

# Reading and Listening Proficiency Test Development for LCTLs

Arabic, Hebrew, Persian and Turkish as a Case Study

Interagency Language Roundtable Washington, D.C.
9 April 2010

#### NMELRC Leadership

Kirk Belnap, Director, BYU
Maggie Nassif, Administrative Director, BYU

#### Mahmoud Al-Batal

UT - Austin
Associate Director
Professional
Development

Ruth Adler Ben-Yehuda, Brown Kristen Brustad, UT Au

Kristen Brustad, UT Austin Muhammad Eissa, Chicago Suzan Oezel, Indiana Vardit Ringvald, Brandeis Martha Schulte-Nafeh, UT Austin Dwight Stephens, Duke



#### Erika Gilson

Princeton Associate Director

> Language Assessment

Roger Allen, Penn
Mahdi Alosh, USMA
Micheline Chalhoub-Deville,
UNC Greensboro
Nihan Ketrez, Yale
Salah-Dine Hammoud,
USAFA
Roberta Micallef, Boston
Vardit Ringvald, Brandeis
Martha Schulte-Nafeh,
UT Austin
Kamran Talattof, Arizona

#### **Shmuel Bolozky**

Massachusetts Amherst
Associate Director
Pathways to
Proficiency

Benjamin Hary, Emory
Ahmad Karimi-Hakkak,
Maryland
Sylvia W. Onder,
Georgetown
Vardit Ringvald, Brandeis
Muhammad Eissa, Chicago
Kamran Talattof, Arizona

# Mandate for Title VI Language Resource Centers

"Improve the Nation's
Capacity to Teach and Learn
Foreign Languages
Effectively"



#### LRC Priorities

(according to Title VI Legislation)

- Research
- Materials Development/Dissemination
- Performance Testing
- Teacher Training
- Assess LCTL Needs, Develop Action Plans
- K-12
- Advanced Summer Intensive Programs



### NMELRC Mission

- reach more students
- increase quality of learning opportunities for all students



#### "Assess LCTL Needs, Develop Action Plans"

- surveys of students, teachers, administrators
- site visits, telephone interviews
- study of hiring/staffing practices, implications
- collect outcomes data from language programs and funding agencies



# How are we doing in terms of outcomes?



#### Critical Languages Scholarship Program

		2008 applied	2008 awarded	2009 applied	2009 awarded	2010 applied	2010 awarded
Arabic	Beginning	1345	73	1495	73	694*	64
	Intermed.	759	65	872	65	561	61
	Advanced	234	65	388	55	264	60
Persian	Intermed.	60	8	77	8	87	10
	Advanced	17	7	29	7	20	10
Turkish	Beginning	293	17	279	25	409	20
	Intermed.	73	19	102	17	51	20
	Advanced	28	14	32	10	21	15

<sup>\*</sup>Starting in 2010, "beginning" requires applicants to have taken one year of college-level Arabic or the equivalent prior to the beginning of the summer institute.



### Case Study

# CASA (The Center for Arabic Study Abroad)



#### CASA Applications and Fellowships

	Number of	<b>Summer-Only</b>	Full-Year
Year	<b>Applications</b>	Fellowships	<b>Fellowships</b>
2000-2001	42	6	23
2001-2002	54	8	18
2002-2003	50	8	15
2003-2004	81	6	25
2004-2005	99	6	28
2005-2006	118	6	32
2006-2007	129	4	36
2007-2008	127	6	35
2008-2009	123	6	38
2009-2010	162	5	42
2010-2011	192	2	42





### This advertisement offers new computers

- A for rent with a low security deposit.
- B for rent for as little as 43 LE per month.
- C for sale with a 2 year warranty plan.
- D for sale with a long payment plan.





#### What is the cow saying?

- A The cow wants to get his CDs.
- B The cow wants to sell his CDs.
- C The cow wants to record his CDs.
- D The cow wants to pay for the CDs.



#### חיפש חבר לעט, ומצא את המלוכה

, שהצמיד לבלון הֶלְיוּם מכתב,4ילד בריטי בן בתקווה שימצא כך חבר לעט, מצא את עצמו מנהל חלופת מכתבים עם המלכה אליזבת. העיתון הבריטי "דיילי מייל" דיווח היום כי הילד, טום רולנד, שחרר את בלון הַהֶּלְיוּם שלו באזור המפשייר, מערבית ללונדון. אך במְקום שהבלון יעופף לצרפת או מסביב לעולם, כפי שקיווה קילומטרים32רונלד, מסעו של הבלון הסתיים משם, כאשר נחת בטירת ווינדזור Level 2, Question 1 of 1

#### This passage says that Tom Roland wanted:

- A to correspond with the queen.
- B to send his balloon to England.
- C his letter to go to another country.
- D to travel in a hot-air balloon.



			44 Ex	ams							
Score legen	d: Rando	om	Emerging	Develop	ing	Sustain	ned	<u>N</u>	More info.		
									Score by lev		
Started	Time		Name		N	umber		1	2	3	
6 Apr 10	1h 8m	Ai						0	53	0	
6 Apr 10	0h 56m	A						0	59	0	
6 Apr 10	0h 49m	an						0	47	0	
6 Apr 10	1h 10m	av						0	47	0	
6 Apr 10	0h 49m	Ва						0	76	0	
6 Apr 10	0h 40m	В						0	71	0	
6 Apr 10	0h 53m	В						0	65	0	
6 Apr 10	0h 57m	Cl						0	76	0	
6 Apr 10	0h 57m	Eı						0	41	0	
6 Apr 10	0h 48m	G						0	71	0	



# Four Challenges Encountered When Testing Proficiency in the Receptive Skills

Ray Clifford



#### Four Challenges

- 1. Defining the proficiency construct to be tested.
- 2. Developing curriculum-independent proficiency tests.
- 3. Applying proficiency-based scoring criteria.
- 4. Reporting proficiency results.



#### Four Challenges

- 1. Defining the proficiency construct to be tested.
  - a. Language ability is a multidimensional trait.
  - b. There are many definitions of "Reading" (and "Listening") comprehension.
- 2. Developing curriculum-independent proficiency tests.
- 3. Applying proficiency-based scoring criteria.
- 4. Reporting proficiency results.



#### What is reading?

- How many variables are involved in defining reading?
  - How many are related to the author and the text s/ he produced?
  - How many are related to the reader's skills and purposes?
- Assuming five variables for the author and 5 variables for the reader, how many types of reading are there?



#### How many types of reading are there?

10 Factors		•	Author	1	1	Reader						
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy		
3 / Superior												
2 / Advanced												
1 / Intermediate												
0 / Novice												

- For 10 factors, each with 4 levels there are 40 cells in which a rating may be assigned.
- With one rating per factor, how many different rating profiles are possible?



### For 10 factors, how many rating profiles are possible?

10 Factors			Author			Reader						
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy		
Superior												
Advanced												
Intermediate												
Novice												

10 factors with 4 levels produces 4<sup>10</sup> ...

or 1,048,576 possible rating profiles.



### How might this unwieldy complexity be made more manageable?

• We can reduce the scoring complexity by aligning the rating factors!

• For instance, it would make sense to align the author "topical domains" with the author "purposes" typically associated with those topics.



### For 9 factors, how many rating profiles are possible?

9 Factors			Author			Reader					
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy	
Superior											
Advanced											
Intermediate											
Novice											

9 factors with 4 levels produces 4<sup>9</sup> ... or 262,144 possible rating profiles.



### How might this unwieldy complexity be made more manageable?

- We can reduce the scoring complexity by further aligning the rating factors!
- For instance, it would make sense to align the author "genre" with the author "purposes" generally associated with those topics.
- (In fact, we could reapply the "alignment" principle across all factors.)



### For 8 factors, how many rating profiles are possible?

8 factors			Author			Reader					
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy	
Superior											
Advanced											
Intermediate											
Novice											

8 factors with 4 levels produces 4<sup>8</sup> ... or 65,536 possible rating profiles.



### For 7 factors, how many rating profiles are possible?

7 Factors			Author			Reader				
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy
Superior										
Advanced										
Intermediate										
Novice										

7 factors with 4 levels produces 4<sup>7</sup> ... or 16,384 possible rating profiles.



### For 6 factors, how many rating profiles are possible?

6 Factors			Author			Reader						
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy		
Superior												
Advanced												
Intermediate												
Novice												

6 factors with 4 levels produces 4<sup>6</sup> ... or 4,096 possible rating profiles.



### For 5 factors, how many rating profiles are possible?

5 Factors			Author			Reader						
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy		
Superior												
Advanced												
Intermediate												
Novice												

5 factors with 4 levels produces 4<sup>5</sup> ... or 1,024 possible rating profiles.



### For 4 factors, how many rating profiles are possible?

4 Factors			Author			Reader						
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy		
Superior												
Advanced												
Intermediate												
Novice												

4 factors with 4 levels produces 4<sup>4</sup> ... or 256 possible rating profiles.



### For 3 factors, how many rating profiles are possible?

3 Factors			Author					Reader				
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy		
Superior												
Advanced												
Intermediate												
Novice												

3 factors with 4 levels produces 4<sup>3</sup> ... or 64 possible rating profiles.



### For 2 factors, how many rating profiles are possible?

2 Factors	Author				Reader					
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy
Superior										
Advanced										
Intermediate										
Novice										

2 factors with 4 levels produces 4<sup>2</sup> ... or 16 possible rating profiles.



### How Many Rating Profiles are Possible?

Aligned Factors	Author				Reader					
Rating Level	Purpose	Topical Domains	Genre	Text Type	Accuracy	Purpose	Topical Domains	Type of Reading	Reading Strategy	Accuracy
Superior										
Advanced										
Intermediate										
Novice										

1 factor with 4 levels produces 4<sup>1</sup> ... or 4 possible rating profiles.



### Benefits of Aligning Reading Factors

- Complexity is reduced.
- Each level becomes a separate "Task, Condition, and Accuracy" ability criterion.
- With aligned criteria:
  - Raters can look for sustained ability.
  - Assign ratings using a floor and ceiling approach.
- This hierarchy of levels defines the construct of what it means to be able to "read".



### What is "reading"?

- For achievement and performance testing, there are a million possible definitions of reading.
- For proficiency testing, reading requires:
  - Automatic comprehension rather than laborious decoding.
  - "Reading to learn" rather than "learning to read".
  - Understanding of a text for the purpose for which it was written.
  - General, transferable ability beyond one's own areas of specialization.

### Our Definition of Reading

- Proficient reading: The active, automatic process of using one's internalized language and culture expectancy system to comprehend an authentic text for the purpose for which it was written.
- Author purpose and reading purpose.
  - Orientation Get necessary information
  - Instruction Learn
  - Evaluation Evaluate and synthesize

#### Four Challenges

- 1. Defining the proficiency construct to be tested.
- 2. Developing curriculum-independent proficiency tests.
  - a. Communication tasks to be performed.
  - b. Contexts and typical topical domains.
  - c. Communication accuracy.
  - d. Associated text types and genres.
- 3. Applying proficiency-based scoring criteria.



## The type of learning expected: There are 3 Types of Learning

- A. Direct Application
- B. Near Transfer
- C. Far Transfer



#### The 1st Type of Learning

- With <u>direct application</u> learning, students...
  - Memorize and practice specific responses.
  - Focus is on the content of a specific course, textbook, or curriculum.
  - Learn only what is taught.



#### The 2nd Type of Learning

- With <u>near transfer</u> learning, students...
  - Go beyond rote responses to rehearsed and semi-rehearsed responses.
  - Focus on a predetermined set of tasks or settings.
  - Apply what they learn within a range of familiar, predictable settings.

#### The 3rd Type of Learning

- With learning for <u>far transfer</u>, students...
  - Develop the ability to transfer what is learned from one context to another.
  - Acquire the knowledge and skills needed to respond spontaneously to new, unknown, or unpredictable situations.
  - Learn how to continue learning and to become independent learners.

### The type of test used: 3 Types of Tests

- A. Achievement
- B. Performance
- C. Proficiency



#### The 1st Type of Test

- Achievement tests measure:
  - Practiced, memorized responses.
  - What was taught.
  - The content of a specific textbook or curriculum.



#### The 2nd Type of Test

- Performance tests measure:
  - Rehearsed and semi-rehearsed responses.
  - Ability to respond in constrained,
     familiar, and predictable settings.
  - Whether learning transfers to similar situations.



#### The 3rd Type of Test

- Proficiency tests measure:
  - Whether skills are transferable to new tasks.
  - Spontaneous, unrehearsed abilities.
  - General ability to accomplish tasks across a wide variety of real-world settings.



#### Aligning Learning and Testing

- Direct Application <=> Achievement
  - Memorized responses using the content of a specific textbook or curriculum.
- Near Transfer <=> Performance
  - Rehearsed ability to communicate in specific,
     familiar settings.
- Far Transfer <=> Proficiency
  - Unrehearsed general ability to accomplish
     real-world communication tasks across a wide range of topics and settings.

#### **General Proficiency Requires a Transfer of Learning**

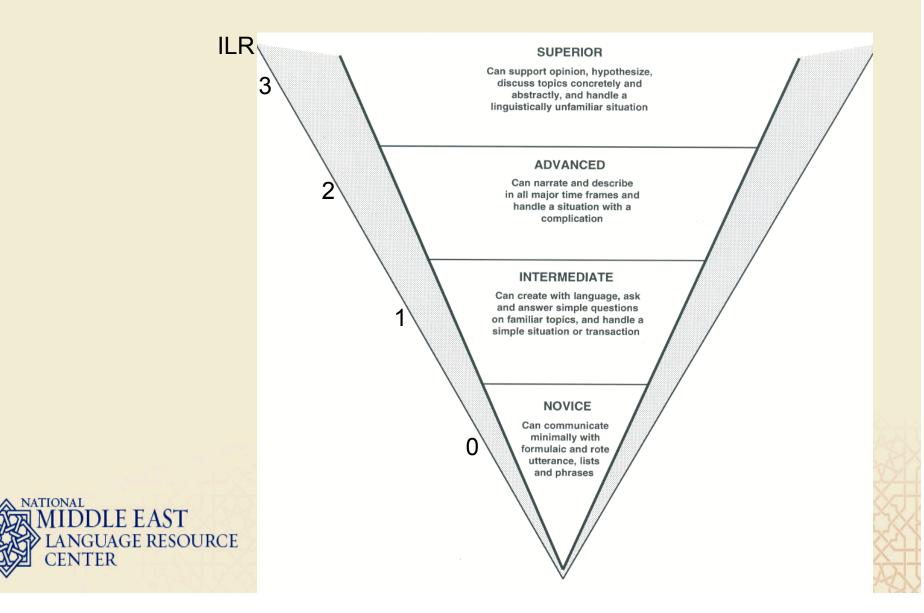
A By-Level Proficiency Summary with General Text Characteristics and Learning Types

[Text Type] + Green = Far Transfer, Blue = Near Transfer, Red = Direct Application

ILR LEVEL	FUNCTION/TASKS	CONTEXT/TOPICS	ACCURACY
5	All expected of an educated NS [Books]	All subjects	Accepted as a well-educated NS
4	Tailor language, counsel, motivate, persuade, negotiate [Chapters]	Wide range of professional needs	Extensive, precise, and appropriate
3	Support opinions, hypothesize, explain, deal with unfamiliar topics [Multiple pages]	Practical, abstract, special interests	Errors never interfere with communication & rarely disturb
2	Narrate, describe, give directions [Multiple paragraphs]	Concrete, real- world, factual	Intelligible even if not used to dealing with non-NS
1	Q & A, create with the language [Multiple sentences]	Everyday survival	Intelligible with effort or practice
O NATIONAL	Memorized [Words and Phrases]	Random	Unintelligible



### The Major ACTFL Levels



# The Phenomenon of Shrinking Educational Expectations

- Students don't want to waste their time studying what is not going to "needed."
- For students (and often teachers, parents, and administrators); the **tests** used and not a course's stated **learning objectives** define what is "needed." Therefore,
  - Limited-scope tests reduce the **breadth** of learning.
  - Simple tests reduce the **level** of learning.

### Shrinking Educational Expectations (When only Achievement Tests are Used)

- 1. High academic goals are set and learner outcomes are defined.
- 2. Developers include examples of the most important goals in a textbook.
- 3. Teachers present as much of the textbook as time allows.
- 4. Students are tested on some of the items taught.

Test

Instructional Goals and Learning Outcomes Textbook



### Expanding Educational Expectations (When Proficiency Tests are Added)

Textbook

**Teacher** 

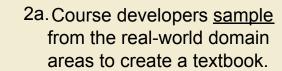
Students

1. Set instructional goals and define expected learner outcomes.

Real-world Instructional Domains: cognitive understanding, psychomotor skills, and affective insights.

Test

2b. Test developers use an independent sample of the real-world domain areas to create proficiency tests that are not based on the textbook.



- 3. Teachers adapt text materials to learners' abilities, diagnose learning difficulties, adjust activities and add supplemental materials to help students apply new knowledge and skills in achievement and constrained performance areas, and then in real-world proficiency settings.
- 4. Students practice, expand, and then demonstrate their unrehearsed, extemporaneous proficiency across a broad range of real-world settings that are not in the textbook.



### Adding Proficiency Expectations and Proficiency Testing Requires:

- Improved assessment literacy for everyone: Teachers, Administrators, Students, and Parents.
- Ongoing communication among stake holders.
- A tolerance for formative assessment that allows programs to "fail forward."
- Clearly stated **E**xpected **L**earner **O**utcomes.
- Assessment practices that match the <u>ELO</u>s.



#### Four Challenges

- 1. Defining the proficiency construct to be tested.
- 2. Developing curriculum-independent proficiency tests.
- 3. Applying proficiency-based scoring criteria.
  - a. Ratings are Criterion-Referenced rather than Norm-Referenced.
  - b. Ratings assigned are non-compensatory.
- 4. Reporting proficiency results.



#### ILR Proficiency Level Summary

LEVEL	FUNCTION/TASKS	CONTEXT/TOPICS	ACCURACY	
5	All expected of an educated NS	All subjects	Accepted as an educated NS	
4	Tailor language, counsel, motivate, persuade, negotiate	Wide range of professional needs	Extensive, precise, and appropriate	
3	Support opinions, hypothesize, explain, deal with unfamiliar topics	Practical, abstract, special interests	Errors never interfere with communication & rarely disturb	
2	Narrate, describe, give directions	Concrete, real- world, factual	Intelligible even if not used to dealing with non-NS	
1 NATIONA MUI	Q & A, create with the language	Everyday survival	Intelligible with effort or practice	
LAN	NGUAGE RESO <b>Memorized</b>	Random	Unintelligible	

#### **ORAL PROFICIENCY INTERVIEW**

© ACTFL, Inc., 1999 All rights reserved.

PROFICIENCY LEVEL*	GLOBAL TASKS AND FUNCTIONS	CONTEXT / CONTENT	ACCURACY	Техт түре
Superior	Discuss topics extensively, support opinions and hypothesize. Deal with a linguistically unfamiliar situation.	Most formal and informal settings / Wide range of general interest topics and some special fields of interest and expertise.	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.	Most informal and some formal settings / Topics of personal and general interest.	Understood without difficulty by speakers unaccustomed 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 / Predictable, familiar topics related to daily activities.	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 / Most common aspects of daily life.	May be difficult to understand, even for speakers accustomed to dealing with non-native speakers.	Individual words and phrases

[\*A rating at any major level is arrived at by the sustained performance of the functions of the level, within the contexts and content areas for that level, with the degree of accuracy described for the level, and in the text type for the level. The performance must be sustained across ALL of the criteria for the level in order to be rated at that level.]

# Applying C-R Rating Criteria Requires the Use of "Fuzzy Logic"

- 1. Success is defined as a "standard" with task, condition, and accuracy criteria.
- 2. Base-level ratings are non-compensatory.
- 3. The hierarchy of level-specific standards forms a "Guttmanesque" scale.
- 4. Ratings are based on unrehearsed ability to accomplish communication tasks in "real world" settings, rather than on rehearsed abilities in curriculum-dependent situations.



#### What is Fuzzy Logic?

- A form of multi-valued logic derived from fuzzy set theory. (Latfi Zadeh, 1965).
- The logic used in many real-world, "soft-computing" applications.
  - Robotics.
  - Artificial intelligence.
  - Neural networks.
  - Natural Language Processing.
- Fuzzy logic is the way humans think.



#### What is Fuzzy (Human) Logic?

- In contrast with binary logic (where sets use 0/1 or "either/or" logic), fuzzy logic variables may have a range of values representing their <u>degree of membership</u> in a given set or class.
  - Binary logic: Is the greenish color blue or yellow?
  - Fuzzy logic: How much blue and how much yellow does the greenish color contain?



#### What is Fuzzy (Human) Logic?

- In contrast with "probability logic" (where one estimates the likelihood that something is true), fuzzy logic variables represent the degree of membership an item has in a given set or class.
  - Probability: What is the chance that a greenish color is really blue? (Or really yellow?)
  - Fuzzy logic: How much blue and how much yellow does the greenish color contain?



### Proficiency Ratings and Fuzzy Logic

- Raters DON'T use binary or probability logic to determine whether an individual meets the standard for a given level.
- Raters DO assess how many of the levelspecific proficiency criteria are met.
  - No evidence.
  - Some random evidence.
  - Emerging evidence
  - Evidence of developing skills.
  - Evidence of sustained ability.

Level 2

Level 3

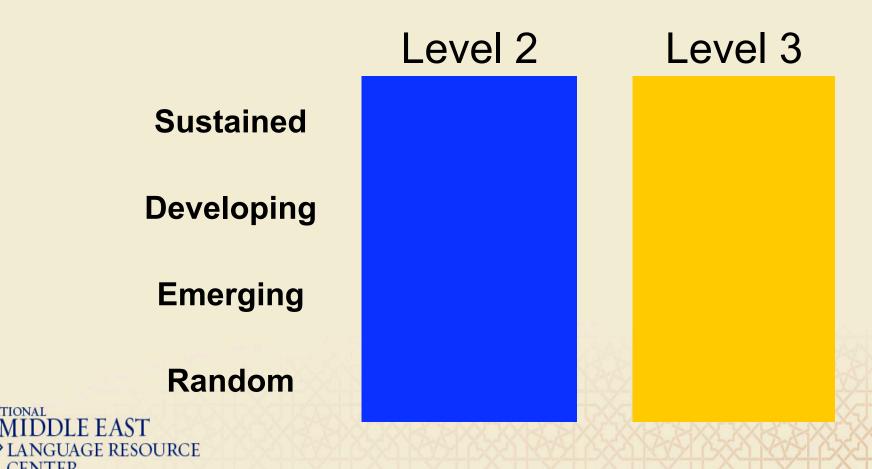
**Sustained** 

**Developing** 

**Emerging** 

Random





Level 2 Level 3
Sustained

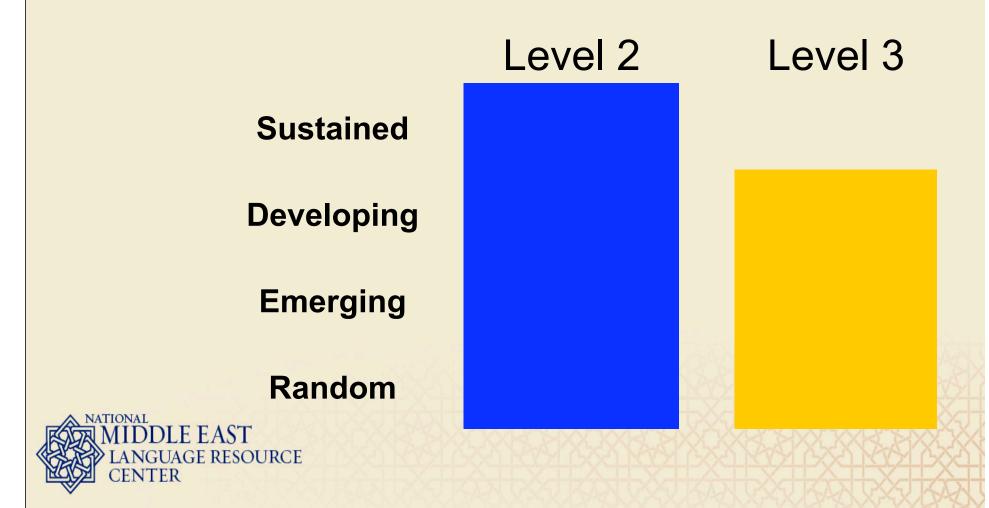
**Developing** 

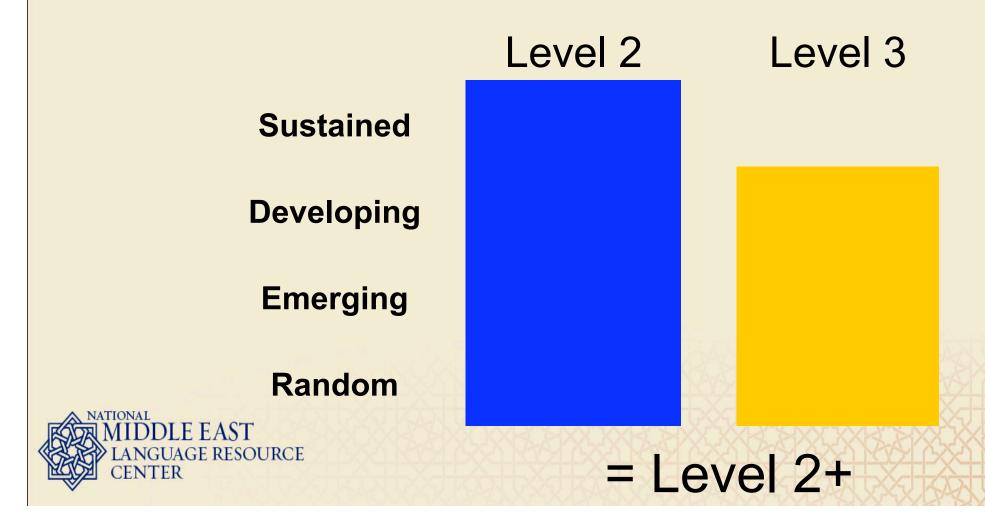
**Emerging** 

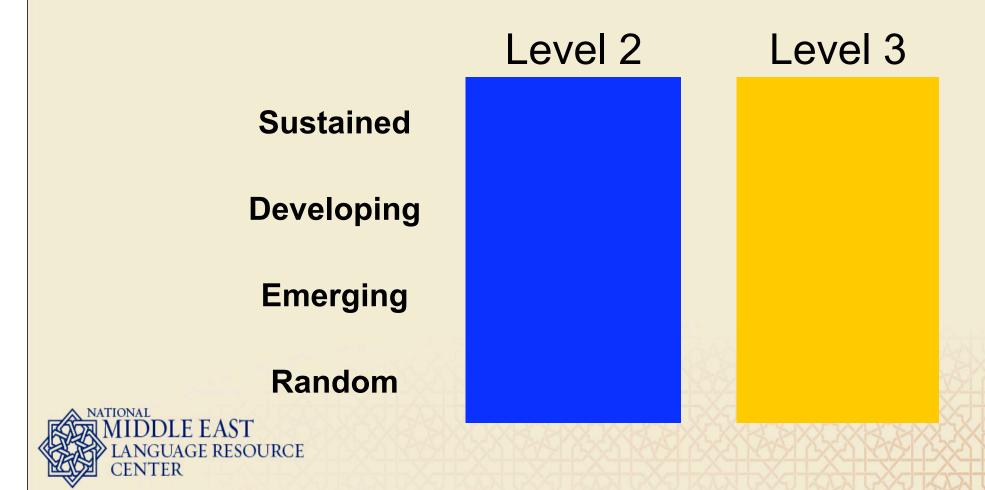
Random



Level 2 Level 3 **Sustained Developing Emerging** Random JAGE RESOURCE







### Human Logic, Ratings, and Total Scores

- 3 Job candidates (Alice, Bob, and Carol) have just taken a reading proficiency test covering Levels 1, 2, and 3.
- All three have different reading abilities.
- All three received the same total score of 195. (65% of a possible 300 points).
- How is that possible?



### Criterion-Referenced Scoring Provides the Answer

- Test each proficiency level separately.
- Check for sustained ability at each level.
  - Sustained (consistent evidence)  $\approx 67\%$  to 100%
  - Developing (a lot, not enough)  $\approx 51\%$  to 66%
  - Emerging (some evidence)  $\approx 26\%$  to 50%
  - Random (no evidence)  $\approx 0\%$  to 25%
- Apply the non-compensatory, "floor" and "MICELLING" criteria used in rating speaking skills.

# What is Alice's Reading Proficiency Level?

Total Score 195 (65%)

Level 1

Level 2

Level 3

Sustained 85% 70%

Developing 40%

Random



# What is Alice's Reading Proficiency Level?

**Total Score** Level 2 Level 3 Level 1 195 (65%) **Sustained** 85% 70% **Developing** 40% **Emerging** Random



Threshold 2 (Advanced Low)

# What is Bob's Reading Proficiency Level?

Total Score 195 (65%)

Level 1

Level 2

Level 3

Sustained 90% 85%

Developing Emerging

Random 20%

Level 2

Level 3

4 20%



## What is Bob's Reading Proficiency Level?

Total Score 195 (65%)

Level 1

Level 2

Level 3

Sustained 90% 85%

Developing Emerging Random

Level 1

Level 2

Level 3

A Sustained 90% 85%

A Sustained 90%

A Sustained 90% 85%

A Sustained 90% 90%

A Sustained 90%



Solid 2 (Advanced Mid)

# What is Carol's Reading Proficiency Level?

**Total Score** Level 1 Level 2 Level 3 195 (65%) 90% **Sustained** 60% **Developing** 45% **Emerging** Random



# What is Carol's Reading Proficiency Level?

**Total Score** Level 2 Level 3 Level 1 195 (65%) **Sustained** 90% 60% **Developing** 45% **Emerging** Random



1+ (Intermediate High)

#### Four Challenges

- 1. Defining the proficiency construct to be tested.
- 2. Developing curriculum-independent proficiency tests.
- 3. Applying proficiency-based scoring criteria.
- 4. Reporting proficiency results.
  - a. Level of sustained ability.
  - b. Status within the "non-sustained" level.



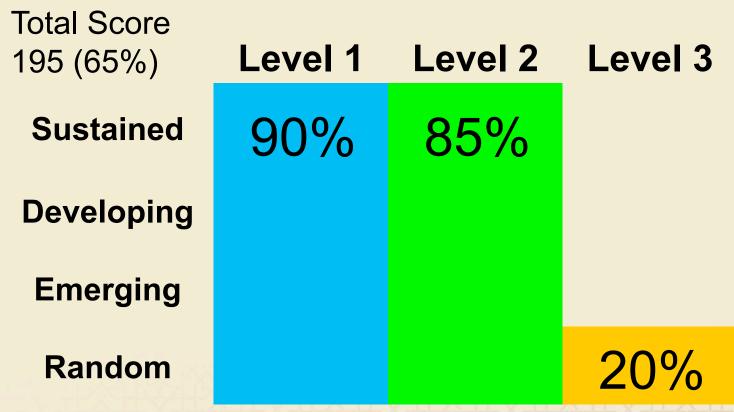
# What is Alice's Reported Reading Proficiency Level?

**Total Score** Level 2 Level 3 Level 1 195 (65%) **Sustained** 85% 70% **Developing** 40% **Emerging** Random



Threshold 2 (Advanced Low) with Emerging abilities at Level 3 (Superior)

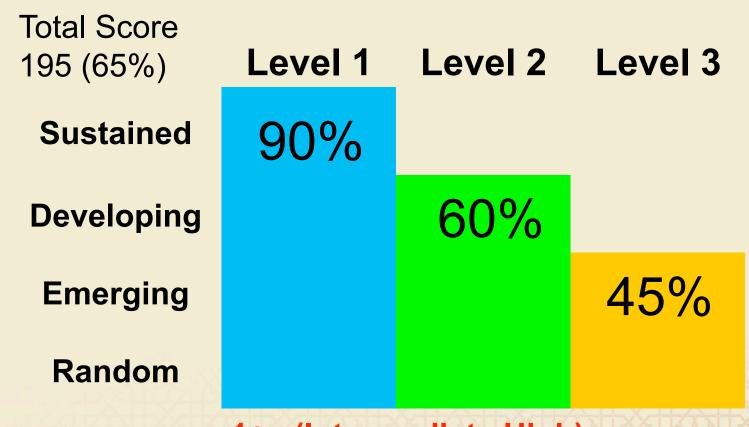
## What is Bob's Reported Reading Proficiency Level?





Solid 2 (Advanced Mid) with Random abilities at Level 3 (Superior)

# What is Carol's Reported Reading Proficiency Level?



OLE EAST
1+ (Intermediate High)
UAGE RESOURCE With Developing abilities at Level 2 (Advanced)

#### Summary

We are Addressing the "Four Challenges" By

- 1. Defining the proficiency construct to be tested.
- 2. Developing curriculum-independent, Criterion-Referenced proficiency tests.
- 3. Applying non-compensatory, proficiency-based scoring criteria.
- 4. Reporting results for both the sustained level and status at the next higher level.



#### **ACTFL**

#### Building Capacity for Testing Reading and Listening



#### Did You Know?

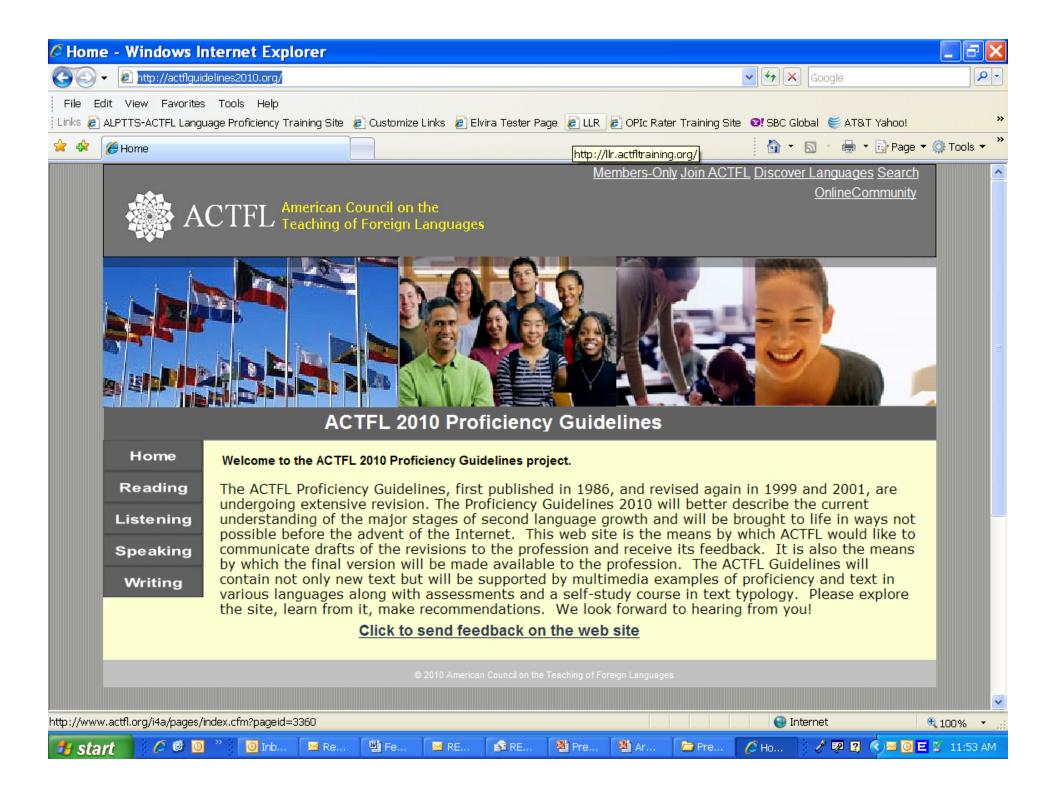
- NATO Benchmark Advisory Tests
  - 4 skills
- Item development and review (ILR)
- Pro-chievement Tests
- AAPPL
  - Interpretive Reading and Listening assessment of academic language
- Arabic Consensus Project

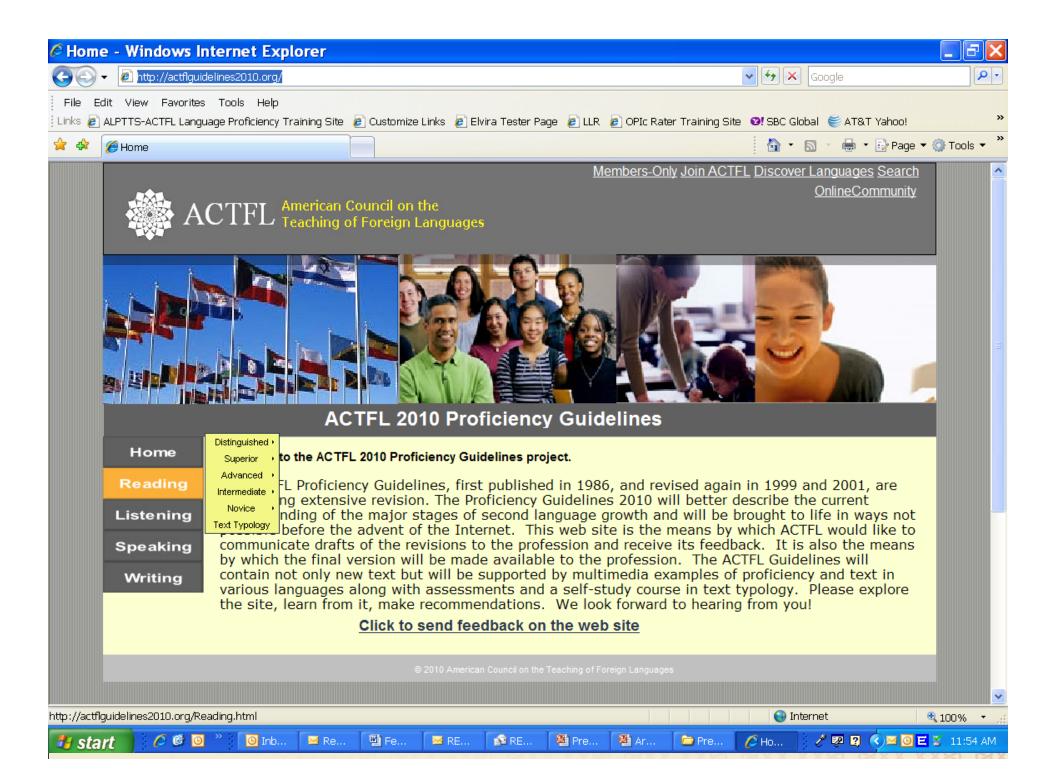


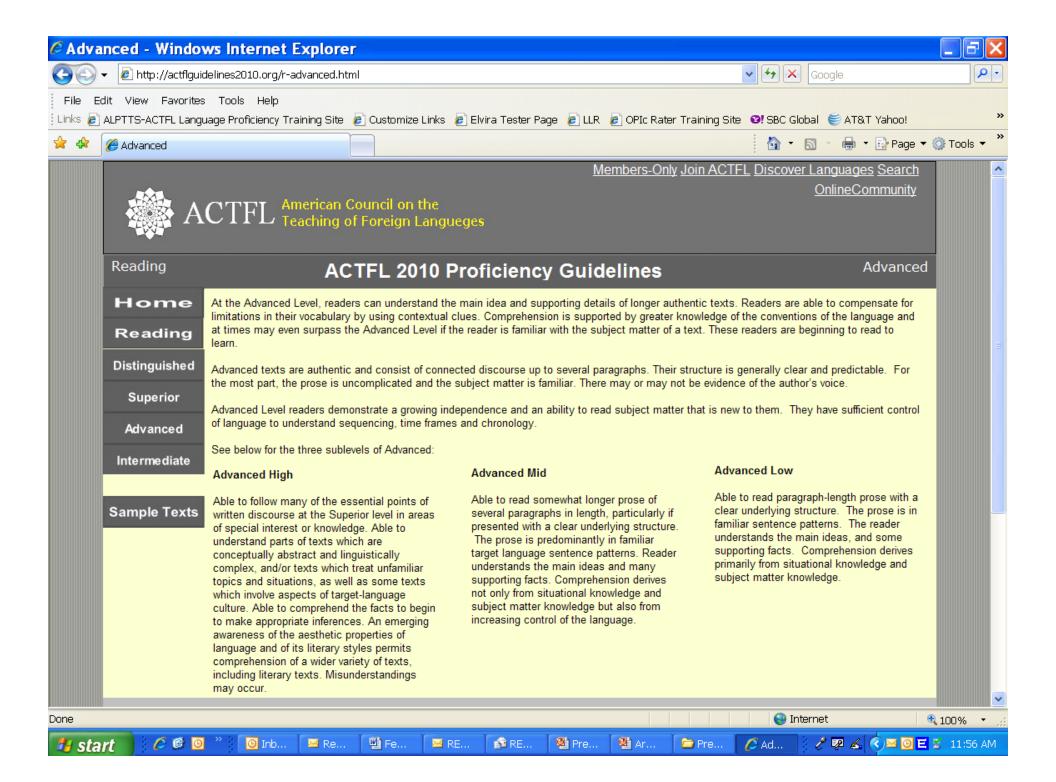
### ACTFL Proficiency Guidelines 2010 Revisions

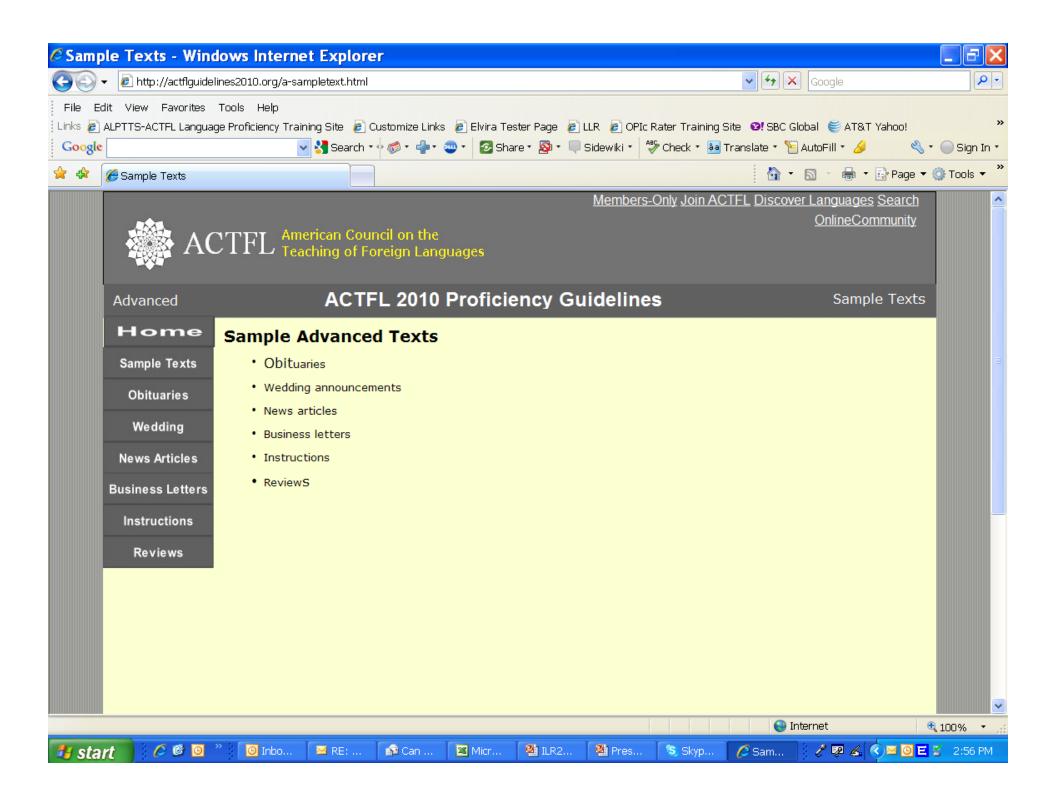
- Reading and Listening
  - First major revision since 1986 publication
  - Addition of sub levels at Advanced
    - Advanced Low
    - Advanced Mid
    - Advanced High
- Paperless distribution via a dedicated web site
- Level-specific descriptions enhanced by sample texts

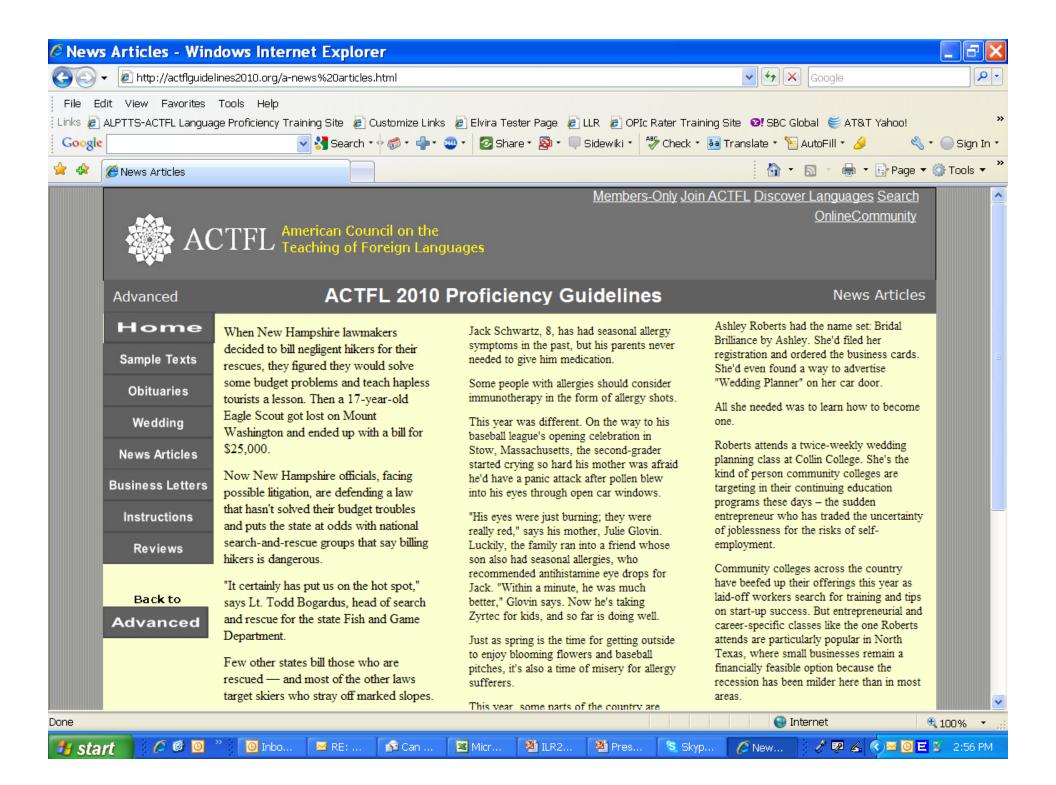


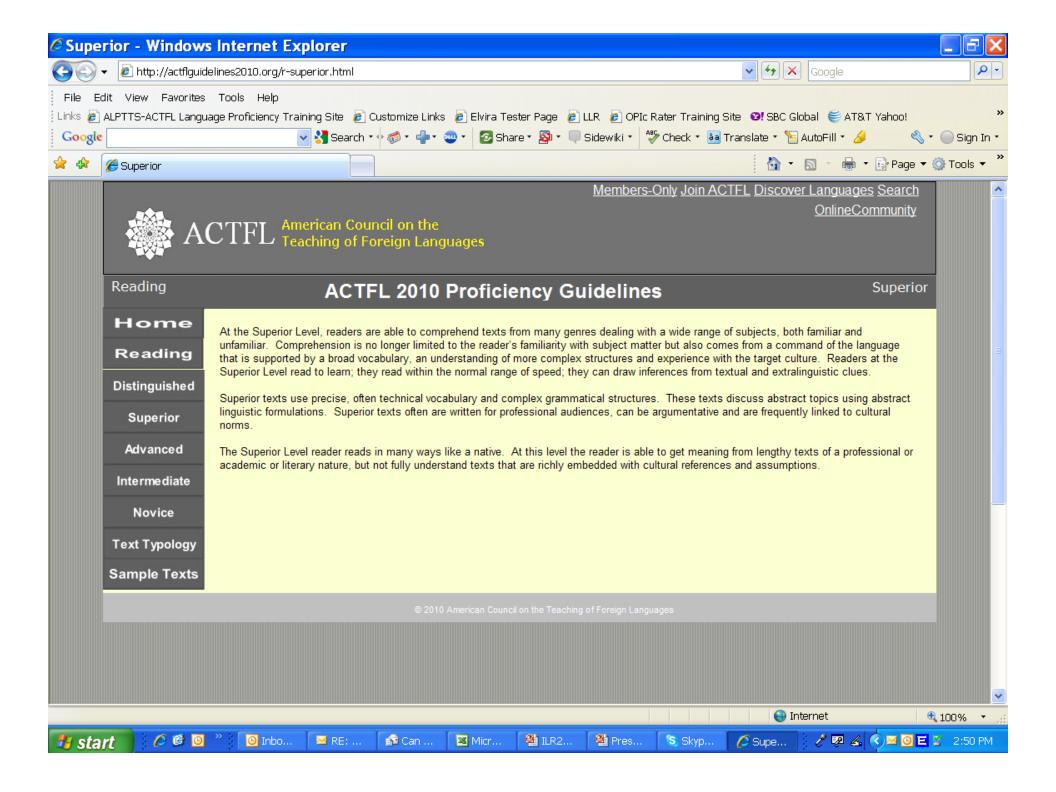


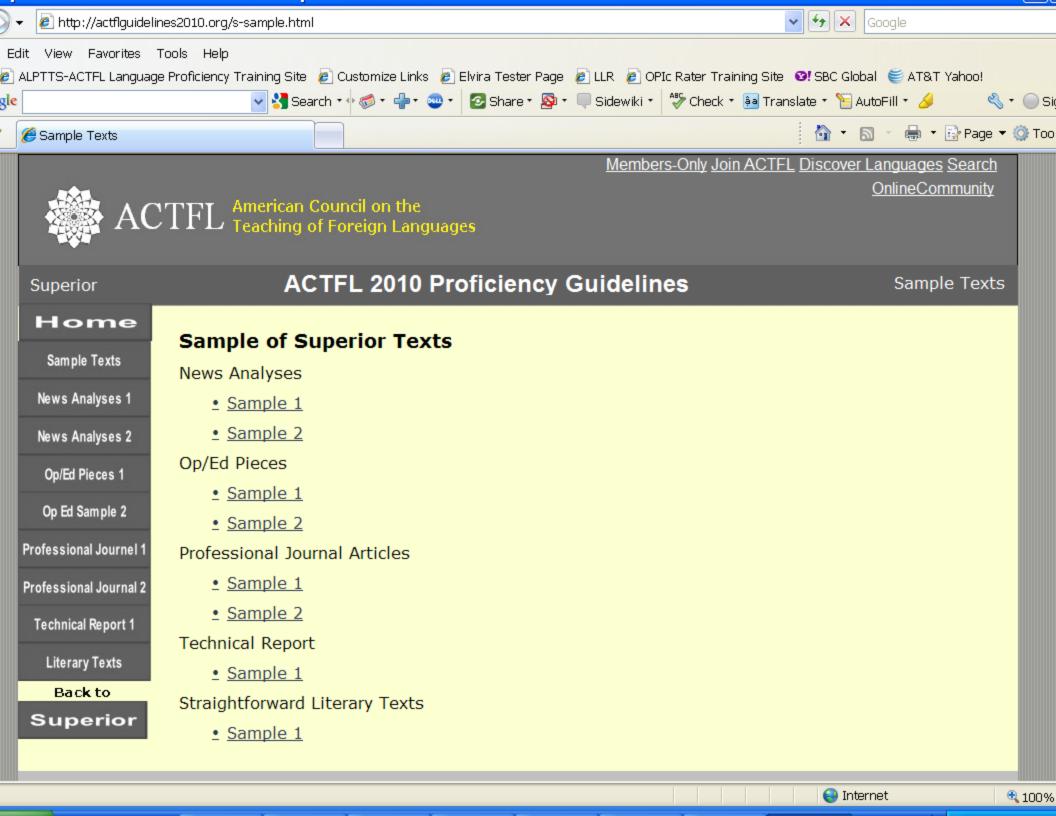


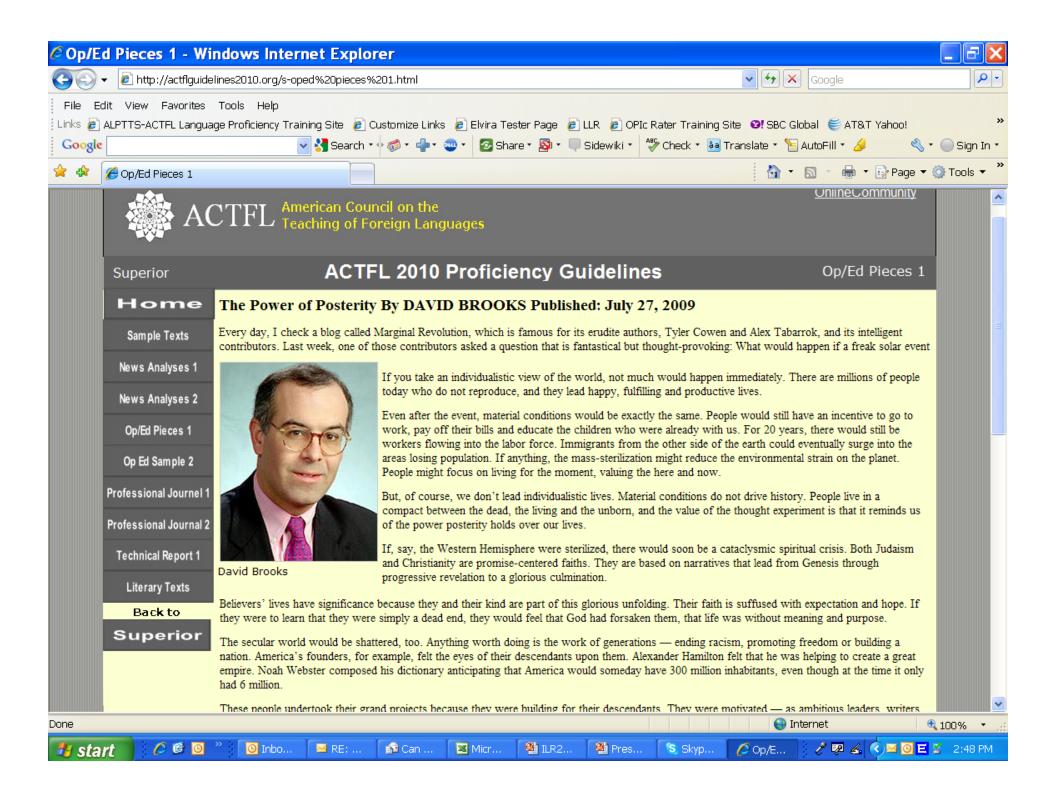


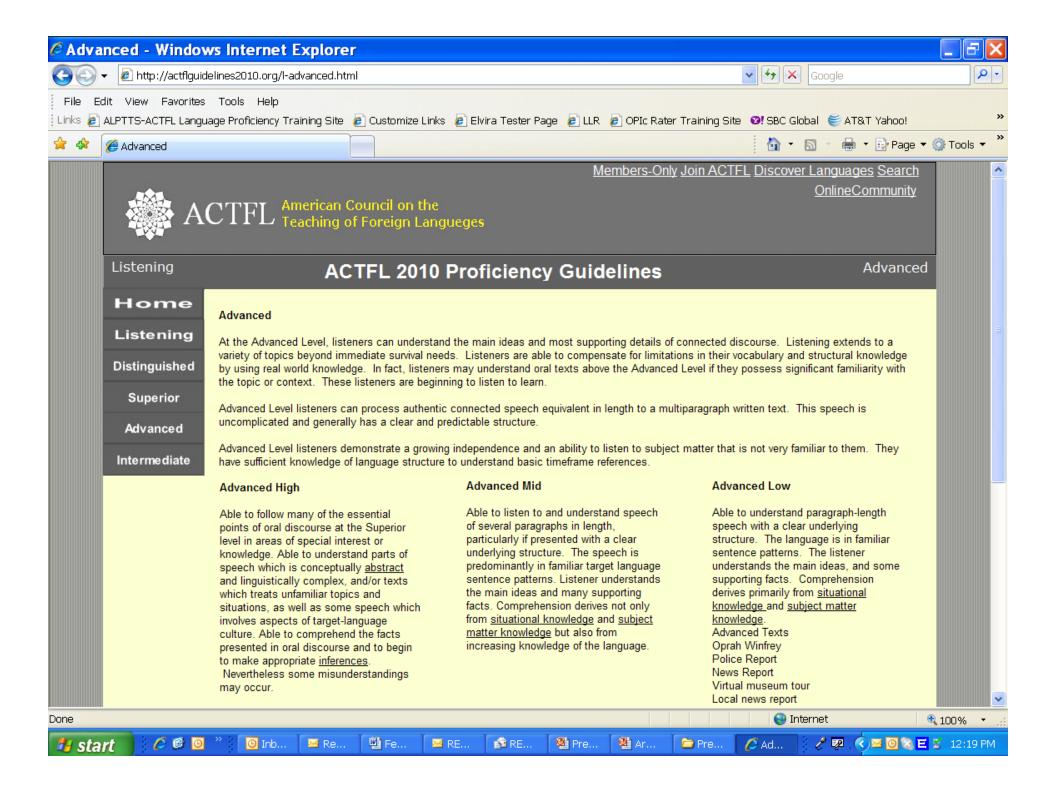


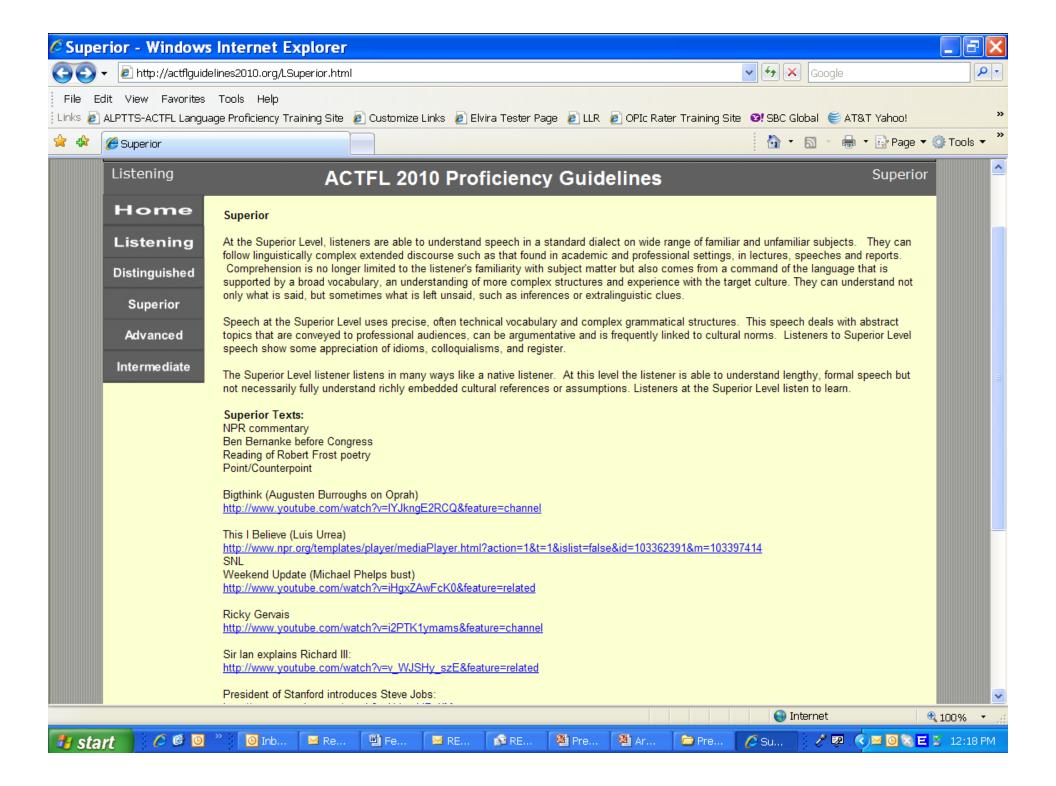












### Reading and Listening Test Delivery

- PATH tests developed by NMERC to be delivered from the ACTFL/LTI Test Delivery Platform
- Internet delivered
- Can be delivered individually or bundled in any combination
  - With OPIs, OPIcs and WPTs
  - No minimum or maximum number
- ACTFL ratings assigned based on Revised Proficiency Guidelines
- Machine scored
  - Official ACTFL certificates issued



#### Next steps

- ACTFL Professional Development
  - Familiarization with Revised Guidelines
  - Workshops in Text Selection and Rating
  - Implications of the Guidelines for Instruction
- Potential for ACE Credit Recommendation
- Adaptive Reading and Listening Tests
  - Self assessment up front
  - Branching
- Development underway in CTL's
  - English, French, German, Latin, Spanish
- Critical languages are leading the way



### nmelrc.org

