Unit 6

Historical perspectives. Integrative, discrete-item & communicative tests

---

Historically, lg testing theory & practice have followed the changing winds of lg teaching methodology. E.g.:

- 1950s: behaviorism, special attention to contrastive analysis. Lg testing focused on specific lg elements & contrastive L1 & L2 analysis.

- 1970s & 1980s: CLT brought along a more integrative approach to L2 teaching; "the whole of the communicative event is greater than the sum of its linguistic elements (Clark 1983)."

Two major approaches to lg testing that still prevail today (though in 'mutated' form): discrete-point & integrative testing techniques and procedures.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

2

These 2 approaches are related to the **two main periods** in the history of lg testing: the **psychometric-structuralist** period and the **integrative-sociolinguistic** period. Thus:

Psychometric-structuralist period   Discrete-point testing

Discrete-point testing

Psycholinguistic-sociolinguistic period   Integrative testing

Integrative testing

Psychometric-structuralist period

1950s-1960s

Linguistics: Structuralism

Psychology: Behaviorism

Testing focused on specific lg elements (phonological, lexical, or grammatical).

Test-developers focused on testing single elements of lg in a test, such that each test item tested a single lg element.

This approach came to be known as the discrete point approach to lg testing.

Discrete-point tests constructed on assumptions that

- lg can be broken down into component parts &

- those parts can be adequately tested in isolation.

Those components are:

-the lg skills

-the various hierarchical units of lg (phonology / graphology, morphology, lexicon, syntax, discourse) within each skill, &

-the subcategories within those units (phonemes, morphemes, words, etc.)

E.g.: A proficiency test with m-c qs divided into grammar, vocab, reading, …, with some items attending to smaller units & other to larger units, can measure discrete points of lg.


Advantages of discrete-point testing

D-p ts…

☞ yield data which are easily quantifiable,

☞ allow a wide coverage of items,

☞ are efficient,

☞ have the scoring reliability that is commonly associated with objective scoring.

Problem with discrete-point testing

D-p t assumes proficiency can be quantified in this way:

| sum of parts | = | total |

lg skills **+** lg components g^{al} lg proficiency
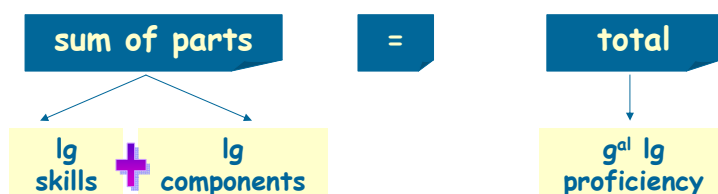
Oller (1976, 1979) argues that lg competence is a unified set of interacting abilities which can't be separated & tested adequately in isolation—i.e., communicative competence is so global & requires such integration (hence 'integrative testing') that it can't be captured in additive tests of grammar & reading & vocabulary & other d-ps of lg

Oller (1979:37): 'If discrete items take lg skill apart, integrative tests put it back together. Whereas discrete items attempt to test knowledge of lg one bit at a time, integrative tests attempt to assess a learner's capacity to use many bits at the same time'.

Morrow (1979:145): '…knowledge of the elements of a lg in fact counts for nothing unless the user is able to combine them in new and appropriate ways to meet the linguistic demands of the situation in which he wishes to use the lg'.

Oller, Morrow, Savignon, Spolsky, Rea… are all promoters of the communicative approach to lg teaching, and inextricably related to this approach to teaching is the integrative approach to testing

4

Psycholinguistic-sociolinguistic period

In this period, global testing is preferred to d-p t. According to Oller, global integrative tests like the cloze procedure & dictation

☞ went beyond the measurement of a limited part of lg competence achieved by d-p ts (biased towards the testing of receptive skills), and

☞ could measure the ability to integrate disparate lg skills in ways which more closely approximated the actual process of lg use.

Why is this period called 'psycholinguistic-sociolinguistic'?

From a psycholinguistic perspective, lg came not so much a well-defined taxonomic structure (structuralism) but a dynamic, creative, functional system. The sociolinguistic contribution centers on the concept of communicative competence (C.C.).

C.C. is a broader notion than Chomsky's 'competence' (vs. 'performance'): it covers knowledge of rules to form gram. sentences AND of rules for using those sentences appropriately with different contexts (pragmatics).

Two types of tests have been held as prime examples of integrative tests: cloze tests and dictations.

For features & formats of cloze tests, see previous unit.

☛ Why are cloze tests results said to be a good measure of overall lg proficiency?

Bc ability to supply appropriate words in blanks requires a # of abilities that lie at the heart of competence in a lg: knowledge of vocab, gram structure, discourse structure, reading skills & strategies, & an internalized expectancy grammar (that enables one to predict an item that will come next in a sequence).

Features & formats of dictations don't need explanation here, do they?

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo   11

☛ Why are dictation results said to be good a measure of overall lg proficiency?

Savignon (1982) & Oller (1979), claim that dictation taps into certain gram & discourse competences & that dictation test results tend to correlate strongly with other tests of proficiency.

Success on a dictation test requires careful listening, reproduction in writing of what is heard, efficient short-term memo & some expectancy rules to aid short-term memo.

Dictation tests, however, are + classroom centered since large-scale administration is quite impractical from a scoring standpoint.

☞ Reliability of scoring criteria is also a problem that is not present in m-c qs or exact-word cloze test scoring.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo   12

Debate: Discrete-point testing or integrative testing?

The q of whether integrative tests of lg proficiency really exist is directly related to the q of whether or not lg ability can be divided up into separately testable components.

Proponents of the indivisible view have proposed the unitary trait hypothesis (Oller 1979), which suggests that…

➲ lg proficiency is more unitary than the d-p testers have claimed –i.e. vocab, gram, phonology, the 4 skills, & other discrete points of lg cannot actually be distinguished from each other. It contends that there's a general factor of lg proficiency such that all the d-ps do NOT add up to that whole.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

13

What is communicative lg testing (c.lg.t)?

C.lg.t is intended to provide testers with info about the testee's ability to perform in the target lg in certain context-specific tasks.

It has to be recognized…

- that given the constraints of time & practicality, only a small sample of the t-ts's lg can be collected, &

- that however realistic the tasks may be intended to be, the student's performance will inevitably reflect the fact that s/he was performing under tests conditions.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

14

What does communicative lg testing try to assess? Communicative competence in the L2.

Thus, **communicative competence** is the construct that communicative lg tests are supposed to measure.

What is this construct like? –i.e. what is understood by communicative competence?

Many are the definitions of communicative competence, and by quite a few authors, but it has been the work of **Canale & Swain (1980)** that has contributed the most to the definition of the construct.

Communicative competence (c.c.) comprises 4 types of competence, in Canale & Swain's model:

☞ Grammatical competence

☞ Sociolinguistic competence (sociocultural rules)

☞ Discourse competence (cohesion & coherence)

☞ Strategic competence (knowledge & skills involved in planning & executing communicative acts efficiently; it's usually put to work when communication requires improvisation bc the other competences are somehow insufficient).

For Canale & Swain…

…communicative lg testing, apart from being concerned with what the learner knows about the form of the lg & about how to use it appropriately in contexts of use (competence), must also deal with the extent to which the learner is actually able to demonstrate this knowledge in a meaningful communicative situation (performance) –i.e. what he can do with the lg, their ability to communicate with ease and effect in specified sociolinguistic settings.

The capacity to use lg communicatively thus involves both competence & ability to use this competence. It is held that the performance tasks candidates are asked to carry out in communicative tests should

☞ be representative of the type of task they might encounter in their own real-life situation &

☞ correspond to normal lg use where an integration of communicative skills is required with little time to reflect on, or monitor lg input and output.

The criteria employed in the assessment of performance on these tasks should relate closely to the effective communication of ideas in that context.

This perspective is consistent with the work of lg testers who support a model of communicative lg ability with a marked shift of emphasis from the linguistic to the communicative dimension –i.e. emphasis no longer on linguistic accuracy but on the ability to function effectively thru lg in particular contexts of situation.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

19

Distinguishing features of communicative lg tests

Prior to test construction

Tests constructors must…

1. identify the skills & performance conditions that are the most important components of lg use in particular contexts. The incorporation of these features indicates the degree to which the test task reflects the attributes of the activity in real life that it was meant to replicate. Otherwise, it would be imprudent to make statements about a student's ability to function in normal conditions in his/her future target situation.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

20

Distinguishing features of communicative lg tests

Prior to test construction

Tests constructors must…

2. ensure the sample of communicative lg ability in tests is as representative as possible. If we are to extrapolate from test data & make statements about comm. lg ability in real life situations, great care needs to be taken over texts & tasks used in tests.

Distinguishing features of communicative lg tests

Prior to test construction

Tests constructors must…

3. bear in mind that tests of comm. lg ability should be as direct as possible (attempt to reflect the real life situation) & the tasks candidates are asked to perform should involve realistic discourse processing.

Direct testing requires integrated performance from the candidate involving communication under realistic linguistic, situational, cultural and affective constraints. Candidates have to perform both receptively and productively in relevant contexts.

Distinguishing features of communicative lg tests

Prior to test construction

Tests constructors must…

4. remember that, when it comes to marking, the holistic and qualititative assessment of productive skills, and the implications of these for test reliability, need to be taken on board.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

23

Criteria for the construction of communicative tests

Merrill & Swain (1984) distinguish 4 main criteria:

1. Start from somewhere

2. Concentrate on content

3. Bias for best

4. Work for washback

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

24

1. START FROM SOMEWHERE

Tests should **build on existing knowledge and principles** –i.e. a theoretical framework is an important foundation for creating communicative tests.

In other words, you need to **have a sound theoretical basis to be able to build communicative tests** –namely, you need to know…

o what the CLT approach brought to the field of lg teaching,

o what is understood by communication in relation to lg teaching,

o what is communicative competence,

o the distinction bt competence and performance, etc.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

25

2. CONCENTRATE ON CONTENT

Traditional lg tests might have tons of unrelated topics which require the testee to jump from topic to topic in a matter of seconds —unlike in natural communicative situations.

Swain suggests that **comm. tests should have motivating, interesting, substantive content.**

E.g.: Students might be given a situation ("job opportunities in two locales"), which forms the setting for the test. The situation is motivating & provides new info. to deal with.

Tests that concentrate on content also need to **be integrated & interactive.** Integration provides context, that all-important element in communication, and by eliciting opinions & the expression of one's own ideas, an interactive element can occur.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

26

3. BIAS FOR BEST

Comm. lg tests should do everything possible to elicit the very best performance from students. How can students do their best at the lg test?

- If their anxieties are lowered

- If they have enough time to complete the task(s)

- If they are made to feel they can indeed succeed

---

3. BIAS FOR BEST

Learners tend to perform best when lg tests have the following 4 phases:

(a) *warm-up*, where they get accustomed to the test format;

(b) *level check*, where items confirm that testees are where they are assumed to be;

(c) *probe*, where the limits of one's abilities are ascertained thru challenging, difficult items; and

(d) *wind down*, where the learner is allowed to relax with some easier questions.

COMMUNICATIVE TESTING — Unit 6

4. WORK FOR WASHBACK

For a definition of washback, see Unit 1.

While we should not teach toward a test, we can use tests as teaching tools. Tests become feedback devices whereby a student perceives elements of communicative performance that need improvement. Our tests, then, should be designed to provide that feedback.

Técnicas y Procedimientos de Evaluación ... 2008/09
Prof. Lourdes Cerezo

29