

Interagency Language Roundtable



Serving the Language Community for over 50 years



# ILR revisions and the way forward

ILR Plenary, October 4, 2019,  
Gerd Brendel, DLIFLC, Rachel Brooks, FBI,  
Clayton Leishman, DLIELC  
The ILR SLD Revisions Committee



# ILR SLD Revisions Initiative

# ILR SLDs Origins



- USG need post WWII to have a consistent way to measure language (Senate Bill 1243, 1959)
- FSI created a speaking scale (the FSI scale) to measure progressive functional language ability, not language acquisition
- 1970s: interagency effort with government test experts produced scales that were useful to missions of all agencies
  - Based on research of the time and experience (Clark & Clifford, 1988; Lowe, 1983)

# ILR SLDs Origins



- USG need post WWII to have a consistent way to measure language: FSI scale
- 1985 ILR SLDs for Proficiency: Speaking, Listening, Reading, and Writing Proficiency, published by OPM
  - Replaces the FSI scale
- After 2000, ILR SLDs for Performance and Competence added:
  - Translation Performance
  - Interpretation Performance (3 modes)
  - Audio Verbatim Translation Performance
  - Intercultural Communicative Competence

# From 1985 to today



- Feedback from users (examinees, test developers, raters, trainers, managers)
- Discussions with ACTFL and CEFR communities
- ILR Speaking Summits: define levels; comparability across agencies
- ILR Listening Summits: participatory and non-participatory; difficulty
  - CASL summary literature review
- Interagency Comparability Study
- Testing and Assessment Expert Group-sponsored Interagency Speaking Summits
- Defense Language Testing Advisory Project (DELTAP)

# What Works



- A clear system of record that is consistent and understood across agencies
- Tests developed according to scale have been proven to produce valid and reliable ratings
- Examinees progress from one level to another
- Tens of thousands of exams are given with consistent results
- Ability to share resources

# The Case for Revisions



Why are we revising the Proficiency SLDs?

- They are over 30 years old
- Language within the descriptions needed to be updated to reflect current use and needs
- Different agency missions (intelligence, diplomatic, defense, judicial)
- A very different test taker population
- Shared interagency task list and linguistic categories
- Feedback from testers/raters/test developers
- Applied linguistics and language testing research has evolved



# ILR Revisions Committee

- Members include representatives from all major USG users of the ILR SLDs
  - CIA, DLIELC, DLIFLC, FBI, FSI, NSA, and others
- Members reviewed draft versions with their testing specialists and other stakeholders and brought valuable comments back to the ILR Revision Committees
- The revision work is done as an ancillary duty on top of regular workload

# Timeline



Early 2014 – June 2015

- Revision of Listening SLDs to add participatory listening and interference

Early 2016 – Oct 2016

- Revision of Speaking SLDs
- Creation of crosswalk matrix

Oct 2016 – Sept 2017

- Revision of Reading SLDs

Sept 2017 - May 2018

- Presented on ILR SLD revisions at ECOLT
- Revisited the Listening SLDs to place in matrix

May 2018 – July 2018

- Revision of Writing SLDs
- ISS Russian study

July 2018 – Dec 2018

- Crosswalking review

Feb 2019 – Sept 2019

- DELTAP report
- Validation study prep/ research-based framework

Oct 2019

- ILR update presentation
- DELTAP review

Nov 2019 – March 2020

- Revising based on DELTAP and research-based frameworks
- Validation study prep

April 2020 – December 2020

- Validation study
- Posting for public comment, approval by ILR



# ILR SLDs Revisions Process



# Goals

- To clarify and update the wording of the SLDs
- To retain the underlying framework without shifting the difficulty of the levels
- To complete the framework with consistency across the modalities and levels
- To incorporate current research and updated language testing concepts



# Process

- Met in-person and virtually over the course of the last five years
- Revised the Listening SLD and posted for feedback on the ILR website
- Current SLDs were fit into a hierarchical matrix and checked for completeness, hierarchy, and clarity
- Listening, Speaking, and Reading were revised
- Began in-depth crosswalk of skills across modalities and levels
- Revised Writing
- Narratives were built from the matrix and edited for readability
- Wrote new general preface and added skill-specific information
- Incorporated ideas from research-based frameworks
- Designed a validation study

# Process



## ILR SLDs 1985 Version

### Functional Trisection

- Functions
- Topics
- Accuracy

Inconsistent within a SLD, across SLDs

- Crosswalking needed

## Revisions via Hierarchical Grid

### Functional Trisection + Combined Rating Categories

- Functions
- Topics
- Accuracy
  - Discourse and Text Type
  - Social/Cultural Appropriateness
  - Fluency
  - Structures
  - Vocabulary
  - Production/Interference

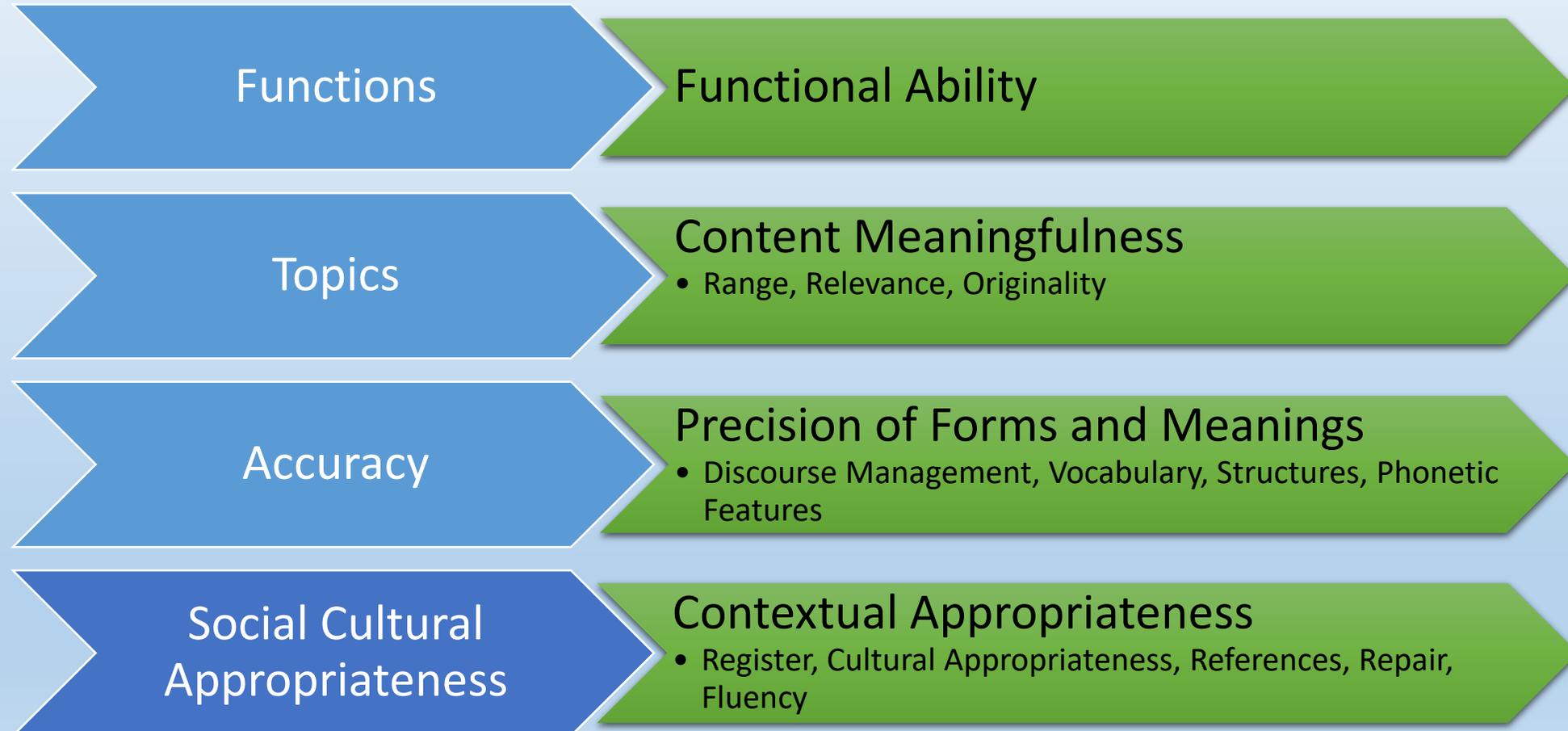
## Incorporating research

### Functional Quatrisection

- Functional Ability
- Precision of Forms and Meanings
- Content Meaningfulness
- Contextual Appropriateness

Shift the focus from trait to ability

# Category Evolution



# Revising Listening



- Clarified functions in both participative and non-participative settings
- Changed some problematic language
- Updated examples of contexts
- Removed references to the native speaker and to learners

# Revising Speaking



- Reorganized information
- Removed references to the native speaker and to learners
- Removed specific comments on grammar to ensure language-neutrality
- Removed references to irrelevant skills, language interpreting
- Edited listening statements to be productive in nature
- Included a hierarchy of terms for a clear progression
- Changed some unclear terminology
- Incorporated examples section



# Revising Reading

- Removed references to the native speaker and to learners
- Considered new uses: texts, social media, etc.
- Clarified functions in both participative and non-participative settings
- Added orthography/text quality category



# Revising Writing

- Removed references to the native speaker and to learners
- Focused on the functional ability (process)
- Added new uses (social media)
- Clarified functions in both participative and non-participative settings
- Adjusted the levels that did not match other skills



# Revisions to the Preface

- Goal of the preface revision
  - Explain the purpose, structure, and concepts of the SLDs
  - Reinforce the importance of reading the preface with development of a skill-specific SLD
  - Address common misconceptions
- Prominent themes explained in the prefaces:
  - SLDs can be used by both learners and those whose primary language is the target language
  - Clarify the definition of a plus level
  - Ability level may differ in a topic area of special interest

# Change: Main statements



**Speaking 3 (General Professional Proficiency)** Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most conversations in practical, social and professional topics. Never

## Main Statements: QUATRASECTION

with sufficient accuracy to participate effectively in formal and social and general professional topics and tasks. Can perform a range of tasks, discuss and compare societal issues and their implications, state

limitations generally restrict the professional contexts of language use to matters of shared knowledge and/or common interest. The individual uses the language clearly and never interferes with communication. Can usually participate in extended, effective communication in most formal and informal contexts. Individual speaking the standard is complete. Although cultural references may not be fully understood, they are not obviously foreign. Individual sounds are accurate: but stress, intonation and pitch control may be faulty. **Examples:** Can typically discuss particular interests and special fields of competence with reasonable ease. Can use the language as part of normal professional duties such as answering objections, clarifying points, justifying decisions, understanding the essence of challenges, stating and defending policy, conducting meetings, delivering briefings, or other extended and elaborate informative monologues. Can reliably elicit information and informed opinion from native speakers. Structural inaccuracy is rarely the major cause of misunderstanding. Use of structural devices is flexible and elaborate. Without searching for words or phrases, the individual uses the language clearly and relatively naturally to elaborate concepts freely and make ideas easily understandable to native speakers. Errors occur in low-frequency and highly complex structures. (Has been coded S-3 in some nonautomated applications.) [Data Code 30]

**Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations in practical, social and professional topics.**

and defend a position or policy, support opinions, hypothesize, give advice, elicit information and informed opinion from native speakers. Structural inaccuracy is rarely the major cause of misunderstanding. Use of structural devices is flexible and elaborate. Without searching for words or phrases, the individual uses the language clearly and relatively naturally to elaborate concepts freely and make ideas easily understandable to native speakers. Errors occur in low-frequency and highly complex structures. (Has been coded S-3 in some nonautomated applications.) [Data Code 30]

**Able to speak with sufficient accuracy to participate effectively in formal and informal contexts on most social and general professional topics and tasks.**

at a rate of speech that does not impede comprehension. Can discuss particular interests and special fields of competence. Nevertheless, linguistic limitations in depth and precision generally restrict the contexts of language use to matters of general knowledge. Can usually control formal and informal registers in various settings. May make some appropriate cultural references and use widely known proverbs and idiomatic expressions. Fills pauses suitably.

# Change: Native speaker removed



## CURRENT

Speaking proficiency is functionally equivalent to that of a highly articulate well-educated native speaker and reflects the cultural standards of the country where the language is natively spoken. (Level 5 Speaking)

## DRAFT

Able to speak eloquently using highly articulate language in virtually any context. ... Uses language that reflects a deep understanding of the culture as well as social and circumstantial knowledge, including idioms, colloquialisms, rhetorical devices, and both literary and popular references. (Level 5 Speaking)

# Change: Participatory context added



## CURRENT

Able to understand face-to-face speech in a standard dialect ... about everyday topics, common personal and family news, well-known current events and routine office matters. ... Understands factual content. (Level 2 Listening)

## DRAFT

Can understand factual content in both participatory and non-participatory situations, such as interviews, presentations, news broadcasts... (Level 2 Listening)

# Change: Contexts modernized



## CURRENT

Texts may include descriptions and narrations in contexts such as news items describing frequently occurring events, simple biographical information, social notices, formulaic business letters, and simple technical material written for the general reader. (Level 2 Reading)

## DRAFT

Can understand minimally cohesive, factual texts, including online or printed materials such as news items describing frequently occurring events; routine business correspondence, email, and text messages; clearly written product and service descriptions; and user comments. (Level 2 Reading)

# Process outcomes



- In addition to the development of revised versions of all four skills, the committee:
  - Developed a common understanding of the ILR SLDs
  - Drafted a glossary of terms to help explain the ILR
  - Partnered with the IC's Testing and Assessment Expert Group to conduct summits on testing speaking and a validation study
  - Tracked issues important to the USG community for further discussion with the ILR
  - Initiated development of online Interagency oral proficiency training modules
  - Developed best practices documents for speaking testing

# Glossary



- Recognized the need for a glossary of terms to clarify meaning in the context of the ILR Community
- Sample terms defined in the glossary:
  - Abstract and concrete
  - Base level and plus level
  - Hierarchy
  - Unfavorable conditions
  - Errors versus mistakes
  - Threshold

# Challenges



- Agreeing on terms across agencies
  - Certain terms (abstract, sporadic, routine, etc.) are important to certain agencies
- Considering agency mission
  - Including or changing phrasing has policy ramifications
- Allowing for future needs
  - Texting and emojis are prominent today, but could disappear soon
- Anticipating unintended consequences



# ILR Speaking Validation Study

- Purpose:
  - To help build a validation argument for the ILR SLD speaking revisions.
  - To examine whether there will be a score shift resulting from the updates of the ILR SLD speaking revisions.
- Questions:
  - To what degree does validity evidence support stakeholder use of the revised ILR SLDs – Speaking in the United States Government context?
  - Will revisions in the ILR SLDs - Speaking lead to a score shift from the current 1985 version?

# Interagency ILR SLD Revisions Implementation Plan



- Consult again with DELTAP for additional input (October 2019)
- Complete drafts ready for comment (October 2019 – April 2020)
- Administer and collect data for ILR SLD Speaking Validation Study (April 2020 – October 2020)
- Publish revisions for public comment (April 2020 – October 2020)
- Committee analysis and review of study data and public comments (October 2020 – November 2020)
- Produce final version (November 2020 – December 2020)
- Approve final SLDs by ILR USG vote (December 2020)

Thank you

