C H A P T E R   III

PREPARATION OF THE TEST-BATTERY

==

I. INTRODUCTION

In light of objectives laid down in the prior chapter, a test-battery is needed. The study of relevant existing tests outcome was that the tests cannot be used as it is. So, it was thought proper to devise new tests rather than directly adapting any old one in view of the following main reasons.

Firstly: It was found that situations included in existing relevant test-batteries were not familiar to our students. The batteries were, namely, (a) Brown-Carlson Listening Comprehension Test (BCLCT (24), (b) STEP Listening Comprehension Test (42) and (c) the tests compiled or adopted for sixth standard students in Australia by Donald Spearritt (123). Secondly, the contents of those tests pertained more to the foreign culture. Thirdly, it was noticed that the content and the language of other factors to be tested (viz., memory, reasoning, attention, etc.) and the level of those tests were not suitable to Gujarati pupils.

Last, but not the least, important the author thinks on the basis of the review of relevant literature and other attempts to test listening comprehension that listening comprehension is something more than simple sensory process of auding, as described and defined in the first chapter. In view of this to begin with tests for the following six situations or purposes taking care of the definition and objectives noted earlier were specially devised in Gujarati:

1. Listening to different specific situations.

2. Listening for vocabulary.

3. Listening for span of Attention.

4. Listening for Memory.

5. Listening for Auditory Resistance.

6. Listening for Reasoning.

II. SOME OTHER POINTS TO BE CONSIDERED WHILE CONSTRUCTING THE TEST

Before we go to the detailed description of the test two more points are worth discussing:

(a) The problem of dialectical element,

(b) The form of language used in the test.

(a) <u>The Problem of Dialectical Element</u>

As the **test** is to be constructed for whole of Gujarat State dialectical differences in vocabulary, pronunciation etc. may have an adverse effect on comprehension. So, it was thought fit to consult a linguistic expert (of Gujarati language) in the field. On such consultation valuable suggestions given by the linguistic expert* are summarised below:

(i)   Some dialectical words were pointed out for change. They were:

(ii)   More refined Sanskrit words were also pointed to be replaced by appropriate words. They were:

(iii) It was also pointed out by him that the test is an oral version of written form of language.(The oral performance differs from written one in juncture, repetitions, speed of delivery, longer pauses between two utterances etc. Here the test no doubt measures the comprehension level of written form put forth orally; yet it is a listening comprehension test.)

---

* Dr. P.B.Pandit, Head of the Linguistic Department, Delhi University.

(b) <u>Why Written Form of Language?</u>

The question now raised is why was spontaneous speech form is not preferred?

As it has been noted earlier the language used in this test is written language read aloud. In fact, written language and spoken language differ in their presentation, syntax, use of words, repetitions, etc. Holliday and others ('62) write:

> By comparison, spoken and written varieties of most modern languages are extremely close. (Yet) The spoken and written English (varieties of most modern languages) are by no means identical.They differ both in grammar and in lexis, as anyone by recording and transcribing conversation can find out.

Wilkinson (137) has raised a question about the legitimacy of using written language for listening comprehension. He has put forth very strong argument against the content of listening comprehension tests. He commented like this:

> The Brown-Carlsen and the tests of large numbers of workers do not differ materially from reading tests, Since they also use written material. The fact that it is read aloud may not be of great importance. Spearitt (1962) isolates a listening factor, but several of his listening tests are of written material spoken. The new studies in linguistics which have focussed our attention on the essential differences between the spoken and written languages, both in their

form and their action, are a product of
the sixties. As the present writer has
demonstrated elsewhere (Wilkinson, 1965)
several of the batteries in Brown-Carlsen
tests are essentially literary in
operation. And yet, it would seem clear
that a test of listening must essentially
be based on the spoken language.

The point raised by Wilkinson may have importance

elsewhere. It does not have that much importance in

connection with the listening comprehension testing

programme mainly for the following reasons:

1. There is much in common between listening

and reading comprehension skills and are both concerned

with receptive part of the language.

2. Linguists (62) mention varieties of language.

But in testing situations such varieties - differences

are minimised. The varieties are: (a) variety according

to users. (Speaker chooses one variety and uses it all

the time), (b) varieties according to use (each speaker

has a range of variety and chooses between them at

different times). The variety (a) is called dialect,

variety (b) is called register. It can be said that

roughly (about) 90 per cent speakers use a particular

variety in a particular area then it is called a dialect.

But, as register is not scientifically defined, we can

roughly say that when 20, 30 or 40 per cent people use a

particular variety of a language then it is called

register. Thus, registers refer to individual differences in using language varieties. It is further explained that every speaker has at his disposal a continuous scale of patterns and items, from which he selects for each situation type the appropriate stock of available harmonies in the appropriate key. He speaks, in other words, in many registers. An example of a lecturer will make the concept more clear. A lecturer uses one kind of registers in classroom, another kind of registers in common room, a third kind at home, so on and so forth. That is he keeps on shifting from one register to another according to situation. The linguistic differences may be slight; but they may be considerable, if the use of language in the new situation differ sharply from that of old. One should choose accordingly.

Thus, there are number of registers. Hence one cannot list the total range of such uses and also one cannot have absolute criterion to identify register. "To sing popular sing in Queen's English is a linguistic error".

The 'style' refers to the relationship among the participants, i.e. colloquial (informal) and polite (formal) as it can be called. In this thesis, however, we are mainly concerned with register.

How these two concepts, viz., register and style

affect**ë** the comprehension of pupils? It is implied in the
foregoing discussion that all the people especially school
children, are not familiar with all types of registers
(may be written or spoken). Unfamiliar (or unknown)
registers and styles used in the test may adversely
affect the comprehension of the pupils. So a test
constructor (for listening comprehension)should use
familiar and common registers for all testees or eliminate
them so as to make it understandable; which will again
come nearer to spoken standard language. Such form of
language, standard language is always depicted in
literature, in writing. So, the language of such test
will be much nearer to the written language - standard
language.

3. Again, when written material is read aloud
in front of pupils it will not be without registers that
of speaking,i.e.without tone, intonation, stress, etc.
The characteristics of spoken form are included. Thus,
difference in the two are minimised.

4. The script of such test can be written accord-
ingly (as it has been done in this research) with the
characteristics of spoken form. Not only that as we find
some sentences of conversations are included in various
writings from actual spoken language. Thus, it can be
minimized most. A written script with serious thoughts,
unknown words and unfamiliar register is not usually put

forth before pupils.

5. Again, to ask the pupils to comprehend
spontaneous speech confronts several problems:

(i) Spontaneous speech is always a speedy
discourse and so is difficult for this age-group
to follow it. (ii) The content of spontaneous
speeches of adults is not always to the level
of high school pupils. (iii) If spontaneous
speech of pupils is recorded, it is difficult to
find out the 'content' from it as it is usually
a chit-chating. (iv) If any planned or pre-
discussed talk etc. is recorded sponteineity is
withheld. This has been visualised by
Wilkinson (137). He described various subtest
(at the end of his article) which are not based
on spoken language. In the same article he
describes a test based on conversation between
two friends, for which a topic was chosen and
the discussion on it was recorded, which, in
turn, were roughly scripted. Care was taken
to remove certain inferences and obscurities;
again while recording on tape, general line
of the rough script was kept in mind and was
not, however, read verbatum. He then observes:
"The result had thus the feature of normal
conversation". Can we accept such discourse as

spontaneous one? Here it seems that care might
be taken to remove the barriers which listener
comes across while writing the script. Thus,
such oral script would resemble more to written
language than spoken one.

Anyway, the present investigator has taken into
the consideration all these issues and tried her best to
evolve a listening comprehension test.

III. DESCRIPTION OF THE CONSTRUCTION OF EACH GROUP OF SUBTESTS

1. Listening to Different Situation Tests

(a) Selection of the Content of the Tests

Various situations which were included in the two
tests BCLCT and STEP Listening - were first carefully
studied for this purpose. Moreover, some story books and
other literature for children in English, Gujarati and Hindi
were scrutinized; topics there in pertaining to science,
religion, stories, poems, geography, etc. satisfying the
following criteria were selected:

1. The content of the topics should be now i.e.
   not known earlier or studied by pupils.

2. It should be interesting or providing
   natural motivation for hearing.

3. It should never be too easy to lose interest
nor too hard to grasp; it should be within
the understanding of capacity of the age
groups under study.

4. It should not involve any social racial
prejudices.

5. It should avoid as far as possible unnecessary
personal involvement of listener of any sex.

6. It should not contain any topical matter or
content which may become obsolete after a
time.

7. It should be typical oral situation common
to all pupils to be tested, and not unfami-
liar situation.

After selecting suitable topics they were re-
written or shortened or lengthened to satisfy the above
mentioned criterion. The number of various situations
which were selected were three times more than what was
finally needed.

All these selected pieces of literature were
then referred to experienced teachers of eight class for
the opinion about their suitability in the light of the
prior criterion and in addition they were asked to bear
in mind the following points:

(i)   The difficulty level should be suitable
        to eighth standard.

(ii)   There should be clarity of ideas.

(iii)   The language should be clear and simple.

The consensus of teachers suggested modifications,
which may be listed as follows:

(i)   Some of the passages were needed to be
        shortened.

(ii)   There were some unfamiliar and difficult
        words. It was suggested that they/be
        should
        dropped.

(iii)   The test item on listening news was found
        difficult.

(iv)   Situations like preparing tea or any
        eatable was thought to be unfamiliar
        especially for boys.

(v)   Two talks viz. one on જવ મુજબ   and
        other on preparation and utility of the
        flannel cloth were found to be difficult
        and hence/be discarded as they were
        to
        thought to be difficult.

(vi)   The poem on " મેઠિયો " was preferred to
        " આવરે સાગર " and " વ્યા કરૂજો ".

(b) <u>Construction of Test Items</u>

While incorporating all these valuable suggestions a varied number of test items were constructed, on the following thirteen subtests for the first of the main test:

Sr. No.	Type of the subtest	No. of items	No.of questions asked
1	Story	2	15 (in each)
2	Short passages (general significance)	15	15
3	Talk (Boys' interest)	1	15
4	Talk (Girls' interest)	1	14
5	Talk (General)	1	15
6	Drama	1	15
7	Poem	1	15
8	Passages (Short) (Noting Details)	10	32
9	Anecdote	1	14
10	Talks for following directions	1	13
11	Short passages (Inference)	15	15
12	Passage (Narration)	1	10
13	Sentences (Accuracy of Ideas)	25	25

The form of the question was multiple choice, with one correct answer and three distractors. ~~(meant to judge the accuracy of ideas)~~. This form was thought fit to be used in the test-battery for the following reasons:

(i)   it reduces the percentage of guessing;

(ii)   measures all types of mental processes
        supposed to be involved in listening
        comprehension;

(iii)   it is widely used in almost all types of
        tests.

After constructing the items on all the subtests
noted above, all the items were read, reviewed by
researcher from the point of view of:

(i)   appropriateness of the stem, as well as
        of questions based on the content,

(ii)   homogeneity of the distractors,

(iii)   unambiguity and correctness of the language
        of the stem, distractors etc., and

(iv)   indefiniteness of statements or any other
        undue provision of a correct answer.

The written copy of the whole test was given to
experts to review the items from the point of view noted
above. Collection of the suggestion from different
experts was the next step. After that a group discussion
on the test was held, which helped mainly in correcting
distractors, language of instruction and the stem
questions.

As the test was to be taped, the reading of the
tests by various persons was also tried to select a voice
with the clarity and suitable speed, accent, tone, pitch,
etc.

Next step was to present the test in live form by
the investigator herself in two classes. Each class
consisted of 45 to 50 students.

Reading session for these tests was of about two
hours, and it took three days to present all the thirteen
tests to the class. The instruction, the questions and the
answers were read to pupils. The difficulties were solved
if any with the help of blackboard whenever necessary.

While presenting the test before the class
the investigator did following things:

(i)   Timings of reading - the passage, the
       question, of the answers given by pupils
       and of instruction given for each subtest
       were noted. The timing was taken by the
       helper; and with the help of stop-watch.

(ii)   The helper also noted ambiguous points,
       words or any difficulty realized by
       students.

(iii) The instruction was given to the students
       to point out any difficulty they felt at the
       end of the each test.

This whole process helped to:

(i)   determine appropriate speed of reading the
content of all the above mentioned tests,

(ii)   determine appropriateness of pitch and
volume of the voice,

(iii) have an idea of clarity of pronunciation
and voice,

(iv) determine appropriate pauses while reading,

(v)   get an idea of maximum time needed to
present each sub-test,

(vi) decide whether oral presentation - the
whole process of presentation and the
content of the test is boring or interest-
ing to pupils.

It was observed during oral presentation that:

(i)   The selected topics and types of tests
were interesting.

(ii)   The speed should be slower than it was
presented.

(iii) The administration of the test on tape
will take more time than it was presented
in live form.

(iv) The lessening of the speed will lenghten the test.

2. Vocabulary Tests

(a) Selection of the Words,Idioms and Proverbs

(i) Various tests were read viz., listening and reading, comprehension tests mentioned earlier in English and reading comprehension tests in Gujarati, silent reading comprehension tests, and vocabulary tests prepared by M.Ed. students were referred to.

(ii) Various text-books for eighth standards studied.

Twenty words for synonyms, twenty words for antonyms, thirteen idioms and sixteen proverbs, were selected.

(b) Construction of Items

After selection, test items were constructed. Review by teachers, colleagues and other experts gave useful suggestions regarding (i) difficulty level of words, (ii) new words.

Incorporating the suggestions the test was orally put forth before pupils. It was observed and suggested by helper and pupils that: (i) the words should be easier, (ii) ambiguous distractors should be removed, (iii) the

time gap for finding correct answers, reading distractors should be lesser than the former test.

〝3. Attention Tests

There are five tests in this section. Of the five, three of the attention tests were constructed on the line of those of Spearritts. They are:

(i)   Triplet Numbers with fifteen items.

(ii) Letter list which included three lists with ten items in each.

(iii) Five Letters   Test: with fifteen items.

Two more tests were newly constructed. It was felt in the beginning that the test "Five Letters" based on Gujarati alphabets will be too hard to solve. So the same mental processes were decided to be tested with the help of five digits (1, 2, 3, 4, 5). Both the tests were included in the battery.

(iv) The Five Digit test with fifteen items.

(v)   The other one added is Listening to phonetic variations. In Gujarati language broad and narrow pronunciations (મોર, મોરુ) work as phonemes. Attentive listener catches these sound variations and derives meanings accordingly: bearing in mind this, the test has been devised. They were given to experts. Suggestions were:

1. Five Letters Test, Letter List tests will be too difficult to solve.

2. Easier items, detailed instructions and number of items for practice should be more than those of other tests.

Incorporating the suggestions, especially that of practice examples the tests were put before the classes.

Oral presentation gave insight into the following matters:

(i)   The preparation of the format of the answer sheet of the Letter List Test and the Listening to Phonetic Variations Test.

(ii) The preparation of instructions of all the five tests to be more explicit and short.

(iii) The test on Five Letters Test was found more difficult.

4. Memory Tests

Spearritt's study consisted of following types of memory tests:

(i)   Memory for Words

(ii) Letter Span (Visual)

(iii) Letter Span (Auditory)

(iv)   Word Number Recall

(v)   Consequences

(vi)   Sentence Completion (Immediate Recall)

Out of these six tests, Letter Span (Visual) was dropped, as the mode of presentation of the test under construction was aural one. The test for Memory for Words' test, and 'Word Number Recall' test and 'Consequences' were made easier by decreasing the amount of text to memorize and changing the method of presentation. Each one had fifteen pairs in each 'Memory for Words' and 'Word Number Recall', and 'Consequences' included 23 items.

Two more tests were added with fifteen items in each. One was 'Memory for Sentences" (Immediate Recall) and the other one was 'Sentence Completion' (Late Memory/Remote one).

Experts opined to include as many items in the test, as no definite concept for the difficulty level of the pupils can be put forth.

After oral presentation to two classes it was found that:

(i)   the amount of text to be memorized should still be decreased in all tests.

(ii)   Word Pair Test was found difficult. So

easier word pairs were constructed.

5. Auditory Resistance

Auditory Resistance Tests are included in Spearritt's study. We find these types of tests in Karlin's study. Spearritt also constructed from the description of the original tests of Karlin. We find this factor is included in Dr. K. G. Desai's Intelligence Test. In all five tests were constructed. The test on Delayed Spelling consisted of 20 words. The tests on Singing, Haphazard Speech, Illogical Grouping had 20 items. One more test is constructed viz. the test on "Correcting the word order". This has been devised from that of Dr. K. G. Desai's study.

Expert opinion was like that of Memory Tests.

It was realized during the oral presentation of the tests that:

(i) While reading the content of the tests pauses should be longer, intonation and pitch should be more distinct and pronunciations should be more clear,

(ii) the sentences were too long,

(iii) Some of the words used were not very familiar to the pupils and words with consonant cluster were found difficult

-ing
in Delayed Spell∠ Test, Haphazard as well
as Illogical Grouping Test,

(iv)   the lines of different songs were too long,

(v)   Pupils found it difficult to grasp the
words when the musical composition was
not simple.

6. Reasoning Tests

Out of six tests of Reasoning, five tests are
that of Spearritt. They are:

1. Letter Grouping (Non-sense Syllables)
2. Letter Series
3. Deductive and Inductive Reasoning (Solving
Problems)
4. Matching Words
5. Arithmatical Reasoning

The test items were constructed. It was felt
while constructing the items on the test "Letter Grouping"
that it will be difficult for the eighth standard pupils
to solve. So the items with meaningful words were
constructed; it is the sixth test, namely, "Word Grouping"
test.

The two tests "Letter Grouping" and "Word
Grouping" had more than 30 items.

The items on problem solving were newly construct-
ed. The items of all the reasoning tests were made
easier.

The oral presentation gave the idea of:

(i)   the difficulty level of each test; they
were found difficult. So more easier items
were to be constructed,

(ii)   the problems to be solved, based on who is
related to whom, finding directions of the
places, were found difficult,

(iii) the rate of reading should be slower,

(iv)   Letter Series Test should be shorter and
            -ed
with mix/letters - vowels and consonants
and the pronunciation of Letters more
distinct,

                                           test
(v)   The length of the sentences of the/Matching
Words should be shorter,

(vi)   the time gap to be kept between two
questions (items); should be more than the
other tests,

(vii) the nature of instruction should be more
clear,

(viii) the practice examples should be increased

to 3 or 4 representing various types of

reasoning.

Thus, the final test-battery for the pre-tryout
was prepared. The arrangement of the final test-battery
for pre-tryout was as follows:

Sr. No.	Name of the test	No. of questions	Type of the sub-test in order
1.	Story	16	Listening to Different Situation (LDS)
2.	Poetry	11	LDS
3.	Letter Grouping	20	Reasoning
4.	Word Grouping	33	Reasoning
5.	Sentence Completion (Immediate)	25	Memory
6.	Triplet Numbers	22	Attention Test
7.	Sentence Completion	15	Memory
8.	Narration	15	LDS
9.	Memory for Words	15(Groups)	Memory
10.	Letter Series	19	Reasoning
11.	Letter List	3 Lists (10 syllables in each)	Memory
12.	Consequences	15 Pairs	Memory
13	General Significance	15 Para	LDS
14	Word Number Recall	14 Groups	Memory

Sr. No.	Name of the test	No. of questions	Type of the sub-test
15.	Illogical Grouping	17 Sentences	Auditory Resistance
16.	General Reasoning	33 Problems	Reasoning
17.	Five Letters	17 Instruction	Attention Test
18.	Anecdote	14 Questions	LDS
19.	Letter Span	22 From 3 to 9 Syllables	Memory
20.	Listening to Note Details	10	LDS
21.	Five Numbers	11	Attention
22.	Singing	20	Auditory Resistance
23.	General Talk	15	LDS
24.	Vocabulary Tests		
	(a) Synonyms	20	
	(b) Antonyms	16	Vocabulary Tests
	(c) Idioms	14	
	(d) Proverbs	16	
25.	Girls' Talk	15	LDS
26.	Drama	15	LDS
27.	Listening to Phonetic Variation	16	LDS
28.	Boys' Talk	15	LDS
29.	Listening for Inference	13	LDS
30.	Arithmetical Reasoning	23	Reasoning

Sr. No.	Name of the test	No. of questions	Type of the sub-test
31.	Correcting the order of the Sentence	13	Auditory Resistance
32.	Matching Words	31	Reasoning
33.	Haphazard Speech	15	Auditory Resistance
34.	Story	15	LDS
35.	Delayed Spelling	34	Auditory Resistance
36.	Listening for Accuracy of Ideas	25	LDS
37.	Sentence Writing	15	Memory
38.	Follow Directions	13	LDS

The preliminary tryout based on these tests and the statistical calculations done on the basis of the data collected will be described in the next chapter.