

ASSESSING SPEAKING

A BRIEF SUMMARY

International English Language Testing System



Compiled by :

Edtwin Sulispriyanto - 11520010
FX. Oki Setyawan

ENGLISH DEPARTMENT AND TEACHER TRAINING FACULTY
UNIVERSITAS SLAMET RIYADI
2014/2015

I. ASSESSING SPEAKING

While speaking is a productive skill that can be directly and empirically observed, those observations are invariably colored by the accuracy and effectiveness of a test-taker's listening skill, which necessarily compromises the reliability and validity of an oral production test. How do you know for certain that a speaking score is exclusively a measure of oral production without the potentially frequent clarifications of an interlocutor? This interaction of speaking and listening challenges the designer of an oral production test to tease apart, as much as possible, the factors accounted for by aural intake.

Another challenge is the design of elicitation techniques. Because most speaking is the product of creative construction of linguistic strings, the speaker makes choices of lexicon, structure, and discourse. If your goal is to have test-takers demonstrate certain spoken grammatical categories, for example, the stimulus you design must elicit those grammatical categories in ways that prohibit the test-taker from avoiding or paraphrasing and thereby dodging production of the target form.

As tasks become more and more open ended, the freedom of choice given to test-takers creates a challenge in scoring procedures. In receptive performance, the elicitation stimulus can be structured to anticipate predetermined responses and only those responses. In productive performance, the oral or written stimulus must be specific enough to elicit output within an expected range of performance such that scoring or rating procedures apply appropriately. For example, in a picture series task, the objective of which is to elicit a story in a sequence of events, test-takers could opt for a variety of plausible ways to tell the story, all of which might be equally accurate. How can such disparate responses be evaluated? One solution is to assign not one but several scores for each response, each score representing one of several traits (pronunciation, fluency, vocabulary use, grammar, comprehensibility, etc.).

All of these issues will be addressed in this chapter as we review types of spoken language and micro and macro skills of speaking, then outline numerous tasks for assessing speaking.

II. BASIC TYPES OF SPEAKING

1. Imitative

At one end of a continuum of types of speaking performance is the ability to simply parrot back (imitate) a word or phrase or possibly a sentence. While this is a purely phonetic level of oral production, a number of prosodic, lexical, and grammatical properties of language may be included in the criterion performance. We are interested only in what is traditionally labeled „pronunciation”; no inferences are made about the test-taker's ability to understand or convey meaning or to participate in an interactive conversation. The only role of listening here is in the short term storage of a prompt, just long enough to allow the speaker to retain the short stretch of language that must be imitated.

2. Intensive

A second type of speaking frequently employed in assessment contexts is the production of short stretches of oral language designed to demonstrate competence in a narrow band of grammatical, phrasal, lexical, or phonological relationships (such as prosodic elements – intonation, stress, rhythm, juncture). The speaker must be aware of semantic properties in order to be able to respond, but interaction with an interlocutor or test administrator is minimal at best. Examples of intensive assessment tasks include directed response tasks, reading aloud, sentence and dialogue completion; limited. Picture-cued tasks including simple sequences; and translation up to the simple sentence level.

3. Responsive

Responsive assessment tasks include interaction and test comprehension but at the somewhat limited level of very short conversations, standard greetings and small talk, simple requests and comments, and the like. The stimulus is almost always a spoken prompt (in order to preserve authenticity), with perhaps only one or two follow up questions or retorts:

- A. Mary: Excuse me, do you have the time?
 Doug: Yeah. Ninefifteen.
- B. T: What is the most urgent environmental problem today?
 S: I would say massive deforestation.
- C. Jeff: Hey, Stef, how's it going?
 Stef: Not bad, and yourself?
 Jeff: I'm good.
 Stef: Cool. Okay, gotta go.

4. Interactive

The difference between responsive and interactive speaking is in the length and complexity of the interaction, which sometimes includes multiple exchanges and/or multiple participants. Interaction can take the two forms of transactional language, which has the purpose of exchanging specific information, or interpersonal exchanges, which have the purpose of maintaining social relationships. (In the three dialogues cited above, A and B were transactional, and C was interpersonal.) In interpersonal exchanges, oral production can become pragmatically complex with the need to speak in a casual register and use colloquial language, ellipsis, slang, humor, and other sociolinguistic conventions.

5. Extensive (monologue)

Extensive oral production tasks include speeches, oral presentations, and storytelling, during which the opportunity for oral interaction from listeners is either highly limited (perhaps to nonverbal responses) or ruled out altogether. Language style is frequently more deliberative (planning is involved) and formal for extensive tasks, but we cannot

rule out certain informal monologues such as casually delivered speech (for example, my vacation in the mountains, a recipe for outstanding pasta primavera, recounting the plot of a novel or movie).

III. MICRO AND MACRO SKILLS OF SPEAKING

Micro and macro skills enumerated the various components of listening that make up criteria for assessment. A similar list of speaking skills can be drawn up for the same purpose: to serve as a taxonomy of skills from which you will select one or several that will become the objective(s) of an assessment task. The micro skills refer to **producing the smaller chunks of language such as phonemes, morphemes, words, collocations, and phrasal units**. The macro skills imply the speaker's focus on the larger elements: **fluency, discourse, function, style, cohesion, nonverbal communication, and strategic options**. The micro and macro skills total roughly 16 different objectives to assess in speaking.

Micro skills

1. Produce differences among English phonemes and allophonic variants.
2. Produce chunks of language of different lengths.
3. Produce English stress patterns, words in stressed and unstressed positions, rhythmic structure, and intonation contours.
4. Produce reduced forms of words and phrases.
5. Use an adequate number of lexical units (words) to accomplish pragmatic purposes.
6. Produce fluent speech at different rates of delivery.
7. Monitor one's own oral production and use various strategic devices – pauses, fillers, self-corrections, back tracking – to enhance the clarity of the message.
8. Use grammatical word classes (nouns, verbs, etc.) systems (e.g., tense, agreement, pluralization), word order, patterns, rules, and elliptical forms.
9. Produce speech in natural constituents: in appropriate phrases, pause groups, breath groups, and sentence constituents.
10. Express a particular meaning in different grammatical forms.
11. Use cohesive devices in spoken discourse.

Macro skills

12. Appropriately accomplish communicative functions according to situations, participants, and goals.
13. Use appropriate styles, registers, implicature, redundancies, pragmatic conventions, conversation rules, floor keeping and –yielding, interrupting, and other sociolinguistic features in face toface conversations.
14. Convey links and connections between events and communicate such relations as focal and peripheral ideas, events and feelings, new information, and given information, generalization and exemplification.

15. Convey facial features, kinesics, body language, and other nonverbal cues along with verbal language.
16. Develop and use a battery of speaking strategies, such as emphasizing key words, rephrasing, providing a context for interpreting the meaning of words, appealing for help, and accurately assessing how well your interlocutor is understanding you.

As you consider designing tasks for assessing spoken language, these skills can act as a checklist of objectives. While the macro skills have the appearance of being more complex than the micro skills, both contain ingredients of difficulty, depending on the stage and context of the test-taker.

There is such an array of oral production tasks that a complete treatment is almost impossible within the confines of one chapter in that book. Below is a consideration of the most common techniques with brief allusions to related tasks. As already noted in the introduction to this chapter, consider three important issues as you set out to design tasks:

1. No speaking task is capable of isolating the single skill of oral production. Concurrent involvement of the additional performance of aural comprehension, and possibly reading, is usually necessary.
2. Eliciting the specific criterion you have designated for a task can be tricky because beyond the word level, spoken language offers a number of productive options to test-takers. Make sure your elicitation prompt achieves its aims as closely as possible.
3. Because of the above two characteristics of oral production assessment, it is important to carefully specify scoring procedures for a response so that ultimately you achieve as high a reliability index as possible.

IV. DESIGNING SPEAKING ASSESSMENT

1. Imitative Speaking Assessments

a. *Word repetition Task (repeat after me...)*

Test-takers hear:

beat/bit bat/vat

I bought a boat yesterday.

The glow of the candle is growing.

Test-takers repeat the stimulus

b. *Phonepass Test (VERSANT Exe)*

It elicits computer-assisted oral production over a telephone. Test-takers read aloud, repeat sentences, say words, and answer questions.

Part A: read aloud selected sentences.

Examples: Traffic is a huge problem in Southern California.

Part B: repeat sentences dictated over the phone.

Example: Leave town on the next train.

Part C: Answer questions with a single word or a short phrase.

Example: Would you get water from a bottle or a newspaper?

Part D: hear three word groups in random order and link them in a correctly ordered sentence.

Example: was reading/my mother/a magazine

Part E: have 30 seconds to talk about their opinion about some topic that is dictated over the phone.

Topics center on family, preferences, and choices

Scores are calculated by a computerized scoring template and reported back to the test-taker within minutes.

Scoring scale for repetition tasks

- 2 acceptable pronunciation.
- 1 comprehensible, partially correct.
- 0 silence, seriously incorrect.

Pict.2 Scoring scale

2. Intensive Speaking Assessment

a. *Directed Response Tasks*

Directed response Elicits a particular grammatical form/transformation of a sentences :



TEST INSTRUCTIONS

PLEASE READ THIS BEFORE TAKING THE TEST

Versant tests are automated spoken language tests that are taken on the telephone or computer. If you would like to listen to a sample test, purchase a practice test, or view the test score after taking the test (if applicable), please visit www.VersantTest.com

PART	INSTRUCTIONS
Before the Test	<ul style="list-style-type: none"> • Carefully read this instruction page and the test paper. You may use a dictionary or ask someone for help if there are words or sentences that you don't understand. • Choose a quiet location with a landline phone where you will not be interrupted during the test. • Do not use a cordless phone, cellular phone, or VoIP phone (e.g., Skype™ or PC-to-phone services). Newer phones are generally better than older phones. Make sure that the phone is set to tone and not pulse.
Beginning the Test	<ul style="list-style-type: none"> • To begin the test, call the phone number on the test paper using a landline push-button telephone. • A recorded examiner's voice will guide you through each section of the test. • Enter your Test Identification Number using the telephone keypad when the examiner's voice asks you to do so. This number is printed on the top right of your test paper. • The examiner's voice will then ask you two questions: your name, and the country you are calling from. If you are speaking too loudly or too quietly, the examiner's voice will tell you. • The test begins when you say your name. If you hang up before you complete the test, the test cannot be graded. You cannot reuse the Test Identification Number.
During the Test	<ul style="list-style-type: none"> • Hold the phone close to your mouth as shown in the picture below. <div style="text-align: center;">  </div> <ul style="list-style-type: none"> • Answer all questions smoothly and naturally in a clear, steady voice. • If you don't know the proper way to respond to a test item, you can remain silent or say, 'I don't know.' • Do not take notes or write during the test. • When you hear, "Thank you for completing the test", you may hang up. • If you wish, you may answer the optional questions at the end of the test. Your personal information will be kept anonymous.

PEARSON

© 2008 Pearson Education, Inc. or its affiliate(s). All rights reserved. Ordinate and Versant are trademarks in the U.S. and/or other countries, of Pearson Education, Inc. or its affiliate(s).

Pict.1 Versant speaking Assessment

Tell me he went home.
 Tell me that you like rock music.
 Tell me that you aren't interested in tennis.
 Tell him to come to my office at noon.
 Remind him what time it is.

b. Read-Aloud task

This technique is easily administered by selecting a passage that incorporates test specs and by recording the test-takers' output.

The scoring scales are :

- Pronunciation Points:
 - 0.0—0.4 frequent errors and unintelligible.
 - 0.5—1.4 occasionally unintelligible.
 - 1.5—2.4 some errors but intelligible.
 - 2.5—3.0 occasional errors but always intelligible
- Fluency
 - 0.0– 0.4 slow, hesitant, and unintelligible.
 - 0.5– 1.4 non-native pauses and flow that interferes with intelligibility.
 - 1.5--2.4 non-native pauses but the flow is intelligible.
 - 2.5—3.0 smooth and effortless.

The variations could be :

- a. Reading a scripted dialogue.
- b. Reading sentences containing minimal pairs. Examples: Try not to heat/ hit the pan too much.
- c. Reading information from a table or chart

c. Sentence/Dialogue Completion Task and Oral Questionnaires

- First, test-takers are given time to read through the dialogue to get its gist (main point), then the tape/teacher produces one part orally and the test-taker responds.
- Example

Salesperson : May I help you
 Customer : _____ (test takers respond with appropriate lines)
- **Advantage:** more time to anticipate an answer, no potential ambiguity created by aural misunderstanding (oral interview).
- **Disadvantage :** It is inauthentic, except in situations such as parent reading to a child, sharing a story with someone, giving a scripted oral presentation. It is not communicative in real contexts.

d. Picture-cued task

A picture-cued stimulus requires a description from the test-taker. It may elicit a word, a phrase, a story, or incident.

The types are :

- Picture-cued elicitation of minimal pairs
- Picture-cued elicitation of comparatives
- Picture-cued elicitation of future tense
- Picture-cued elicitation of nouns, negative responses, numbers, location
- Picture-cued elicitation of responses and description
- Picture-cued elicitation of giving directions
- Picture-cued elicitation of multiple choice description for two tests takers

e. *Translation*

Translation is a communicative device in contexts where English is not a native language. English can be called on to be interpreted as a second language. Conditions may vary from an instant translation of a native word, phrase, or sentence to a translation of longer texts.

Advantages: the control of the output & easily specified scoring.

3. Responsive Speaking Assessment

a. *Question and Answer*

Question and answer tasks can consist of one or two questions from an interviewer, or they can make up a portion of a whole battery of questions and prompts in an oral interview.

- They can vary from simple questions to complex questions :

- * “What is this called in English?” → *simple question*

- * “What are the steps governments should take, if any, to stem the rate of deforestation in tropical countries?” → *complex question*

- The first question is **intensive** in its purpose. It is a **display question** intended to elicit a predetermined correct response.

Questions at the **responsive level** tend to be **genuine referential questions** in which the test-taker is given more opportunity to produce meaningful language in response.

Responsive questions may take the following forms:

- Questions eliciting open-ended responses
- Elicitation of questions from the test-taker
- Test-takers respond with questions

b. *Giving Instructions and Directions*

The technique is simple: the administrator poses the problem, and the test-taker responds. Scoring is based primarily on comprehensibility and secondarily on other specified grammatical or discourse categories.

c. *Paraphrasing*

The test-takers read or hear a limited number of sentences (perhaps two or five) and produce a paraphrase of the sentence. The forms are :

1) **Paraphrasing a story**

Test-takers hear: Paraphrase the following little story in your own words.

My weekend in the mountains was fabulous. The first day we backpacked into the mountains and climbed about 2.000 feet. The hike was strenuous but exhilarating. By sunset we found these beautiful alpine lakes and made camp there. The sunset was amazingly beautiful. The next two days we just kicked back and did little day hikes, some rock climbing, bird watching, swimming, and fishing. The hike out on the next day was really easy – all downhill – and the scenery was incredible.

Test-takers respond with two or three sentences.

2) **Paraphrasing a phone message**

Test-taker hear:

please tell Jeff that I'm tied up in traffic so I'm going to be about a half late for the nine o'clock meeting. And ask him to bring up our question about the employee benefits plan. If he wants to check in with me on my cell phone, have him call 415-338-3095. Thanks.

d. *TEST OF SPOKEN ENGLISH (TSE)*

- The TSE is a 20 minute audio-taped test of oral language ability within an academic or professional environment.
- Lazaraton and Wagner (1996) examined 15 different specific tasks in collecting background data from native and non-native speakers of English.
 1. giving a personal description
 2. describing a daily routine
 3. suggesting a gift and supporting one's choice
 4. recommending a place to visit and supporting one's choice
 5. giving directions
 6. describing a favorite movie and supporting one's choice
 7. telling a story from pictures
 8. hypothesizing about future action
 9. hypothesizing about a preventive action
 10. making a telephone call to the dry cleaner
 11. describing an important news event
 12. giving an opinion about animals in the zoo
 13. defining a technical term
 14. describing information in a graph and speculating about its implications
 15. giving details about a trip schedule

From their findings, the researchers were able to report on the validity of the tasks, especially the match between the intended task functions and the actual output of both native and non-native speakers.

TSE Rating Scale

- 60.** Communication almost always effective: task performed very competently; speech almost never marked by non-native characteristics
- 50.** Communication generally effective: task performed competently, successful use of compensatory strategies; speech sometimes marked by on-native characteristics
- 40.** Communication somewhat effective: task performed somewhat competently, some successful use of compensatory strategies; speech regularly marked by non-native characteristics
- 30.** Communication generally not effective: task generally performed poorly, ineffective use of compensatory strategies; speech very frequently marked by non-native characteristics
- 20.** No effective communication: no evidence of ability to perform task, no effective use of compensatory strategies; speech almost always marked by non-native characteristics

4. Interactive Speaking Assessment

Interactive speaking is more interpersonal described and more transactional focused speech event

a. INTERVIEW

Oral production assessment that is words. This effective interview contains a number of mandatory stages (Michael Canale: 1984). The steps are :

1) WARM – UP

- a. The interviewer direct mutual introductions, helps the test-taker become comfortable with the situation, appraises the test – taker of the format and allays anxieties.
- b. No scoring takes place.

2) LEVEL CHECK

- a. The interviewer stimulates the test taker to respond using expected or predicted forms and functions.
- b. Question are design to elicit grammatical categories, discourse structure, vocabulary usage and/or sociolinguistic factors.
- c. This stage give the interviewer a picture of the test-taker's extroversion, readiness to speak and confidence.
- d. Linguistic target criteria are scored in this phase.

3) PROBE

- a) Challenge test-takers to go to the heights of their ability, to extend beyond the limits of the interviewer's expectation through increasingly difficult questions.

- b) Through probe items, the interviewer discovers the ceiling or limitation of the test-taker proficiency.
- c) At the lower level of proficiency : It may simply demand higher range of vocabulary and grammar.
- d) At the higher level of proficiency : It will typically ask the test taker to give an opinion or a value judgment, to discuss his/her field of specialization, to recount a narrative, or to respond to questions that are wondered in complex form.
- 4) **WIND-DOWN**
- a) The interviewer encourages the test-taker to relax with some easy questions
- b) This part is not scored.

Two kinds of interview

- a) Placement Interview : Designed to get a quick spoken sample from a student in order to verify placement into course

Level	Grammar (20%)	Vocabulary (20%)	Comprehension (20%)	Fluency (10%)	Pronunciation (10%)	Task (20%)
E (0-54)	Errors in grammar are frequent, but speaker can be understood by a native speaker used to dealing with foreigners attempting to speak his language	Speaking vocabulary inadequate to express anything but the most elementary needs	Within the scope of his very limited language experience, can understand simple questions and statements if delivered with slowed speech, repetition, or paraphrase.	(No specific fluency description. Refer to other four language areas for implied level of fluency.)	Errors in pronunciation are frequent but can be understood by a native speaker used to dealing with foreigners attempting to speak his language	Can ask and answer questions on topics very familiar to him. Able to satisfy routine travel and minimum courtesy requirements. (Should be able to order a simple meal, ask for shelter or lodging, ask and give simple directions, make purchases, and tell time.)
D (55-64)	Can usually handle elementary constructions quite accurately but does not have thorough or confident control of the grammar.	Has speaking vocabulary sufficient to express himself simply with some circumlocutions.	Can get the gist of most conversations of non-technical subjects (i.e., topics that require no specialized knowledge).	Can handle with confidence but not with facility most social situations, including introductions and casual conversations about current events, as well as work, family, and autobiographical information.	Accent is intelligible though often quite faulty.	Able to satisfy routine social demands and work requirements; need help in handling any complication or difficulties.

C (65-74)	Control of grammar is good. Able to speak the language with sufficient structural accuracy to participate effectively in most formal and informal conversations on practical, social, and professional topics.	Able to speak the language with sufficient vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Vocabulary is broad enough that he rarely has to grope for a word.	Comprehension is quite complete at a normal rate of speech.	Can discuss particular interests of competence with reasonable ease. Rarely has to grope for words.	Errors never interfere with understanding and rarely disturb the native speaker. Accent may be obviously foreign.	Can participate effectively in most formal and informal conversation on practical, social, and professional topics.
B (75-84)	Able to use the language accurately pertinent to professional needs. Errors in grammar are quite rare.	Can understand and participate in any conversations within the range of his experience with a high degree of precision of vocabulary.	Can understand any conversations within the range of his experience	Able to use the language fluently on all levels normally pertinent to professional needs. Can participate in any conversations within the range of his experience with a high degree of fluency.	Errors in pronunciation are quite rare.	Would rarely be taken for a native speaker but can respond appropriately even in unfamiliar interpreting from and into language.
A (85-100)	Equivalent to that of an educated native speaker.	Speech on all levels is fully accepted by educated native speakers in all its features including breadth of vocabulary and idioms, colloquialisms, and pertinent cultural references.	Equivalent to that of an educated native speaker	Has complete fluency in the language such that his speech is fully accepted by educated native speaker.	Equivalent to and fully accepted by educated native speakers.	Speaking proficiency equivalent to that of an educated native speaker

Score	On functional score tracking
0	Unable to function in the spoken language
0+	Able to satisfy immediate needs using rehearsed utterances
1	Able to satisfy minimum courtesy requirements and maintain very simple face to face conversation on familiar topics
1+	Can initiate and maintain predictable face-to face conversations and satisfy limited social demands
2	Able to satisfy routine social demands and limited social demands and limited work requirements
2+	Able to satisfy most work requirements with language usage that is often, but not always, acceptable and effective
3	Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics
3+	Often able to use the language to satisfy professional needs in a wide range of sophisticated and demanding tasks
4	Able to use language fluently and accurately on all levels normally pertinent to professional needs
4+	Speaking proficiency is regularly superior in all aspects, usually equivalent to that of a well-educated, highly articulate native speaker
5	Speaking proficiency is functionally equivalent to that of a highly accurate, well-educated native speaker and reflects the cultural standards of the country where the language is spoken

The success of an oral interview will depend on :

1. Clearly specifying administrative procedures of the assessment (practicality)
2. Focusing the questions and probes on the purpose of the assessment (validity).
3. Appropriately eliciting an optimal amount and quality of oral production from the test-taker (biased for best performance).
4. Creating a consistent, workable scoring system (reliability).

b) Comprehensive Interview (OPI) : Designed to cover predetermined oral production contexts and may require the better part of an hour.

b. *Oral Proficiency Interview (OPI)*

- Originally known as the Foreign Service Institute (FSI) test.
- OPI is designed to elicit pronunciation, fluency and integrative ability, sociolinguistic and cultural knowledge, grammar, and vocabulary.
- Performance is judged by the examiner to be at one of ten possible levels on the ACTFL- designed proficiency guidelines for speaking : superior – advanced – intermediate – Novice

Chart			
SUMMARY HIGHLIGHTS ACTFL PROFICIENCY GUIDELINES—SPEAKING (REVISED 1999)			
SUPERIOR	ADVANCED	INTERMEDIATE	NOVICE
<p>Superior-level speakers are characterized by the ability to:</p> <ul style="list-style-type: none"> • participate fully and effectively in conversations in formal and informal settings on topics related to practical needs and areas of professional and/or scholarly interests • provide a structured argument to explain and defend opinions and develop effective hypotheses within extended discourse • discuss topics concretely and abstractly • deal with a linguistically unfamiliar situation • maintain a high degree of linguistic accuracy • satisfy the linguistic demands of professional and/or scholarly life 	<p>Advanced-level speakers are characterized by the ability to:</p> <ul style="list-style-type: none"> • participate actively in conversations in most informal and some formal settings on topics of personal and public interest • narrate and describe in major time frames with good control of aspect • deal effectively with unanticipated complications through a variety of communicative devices • sustain communication by using, with suitable accuracy and confidence, connected discourse of paragraph length and substance • satisfy the demands of work and/or school situations 	<p>Intermediate-level speakers are characterized by the ability to:</p> <ul style="list-style-type: none"> • participate in simple, direct conversations on generally predictable topics related to daily activities and personal environment • create with the language and communicate personal meaning to sympathetic interlocutors by combining language elements in discrete sentences and strings of sentences • obtain and give information by asking and answering questions • sustain and bring to a close a number of basic, uncomplicated communicative exchanges, often in a reactive mode • satisfy simple personal needs and social demands to survive in the target language culture 	<p>Novice-level speakers are characterized by the ability to:</p> <ul style="list-style-type: none"> • respond to simple questions on the most common features of daily life • convey minimal meaning to interlocutors experienced with dealing with foreigners by using isolated words, lists of words, memorized phrases and some personalized recombinations of words and phrases • satisfy a very limited number of immediate needs
© ACTFL, Inc., 1999			

Criticism on OPI

● Valdman (1988,p.125)

From a Vygotskyan perspective, the OPI forces test-takers into a closed system where, because the interviewer is endowed with full social control, they are unable to negotiate a social world. For example, they cannot nominate topics for discussion, they cannot switch formality levels, they cannot display a full range of stylistic maneuver. The total control the OPI interviewers possess is reflected by the parlance of the test methodology... In short, the OPI can only inform us of how learners can deal with an artificial social imposition rather than enabling us to predict how they would be likely to manage authentic linguistic interactions with target-language native speaker.

● Bachman (1988,p.149)

pointed out that the validity of the OPI simply cannot be demonstrated “because it confounds abilities with elicitation procedures in its design, and it provides only a single rating, which has no basis in either theory or research.”

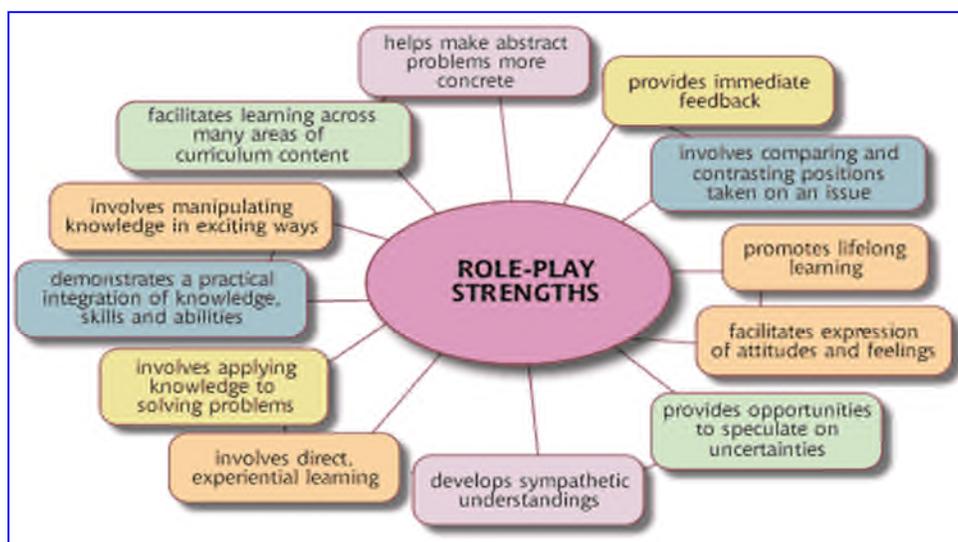
The Phases of the OPI

The Warm Up	Iterative Process		The Wind Down
	The Level Checks	→ The Probes	
<p>This first phase of the OPI serves as the introduction to the interview. It consists of greetings, informal exchanges of pleasantries, and conversation openers pitched at a level which appears comfortable for the speaker. Every OPI begins with the assumption that a conversation will take place (Intermediate Level).</p>	<p>When the speaker has settled into the interview and appears to be reasonably comfortable using the target language, the interviewer moves to the next phase of the OPI, the level checks. The interviewer engages the speaker in conversation on several topics of interest so that the tasks characterizing any given level can be performed. Level checks are questions that elicit the performance floor, the linguistic tasks and contexts of a particular level which can be handled successfully.</p>	<p>Once the interviewer has begun to establish that the speaker can handle the tasks and topics of a particular level, the interview proceeds to the next phase, the probes. The purpose of the probes is to discover the ceiling or limits of the speaker's proficiency, i.e. the patterns of weakness. This is done by raising the level of the interview to the next higher major level in an attempt to discover the level at which the speaker can no longer sustain functional performance.</p>	<p>The final phase returns the speaker to a comfortable level of language exchange and ends the OPI on a positive note.</p>
<p>↑ ————— The Role Play ————— ↑ A transactional or social situation can serve as either an additional level check or probe as needed in a particular interview.</p>			

Proficiency Level*	Global Tasks and Functions	Context / Content	Accuracy	Text Type
Superior	Discuss topics extensively, support opinions and hypothesize. Deal with a linguistically unfamiliar situation.	Most formal and informal settings. <i>Wide range of general interest topics and some special fields of interest and expertise.</i>	No pattern of errors in basic structures. Errors virtually never interfere with communication or distract the native speaker from the message.	Extended discourse
Advanced	Narrate and describe in major time frames and deal effectively with an unanticipated complication.	Some informal settings and a limited number of transactional situations. <i>Predictable, familiar topics related to daily activities.</i>	Understood, with some repetition, by speakers accustomed to dealing with non-native speakers.	Paragraphs
Intermediate	Create with language, initiate, maintain, and bring to a close simple conversations by asking and responding to simple questions.	Some informal settings and a limited number of transactional situations. <i>Predictable, familiar topics related to daily activities.</i>	Understood, with some repetition, by speakers accustomed to dealing with non-native speakers.	Discrete sentences
Novice	Communicate minimally with formulaic and rote utterances, lists, and phrases	Most common informal settings. <i>Most common aspects of daily life.</i>	May be difficult to understand, even for speakers accustomed to dealing with non-native speakers.	Individual words and phrases

c. *Role Play*

- 1) It frees students to be more creative in their linguistic output.
- 2) It Allows some rehearsal time so that students can map out what they are going to say.
- 3) It has the effect of lowering anxieties as students can.
- 4) As an assessment device, role play opens some windows of opportunity for test taker to use discourse that might otherwise be difficult to elicit.
- 5) Certain personal, strategic, and linguistic factors come into the foreground of the test-taker's oral abilities.
- 6) Role play takes test-takers beyond simple intensive and responsive levels to a level of creativity and complexity that approaches real-world pragmatics.



d. *Discussions and Conversations*

Discussion may be especially appropriate tasks through which to elicit and observe such abilities as:

- 1) Topic nomination, maintenance and termination
- 2) Attention getting, interrupting, floor holding, control;
- 3) Clarifying, questioning paraphrasing
- 4) Comprehension signals (nodding, “uh-uhh”, “hmm”, etc)
- 5) Negotiating meaning
- 6) Intonation patterns for pragmatic effect
- 7) Kinesics, eye contact, proxemics, body language
- 8) Politeness, formality and other sociolinguistic factors.

e. *Games*

Such tasks have wandered away from the traditional notion of an oral production test and may even be well beyond assessment. As assessment, the key is to specify a set of criteria and a reasonably practical and reliable scoring method and technique. (Microsoft Encarta, encyclopaedia Britannica, Rosseta Stone)

5. Extensive Speaking Assessment

Extensive speaking tasks involve complex, relatively lengthy stretches of discourse. They are frequently variations on monologues, usually with minimal verbal interaction.

a. *ORAL PRESENTATION*

The rules for effective assessment must be invoked :

- 1) Specify the criterion
- 2) Set appropriate tasks
- 3) Elicit optimal output

4) Establish practical, reliable scoring procedures

Evaluation of oral presentation	
Assign a number to each box according to your assessment of the various aspects of the speaker's presentation.	
3	Excellent
2	Good
1	Fair
0	Poor
Content:	
<input type="checkbox"/> The purpose or objective of the presentation was accomplished.	
<input type="checkbox"/> The introduction was lively and got my attention.	
<input type="checkbox"/> The main idea or point was clearly stated toward the beginning.	
<input type="checkbox"/> The supporting points were	
• clearly expressed	
• supported well by facts, argument	
<input type="checkbox"/> The conclusion restated the main idea or purpose.	
Delivery:	
<input type="checkbox"/> The speaker used gestures and body language well.	
<input type="checkbox"/> The speaker maintained eye contact with the audience.	
<input type="checkbox"/> The speaker's language was natural and fluent.	
<input type="checkbox"/> The speaker's volume of speech was appropriate.	
<input type="checkbox"/> The speaker's rate of speech was appropriate.	
<input type="checkbox"/> The speaker's pronunciation was clear and comprehensible.	
<input type="checkbox"/> The speaker's grammar was correct and didn't prevent understanding.	
<input type="checkbox"/> The speaker used visual aids, handouts, etc., effectively.	
<input type="checkbox"/> The speaker showed enthusiasm and interest.	
<input type="checkbox"/> [If appropriate] The speaker responded to audience questions well.	

- Oral Presentation picture : a checklist or grid is a common means of scoring or evaluation.
- Holistic scores are tempting to use for their apparent practicality, but they may obscure the variability of performance across several subcategories, especially they two major components of content and delivery.
- Its authenticity can be supported in that all of the items on the list contribute to an effective presentation.
- The washback effect of such a checklist will be enhanced by written comments from the teacher, a conference with the teacher, peer evaluations using the same form, and self-assessment.

b. *Pictured–Cued Story –Telling*

One of the most common techniques for eliciting oral production is through visual pictures, photographs, diagrams, and charts.

Picture-cued story-telling task (H. D. Brown, 1999, p. 29) [L/R, S]

Test-takers see the following six-picture sequence:

Test-takers hear or read: Tell the story that these pictures describe.
 Test-takers use the pictures as a sequence of cues to tell a story.

c. *Retelling a Story, News Event*

- 1) Test-takers hear or read a story or news event that they are asked to retell.
- 2) It is a longer stretch of discourse and a different genre.
- 3) The objective is to produce a number of oral discourse features (communicating sequences and relationships of events, stress and emphasis patterns, "expression" in the case of a dramatic story), fluency, and interaction with the hearer/ listener.

d. *Translation (of Extended Prose)*

- 1) Translation of words, phrases, or short sentences was mentioned under the category of intensive speaking. Here, longer texts are presented for the test-taker to read in the native language and then translate into English.
- 2) Those texts could come in many forms: dialogue, directions for assembly of a product, a synopsis of a story or play or movie, directions on how to find something on a map, and other genres.

- 3) The advantage of translation is in the control of the content, vocabulary, and to some extent, the grammatical and discourse feature.
- 4) The disadvantage is that translation of longer texts is a highly specialized skill