



Evaluating Spoken Language

Part I: Oral Proficiency Interviews (OPI)



Procedures – Interviews

1. **Warm-up** – Purpose: put test-taker at ease and indicate level of proficiency of test-taker
Time: Less than three minutes
2. **Level-check** - Purpose: check functions and levels where test-taker is most accurate.
Time: Varies - depends on time it takes to elicit adequate sample to support level hypothesized by tester. (Tester establishes floor)



Procedures - Interviews

3. **Probing** - Purpose: assess highest level at which candidate can function.
Time varies: depends on time to elicit adequate sample to establish ceiling for test-taker. Probes may result in linguistic breakdown.
(Probes alternate with level checks).
4. **Wind-up** - Purpose: take test-taker to level where he/she is most comfortable.
Time: Same as warm-up. Opportunity to verify rating.

(Total time: varies from 15 to 30 minutes).



Grading Rubric - OPI

Single-format test: Holistic Rating

- **Novice**: Minimal language - mostly memorized material
 - a) formulaic chunks and lists
- **Intermediate**: Low level of survival skills
 - a) combine and recombine learned elements
 - b) initiate and close basic tasks
 - c) ask and answer questions

Details – Hadley pp. 16-19



Grading Rubric - OPI

- **Intermediate-Mid**: communicates about high frequency topics
 - a) talk about self and family
 - b) ask and answer questions and participate in simple conversations
(Personal history and leisure time activities)
 - c) long pauses frequent
 - d) pronunciation strongly influenced by L1
- **Advanced**: converses in a participatory fashion
 - a) initiates, sustains, and closes a variety of tasks using diverse language strategies
 - b) satisfies requirements of school and work
 - c) narrates and describes at paragraph-length connected discourse
(Present, past and future time frames).



Grading Rubric - OPI

- **Superior** – Purpose: participate effectively in formal and informal conversations
 - a) participates in conversations on wide variety of topics.
(Practical, social, professional, abstract topics).
 - b) supports opinions and hypothesizes using native-like discourse strategies

Details – Hadley pp. 16-19

Expected Levels of Proficiency - FSI

Group I: Afrikaans, Danish, Dutch, French, Haitian Creole, Italian, Norwegian, Portuguese, Romanian, Spanish, Swahili, Swedish

Length of Training	Aptitude for Language Learning		
	Minimum	Average	Superior
8 weeks (240 hours)	1	1/1+	1+
16 weeks (480 hours)	1+	2	2+
24 weeks (720 hours)	2	2+	3

Group II: Bulgarian, Dari, Farsi, German, Greek, Hindi, Indonesian, Malay, Urdu

16 weeks (480 hours)	0+	1	1/1+
24 weeks (720 hours)	1+	2	2/2+
44 weeks (1320 hours)	2	2+	3

Expected Levels of Proficiency - FSI

Group III: Amharic, Bengali, Burmese, Czech, Finnish, Hebrew, Hungarian, Khmer, Lao, Nepali, Philipino, Polish, Russian, Serbo-Croatian, Sinhala, Thai, Tamil, Turkish, Vietnamese

Length of Training	Aptitude for Language Learning		
	Minimum	Average	Superior
16 weeks (480 hours)	0+	1	1/1+
24 weeks (720 hours)	1+	2	2/2+
44 weeks (1320 hours)	2	2+	3

Expected Levels of Proficiency - FSI

Group IV: Arabic, Chinese, Japanese, Korean

Length of Training	Aptitude for Language Learning		
	Minimum	Average	Superior
16 weeks(480 hours)	0+	1+	1
24 weeks (720 hours)	1	1+	1+
44 weeks (1320 hours)	1+	2	2+
80-92 weeks (2400-2760)	2+	3	3+

ETS Oral Proficiency Testing Manual

Evaluating Spoken Language

Part II: Israeli National Oral Proficiency Test (*INOPT*)

Israeli National Oral Proficiency Test *INOPT*

Multicomponential format

Each format:

- elicits a different speech style.
- discriminates among different levels of oral proficiency.
- unable to reliably predict performance on another.

Four Formats of *INOPT*

- Test 1: Oral Interview - Same as OPI
- Test 2: Role Play: (L1 = Hebrew)
 - Purpose - simulate conversation derived from situation given on a card.
 - a) Test-taker plays one role and tester plays the other role.
 - b) Test-taker given card in L1 describing situation and his/her role in it.
 - c) They engage in simulated conversation.

(All formats scored on the same rating scale below).

Four Formats of INOPT

- Test 3: Reporting Test: (L1 = Hebrew)
Purpose - stimulate test-taker into L2 monologue
 - a) Test-taker given newspaper article in L1
 - b) Test-taker reads article silently.
 - c) Test-taker reports on contents of article to test-giver in his/her own words.

(All formats graded on the same grading scale below)

Four Formats of INOPT

- Test 4: Group Discussion:
Purpose - stimulate spontaneous discussion of a controversial issue.
Test-takers express views on topical matters, debate and argue about them, defend their opinions and try to persuade the other participants to accept them.
 - a) Members (4) pick up card from table.
 - b) Card gives information on topic and cues for discussion, such as guiding questions with relevant lexical items.
 - c) Test-giver listens and scores test-takers.

(All formats graded on same scale shown below)

Grading Rubric - INOPT

- 10: No limitation whatsoever
Near Native
- 9: Facility of expression
Comfortable, initiating in interaction
Sporadic mistakes
- 8: Almost effortless in expression
Adequate in interaction
Errors: not consistent

Grading Rubric - INOPT

- 7: Responsive in interaction
Slightly more sophisticated language produced
Consistent errors: but do not interfere with fluency
Strong MT (mother tongue) interference (translated patterns, etc.)
- 6: Clearly intelligible
Simple language produced
Interaction possible
Not articulate

Grading Rubric - INOPT

- 5: Hardly intelligible
Very poor language produced
Only simplest, fragmentary interaction possible
- 4: Unintelligible
No language produced
No interaction possible

Evaluating Spoken Language

Part III: Classroom Oral Achievement Interviews

Four Criteria for Designing Tests

- ◆ Economy
- ◆ Relevance
- ◆ Acceptability
- ◆ Comparability

Washback Effect

1. Teachers incorporate activities into classroom instruction that they feel will lead students to success on the test.

Therefore, tests influence what teachers teach and how they teach it.

2. If students know they will be tested on a given instructional activity, they are motivated to participate more fully in the activity

Goals for Paso A – Dos Mundos

Students will be able to:

- ◆ describe themselves or a friend/family member etc.
- ◆ understand and use simple classroom commands.
- ◆ recognize and use numbers from 1-30.
- ◆ participate in brief situations that require greetings and farewells.

Rubric - Oral Achievement Interview

McCluer North High School

- ◆ 10 Excellent: Total comprehension; immediate response;
Conversational, glib, correct
Obvious preparation.
- ◆ 9 Good+ Comprehends - few-prompts; immediate response
Searches a bit for word(s);
Accurate in majority of vocabulary and structures;
Prepared
- ◆ 8 Good Understands - unsure until prompted for context;
Reduced accuracy in vocabulary.

Oral Achievement Interview

McCluer North High School

- ◆ 7 Average Reduced comprehension: depends on evaluator for contexts and prompts.
Inconsistent accuracy in responses with regard to vocabulary and structure.
- ◆ 6 Average Marginal comprehension; requires prompts of all kinds, delayed responses
Unsure; inaccurate vocabulary and structures
- ◆ 5 Poor Misunderstands; unable to follow prompts;
Gross inaccuracies in vocabulary and structure;
Totally unprepared.

Rubric for Pair Interviews

Name _____	Name _____
Notes: _____	Notes: _____
Pronunciation:	Pronunciation:
F D C B A x 4 = _____	F D C B A x 4 = _____
Vocabulary	Vocabulary
F D C B A x 7 = _____	F D C B A x 7 = _____
Grammar	Grammar
F D C B A x 6 = _____	F D C B A x 6 = _____
Fluency	Fluency
F D C B A x 3 = _____	F D C B A x 3 = _____
A = 4.5 - 5.0 Total = _____	A = 4.5 - 5.0 Total = _____
B = 4.0 - 4.4	B = 4.0 - 4.4
C = 3.5 - 3.9	C = 3.5 - 3.9
D = 3.0 - 3.4	D = 3.0 - 3.4