Examining Validity of Oral Proficiency Interview

AATK 2015
Myong-Hee Choi
University of Montana
GOAL

To examine the validity of an ACTFL Oral Proficiency Interview (OPI), which concerns the extent to which a test measures what it claims (Brown 1996)

“The ACTFL OPI is a standardized procedure for the global assessment of functional speaking ability.”
OPI Background

- Originally developed in the 1950s by the FSI (Foreign Service Institute) and completed by ACTFL (the American Council on the Teaching of Foreign Languages)

- Communicative language teaching+ proficiency movement (1980s)

- No alternatives to the OPI to assess the spoken language skills of learners in terms of practicality and accountability (e.g., Salaberry, 2000)

- Served as a catalyst for teaching, testing, and curriculum development.
What is OPI?

A criterion-referenced test (Swender, 2003) to the speaker’s global speaking proficiency—across a variety of languages.

“It is a face-to-face or telephonic interview between a certified ACTFL tester and an examinee that determines how well a person speaks a language by comparing his or her performance of specific communication tasks with the criteria for each of ten proficiency levels described in the ACTFL Proficiency Guidelines 2012”
<table>
<thead>
<tr>
<th>Proficiency Level</th>
<th>Global Tasks/Functions</th>
<th>Context/Content</th>
<th>Accuracy</th>
<th>Text Type</th>
</tr>
</thead>
</table>
| Novice (ILR 0+)   | Communicate minimally with formulaic and rote utterances, lists, and phrases.         | Most common informal settings  
*Most common aspects of daily life* | May be difficult to understand, even for speakers accustomed to non-native speakers.                         | Individual words and phrases                                       |
|                   |                                                                                        |                                                                                |                                                                                                               |                               |
| Intermediate (ILR 1) | Create with language, initiate, maintain, and bring to a close simple conversation by asking and responding to simple questions. | Some informal settings and a limited number of transactional situations  
*Predictable, familiar topics related to daily activities.* | Understood with some repetition, by speakers accustomed to non-native speakers.                             | Discrete sentences          |
|                   |                                                                                        |                                                                                |                                                                                                               |                               |
| Advanced (ILR 2)  | Narrate and describe in major time frames and deal effectively with an unanticipated complication. | Most informal and some formal settings  
*a variety of concrete topics related to events of current, public, and personal interest or individual relevance* | Have sufficient control of basic structures to be understood by native speakers unaccustomed to non-native speech. | Paragraphs                    |
Level 1 (Intermediate) Tasks

- Simple Short Conversation
- Examinee Asks Questions
- Role-Play: Basic Survival Situation
Level 2 (Advanced)

- Narrate in Present, Past, Future (time)
- Physical Description (person, place, object)
- Instruction or Directions
- Report Facts about Current Events
- Role-Play: Survival Situation with a complication
What is OPI?

- An interactive, adaptive, and learner-centered assessment
- A specific set of procedures (warm-up, level-checks, proves, wind-down)
Literature Review: OPI Validation YES?

The OPI is a valid and reliable testing method (Dandonoli & Henning 1991; Thompson 1995; Reed 1992)

- Henning & Cascallar (1992): 79 ESL learners comparing TOEFL, TWE, and TSE tests
- Dandonoli & Henning (1990): OPI data from 60 French and 59 English as a second language (multitrait-multimethod)
- Magnan (1988): 40 novice-mid through advanced-plus speakers studying French; A significant relationship between (grammatical) accuracy and level
- Thompson (1995): interrater reliability on the OPI given to 795 candidates in 5 languages (English, French, German, Russian, and Spanish)
OPI Validity: Criticism

1. Lack of theoretical and empirical support (Bachman & Savignon 1986; Shohamy 1990b; Salaberry 2000), rather “intuitive judgments” (Fulcher 1996; Lantolf & Frawley 1985, 1988)

“Proficiency is derived by policy and not from science or empirical inquiry” (Lantolf & Frawley 1988:10)

2. Problem of circularity; Speaking ability = the method of measurement (the OPI Guideline) (Bachman & Savignon 1986; Lantolf & Frawley 1985)

3. Survey research interview (not natural conversation) (e.g., John & Tyler 1998; Shohamy 1988)

4. Mixture of linguistic and non-linguistic measures (e.g., Bachman & Savignon 1986; Fulcher, 1996; Taylor & Wigglesworth 2009)
Literature Gap: Empirical Validation

- Previous studies centered on a small part of languages such as French, German, Russian and Spanish (along with ESL), which doesn’t necessarily represent the rest of languages (Malone 2008).

- It is worthy testing the transferability of OPI scales across languages, in which an assumption is generally made that descriptors on such scales are likely to be interpreted similarly across languages.
Research Question

Transferability of OPI ratings across the category IV languages: “Did they take the same test?”

- Are there any difference ratings of OPI between Korean and Chinese novice students?

- To what extent do they differ between Korean and Chinese learners?
Study

- **Subjects**: English adult speakers acquiring Chinese (n=25) or Korean (n=24) as a foreign language

- **Data Collection**: from Jan, 2013 to Dec, 2014
  Intensive 4-week refresher course (120 hrs)

- **Instruments**:
  a. Pre-test: face-to-face Interview (Novice)
  b. Post-test: Official OPI telephonic interview
## OPI Results

### Table 1. Novice Learners’ OPI Result

<table>
<thead>
<tr>
<th>OPI Result</th>
<th>0+ (Novice)</th>
<th>1 (Inter.)</th>
<th>1+ (Inter.High)</th>
<th>Unpaired t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean (n=24)</td>
<td>10 (42%)</td>
<td>14 (58%)</td>
<td>0</td>
<td>$t = 5.47$</td>
</tr>
<tr>
<td>Chinese (n=25)</td>
<td>0</td>
<td>16 (64%)</td>
<td>9 (36%)</td>
<td>$P &lt; .0001$</td>
</tr>
<tr>
<td>total</td>
<td>10</td>
<td>30</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>


Figure 1. OPI Results

- 0+ (Novice)
- 1 (Intermediate)
- 1+ (Inter.High)
# Table 2. Listening & Reading Test Results

<table>
<thead>
<tr>
<th>DLPT</th>
<th>Listening</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>ILR 2 (Advanced)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korean (n=10)</td>
<td>2.00</td>
<td>2.15</td>
</tr>
<tr>
<td>Chinese (n=14)</td>
<td>2.07</td>
<td>2.21</td>
</tr>
<tr>
<td>Unpaired t-tests</td>
<td>$t = 0.04$</td>
<td>$P=0.97$</td>
</tr>
</tbody>
</table>
Table 2. Listening and Reading Test Results: Advanced L2ers

<table>
<thead>
<tr>
<th>Listening</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean L2ers</td>
<td>Chinese L2ers</td>
</tr>
<tr>
<td>0.15</td>
<td>-0.05</td>
</tr>
<tr>
<td>0.16</td>
<td>-0.15</td>
</tr>
<tr>
<td>0.17</td>
<td>-0.2</td>
</tr>
<tr>
<td>0.18</td>
<td>-0.25</td>
</tr>
<tr>
<td>0.19</td>
<td>-0.3</td>
</tr>
<tr>
<td>0.20</td>
<td>-0.35</td>
</tr>
</tbody>
</table>

-0.05
Findings & Discussion

1. The Korean learners’ OPI results significantly differed from the Chinese learners’ counterparts.

2. There is no case for the Chinese learners’ failure in achieving the level 1 (Intermediate)

3. There is no case for the Korean learners’ success in achieving the level 1+ (Intermediate-High)

4. The Korean and Chinese learners did not significantly differ from each other in terms of reading and listening tests.
Conclusion

- OPI might not be necessarily equivalently valid to the assessment of learners’ proficiency across languages.

- This findings represent an invitation for more research.