CHAPTER ONE

INTRODUCTION

1.1 General Background

Language is the most advanced and powerful means and a form of communication which enable us to exchange, share, express ideas, information and feelings through the medium of either graphic or spoken systems. It is species specific and species uniform possession of human beings. It is one of the most valuable gifts of god to human being which distinguishes men from other animals. Language can also be taken as social phenomenon which is used in our society to establish the relationship among the human beings. Through the interaction among the members of the community by means of language a society is alive. It is not just the juxtaposition of words or noises of sounds but a systematic arrangement of these components so as to systematize a message to establish relationship between and among the people not only of society or community but also of entire globe.

According to Lyons (1970, p.3) "The principal systems of communication used by particular groups of human beings within particular society of which they are the members." In the same way, Gimson (1974, p.3) defines language as "A language is a system of conventional signal used for communication by a whole community." Similarly, Bhattarai (1994, p.1) says:

> Universal medium of conveying facts including complete thoughts, emotions and feelings of everyday life. Language has enabled man to establish great civilization. Man differs from the other species on this earth only because s/he possesses a unique faculty of speech. Man expresses his personality through language.

Language is defined variously. No single definition of language is perfect in itself. But it is widely accepted that language is a complex human phenomenon whose main function is to communicate by means of speaking or writing.

Language makes a life bitter as sweet when we use. It is our ability to communicate through words that makes human beings different from animals. It is a very complex psychological and social phenomenon in human life. So, it is common to all and only human beings.

In the field of language teaching, testing plays an important role. Through testing, teachers can evaluate the language proficiency, diagnose the students' weaknesses and strengths and provide the remedial feedback to the students. So, test items should be designed keeping its qualities in mind. There are many qualities to be included for making a test good. Among them, validity is the most one that refers to the extent to which it measures what is intended to measure. We design the test items having certain objectives which should be reflected in tests. Thus, it is based on objectives of course. Heaton (1975, p.153) says "Briefly, the validity of test is the extent to which it measures what it is supposed to measure and nothing else."

1.1.1 English Language Teaching in Nepal

There are several languages spoken in the world among which English is the most dominant one used as an international lingua franca in many parts of the world. It is also used extensively as a link language. It is a global language which at present is most widely taught as a foreign language in over 100 countries. It is one of the important means to access to new scientific, medical, educational, technological information and so on.

Nowadays, English has become an inseparable parts of people and society in the world. It has become an inevitable source of knowledge for non-native speakers because most of the important books of the world are written and

translated in English. At the age of scientific discoveries and development, it is the gateway to knowledge.

Bhattarai (1995, p.226) says:

> English has become indispensable vehicle to the transmission of modern civilization in the nation. It is a passport through which one can visit the whole world and one who knows English can enjoy the advantages of the world citizen. Therefore, it is the only means of preventing our isolation from the world and we will act unwisely if we allow ourselves to be enveloped in the folds of dark curtain of ignorance.

The English language which is known as an international language, entered Nepal in 1910 B.S. when then Prime Minister Jung Bahadur Rana returned from his visit to England. He opened Durbar High School to educate his family members realizing the necessity and importance of the English language. Since then English has been a part of education in Nepal.

In Nepal, English has been teaching from grade 1 to the bachelor's level as a compulsory subject which carries at least 100 full marks and educational curriculum has also managed that any interested students can read English in upper level as a major subject. Thus, English is taught at community school, English medium school, campus, university and so on.

1.1.2 Major English at M.Ed. Level

1.1.2.1 Introduction

Education plays an important role in national development. The extent to which education will be able to support the process of national development depends on the academic and professional strengths of educational planners, managers, educators, headmasters and teachers. As the demand for quality education, the

need for competent professionals become even greater. To meet this need the Faculty of Education (FOE), Tribhuvan University with its two year master programme is committed to provide better and qualified teachers, teacher educators, educational administrators and planners, system analysts and experts in the field of education. The FOE has introduced three-year B.Ed. programme since 1996 with the curriculum focused on meeting varying social needs. The necessities restructuring the M.Ed. curriculum so to make it constant with the newly developed B.Ed. programme. As a result, this change has been prompted by the vision and the experience of teacher educators and the professionals working in the field of education all these years.

No one can be qualified or competent if s/he doesn't hold respective qualification. To produce such type of teacher educators or manpower, M.Ed. programme under T.U. plays a key role. The overall objective of the M.Ed. programme is to produce higher level manpower in the field of education.

1.1.2.2 Course Structure

There are altogether seventeen papers, two papers carry 100 marks and fifteen papers carry 50 marks. Papers are divided into two major groups. Group A for specialization and it carries fourteen papers, and group B for elective and it carries three papers. In elective group out of three papers only one paper will be offered by the concerned subject committee in the campuses for teaching learning purpose.

Table No. 1.1

Specialization of Course in M.Ed.

S.N.	Course No.	Subjects	Marks
1.	Eng. Ed. 516	Phonetics and Phonology	100
2.	Eng. Ed. 517	English Grammar for teachers	100
3.	Eng. Ed. 518	Sociolinguistics	50

S.N.	Course No.	Subjects	Marks
4.	Eng. Ed. 519	Second Language Acquisition	50
5.	Eng. Ed. 520	Research Methodology In Language Education	50
6.	Eng. Ed. 521	Language Testing	50
7.	Eng. Ed. 522	Literature in English Language Teaching	50
8.	Eng. Ed. 523	Pragmatics and Discourse Analysis	50
9.	Eng. Ed. 524	English Language Teacher Development	50
10.	Eng. Ed. 525	Directions and Issues in Applied Linguistics	50
11.	Eng. Ed. 526	Advanced Academic Reading and Writing	50
12.	Eng. Ed. 527	Translation Studies (Elective)	50
13.	Eng. Ed. 528	ELT Curriculum, Materials and Management (Elective)	50
14.	Eng. Ed. 529	Reading in English (Elective)	50
15.	Eng. Ed. 591	Advanced English Language Teaching Methodology	50
16.	Eng. Ed. 598	Thesis Writing	50
17.	Eng. Ed. 599	Practicum	50

It is clear from the abovementioned data and description that at M.Ed. first year four specialization papers: Eng. Ed. 516-Phonetics and Phonology, Eng. Ed. 517-English Grammar for teachers, Eng. Ed. 518-Sociolinguistics, and Eng. Ed. 519-Second Language Acquisition are adopted. In second year ten specialization papers: Eng. Ed. 520-Research Methodology In Language Education, Eng. Ed. 521-Language Testing, Eng. Ed. 522-Literature in English Language Teaching, Eng. Ed. 523-Pragmatics and Discourse Analysis, Eng. Ed. 524-English Language Teacher Development, Eng. Ed. 525-Directions and Issues in Applied Linguistics, Eng. Ed. 526-Advanced Academic Reading and Writing, Eng. Ed. 591-Advanced English Language Teaching Methodology, Eng. Ed. 598-Thesis Writing and Eng. Ed. 599-Practicum are adopted. In second year any one paper either Eng. Ed. 526-Advanced Academic Reading and Writing or Eng. Ed. 598-Thesis Writing is adopted as an option. There are

three elective papers in second year. They are Eng. Ed. 527-Translation Studies, Eng. Ed. 528-ELT curriculum, Materials and Management and Eng. Ed. 529-Reading in English. Among them, any one paper from elective group is adopted.

There are prescribed textbooks, reference materials, syllabus, curriculum etc. to meet the objectives of the courses. In the context of T.U., there is only the provision of annual examination system. According to the nature of course, written, oral and practical examinations are administered. But in M.Ed. first year in the subject Second Language Acquisition only annual written test is administered to measure the students' competence.

1.1.2.3 The Introduction of Second Language Acquisition

The Second Language Acquisition is a specialized subject for M.Ed. first year. It occupies 50 full marks. This is an introductory course on Second Language Acquisition (SLA). It is divided into five units. The first unit deals with the basic concepts of SLA including its history. The second unit incorporates various approaches and perspectives of SLA. The third unit deals with the role of input, interaction and output. The fourth and fifth units include the topic like instructed second language learning and non-language factors in SLA respectively.

Objectives of the Course:

   a. To acquaint the students with the basic concepts of SLA including its history.

   b. To familiarise them with various approaches and perspectives of SLA.

   c. To acquaint them with the role of input and interaction to yield desired output.

   d.   To make them familiar with the roles of non-language factors in SLA.

Contents of the Course

Unit I: Basic concepts of SLA

Unit II: Approaches and perspectives of SLA

Unit III: Input, Interaction and output in SLA

Unit IV: Instructed Second Language Learning

Unit V: Non-language Factors in SLA

The more detail of this course contents is given in the appendix A.

The 'Second Language Acquisition' carries 50 full marks and 20 pass marks. From the external written examination and final examination, the students' competence is tested. Both types of questions i.e. subjective (short and long) and objective questions are administered in the test. 8 objective questions (multiple choice items) carrying 8 marks (i.e. 1 mark for each question) are asked from group A. In case of subjective questions, 5 with 2 'or' short questions are asked carrying 6 marks for each question from group B and only 1 long question carrying 12 marks is asked as obligatory question from group C.

1.1.3 Teaching and Testing

Testing and assessment are important part of education. Testing plays an important role in language teaching in classroom settings. Testing is used as a process of scrutinizing how far learners have learned what the teacher whishes them to learn. To ensure that the students have achieved some or whole of what has been taught is thus purpose of testing. In order to ensure that the teaching is effective, and if not, put more efforts to make it effective, testing is used in the classroom or after the classroom teaching. In many classroom situations, thus teaching and testing are hardly separated.

Testing plays an important role in our daily life. Where there is teaching, then testing exists. So, testing is an inherent part of teaching. Teaching is geared to

the test i.e. to assess the students' performance in language. With the help of testing we can find the areas that need more attention and we can do the remedial works in teaching. Davies (1968, p.5) says" The good test is an obedient servant since it follows and apes the teaching." However, Hughes (2010, p.2) doesn't agree this. He argues that the proper relationship between teaching and testing is surely that of partnership. He views that in some occasions the teaching may be good and appropriate and the testing leads to harmful washback and in some occasions teaching may be poor or inappropriate and testing is able to exert beneficial influence. We cannot expect testing only to follow teaching. Rather, we should demand of it that it is supportive of good teaching and, where necessary, exerts a corrective influence on bad teaching. If testing always had a beneficial backwash on teaching, it would have a much better reputation among teachers.

Teaching and testing are interrelated as two sides of a coin in the sense that one would be meaningless and incomplete in the absence of others. Khaniya (2005, p.1) says "Testing in a broad sense has always been an inherent part of teaching. Assessment of learning is as old as education itself. "It is virtually impossible to work in either field without being constantly concerned with the other. The test may be constructed as device to reinforce learning and motivating students to measure the performance of the students' as well as the ability in education system itself. Teaching without testing does not guarantee the learning on the part of the learners. Testing affects teaching or vice versa. Such effect is what we call backwash effect .Teaching and testing are thus regarded as an integral part of education.

1.1.4 Significance of Testing

Teaching is, after all, the primary activity; if testing comes in conflict with it, then it is testing that should go, especially when it has been admitted that so much testing provides inaccurate information. Information about one's language ability is very important and sometimes necessary too. British and

American universities accept students form different overseas only after measuring language proficiency in English. The same is true for different organizations. They hire translators or interpreters on the basis of proficiency they performed in concerned genres. Within teaching systems, too, so long as it is thought appropriate for individuals to be given a statement of what they have achieved in a second or foreign language, tests of some kind or another will be needed. They will also be needed in order to provide information about achievement of groups of learners, without which it is difficult to see how rational education decisions can be made. Heaton (1975, p.3) says:

> Evaluation for the purpose of selection, the classroom test is concerned with evaluation for the purpose of enabling the teacher to increase his own effectiveness by making adjustment in his teaching to enable certain group of students or individual in the classroom to benefit more.

A classroom test can help to locate precise area of difficulty encountered by the individuals or class. It can be comparable to doctor's treatment as in the process of diagnosing his/her patients' illness. Thus, it is equally necessary to diagnose students' weaknesses and strengths. In this way, a teacher should find out the errors and mistakes committed by his/ her learners and provide appropriate feedback. Test is equally significant to language programme to find out weakness. Heaton (1975, p.3) further states" The teacher can evaluate effectiveness of syllabus as well as methods and materials he is using."

In nutshell, test is significant from different perspectives, evaluating learners' ability for selecting appropriate level, diagnosing errors and mistakes, provide accurate information, selecting syllabus, methods and techniques as well as motivating students. Testing is applied for different purposes, There are many objectives of it. Hughes (2010, p.8) writes the following purposes of testing :

  i. To measure language proficiency.

ii. To discover how successful students have been in achieving the objectives of a course of study .

iii. To diagnose students' strengths and weaknesses, to identify what they know and what they don't know.

iv. To assist placement of students by identifying the stage or part of a teaching programme most appropriate to their ability.

1.1.5 Qualities of a Test

The factors which are responsible for making a test good are the qualities of a good test. The test which is constructed to conduct the exam must be of good quality. So, the usefulness of the test or the quality of the test should be considered while designing a test. What are the qualities of a test? What makes a test useful? What is test efficiency? These are some issues to make the test good. There are different views on what makes a test good. Some expert says that there are three constituents of exam efficiency: validity, reliability and practicality. In this sense, validity, reliability and practicality should be seen as relativistic concepts. The whole idea of considering the three constituents of exam efficiency is to build-up a framework for designing a good test. Bachman and palmer (1996) argue that test usefulness involves reliability, construct validity, authenticity, interactiveness, impact and practicality. Though the presentation of Bachman and palmer may appear to be a bit elaborative, in essence, validity, reliability and practicality constitute the quality of the test.

1.1.5.1 Validity

According to Hughes (2010), "A test is said valid if it measured accurately what it is intended to measure." (p.26). Thus validity refers to the degree to which a test measures what it is supposed to measure, if it can be used successful for the purposes for which it is intended. A number of different statistical procedures can be used to a test to estimate its validity. Such procedures generally seek to determine what the test measures, and how well it

does so. In other words, the test should seek to achieve that aim which is in the mind of the constructor. Any test will be valid when it is reliable. Content validity, construct validity, concurrent validity, predictive validity and face validity are major types of validity.

1.1.5.2 Reliability

Reliability is one of the qualities of a good test. In other words, a good test must be reliable. A reliable test is a test that is consistent and dependable. Reliability is basically related to scoring: no matter who is the examiner. If the scoring is same then it is said to be reliable. A test must be consistent in its measurement in order to be reliable. Khaniya (2005, p.116) mentions, "Reliability is another essential quality of a test which refers to the consistency of scores or performance of the same or similar test administered within a reasonable time." He further says that it is a matter of the extent to which we can believe that performance is true, how likely it is that the performance will be repeated next time. It can be dealt with at two levels: test and retest of the students and marking and remarking of the examiners. It is reported in terms of correlation coefficient. In short it is defined as consistency of measurement. Reliability is concerned with examining consistency in the performance of the examinee.

There are basically three methods of determining reliability of the exam. They are test-retest method, parallel test method and internal consistency method.

1.1.5.3 Practicality

Practicality is slightly different to the other qualities as it relates to the implementation of the test and whether it will be developed and used at all, rather than how the test scores will be used. Heaton (1975, p.158) explains practicality is that the exam "must be fairly straightforward to administer." Generally, practicality involves the cost, ease of administration and scoring.

In order to test with a reasonable degree of practically, it is necessary to pay attention to the following issues; human resource, material resource and time. According to Bachman and Palmer (1996, p.37), human resources, material resources and time have pivotal role to make the test practicable.

Human resource refers to test writers, scorers or raters, test administrators and clerical support. Similarly, material resource refers to space, equipment and materials etc. Time is also very important. It refers to the time for development and specific task.

1.1.6 Validity

The measurement of appropriateness and relevancy of the test is regarded as validity of the test. In other words, the validity of the test depends upon the appropriateness, effectiveness, usefulness and relevance of the test. It refers to degree to which a test actually measures what it is designed to measure. For example, if we test the student's writing skills giving them a composition test on ways of cooking, we cannot denote such test as valid, for it can be argued that it tests not our abilities to write, but the knowledge of cooking as skill.

The test is valid if it is appropriate, accurate and relevant in the light of purpose. "The validity of a test is measured on the basis of how far the information is accurate, concrete and representative in the light of the purposes for which it is administered"(Khaniya,2005, p.103). Both validity and reliability are interrelated and crucial qualities of a test. Reliability is prerequisite for validity. To be valid a test must provide consistently accurate measurements. It must, therefore, reliable. A reliable test may not be reliable at all. As a writing test we want learners to write an essay on 'Dowry system' in 150 words in their own language. This could well be reliable test but unlikely to be a valid test of writing. Hughes (2010, p.8) mentions "A test that proves ideal for one purpose may be quite useless for another; a technique that may work well in one situation can be entirely inappropriate in another." In this way, when we write a test, we have an intention to measure something that is

real and that validity enquiry concerns finding out whether a test actually does measure what is intended. There are different kinds of validity-content validity, construct validity, concurrent validity, predictive validity and face validity.

1.1.6.1 Face Validity

A test is said to have face validity if it looks as if it measures what it is supposed to measure. For example, a test that pretended to measure pronunciation ability but which did not require the test taker to speak might to lack face validity. In other words, it is said as the surface credibility or public acceptability. A test which does not have face validity may not be accepted by candidates, teachers, education authorities or employers. It may simply not be used; and if it is used, the candidates' reaction to it may mean that they do not perform on it in a way that truly reflects their ability.

Face validity is hardly a scientific concept, yet it is very important. Regarding its importance in testing, Anderson et al. (2010, p.173) say:

> For one thing tests that do not appear to be valid to the users may not be taken seriously for their given purpose. For another, if test takers consider a test to be face valid we believe that they are more likely to perform to the best of their ability on that test and to respond appropriately to items.

Thus, if a test item looks right to other testers, teachers, moderators and testees it can be described as having at least face validity. Although, face validity is considered as pseudo validity, it is believed that if the examiner does not consider on exam valid one, the information collected from it may not be genuine.

1.1.6.2 Criterion-Related Validity

Criterion-related validity refers to the type of validity where validity is established employing a process of comparing the results of a test with the

results of some criteria already set or the later performance of the students. It is also called the empirical validity or statistical validity. This validity relates to how test taker's scores compare with an external criterion, for example, another test taken at approximately at the same time, a candidate's or teacher's assessment of an individual's language ability, or a different version of the same test. Entrance examinations and final examinations are some examples of it. Because both of them evaluate the present behavior and speculate the future achievement of the students that is why in terms of the result of entrance examination and final examination students may be qualified or disqualified to get admitted in certain academic level. Criterion related validity can be discussed under two types:

(i) Concurrent Validity

Concurrent Validity refers to the process of determining the validity against the set criterion at the same time. Test developers tend to establish the validity of the new test comparing the student's performance on this test against their performance on a test of similar kind already established. The correlation between the two tests is said to be the concurrent validity of the new test. For instance, if a talent student in the class gets the highest marks in the examination, it is known as concurrent validity. For concurrent validity, a set of question for the examination must process some characteristics like relevancy, non-biasness, reliability, etc.

(ii) Predictive Validity

Predictive validity is the extent to which the scores on a test correlates with some later criterion. In other words, predictive validity refers to the degree to which a test can predict candidates future performance. Khaniya (2005, p.108) mentions "Predictive validity of a test is concerned with the extent to which the test can predict the future performance of the testees." We can establish the predictive validity through comparison of the test results and other criteria such as success in particular jobs or higher education. For example, a person who

has passed B.Ed. in the first division in English then it is predicted that he will be a good English teacher at secondary level. When he enters in teaching, teaches very well then the test is said to have predictive validity. Anderson et al. (2010, p.180) say "Test which are intended to predict how well somebody will perform in the future." In this way a test should have a quality of prediction to suit predictive validity.

1.1.6.3 Construct Validity

Construct Validity refers to a kind of validity which is based on the degree to which the items in a test reflects the essential aspect of the theory on which the test is based. In simple language, a test can be said to have construct validity if it measures just the ability which it is expected to measures and nothing else. Here the ability refers to theoretical construct or the theoretical explanation or proposition of a trait. Further explanation is that, as writing ability involves, in addition to other things, socio-cultural norms and cognitive processes, a test which requires testees to demonstrate such traits can be said to have construct validity. "If a test has construct validity, it is capable of measuring certain specific characteristics in accordance with theory of language behaviour and learning " (Heaton, 1975, p.154). Thus, it assumes the existence of certain learning theories or construct underlying, the acquisition of abilities and skills.

Any theory of language learning or behaviour or any assumption has certain type of mental ability, concept which a learner acquires in the process of language learning. This concept should be made operational or measurable through a test. If a test has such quality, i.e. measuring mental concept then it is said to have construct validity. It can be established through comparison with language theory- assessing to what extent the test is successfully based upon in its underlying theory; internal correlation-correlate different test components with each other's; comparison with bio-data and psychological characteristics and factor analysis. In fact, what extent the test is an indicator of construct or trait.

1.1.6.4 Content Validity

A test is said to have a content validity if its content constitutes a representative sample of the language skills, structures etc. with which it is meant to be concerned. In other words, the degree to which the test accurately reflects the syllabus is content validity. When we write a test the test should represent the contents of the course. The test is said to be valid if it represents the whole contents. Brown (1996, pp.122-123) says "whether the items composing the test do, in fact, constitute representative sample of the content domain of concern." To have content validity if its test items are considered to be a representative sample of the tasks which can be seen in the course objectives. "The test would have content validity only if it included a proper sample of the relevant structures" (Hughes, 2010, p.26). Likewise, Anastasi (1982, p.131 as cited in Weir, 1990, p.25) mentions "Essentially the systematic examination of the test content to determine whether it covers a representative sample of a behaviour domain to be measured."

Content validity is important from a washback point of view. It is clear that influence of examination on teaching and learning can't be avoided because examination requires the examinees to exhibit the ability to envisage in the course objective, the examination can be beneficial. An exam based on communicative tasks will encourage the students to use language by providing learning opportunity as well as tool for education change. "First, the greater a test's content validity, the most likely it is to be an accurate measure of what it is supposed to measure i.e. to have construct validity." (Hughes, 2010, p.27). He further says that a test in which major ideas identified in specification are under-represent or not represent at all-is likely to be accurate. As a result, such test can have harmful backwash. Areas which are not tested are likely to become areas ignored in teaching and learning. Too often, in our education system the content of test is determined by what is easy to test rather than what is important to test. The best safeguard against this is to write full test

specifications and to ensure that the test content is a fair reflection of these. Heaton (1975, p.154) says:

> Content validity is analysis of language being tested and of particular course objectives. The test should be so constructed as to contain a representative sample of the course, the relationship between the test items and the course objectives always being apparent.

Thus, content validity is also obtained through careful analysis of language items to be tested and objectives of the course.

1.1.7 Guidelines for Establishing Content Validity

A good test designer should analyze the behaviour domain to be tested systematically so that the test items in the correct proportions would cover all major aspects. The behaviour domain under consideration should be described in detail before designing the test, and the relevance of the responses made by the individual test to the behaviour area under consideration is the key while analyzing content validity. Khaniya (2005, p.104) "Content validity of test must reflect not only content of course but also demonstrate the balance of test items in terms of weighting to each unit or area." He further adds content validity of exam is examined also in relation to its relevance to given course study. This means whether or not the tasks included in exam are relevant to the language activities that are expected to be exercised under the given course. For it, we need a specification of skills or structures etc.

It is not to be expected that everything in the test specification will always appear in the test ; there may be too many things for all of them to appear in single test. But it will provide test constructor on the basis for making a principled selection of elements for inclusion in the test. A comparison of test specification and test content is the basis for judgement as to content validity. Ideally this judgement should be made by people who are familiar with

language teaching and testing but who are not directly concerned with the production of the test in question. Anastasi (1982, p.132 as cited in weir, 1990, p.25) provides following useful guidelines for establishing content validity:

(i) The behaviour domain to be tested must be systematically analyzed to make certain that all major aspects are covered by the test items, and in the correct proportion.

(ii) The domain under consideration should be fully described in advance, rather than being defined after the test has been prepared.

(iii) Content validity depends upon the relevance of the individual's test responses to the behaviour area under consideration rather than on the apparent relevance of item content.

a. Content Coverage

Comparison between course content and test content help to find out content validity. Content of the test items should represent the content of course under the prescribed objectives. Content coverage is the extent to which the tasks required in the test adequately represents the behavioural domain in questions. Fulcher & Davidson (2007, p.6) mentions, as cited in Fulcher (1999, pp.222-225) "In early approaches to communicative language testing the central issue in establishing content validity was how best to 'sample' from needs and target domain." Thus test in judged as having content validity when the test items represent the course contents and course objectives. The more test items are constructed, the more chances of having content validity. It is supposed to have content validity if all teaching units are covered in test.

1.2 Review of Related Literature

Review of literature deals with research studies on relevant proposition in the related area of the study so that all the past studies, their conclusions and deficiencies may be known and further research can be conducted. To this date

a number of research works have been carried out on validity under the department of English language Education, T.U. at M.Ed. level. Some of them as follows:

Gurung (2012) carried out a research entitled "Content Validity of Compulsory English Question Papers" to examine the content validity of question papers of compulsory English in terms of content coverage, component weighting and skill weighting. He used secondary sources of data for his study i.e. the question papers administered in annual examination of compulsory English at class 12 from 2063 to 2068 B.S. Finally, he concluded that the questions papers have low content validity in terms of content coverage and good content validity in terms of component weighting and skill weighting.

Sharma (2011) conducted a survey research on "Content Validity of Examination Papers: A case of Reading Writing and Critical Thinking Course." The objective of his study was to determine the validity of the year end examination of B.Ed. first year's course (Eng. Ed.317). For this, he collected and analyzed only subjective question papers administered from 2066 to 2067 B.S. using only secondary sources of data in terms of content coverage and content weighting. He concluded that the examination papers have good content validity in terms of content coverage but low in terms of weighting.

Khatri (2007) carried out a research on "The Content Validity of Translation: Theory and practice Exam Papers at M.Ed. Level." The study aimed to examine content validity of Translation: Theory and practice exam at M.Ed. level. He collected five years exam papers of Translation administered in annual exam in T.U. Finally, he concluded that the test papers have neither high nor low content validity in terms of content coverage and content weighting.

Timilsina (2007) studied a research on "Content Validity of Phonetics and Phonology Exam at M.Ed. Level" having the objective to examine the content validity question papers of Phonetics and Phonology at M.Ed. first year

administered during 2057 to 2062 B.S. in terms of content coverage and content weighting. He used the collection of six years question papers and analysed and interpreted on the basis of course language contents. Finally, he concluded that the Phonetics and Phonology tests administered during six years have high content Validity in terms of content coverage and low content validity in terms of content weighting.

Adhikari (2006) studied a survey research on "The Content Validity of the English Textbook for Grade IX." The objective of his study was to determine the content validity of English textbook of grade IX on the basis of content set out in curriculum and content in textbook using questionnaires for teachers, students of grade IX and subject experts. Finally, he concluded that content coverage language functions are valid but other aspects like selection and gradation, language skills, interest level and language structures are less valid.

Nepal (2006) carried out a research on "Content Validity of Examination: A case of Fundamentals of Language and Linguistics at B.Ed. Level." The objective of his study was to determine the content validity of Fundamentals of Language and Linguistics at B.Ed. level collecting six years question papers (from 2057 to 2062 B.S.) administered in annual exam of respective paper in T.U. He found out that the test papers have high content validity in terms of content coverage and low content validity in terms of weighting.

Ojha (2005) carried out a research on "Content Validity of ELT Theories and Methods Exam at B.Ed. Level." The objectives of his study was to examine the content validity of ELT Theories and Methods tests at B.Ed. second year in terms of content coverage and content weighting. He used the test papers administered in annual exam of ELT Theories and Methods from 2057 to 2061 B.S. Finally, he concluded that the test papers have good content validity in terms of content coverage and low content validity in terms of content weighting.

Khanal (1997) conducted a research on "A Study on the Effectiveness of Cloze Test over Conventional Objective Test in Testing Reading Comprehension in English." The objective of his study was to compare the effectiveness of the two types of tests viz. close test vs. objective test in testing reading comprehension using the two types of test. He found out that the private school students' performance in both objective and close test was better than public school students' performance and that the close test was far better than the objective test in testing reading comprehension.

Although many researches have been carried out on Validity, this study is different in the sense that no study has been carried out on "Content validity of Test papers: A case of SLA" in this department. In this way, it is a new venture in itself.

1.3 Objectives of the Study

The study had the following objectives:

   i. To examine the content validity of question papers of SLA at M.Ed. first year administered during 2066 to 2068 B.S. in terms of content coverage.

   ii. To suggest some pedagogical implications.

1.4 Significance of the Study

This study will be a distinct research work in the department of English language education because no research has been carried out yet in the present topic. So, it will be most importantly, useful for the department. Similarly, this study will provide information on whether the administered tests of second Language Acquisition have content validity or not. For the test designers, it will give some guidelines to make a test valid. This study will be beneficial to all who are concerned with language teaching and testing, more particularly teachers and the test designers who are involved in the teaching and testing of SLA at M.Ed. first year. It will also help to set valid question papers having content validity. If the teachers and test designers are familiar with test

construction process, they will be aware of the variations and mistakes hidden in constructing test and try to minimize them. It will be fruitful to make the test more valid. Likewise, this research work will be significant to everyone to get the idea for further research in this field.

CHAPTER TWO
METHODOLOGY

Methodology is an important part of a research work. So, it is designed in such a way that it helps to carry out the research more systematically and scientifically. Here, the researcher adopted the following methodology to fulfill the objectives of the study:

2.1 Sources of Data

The researcher used only secondary sources of data to conduct the study.

2.1.1 Primary sources of Data

The researcher did not use the primary sources of data for the study.

2.1.2 Secondary sources of Data

The researcher used the subjective and objective question papers administered in the annual examination of Second Language Acquisition at M.Ed. first year from 2066 to 2068 B.S. as the secondary sources of data. He also consulted different books on Second Language Acquisition, Language Testing, articles, research reports available in related topic such as Hughes (2010), Fulcher & Davidson (2007), Kumar (2005),Weir (1990), Bachman (2010), Heaton (1975), Nunan (1992), Khaniya (2005) and so on.

2.2 Sampling Procedure

The old question papers of Second Language Acquisition administered under T.U. were collected and sampled through judgmental non-random sampling procedure to meet the objectives of this study. For this, only three years question papers (2066-2068B.S.) were selected.

2.3 Tools for Data Collection

This study was totally based on the secondary sources. The researcher studied only three years question papers before analyzing and interpreting the facts. So, he used no tools other than the study of questions.

2.4 Process of Data Collection

The researcher collected the three years (2066-2068 B.S.) question papers on the subject entitled Second Language Acquisition at M.Ed. first year which were administered under T.U. containing both subjective and objective questions. Then he categorized all items in question papers unit-wise based on the course content in the syllabus. For the purpose of examining content coverage he matched the test content language items with the course content language items.

2.5 Limitations of the Study

Limitations of this study were as follows:

a. The area of study as indicated by the title was limited to a small portion of language testing viz. content validity.

b. The data was analyzed only in terms of content coverage.

c. The study was confined to only the question papers administered in the annual written examination of Second Language Acquisition at M.Ed. first year in T.U. from 2066-2068 B.S. for data collection to meet specified objectives.

d. Only charts, percentage and tabulations were used as the statistical tools for data analysis.

CHAPTER THREE

ANALYSIS AND INTERPRETATION

This chapter deals with analysis and interpretation of data. This is an important part of the research. It is concerned with the analysis and interpretation of collected raw data from secondary sources to examine and evaluate the content validity of Second Language Acquisition at M.Ed. first year administered under T.U. This chapter deals with the analysis of content validity of the Second Language Acquisition tests in terms of content coverage. With the view of making the study more objective and effective, the analysis and interpretation of the data was carried out using the statistical tools such as percentage, tabulation and charts.

3.1 Analysis and Interpretation of Content Coverage

This sub-chapter deals with the analysis and interpretation of content coverage of Second Language Acquisition question papers administered in annual examination under T.U. from 2066 to 2068 B.S. Both subjective and objective questions are analyzed in terms of content coverage. The content coverage is drawn by doing comparison between the course content items and test items. Here, the researcher examined the question papers of Second Language Acquisition at M.Ed. first year to find out whether test contents had representative sample of the course contents or not.

Content coverage should be considered while designing a test. A good test designer should prepare the test items as the representative sample of the course content as far as possible. It is assumed that above 50 percent of the coverage of content of test items is nearer to content validity. If it is below 50 percent, then it is supposed to low content validity. Likewise, if more than 60 percent course contents are represented in test papers, and then it is supposed to have high content validity.

3.1.1 Analysis of Content Coverage of SLA

The analysis of contents in syllabus and question papers has done by comparison between them. For it, the researcher tabulated as well as analyzed the data on the basis of the question papers administered under T.U. during 2066 B.S. to 2068 B.S. as follows:

Table No. 3.1

Coverage of Question Contents in terms of Course Contents.

S.N.	Course Contents	Test Contents		
		Test items represented		
		2066	2067	2068
1	Basic concept of SLA			
1.1	Introduction			
1.1.1	Definitions		SS^1	
1.1.2	The nature of language			
1.1.3	Views on the language learner			SS^1
1.1.4	The language learning process	SS^{2b}		
1.1.5	Interlanguage development			
1.1.6	Universal grammar and language learning		SS^1	SS^4
1.2	SLA and related discipline			
1.2.1	Third language acquisition/multilingualism			
1.2.2	Heritage language acquisition	O^1		
1.2.3	Bilingual acquisition			
1.2.4	First language acquisition			
1.3	History of SLA			

S.N.	Course Contents	Test Contents		
		Test items represented		
		2066	**2067**	**2068**
1.3.1	Behaviorism			
1.3.2	Contrastive analysis hypothesis	SS^1		
1.3.3	Error analysis	SS^1	O^1	O^2
1.3.4	Recent perspectives			
	Morpheme order studies			
	Role of native language			
1.4	Transfer and fossilization			O^1
2	Approaches and perspectives of SLA			
2.1	Cognitive approaches to SLA			
2.1.1	Processing approaches		$SL^6 O^3$	SS^3
2.1.2	Connectionism	O^3		
2.2	Functional/Pragmatic perspectives of SLA			
2.2.1	Early functionalist studies in SLA			
2.2.2	The aspect hypothesis	SS^{2a}		
2.3	Sociocultural theory of SLA			$SS^{2a} O^3$
2.3.1	Sociocultural theory and Vygotsky	O^2	$O^2 \ O^5$	SS^{2a}
2.3.2	Application of sociocultural theory of SLA			
2.4	Sociolinguistic perspective of SLA			
2.4.1	Variability in SLA		SS^3	$SL^6 \ O^4$
2.4.2	Second language socialization			

S.N.	Course Contents	Test Contents		
		Test items represented		
		2066	2067	2068
2.4.3	Community of practice and situated SLA			
2.4.4	Affect and investment in SLA			
2.5	Interlanguage and social context			
2.5.1	Variation			
2.5.2	Social interactional approaches			
2.5.3	Communication strategies	SS^{5a}		
2.5.4	Interlanguage pragmatics			
3	Input, interaction and output in SLA			
3.1	Input in SLA (Krashen's hypothesis)		O^4	
3.2	Interaction in SLA (Long's interaction hypothesis)		SS^{2a}	SS^{2b} O^6
3.3	Rethinking output hypothesis			
3.4	Output in SLA (Swain's output hypothesis, Gass's ideas)	SS^3	SS^{2b}	
3.5	Feedback, recasts and negative evidence	O^4	O^6	
3.6	Attention and consciousness raising			SS^4 O^5
3.7	The role of input and interaction in language learning	SL^6		
4	Instructed SLL	O^5		
4.1	Classroom language	O^6	SS^{5a}	

S.N.	Course Contents	Test Contents		
		Test items represented		
		2066	2067	2068
4.2	Processing instruction	SS^4	SS^{5b}	O^7
4.3	Teachability/learnability			
4.4	Focus on form			
4.5	Effectiveness of instruction		O^7	
5	Non-language factors in SLA			
5.1	Affect			SS^5
5.2	Social distance			
5.3	Age difference	SS^{5b}	SS^4	
5.4	Aptitude	O^7	O^8	
5.5	Motivation			
5.6	Personality and learning style	O^8		O^8
5.7	Learning strategies	SS^{5a}		

Note:

SS^1= Subjective short question no. 1

SS^{2a}= Subjective short question no. 2a which is optional

SL^6= Subjective long questions no. 6

O^1=Objective question no. 1

3.1.1.1 Coverage of Contents in Unit One

The table 3.1 shows that there are altogether 5 units. Each unit contains many sub-units. If we see diachronically, there are 21 language items spread over 1 to 1.4 in unit one. Among 21 language items, only 8 language items were represented during the period of there years (2066 to 2068). The most

represented one is 1.3.3 'Error analysis' which was represented in each and every years' examination from 2066 to 2068. 1.1.6 'Universal grammar and language learning' was represented in two years. The others 1.1.1 'Definitions', 1.1.3 'Views on the language learner', 1.1.4 'The language learning process', 1.2.2 'Heritage language acquisition', 1.3.2 'contrastive analysis hypothesis' and 1.4 'Transfer and fossilization' were represented in only one year.

If we see synchronically, in 2066, one objective question and two subjective short questions were asked from unit one. In case of objective questions, only one objective question (O^1) was asked from 1.2.2 'Heritage language acquisition'. In case of subjective questions, the first subjective short question (SS^{2b}) was from 1.1.4 'The language learning process' in optional position and the second subjective short questions (SS^1) was collectively asked from 1.3.2 'Contrastive analysis hypothesis' and 1.3.3 'Error analysis.'

In 2067, only one objective question and one subjective short question were asked from this unit. In case of objective question, (O^1) was asked from 1.3.3 'Error analysis'. In case of subjective question, the only one subjective short question (SS^1) was collectively asked from 1.1.1 'Definitions' and 1.1.6 'Universal grammar and language learning.'

In 2068, two objective questions and two subjective short questions were asked from this unit. In case of objective questions, the first objective question (O^1) was asked from 1.4 'Transfer and fossilisation' and the second (O^2) was asked from 1.1.3 'Error analysis'. In case of subjective questions, the first subjective short question (SS^1) was asked from 1.1.3 'Views on the language learner' and the Second (SS^4) was asked from 1.1.6 'Universal grammar and language learning.'

Finally, the abovementioned table and descriptions can be clear from the following pie chart

Figure No. 3.1 Coverage of contents in unit one
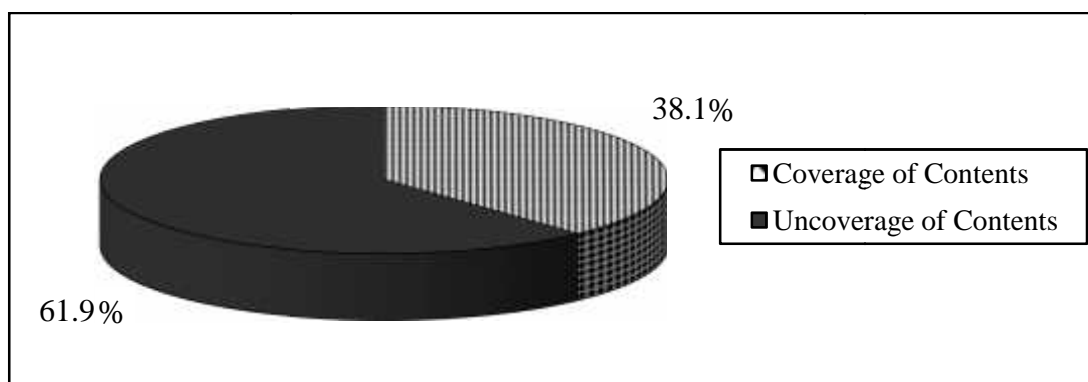
The above pie chart shows that there are altogether 21 language items but only 8 language items were represented during three years. Other 13 language items were neglected. It means the coverage of content of unit one was 38.1% and 61.9% contents were not represented. That is to say, the content validity of the test papers in unit one was low and unsatisfactory.

3.1.1.2 Coverage of Contents in Unit Two

The table 3.1 shows that there are 20 language items in unit two. Among 20 language items, only 7 language items were represented within three years i.e. 2066 to 2068. Other language items were neglected in this unit. The most represented one is 2.3.1 'Sociocultural theory and Vygotsky' which was represented in each and every years' examination from 2066 to 2068. Similarly, 2.1.1 'Processing approaches' and 2.4.1 'Variability in SLA' were represented in two years. The others 2.1.2 'Connectionism', 2.2.2 'The aspect hypothesis', 2.3 'Sociocultural theory of SLA' and 2.5.3 'Communication strategies' were represented in only one year and other language items had been neglected.

The question papers can be analyzed synchronically here. In 2066, two objective questions and two subjective short questions were asked from unit two. In case of objective questions, the first objective question (O^2) was asked from 2.3.1 'Sociocultural theory and Vygotsky' and the second (O^3) was asked from 2.1.2 'Connectionism'. In case of subjective questions, the first subjective short question (SS^{2a}) was asked from 2.2.2 'The aspect hypothesis' and the

second subjective short question (SS[5a]) was asked from 2.5.3 'Communication strategies' in optional positions.

In 2067, three objective questions, one subjective short question and one subjective long question were asked from this unit. The first objective question (O[3]) was asked from 2.1.1 'Processing approaches' and the second (O[2]) and the third (O[5]) were asked from 2.3.1 'Socio cultural theory and Vygotsky', Likewise, the subjective short question (SS[3]) was from 2.4.1 'Variability in SLA' and the subjective long question (SL[6]) was asked from 2.1.1 'Processing approaches'.

In 2068, two objective questions, two subjective short questions and one subjective long question were asked from this unit. The first objective question (O[3]) was asked from 2.3 'Sociocultural theory of SLA' and the second (O[4]) was asked from 2.4.1 'Variability in SLA'. Similarly, the first subjective short question (SS[3]) was asked from 2.4.1 'Variability in SLA' and the second (SS[2a]) was collectively asked from 2.3 'Sociocultural theory of SLA' and 2.3.1 'Sociocultural theory and Vygotsky' in optional position. The subjective long question (SL[6]) was asked from 2.4.1 'Variability in SLA'.

To be clear, take the help of pie chart.

Figure No. 3.2 Coverage of Contents in Unit Two
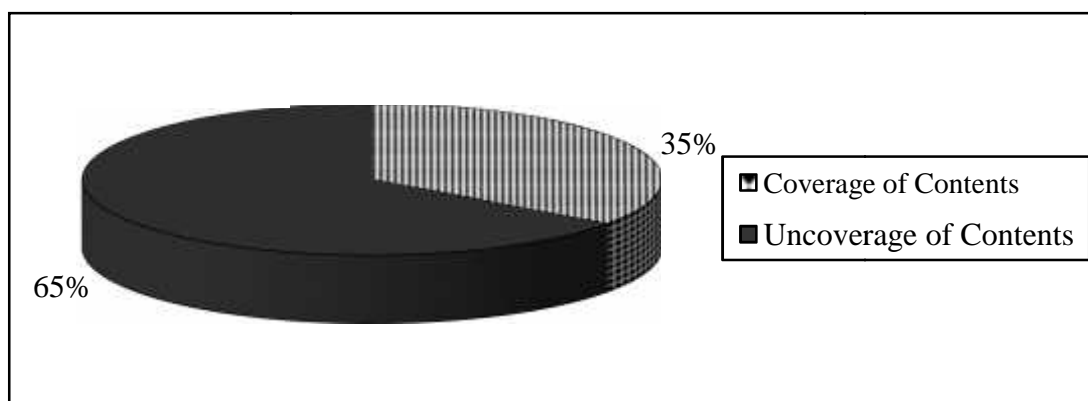
From the above table and description, it is clear that there are altogether 20 language items in unit two spread over 2 to 2.5.4 according to the course contents. However, only seven test items were asked from this topic during

three years (2066 to 2068) and other 13 language items were neglected. This shows that the content coverage of test items from unit two was 35% whereas 65% of content was not included in question papers. That's why, it can be said that the content validity in unit two was extremely low.

3.1.1.3 Coverage of Contents in Unit Three

In unit three, there are altogether 8 language items. Among 8 language items, 6 language items were represented during three years question papers (2066 to 2068). The most represented language items of this unit are 3.2 'Interaction in SLA (Long's interaction hypothesis)', 3.4 'Output in SLA (Swains output hypothesis, Gass's ideas)' and 3.5 'Feedback, recasts and negative evidence' which were represented in two years. 3.1 'Input in SLA (Krashen's hypothesis)', 3.6 'Attention and consciousness raising' and 3.7 'The role of input and interaction in language learning', were represented in only one year. Only two language items were neglected in this unit.

If we see synchronically, in 2066, only one objective question, one subjective short question and one subjective long question were asked from unit three. The objective question (O^4) was asked from 3.5 'Feedback, recasts and negative evidence.' The subjective short question (SS^3) was asked from 3.4 'Output in SLA (Swain's output hypothesis, Gass's ideas)' and the subjective long question (SL^6) was asked from 3.7 'The role of input and interaction in language learning.'

In 2067, two objective questions and two subjective short questions in optional positions were asked from this unit. In case of objective questions, the first question (O^4) was asked from 3.1 'Input in SLA (Krashen's hypothesis)' and the second objective question (O^6) was asked from 3.5 'Feedback, recasts and negative evidence '. In case of subjective questions, the first subjective short question (SS^{2a}) was asked from 3.2 'Interaction in SLA (Long's interaction hypothesis)' and the Second (SS^{2b}) was asked from 3.4 'Output in SLA (Swain's output hypothesis, Gass's ideas)' in optional positions.

In 2068, two objective questions and two subjective short questions were asked from this unit. In case of objective questions, the first objective question (O^6) was asked from 3.2 'Interaction in SLA (Long's interaction hypothesis)' and the second objective question (O^5) was asked from 3.6 'Attention and consciousness raising'. In case of subjective questions, the first subjective short question (SS^{2b}) was asked from 3.2 'Interaction in SLA (Long's interaction hypothesis)' in optional position and the second subjective short question (SS^4) was asked from 3.6. 'Attention and consciousness raising'.

In conclusion, the abovementioned table and description can be shown in the following pie chart.

Figure No. 3.3 Coverage of Contents in Unit Three
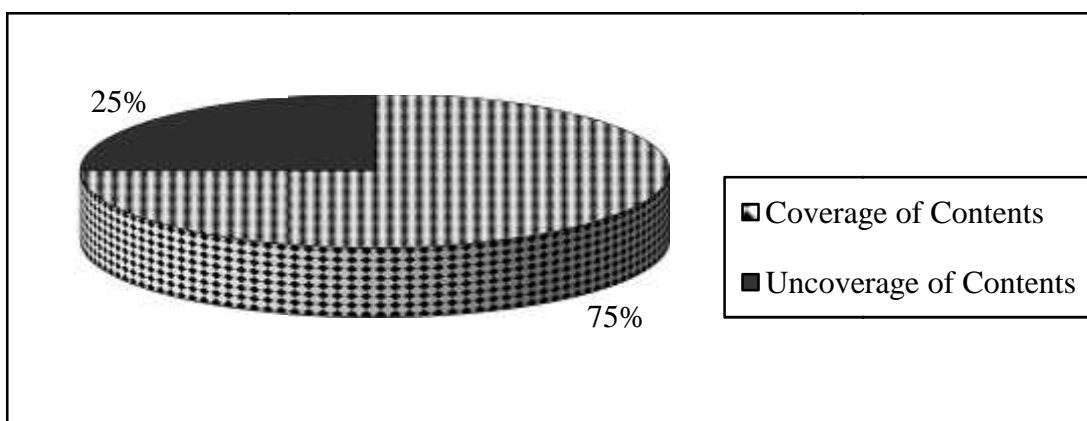
The pie chart shows that there are 8 language items in unit three over 3 to 3.7. Out of 8 language items, 6 language items were represented during three years. Two language items were neglected while constructing test items. It means the coverage of course items in test contents in unit three was 75%. Likewise, 25% contents were not covered in the test papers. That is to say, the content validity of test papers in unit three was highly satisfactory.

3.1.1.4 Coverage of Contents in Unit Four

The table 3.1 shows that there are altogether 6 language items spread over 4 to 4.5 in unit four. Out of 6 language items, 4 language items were represented during the period of three years (2066 to 2068). Among them, the most

represented one is 4.2 'Processing instruction' which was represented in each and every years' examination from 2066 to 2068 B.S. Similarly, 4.1 'Classroom language' was represented in two years. The remaining two language items i.e. 4 'Instructed SLL' and 4.5 'Effectiveness of instruction' were represented in only one year. In this unit, only two language items had been neglected.

If we see synchronically, in 2066, two objective questions and one subjective short question were asked from this unit four, In case of objective questions, the first objective question (0^5) was asked from 4 'Instructed SLL' and the second (0^6) was asked from 4.1 'Classroom language '. In case of subjective question, the only one subjective short question (SS^4) was asked from 4.2 'Processing instruction'.

In 2067, one objective question and two subjective short questions were asked from this unit. The objective question (0^7) was asked from 2.4 'Processing instruction'. The first subjective short question (SS^{5a}) was asked from 4.1 'Classroom language' and the second (SS^{5b}) was from 4.2 'Processing instruction'. Both subjective short questions were optional in position.

In 2068, no subjective question was asked from this unit. Only one objective question (0^7) was asked which was from 4.2 'Processing instruction'.

Finally, it can be made clear using the following pie chart.

Figure No. 3.4 Coverage of Contents in Unit Four
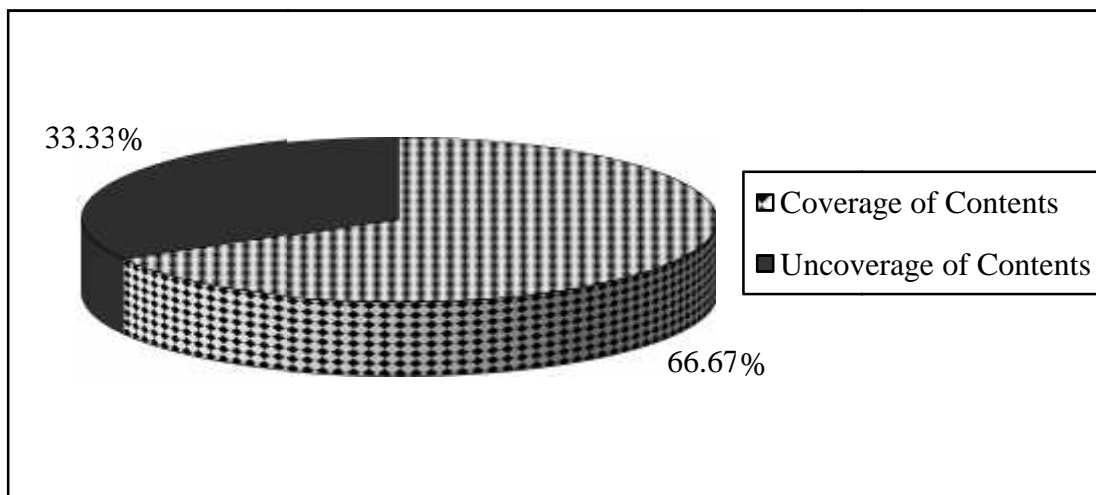
From the above description and chart, it is clear that there are 6 language items in unit four spread over 4 to 4.5 according to the course contents. The represented test items were 4 language items during three years. Only two language items were neglected. It means that the coverage of contents in test contents in unit four was 66.67% and 33.33% was not included. That's why, the content validity in unit four was highly satisfactory.

3.1.1.5 Coverage of Contents in Unit Five

In unit five, the table 3.1 shows that there are altogether 8 language items spread from 5 to 5.7. Among 8 language items, 5 languages items were represented during the period of three years (2066 to 2068). Among them, the most represented ones were 5.3 'Age difference', 5.4 'Aptitude' and 5.6 'Personality and learning style'. The others 5.1 'Affect' and 5.7 'Learning strategies' were represented in only one year. The remaining 3 language items 5 'Non–language factors in SLA', 5.2 'Social distance' and 5.5 'Motivation' were neglected in this unit.

If we see synchronically, in 2066, two objective questions and two subjective short questions were asked from unit five. In case of objective questions, the first objective question (0^7) was asked from 5.4 'Aptitude' and the second objective question (0^8) was asked from 5.6 'Personality and learning style'. In case of subjective questions, the first subjective short question (SS^{5a}) was asked from 5.7 'Learning strategies' and the second subjective short question (SS^{5b}) was asked from 5.3 'Age difference'. Both subjective short questions were optional in nature.

In 2067, one objective question and one subjective short question were asked from this unit. In case of objective question, the objective question (0^8) was asked from 5.4 'Aptitude'. In case of subjective question, the subjective short question (SS^4) was asked from 5.3 'Age difference'.

In 2068, one objective question and one subjective short question were asked from this unit. The objective question (0^8) was asked from 5.6 'Personality and learning style'. Similarly, the subjective short question (SS^5) was asked from 5.1 'Affect'.

In conclusion, the abovementioned table and description can be shown in the following pie chart.

Figure No. 3.5 Coverage of Contents in Unit Five
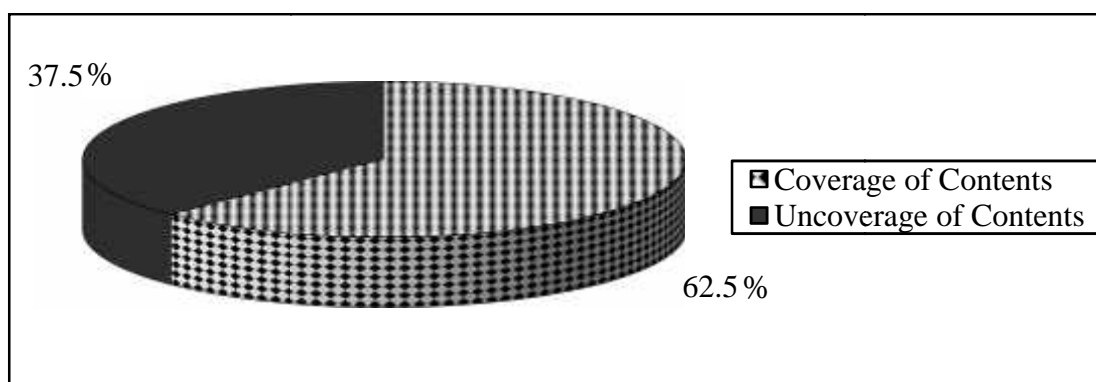
From the above table and description, it is clear that there are altogether 8 language items. Out of 8 language items, 5 language items were represented within three years (2066 to 2068). Three language items were neglected while constructing test items. It means that the coverage of course contents in test contents in unit five was 62.5%. Similarly, 37.5% contents were not covered in the test papers. In this way, the content validity of the test papers in unit five was good and satisfactory.

3.1.2 Repeated Test Items

After analyzing the above table, the researcher found that from unit one within three years (2066–2068) one subjective question and one objective question were repeated or same. In case of subjective questions, the subjective short question (SS^1) of 2067 was the same as the question (SS^4) of 2068. They were asked from 1.1.6 'Universal grammar and language learning'. Likewise, the objective question (0^1) of 2067 was the same as the (O^2) of 2068 which were from 1.3.3 'Error analysis'.

In unit two, the subjective short question (SS^3) of 2067 was partially similar to the subjective long question (SL^6) of 2068. They were from 2.4.1 'Variability in SLA'. The objective question (O^2) of 2066 was the same as the question (O^2) of 2067 which were from 2.3.1 'Socio-cultural theory and Vygotsky'.

In unit three, no objective test item was repeated. But in case of subjective test items the researcher found that the subjective short question (SS^{2a}) of 2067 was partially similar to the subjective short question (SS^{2b}) of 2068. They were asked from 3.2 'Interaction in SLA (Long's interaction hypothesis)'. Like this, the subjective short question (SS^3) of 2066 was exactly the same as the subjective short question (SS^{2b}) of 2067 which were asked from 3.4 'Output in SLA (Swain's output hypothesis, Gass's ideas)'.

In unit four, no questions were repeated during three years (2066–2068).

In unit five, no objective test item was repeated. But in case of subjective questions, the subjective short question (SS^{5b}) of 2066 was similar to the subjective short question (SS^4) of 2067 which were asked from 5.3 'Age difference'.

3.1.3 Examining Content Validity of the Test Papers on the Whole in terms of Content Coverage
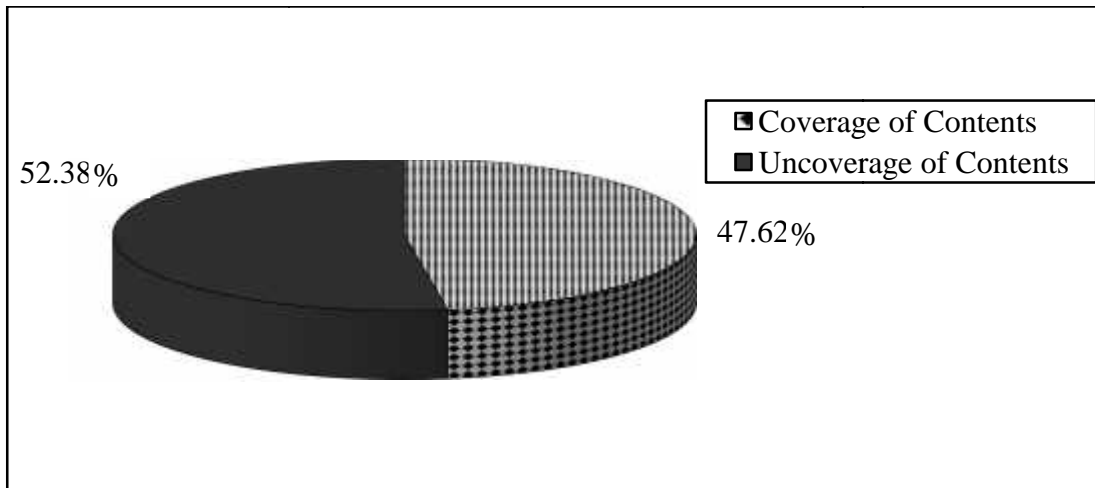
Table No. 3.2

Examining Content Validity of the Test Papers on the Whole in terms of Content Coverage

S.N.	Units	Course Contents language items	Test Contents language items	Test Coverage in percentage
1.	1	21	8	38.1%
2.	2	20	7	35%
3.	3	8	6	75%
4.	4	6	4	66.67%
5.	5	8	5	62.5%
Total		63	30	47.62%

The above table shows that there are altogether 63 language items within five units in the Second Language Acquisition course at M.Ed. first year. Out of 63 language items, the representation of the test content language items were 30 during three years (2066 to 2068). Similarly, 33 language items were neglected while constructing the test items. It means the coverage of course contents in test contents on the whole was 47.62% whereas 52.38% of the course contents was not covered in the test papers.

In Conclusion, the above presented table and description can be shown in the following pie chart.

Figure No. 3.6 Coverage of Contents on the Whole

After analyzing the above table and description, the researcher concludes that the content validity of the test papers on the whole in terms of content coverage is low and unsatisfactory as it covers only 47.62% of the course content language items as a whole. In other words, the test papers haven't really tested what they have claimed to test on the part of the testees and they can't cover the representative sample of the course contents. So, the test papers of SLA have low content validity in terms of coverage during three years (2066 to 2068).

CHAPTER FOUR

FINDINGS AND RECOMMENDATIONS

This is the final chapter of this research work. The major focus of this research was to find out the content validity of test papers of Second Language Acquisition at M.Ed. I year. In order to fulfill the objectives, the researcher collected and analyzed the test papers on the subject entitled Second Language Acquisition at M.Ed. first year which were administered under T.U. containing both subjective and objective questions within three years (2066–2068 B.S.).

This chapter is basically subdivided into two parts where the first part deals with the findings made by the researcher after analyzing the data and the second part deals with the recommendations made on the basis of the collected and analyzed data and the findings. At first, the findings of the presented study are summarized in the following points.

4.1 Findings

The following findings have been listed on the basis of the analysis and interpretation of the data:

4.1.1 General Findings

a. In terms of the content coverage, the Second Language Acquisition test papers have low content validity because out of 63 language items in totality of the course, only 30 language items (i.e. 47.62%) were represented within three years question papers (2066–2068).

4.1.2 Unit Specific Findings

The unit specific findings of this research work are as follows:

1. The content validity of question papers in terms of content coverage and its findings is given below:

a. In unit one, there are altogether 21 language items but only 8 language items were represented during three years. It means the coverage of course contents in question contents is 38.1%, which shows that the content validity of the unit one has low and unsatisfactory.

b. In unit two, there are altogether 20 language items out of which only 7 language items were represented during three years (2066–2068). It means, the coverage of course contents in this unit is 35%. Therefore, unit two has extremely low content validity.

c. In unit three, there are altogether 8 language items. Among them 6 language items were represented during three years. This means, the coverage of course contents in unit three is 75%. Hence, unit three has high content validity.

d. In unit four, there are altogether 6 language items where 4 language items were represented during three years. It means, the coverage of the course contents in unit four is 66.67%. Thus, unit four has also good content validity.

e. In unit five, there are altogether 8 language items out of which 5 language items were represented within three years. It means, the coverage of course contents in unit five is 62.5%. Therefore, unit five has also good content validity.

2. On the whole, it has been found that Second Language Acquisition test papers have low content validity in terms of content coverage.

3. It has also been found that some language items e.g. 1.3.3 'Error analysis', 2.3.1 'Sociocultural theory and Vygotsky' etc. were represented in all three years and some language items were neglected e.g. 2.1 'Cognitive approaches to SLA', 2.5.1 'Variation', 5.2 'Social distance', etc.

4. Likewise, it has been found that there is variation in asking all types of questions (objective, subjective short, subjective long) in different years.

5. Some questions were repeated during the period of three years. For example, the subjective short question (SS^1) of 2067 was the same as the subjective short question (SS^4) of 2068, the objective question (O^2) of 2066 was the same as the objective question (O^2) of 2067. Likewise, the subjective short question (SS^{2a}) of 2067 was partially similar to the subjective short question (SS^{2b}) of 2068 and the subjective short question (SS^{5b}) of 2066 was similar to the subjective short question (SS^4) of 2067 and so on. Such kind of repetition of the same test items encourages guessing and reduces content coverage and which ultimately reduces the content validity of the test papers.

4.2 Recommendations

On the basis of findings, the following recommendations have been suggested for pedagogical implications:

1. As the Second Language Acquisition had low content validity in test papers in terms of content coverage, the test designers should design the question papers ensuring more content representation as far as possible.

2. It was found that some units and language items were given more emphasis and some others were neglected in designing the tests. Therefore, the test setters should give equal emphasis to all the units and language items. The test items should be represented widely from the whole area of contents.

3. It was found that repetitions of test items were made during the period of three years. Such repetitions encourage guessing and reduce content validity of the test papers. In this way, the repetition of the same test items time and again should be primarily avoided and discouraged. Instead of repeating the same test items, the priority should be given to other language items so as to increase content validity.

4. There should be fixed criteria on how many and what types of test items to be asked from each unit in each year and so on. For this, test specification should be prepared before designing the test and should strictly follow this.

5. The office of the controller of examinations T.U. has the sole authority in conducting examinations. So, it should play vital role in improving the examination of Second Language Acquisition as a whole. To obtain this aim, the concerned authority should organize seminars, workshops, conferences and so on.

6. Trained and highly experienced test designers, teachers, stakeholders should design the test papers. If feasible the question should be piloted to achieve good content validity.

7. To achieve high content validity in terms of content coverage, the semester system should be applied.

8. The concerned authority should provide adequate orientation and trainings to the test designers and should consider their problems occurred while designing the test.

Finally, the researcher wants to request the concerned authority to take the abovementioned recommendations into consideration so as to obtain good content validity in Second Language Acquisition.

REFERENCES

Adhikari, R.P.(2006). *The content validity of the English text book for grade IX.* An unpublished M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Bachman, L.& Palmer, A.(1996). *Language testing in practice.* Oxford: OUP.

Bhattarai, G.R.(1994). *An introduction to linguistics.* Kathmandu: Ratna Pustak Bhandar.

Bhattarai, G.R.(2003). *A thematic analysis of research reports.* Kathmandu: Ratna Pustak Bhandar.

Brown, F.G.(1976). *Principles and educational and psychological testing.* New York*:* Holt, Rinehart and Winston.

Fulcher, G. & Davidson, F.(2007). *Language testing and assessment.* London and New York: Routledge.

Gurung, D.B. (2012). *Content validity of compulsory English question papers.* An unpublished M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Heaton, J.B. (1975). *Writing English language tests.* London: Longman.

Hughes, A. (2010). *Testing for language teachers.* Cambridge: CUP.

Khanal, K.P. (1997). *A study on effectiveness of the cloze test over conventional objective tests.* An unpublished, M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Khaniya, T.R. (2005). *Examination for enhanced learning.* Lalitpur: Millennium Publication.

Khatri, R. (2007). *The content validity of translation: Theory and practice exam at M.Ed. level.* An unpublished M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Kumar, R.(2005). *Research methodology.* India: Porling Kindersley.

Nepal, P. (2006). *The content validity of examination: A case of fundamentals of language and linguistics at B.Ed. level.* An unpublished M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Nunan, D. (1992). *Research methods in language learning.* Cambridge: CUP.

Ojha, D.R. (2005). *Content validity of ELT theories and methods at B.Ed. level.* An unpublished M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Sharma, S.D. (2011). *Content validity of examination papers: A case of reading writing and critical thinking course.* An unpublished M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Timilsina, T.P. (2007). *Content validity of phonetics and phonology exam at M.Ed. level.* An unpublished M.Ed. thesis, Tribhuvan University. Kathmandu, Nepal.

Weir, C.J. (1990). *Communicative language testing*. Hemel Hempstead: Prentice Hall.

APPENDIX-A
LIST OF SYLLABUS

Second Language Acquisition

Course title: Second Language Acquisition   Full Marks: 50

Course No: Eng.Ed. 519   Pass Marks: 20

Nature of Course: Theory   Period per week: 3

Level: M.Ed.   Total periods: 75

Year: First   Total per periods: 55 minutes

1. Course Description

This is an introductory course on Second Language Acquisition (SLA). It is divided into five units. The first unit deals with the basic concepts of SLA including its history. The second unit incorporates various approaches and perspectives of SLA . The third unit deals with the role of input, interaction and output. The fourth and fifth units include the topics like instructed second language learning and non-language factors in SLA respectively.

2. Course Objectives

The general objectives of this course are as follows:

- To acquaint the students with the basic concepts of SLA including its history.
- To familiarise them with various approaches and perspectives of SLA.
- To acquaint them with the role of input and interaction to yield desired output.
- To make them familiar with the role of instruction in SLA.
- To help the students to learn the roles of non-language factors in SLA.

3. Specific Objectives and Contents

Specific objectives	Contents
- Define SLA and state various disciplines related to it. - State the history of SLA. - Explain the role of transfer and fossilization in SLA	**Unit I: Basic Concepts of SLA (20)** 1.1 Introduction 　　1.1.1 Definitions 　　1.1.2 The nature of language 　　1.1.3 Views on the language learner 　　1.1.4 The language learning process 　　1.1.5 Interlanguage development 　　1.1.6 Universal grammar and language learning 1.2 SLA and Related Discipline 　　1.2.1 Third Language acquisition \multilingualism 　　1.2.2 Heritage Language Acquisition 　　1.2.3 Bilingual Acquisition 　　1.2.4 First Language Acquisition 1.3 History of SLA 　　1.3.1 Behaviourism

	1.3.2 Contrastive analysis hypothesis
	1.3.3 Error analysis
	1.3.4 Recent perspectives
	    ⌡ Morpheme order studies
	    ⌡ Role of the native language
	1.4 Transfer and fossilisation

⌡ Overview different approaches to and perspectives of SLA
⌡ Carry out small studies on interlanguage development
	Unit II: Approaches and Perspectives of SLA (20)
	2.1 Cognitive approaches to SLA
	    2.1.1 Processing approaches
	    2.1.2 Connectionism
	2.2 Functional\pragmatic perspectives of SLA
	    2.2.1 Early functionalist studies in SLA
	    2.2.2 The aspect hypothesis
	2.3 Sociocultural theory of SLA
	    2.3.1 Sociocultural Theory and Vygotsky
	    2.3.2 Application of sociocultural theory to SLA
	2.4 Sociolinguistic perspective of SLA
	    2.4.1 Variability in SLA
	    2.4.2 Second Language Sociolisation
	    2.4.3 Community of practice and situated SLA
	    2.4.4 Affect and investment in SLA
	2.5 Interlanguage and social context
	    2.5.1 Variation
	    2.5.2 Social interaction approaches
	    2.5.3 Communication strategies
	    2.5.4 Interlanguage pragmatics

⌡ Overview of various hypothesis and models of SLA. ⌡ Highlight the role of input, interacton, feedback, recast and negative evidence in SLA.	**Unit III: Input, Interaction and Output in SLA (15)** 3.1 Input in SLA (Krashan's Hypothesis) 3.2 Interaction in SLA (Long's interaction hypothesis) 3.3 Rethinking output hypothesis 3.4 Output in SLA (Swain's output hypothesis, Gass's ideas) 3.5 Feedback, recasts and negative evidence 3.6 Attention and consciousness raising 3.7 The role of input and interaction in language learning
⌡ Mention the importance of classroom instruction in the development of SLA	**Unit IV: Instructed Second Language Learning (10)** 4.1 Classroom language 4.2 Processing instruction

	4.3 Teachability\Learnability
	4.4 Focus on form
	4.5 Effectiveness of instruction
⌇ Explain the role of non-language factors in SLA. ⌇ Facilitate SLA process keeping in mind the various non-linguistic factors.	**Unit V: Non-Language Factors in SLA** **(10)** 5.1 Affect 5.2 Social distance 5.3 Age difference 5.4 Aptitude 5.5 motivation 5.6 Personality and learning style 5.7 Learning strategies

Note: The figures in the parentheses indicate the approximate periods for the respective units.

4. Instructional Techniques

The instructional techniques for this course are divided into two groups. The first group consists of general instructional techniques applicable to most of the units. The second group consists of specific instructional techniques applicable to specific units.

4.1 General Instructional techniques

⌇ Lecture and discussion
⌇ Demonstration
⌇ presentation
⌇ self-study

4.2 Specific Instructional techniques

Unit III: Conducting Study\Research

Unit II & IV: Project Work

Unit V: Individual and Group work

5. Evaluation

This is a theoretical course. Hence, the learning of the students will be assessed through the annual examinations held by the Office of the Controller of examinations. The types and the number of questions in the annual examination paper are given in the following table:

Type of questions	Total questions to be asked	Number of questions to be answered and marks allocated	Total marks
Group A: Multiple choice items	8 questions	8×1 marks	8
Group B: Short answer questions	5 with 2 'or' questions	5×6 marks	30
Group C: Long answer questions	1 question	1×12 marks	12

6. Recommended Books and References

Recommended Books

Gass, S.& Selinker, L. (2009). *Second language acquisition: An introductory course*. New York: Routledge. (For units I to V)

Cook, V.(2008). *Second language learning and language teaching*. London: Arnold. (For unit I)

Ellis, R. (1995). *The study of second language acquisition*. Oxford: OUP. (For units I to V)

Ellis,R. (1986). *Understanding second language acquisition*. Oxford: OUP. (for unit I)

Mitchell, R.& Myles, F. (2004). *Second language learning theories.* Britain: Hodder Arnold. (For units I to III)

References

Ellis,R. (1992). *Instructed second language acquisition*. Cornwall: Blackwell

Ellis,R. (2000). *Second language acquisition*. Oxford: OUP

Chaudron,C. (1998). *Second language classroom research*. Cambridge: CUP.

Doughty, C.J.& Long, M.H. (Eds.). (2003). *The handbook of second language acquisition*. New York: Blackwell.

Gass, S.M. & Schachter,J.(Eds.). (1990). *Linguistic perspectives on second language acquisition*. Cambridge: CUP.

Saville, T.M. (2006). *Second language acquisition*. Cambridge: CUP.

APPENDIX-B
LIST OF QUESTIONS

TRIBHUVAN UNIVERSITY
2066

Master Level /I Year/EDUCATION   Roll No.: ....
Second Language Acquisition (Eng.519)

Group "A"   8
Attempt ALL the questions.   Tick (√) the best answers.

1. Which of the following statements does not characterize heritage language acquisition?
   a. It is a form of SLA.
   b. It is the historical and personal connection.
   c. Heritage language learners form a heterogeneous group.
   d. It is the dominant language.

2. Which of the following is not a type of private speech?
   a. repetition   b. vicarious responses
   c. manipulation   d. inner speech

3. Emergentism in SLA is also known as
   a. connectionism   b. functionalism
   c. constructivism   d. all of the above

4. A NS's reformulation of a NNS's incorrect utterance that maintains the original meaning of the utterance is called
   a. recast   b. prompt
   c. negotiation   d. pushed output

5. Which of the following statements is not true of instructed SLA?
   a. Formal instruction appears to have no major effect on the route of SLA
   b. Formal instruction does have relative utility in the rate of SLA.
   c. Teachers typically ask a lot of questions.
   d. Learners' errors are always picked up.

6. What are the sources of input in classroom SLA?
   a. teacher   b. materials
   c. other learners   d. classroom

7. Which of the following is not a component of language aptitude?
   a. phonemic coding ability   b. grammatical sensitivity
   c. deductive language learning ability   d. memory

8. Which of the following statements indicates a kinaesthetic perceptual modality?
   a. I need oral direction for a task.
   b. I like to listen to music when I study.
   c. I think better when I move around.
   d. Charts and diagrams help me understand what someone says.

TRIBHUVAN UNIVERSITY
2066

Master Level /I Year/EDUCATION   Full Marks: 50
Second Language Acquisition (Eng.519)   Time: 3 hrs.

Candidates are required to give their answers in their own words as far as practicable.
The figures in the margin indicate full marks.

Attempt ALL the questions.

Group "B"   5×6=30

1. How has SLA emerged with the practice of Contrastive Analysis and Error Analysis?

2. Briefly discuss the 'Aspect Hypothesis' in SLA.
   OR
   How do L₂ learners self-regulate their language learning process?

3. Explain the 'Output Hypothesis' after Swain.

4. How do garden path studies link the relationship between corrective feedback and SLA?

5. Distinguish between communication strategies and learning strategies.
   OR
   Argue for or against the 'critical period hypothesis' in SLA.

Group "C"   1×12=12

6. What is the role of Input and Interaction in SLA?

TRIBHUVAN UNIVERSITY
2067

Master Level /I Year/EDUCATION

Second Language Acquisition (Eng.519)

Full Marks: 50

Time: 3 hrs.

Candidates are required to give their answers in their own words as far as practicable.

The figures in the margin indicate full marks.

Attempt ALL the questions.

Group "B"
5×6=30

1. Define SLA, and briefly explain the role of UG (Universal Grammar) in SLA.

2. Briefly discuss Long's 'interaction hypothesis' in SLA.

   OR

   How does Swain define 'comprehensible output'? Explain the role of the 'comprehensible output' in SLA.

3. What do you mean by 'variability in SLA'? What are three different perspectives of explanation for variability?

4. Do you think 'age' has influential role in L_2 acquisition? Elaborate your answer in the light of critical period hypothesis.

5. What are the distinctive features of teacher talk? Discuss them in brief.

   OR

   Explain the role of 'processing instruction' in L_2 acquisition.

Group "C"
1×12=12

6. Discuss in brief McLanglin's IP Model and Anderson's ACT Model of SLA.

Group "A"
8

Attempt ALL the questions.   Tick ( ) the best answers.

1. __ Compares between the L_2 interlanguage and the standard L_2 forms
   - (a) Contrastive analysis
   - (b) Error analysis
   - (c) UG theory
   - (d) Discourse analysis

2. Which of the following is the most important for the regularity of learners inner thought?
   - (a) Private speech
   - (b) ZPD
   - (c) Scaffolding
   - (d) Learner attitude

3. According to ____ L_2 acquisition is viewed as the movement from controlled processing to automatic processing
   - (a) McLaughlin's IP model
   - (b) Anderson's ACT model
   - (c) Anderson Aspect Hypothesis
   - (d) Vygotsky's sociocultural theory

4. 'i+1' is the symbol that is used to refer to ____
   - (a) Comprehensible Output
   - (b) Comprehensible input
   - (c) explicit feedback
   - (d) selective attention

5. The process of supportive dialogue which directs the attention of the learner towards the key features of the target language has come to be known_____
   - (a) Private speech
   - (b) Self regulation
   - (c) Scaffolding
   - (d) Mediation

6. From _____ the L_2 learner may know and realize the area of errors that they notice in their L_2 performance
   - (a) Positive evidence
   - (b) corrective feedback
   - (c) Comprehensive input
   - (d) Negative evidence

7. Which of the following is not true about implicit instruction?
   - (a) It involves treatment for a long period of time
   - (b) It minimizes explanation
   - (c) The example and illustration follow the explanation
   - (d) It engages learners in real and real like situation

8. _____ refers to the ability to acquire an additional language
   - (a) Attitude
   - (b) Aptitude
   - (c) Learning strategy
   - (d) Learning style

TRIBHUVAN UNIVERSITY
2068

Master Level /I Year/EDUCATION   Full Marks: 50
Second Language Acquisition (Eng.519)   Time: 3 hrs.

Candidates are required to give their answers in their own
words as far as practicable.
The figures in the margin indicate full marks.

Attempt ALL the questions.

Group "B"   5×6=30

1. How is an L_2 learner viewed from different schools of thought? Discuss in brief.

2. According to socio-cultural theory of SLA, how is a second language acquired? Explain the SLA process in short in terms of the concepts mediation, scaffolding and regulation.
OR
Discuss in brief how Long's Interaction Hypothesis is an extended version of Krashen's Input Hypothesis.

3. State and explain McLaughlin's Information Processing Model of SLA in terms of two basic components automaticity and restructuring.

4. Briefly explain the notion of selective attention and its role in L_2 acquisition.
OR
What is Universal Grammar? Discuss the role of Universal Grammar in second language acquisition.

5. Discuss the influence of language shock and culture shock in L_2 acquisition.

Group "C"   1×12=12

6. How is variability explained from sociolinguistic perspective of SLA? Explain the notion of variability in SLA in terms of internal, external and sociolinguistic features.

TRIBHUVAN UNIVERSITY
2068

Master Level /I Year/EDUCATION   Roll No.: ....
Second Language Acquisition (Eng.519)

Group "A"
8
Attempt ALL the questions.   Tick (√) the best answers.

1. _____ is the state of SLA where learning ceases to develop
   a. L_1 transfer   b. cross linguistic influence
   c. fossilization   d. interlanguage

2. _____ compares between the learner's performance and the standard norms of the target language
   a. Contrastive analysis (CA)   b. Error analysis (EA)
   c. Generativism   d. Monitor model

3. _____ defines L_2 acquisition as the mediation of learner's internal mechanisms and external world context through language
   a. cognitive approach   b. socio-cultural theory
   c. functionalist approach   d. socio-international approach

4. Which of the following concepts does not explain the internal variability in SLA?
   a. Linguistic markedness   b. Universal developmental patterns
   c. Language transfer   d. Stylistic variation

5. 'Selective attention' refers to the awareness about the _____ features of a language
   a. common   b. comprehensible
   c. marked   d. unmarked

6. Long's reformulation of Interaction Hypothesis in 1996 included
   _____
   a. language internal factors   b. learner internal factors
   c. socio-cultural factors   d. comprehensible input and output

7. _____ is a way of employing mental and cognitive operations for comprehending and internalising the L_2 input
   a. input processing   b. instruction
   c. consciousness - raising   d. instructed SLA

8. When the learners involve in 'risk taking' tasks in L_2 communication, they _____
   a. develop extroversion in them
   b. develop introversion in them
   c. lack language acquiring personality   d. lack scaffolding