
		optimism	1	3.5	the	1	
		operations	1	3.5	development	1	
		passage	1	3.5	century	1	
		use	1	3.5	and	1	
		navigation	1	3.5	no answer	36	
		movements	1	3.5			
49	to	x to	24	96	x to	9	20
		governments	1	4	a	2	
					is	2	
					was	1	
					they	1	
					does	1	
					no answer	29	
50	of	x of	24	96	x of	3	6.6
		residents	1	4	countries	3	
					things	2	
					services	2	
					aircraft	1	
					other	1	
					airlines	1	
					facilities	1	
					planes	1	
					no answer	31	

Table 2 Significance of Kendall's tau

N = number of pairs of scores
The τ values are the smallest values of τ significant at the 0·05 level for different values of N

N	τ	N	τ
5	0·80	21	0·31
6	0·69	22	0·29
7	0·63	23	0·29
8	0·57	24	0·28
9	0·53	25	0·27
10	0·49	26	0·27
11	0·45	27	0·26
12	0·43	28	0·25
13	0·41	29	0·25
14	0·39	30	0·25
15	0·37		
16	0·35		
17	0·35		
18	0·33		
19	0·33		
20	0·31		

Adapted from B. F. Anderson, *The Psychology Experiment*, Brooks-Cole, 1966.