MEASURING BILINGUALISM
DR. A. POUSADA (2014)
Testing and Assessment

- **Assessment**: tests, projects, portfolios, observation of performance, etc.
- **Formative assessment**: checks progress of learning
- **Summative assessment**: checks results at end of program
Types of tests (purposes)

- Achievement tests
- Admission tests
- Aptitude tests
- Diagnostic tests
- Language dominance tests
- Placement tests
- Proficiency tests
- Progress tests
Direct vs. indirect testing

- **Direct testing:** Ss perform the skill to be measured (e.g., A translation test in which Ss translate a text.)

- **Indirect testing:** Measures the abilities underlying the skills to be measured (e.g., A writing test in which Ss identify grammatical errors in sentences)

- **Semi-direct testing:** (e.g., tape recorded speaking test)
Problems & limitations

- **Direct testing:**
  - only limited tasks can be measured
  - may take long time
  - may require special materials

- **Indirect testing:**
  - trait being measured may not be good indicator of overall skill
Discrete point vs. integrative tests

- Discrete point tests:
  - Assume that language can be broken down into separate elements
  - Focus on one linguistic element at a time
  - Tend to be indirect

- Integrative (holistic) tests:
  - Require Ss to address many linguistic elements simultaneously
  - Tend to be direct
Criterion-referenced tests

- Seek to classify Ss according to specific criteria or tasks
- Yield direct info on what Ss can do.
- Yield percentages and cut-off scores.
- Unaffected by other test-takers’ performance.
- Tend to be “low stakes” tests.
Norm-referenced tests

- Indicate how Ss’ performance compares to that of other groups of test-takers
- Usually yield percentiles or grade equivalence scores
- May be affected by performance of other Ss.
- Tend to be “high-stakes” tests.
Objective vs. subjective tests

- **Objective tests:**
  - require no judgment by the scorer (e.g., multiple choice, T/F tests)
  - mechanical cut-offs lack nuance

- **Subjective tests:**
  - require judgment by the scorer (e.g., essay questions, compositions)
  - scoring can vary tremendously
History of language testing

- Prior to 1950s: Grammar Translation Method, reading-oriented methods
- 1950s-1960s: reliance on structural linguistics, behavioral psychology, discrete point tests
- After 1960s: concern with testing communicative language ability and more holistic measures
Communicative competence

- The ability to use language effectively to communicate in real-life interactions

- Components
  - Structural competence
  - Discourse competence
  - Sociolinguistic competence
  - Strategic competence
Communicative language testing

- Requires tasks that are genuinely communicative and authentic
- May be more time-consuming
- Ultimately yield richer data
- Should produce better assessments that can lead to curricular improvement
Problems of measuring communicative competence

- Language competence is multidimensional and difficult to operationalize.
- Testers assume that it’s sufficient to measure 2 separate monolingual states to obtain an adequate measure of a bilingual
- An adequate methodology to capture the specific nature of bilingual behavior is lacking.
Comparative measures

- Approach: take measures in each of bilingual’s 2 lgs and compare them.
- However, wide variations exist between competence of native speakers of same lg., so it’s extremely difficult to identify and operationalize the salient features of native competence.
Solution to comparisons

- Compare monolingual competences of bilingual speaker to monolingual standards in each lg.

- Measures need not be similar for each lg since the comparison occurs at the level of a statistical distribution of competences of native speakers

- Avoids problem of directly comparing behavior in one lg. with that in another.
Problem in comparing vocabularies

- Not fair to compare a bilingual’s vocabulary in one language and compare it to a monolingual’s.
- Need to look at total vocabulary and total conceptual vocabulary.
- Best solution is to add up bilingual’s words in each lg. and then subtract the vocabulary shared between the 2 lgs.
Mother tongue competence

- Peabody Picture Vocabulary Test -- measures receptive vocabulary

- Reynell’s Syntactic Complexity Test & numerous lg. tests in traditional lg. batteries -- all measure 1 aspect of mother tongue lg. competence, but not clear that aspect measured is most relevant dimension -- is it justifiable to separate skills which make up lg. competence?

- Holistic measures include: cloze tests which tap learner’s internalized grammatical knowledge -- have predictive power and correlate well with other tests
Tests of L2 competence

- cannot compare them with tests of mother-tongue competence
- do enable us to define levels of dominance in bilinguals
- useful for identification of developmental stages
Behavioral measures of L2 competence

- Reaction or latency-time measures (psychological experiments)
- Completion and word-detection tests (*dansonodent*)
- Verbal association tests
- Interlingual verbal flexibility (translation or switching ability)
- Use of ambiguous stimuli (*once, dime*)
Language biographies & self-evaluations

- **Lg. biographies**--provide info on age and context of acquisition, past and present use, number of varieties spoken, degree of literacy, etc.--depend on declared behavior, not observed

- **Self-evaluations and judgments by native speakers of proficiency**--generally done by assessing lg. skills on 3, 5, or 7 point scale ranging from ‘nil’ to ‘native-like’--very subjective
Measures of bilingual specificity

- Grosjean (1985): bilinguals have patterns of behavior that are unique to them—like borrowing and code switching.
- A bilingual’s full repertoire is fully exploited only in situations in which both lgs. can be used.
- Specific bilingual behavior often mistaken for interference.
- Mixed code is only inappropriate in terms of monolingual norms.
Measures of cognitive effects of bilingualism

- Lots of evidence of correlation between development of bilinguality and cognition exists.
- However, studies are contradictory, showing both advantages and disadvantages.
- Usually involve use of tests of verbal and non-verbal intelligence, verbal creativity, divergent thinking, verbal transformations, symbol substitutions, etc.
Measures of affective correlates of bilinguality

- **Value judgments toward lgs. and speakers** -- Lickert-type evaluation scales in which subjects express degree of agreement or disagreement with statements relating to lgs. and speakers (cf. Gardner and Lambert)

- **Semantic differential** -- matched guise technique -- Lambert et al. -- speaker listens to tape recording of one speaker in two guises and makes judgments about speakers based on semantic oppositions (good-bad, smart-stupid, beautiful-ugly, etc.) -- problem is that it ignores all elements relevant to communication except voice characteristics

- **Measures of bilingual’s cultural identity** -- multidimensional scaling, ethnic dolls, role playing, questionnaires
Psychometric tests of academic prof. are not appropriate for minority children because they have not reached level of development required for tests to be valid.

Cultural differences can skew results—if child is presented with pictures of objects that are not part of that culture, lack of response by child has no assessment value.
Measurement of societal bilingualism

- Census
- Index of linguistic diversity
- Index of communicativity
- Surveys
  - Geolinguistic surveys
  - Linguistic atlases
  - Ethnolinguistic studies
- Government inquiries
Sociolinguistic & ethnographic measures

- **Sociolinguistic studies** (cf. Labov) examine social distribution of variants of linguistic variables—LePage & Tabouret-Keller (1985) investigated lg. use and attitudes in Belize and St. Lucia—used questionnaires, recorded interviews, correlated linguistic variables with sociological variables—

- **Ethnographies of communication** observe small, well-defined multilingual communities in minute detail through participation observation of small groups and networks, descriptive analyses, case studies, etc.—used by anthropological linguists and sociolinguists in many societies around the world
Measuring bilinguality is far from simple.

Most important lesson is that a bilingual does not equal two monolinguals squashed together.

Bilingual measurement must assess language competence in each language separately.

Bilinguals may have less vocabulary in each language compared to monolinguals, but their total vocabulary is bigger.

Decisions about bilingual children should be based on the best possible measures of their communicative skills.
THE END