Distinguishing Great from Brilliant:

Investigating the features of high-level speakers

Rachel Brooks, PhD
Tanner Call, MA
ILR Testing Committee, September 17, 2021

Purpose

- USG speaking testers have trouble agreeing on what is required to get an ILR Level 5 in speaking.

- The description of ILR Level 5 and the use of terms like "native speaker" and "well educated" offer little clarity.

- There is a need to determine what really differentiates Leve 4 and Level 5 speakers.

Speaking Testing at the FBI

- The Language Testing and Assessment Unit (LTAU) Speaking Proficiency Test (SPT) Program:
    - 2,000 – 3,000 speaking tests annually
    - 85% SPTs (max rating 5); 15% MSPTs (max rating 4)
    - FBI examinees:
        - Level 4 and 4+: 4%
        - Level 5: 1%

- Why is it important to distinguish Level 4 from 5?
    - Level 5 speakers exist.
    - Level 5 is required to be eligible to test speaking.
    - Level 5 provides a cap for the scale.
    - Removing Level 5 may create score shift.

Speaking Level 4 versus Level 5

ILR Level 4

- Fluent and accurate on all levels normally pertinent to professional needs
- Organizes discourse well
- Uses appropriate rhetorical speech devices and native cultural references
- A high degree of effectiveness, reliability, and precision
- Can perform extensive, sophisticated language tasks
- Can discuss concepts in detail
- Reliably produces shifts of both subject matter and tone

ILR Level 5

- Functionally equivalent to a highly articulate, well-educated native speaker
- Complete flexibility and intuition
- Speech on all levels is fully accepted by well-educated native speakers
- Has a breadth of vocabulary, idioms, colloquialisms, and pertinent cultural references

What is Level 5?

Often operationalized as:

- No grammatical errors and uses highly complex structures
- Few to no fillers
- A lot of highly sophisticated, low frequency vocabulary
- A lot of cultural references, proverbs, and idioms

Research Questions

- What language features distinguish examinees at Levels 4 and 5?

- What sociocultural functions, if any, can we expect to be uniquely achievable for Level 5 speakers?
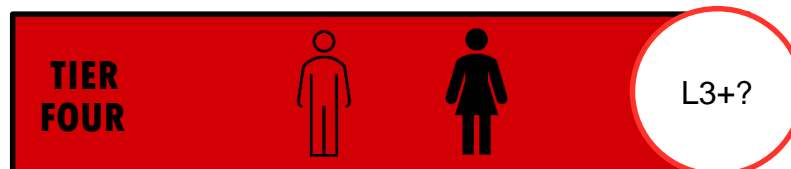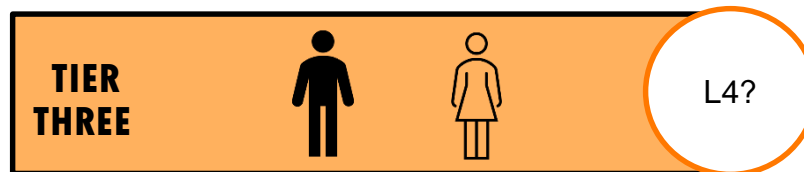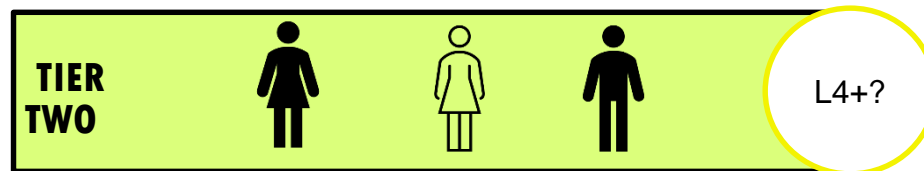
The Process

1. Selected 15 SPT recordings
   - 5 each rated at Levels 4, 4+, and 5

2. Had 11 English raters rank them from 1–15, best to worst
   - Raters provided detailed comments justifying their ranking choices

3. Conducted quantitative and qualitative analyses
   - Sociolinguistic analyses
   - Corpora analyses
   - Grammar counts
   - Filler counts

Rankings

High-level speakers

Quantitative analysis of rankings indicated four tiers of speakers*

TIER ONE — L5?

TIER TWO — L4+?

TIER THREE — L4?

TIER FOUR — L3+?

Low(er)-level speakers

*Four speakers could not be reliably categorized

Sociolinguistic Analysis

Participants' comments seem to center around two main categories:

Language Form	Language Function
How "correct" is the language?	How "effective" is the language?
Grammatical correctness	Engaging discourse
Semantic coherence	Persuasive arguments
Lexical accuracy	Topical relevance
Syntactic organization	Task completion

- "People are equally able to speak meaningfully" but "people are not equally able to speak effectively" (Sanders, 2015).
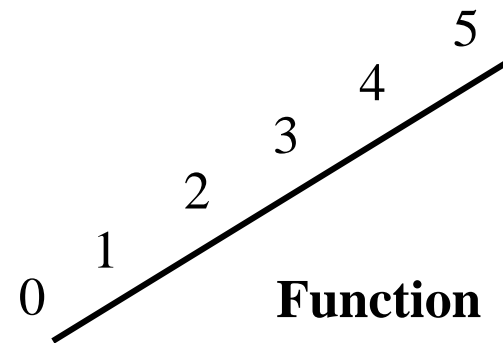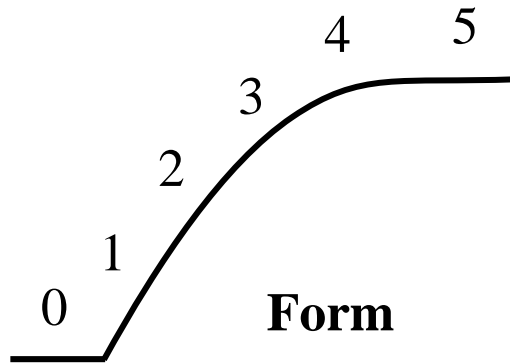
- Linguistic forms *and* linguistic functions should be considered

Form and Function in the ILR

Form: Correctness

- Compared to L4, does L5:
  - make fewer grammatical errors?
  - use more sophisticated words?
  - have more lexical diversity?
  - say more?
  - use less fillers?

Precision of Forms and Meanings

Function: Meaningfulness

- Compared to L4, does L5:
  - complete any additional tasks?
  - stay on topic better?
  - speak more succinctly?
  - speak in a more engaging manner?

Functional Ability

Content Meaningfulness

Contextual Appropriateness

Form Analysis

Level 5:

Has almost no grammatical errors
Speaks extensively

Has almost no fillers

Has highly sophisticated vocabulary

Grammatical Errors

Tier	Mean (SD) grammatical errors	Mean (SD) words produced	Mean (SD) % of grammatical errors
1 n=4	7 (3.20)	4792 (2244)	0.18%
2 n=3	13 (5.92)	3756 (929)	0.33%
3 n=2	10 (1.77)	3624 (210)	0.27%
4 n=2	15 (16.97)	3090 (1226)	0.41%

There are no significant differences between any of the four tiers in grammatical errors.

Elaboration

Tier	Mean (SD) words produced	Mean (SD) words per response
1 n=4	4792 (2244)	580 (384)
2 n=3	3756 (929)	312 (48)
3 n=2	3624 (210)	276 (25)
4 n=2	3090 (1226)	223 (30)

Tiers 2 + 3 (M=294, SD=37) had significantly longer responses than tier 4 (M=223, SD=30), at p=.04.

Fillers

Tier	Mean (SD) fillers	Mean (SD) words produced	Mean (SD) % of fillers
1 n=4	185 (128)	4792 (2244)	4% (2%)
2 n=3	139 (75)	3756 (929)	4% (2%)
3 n=2	222 (55)	3624 (210)	6% (2%)
4 n=2	79 (78)	3090 (1226)	2% (2%)

There are no significant differences between any of the four tiers in fillers.

Lexical Diversity: Type/Token Ratio

Tier	Mean (SD) tokens (words produced)	Mean (SD) types (unique words)	Mean (SD) Guttman type/token ratio
1 n=4	4792 (2244)	1060 (278)	15.51 (1.23)
2 n=3	3756 (929)	909 (243)	14.73 (2.10)
3 n=2	3624 (210)	931 (15)	15.46 (0.20)
4 n=2	3090 (1226)	763 (38)	14.08 (2.16)

There are no significant differences between any of the four tiers in lexical diversity.

Lexical Sophistication: Frequency

Tier	% high frequency terms (1-500)	% mid frequency terms (501-3,000)	% low frequency terms (>3,000)
1 n=4	69% (3%)	10% (1%)	7% (1%)
2 n=3	73% (2%)	10% (1%)	7% (1%)
3 n=2	67% (1%)	11% (1%)	7% (0%)
4 n=2	70% (6%)	10% (4%)	8% (3%)

There are no significant differences between any of the four tiers in use of low frequency terms.

Vocabulary Frequency Examples

High frequency (1 – 500)

the, be, and, a

strong, reach, remain, explain

Mid frequency (501 – 3,000)

site, hit, pull, church

mixture, murder, assistant, retain

Low frequency (>3,000)

Tomato, Indian, testify, ingredient

Vocabulary Frequency Examples

High frequency (1 – 500)	Mid frequency (501 – 3,000)	Pretty low frequency (3,001 – 6,000)	Very low frequency (6,0001 – 10,000)	Extremely low frequency (>10,000)
the, be, and, a	site, hit, pull church	tomato, Indian, testify, ingredient	kidnap, kind, realization, bipartisan	crushed, superior, firsthand, speculative
strong, reach, remain, explain	mixture, murder, assistant, retain	ethnicity, middle-class, total, sympathetic	robotic, automotive, disdain, prioritize	startup, mandatorily, shackled, egocentrism

Lexical Sophistication: Frequency

Tier	% high frequency terms (1-500)	% mid frequency terms (501-3,000)	% pretty low frequency terms (3,000)	% very low frequency terms	% extremely low frequency terms
1 n=4	69% (3%)	12% (3%)	2% (1%)	1% (0%)	1% (0%)
2 n=3	73% (2%)	12% (2%)	2% (0%)	1% (0%)	1% (1%)
3 n=2	67% (1%)	13% (1%)	2% (0%)	1% (0%)	1% (0%)
4 n=2	70% (6%)	11% (6%)	2% (1%)	1% (1%)	2% (1%)

There are no significant differences between any of the four tiers.

The Cooperative Principle

How interlocutors speak and interact with one another in order to achieve effective communication

Gricean Maxims

Quality	Quantity	Relation	Manner
Give true and evidence-supported information	Be informative and give as much information as needed	Be relevant and say things pertinent to the discussion	Be clear, brief, and orderly and avoid ambiguity or obscurity

Sociocultural Functions in Context

1. Quality
   Is the examinee's thesis clear and on topic?

2. Quantity
   Is the examinee on topic of the original question?

3. Relation
   Does the examinee explore related but novel ideas?

4. Manner
   Are the examinee's ideas clear and easy to follow?
   Does the examinee summarize their main point(s)?

1. Is the examinee's thesis clear and on topic?

"Some believe that there should be no such designation as a 'hyphenated American' and that such usage has altered the concept of a unified American people. **To what degree has the promotion of cultural diversity affected our national identity?"**

Responses

LT
(Tier 1)

Well, first let me say that I do agree, to a certain extent, that including the hyphen between one's heritage and "American" has absolutely played a role in the differences that we wear on our sleeves in the population.

Addresses nuance
(cultural differences)

Clear thesis

SM
(Tier 2)

That is an interesting question as well. I feel like in times of crisis we tend to be more united as a country, and so the sense of "I'm an American" comes out very strongly, unfortunately, only after something really tragic maybe has occurred.

Clear thesis but missing nuance
(no cultural component)

More Examples

Question "If the U.S. were to retain our foreign aid funds or redirect it to domestic projects, **how would that alter the foreign perception of Americans**?"

MD
(Tier 1)

If we maintain our foreign aid programs, nobody's going to notice. If we pull them, everybody will see it as a sign of weakness. And if we increase them, there will be a little attention, but nobody will really care.

Thesis is clear and addresses the question's nuance
(foreign aid *and* foreign perception)

ML
(Tier 2)

It's hard to give you a direct answer without looking at the numbers. From what I know, the percentage of our budget that is spent on both of those things is, honestly, almost negligible. So, taking away the foreign aid and all of that, and putting it somewhere else, really doesn't make a huge swing on the budget.

Thesis is clear but not fully on topic
(foreign aid but *no* foreign perception)

"Quality" in High-level Speakers

- Higher-level examinees often provide a clear thesis that is fully on topic

- The thesis usually occurs somewhere towards the beginning of the response and may accomplish the following functions:
    a) Indicate to the raters that the examinee understood the question and will address its nuance(s)
    b) Clearly outline the overall argument to be made
    c) Justify why the examinee is taking a certain stance

2. Is the examinee on topic of the original question?

Components of "Quantity"

1. Does the examinee remain on topic throughout most of their response?

2. Is their response within the same semantic realm as the question?

3. Does their full response sufficiently address the nuance(s) of the question?

"Quantity" by Tier

Percentage of Responses that Accomplish the Sociocultural Function

	Tier One	Tier Two	Tier Three	Tier Four
Is the examinee on topic of the original question?	100%	79%	79%	75%

Question: "Do you think smart photography is as legitimate a field of art as traditional photography?"

Semantic realm

Photographs in the Soviet Union

Photographs in history books

History of political powers editing photos

Photo editing before Photoshop

No discussion of the legitimacy of smart photography

Question's specific topic(s)

- Higher-level examinees usually remain in the established semantic realm AND address the specific topics/nuances provided in the question

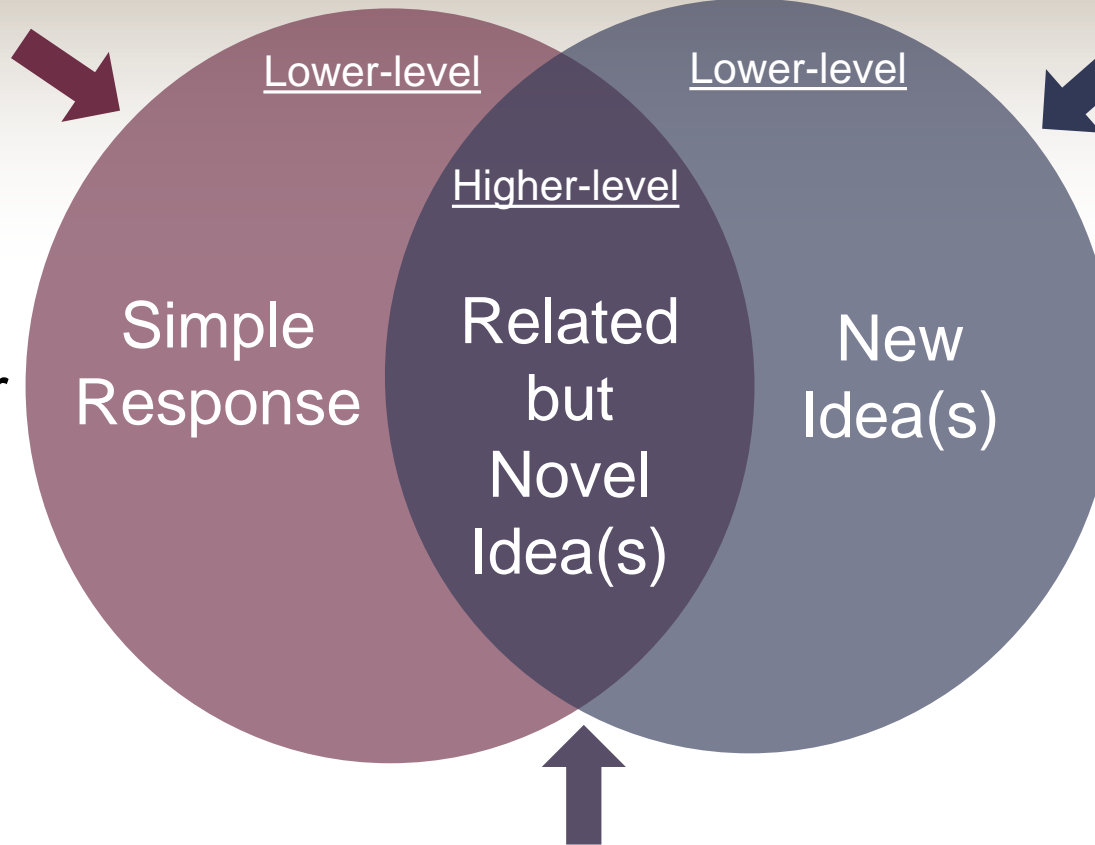- If higher-level examinees diverge from the question's topic, they tend to return to it

3. Does the examinee explore related but novel ideas?

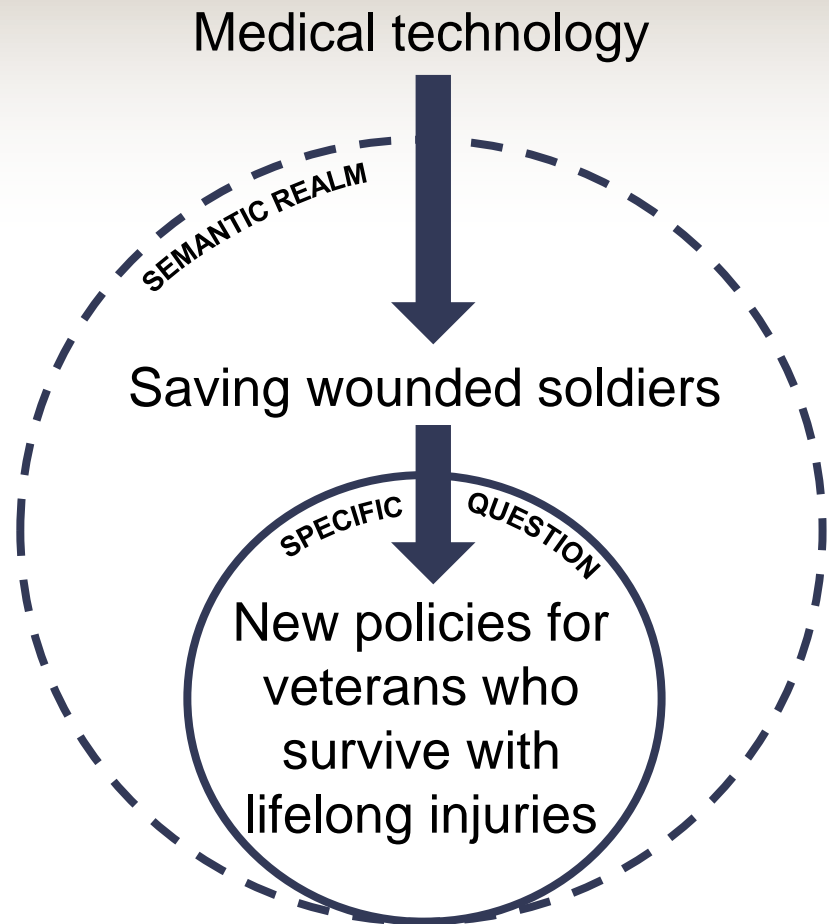Examinee remains on topic but does not provide an elaborate or nuanced response

Examinee explores new topics but does not connect them to the original question

Lower-level

Lower-level

Higher-level

Simple Response

Related but Novel Idea(s)

New Idea(s)

Examinee brings in new topics AND connects them to the original question

Tier One Response

Question:
"How will technological advancements affect global military engagement?"

Medical technology

SEMANTIC REALM

Saving wounded soldiers

SPECIFIC QUESTION

New policies for veterans who survive with lifelong injuries

"Relation" in High-level Speakers

- Examinee brings up a novel idea AND appropriately ties said topic to the original intent of the question

- Examinee weaves two (or more) seemingly distinct ideas together to approach the question in an engaging way

4. Are the examinee's ideas clear and easy to follow?

Components of "Manner"

1. Are intended tone and register accurately conveyed? (e.g., sarcasm, joking, formality)

2. Is the overall argument organized well?

3. Are they articulate and "fluent"?

4. Do they "paint a picture" with their words?

5. Does their speech meet rhetorical demands?

Example Elicitation

Interactive Situation:

The examinee is asked to give a brief speech advocating for or against local police officers wearing body cameras

Tier One Response

1 Hello Mr. Mayor, City Councilors, my fellow citizens.
2 I'm here to discuss today the issue of body cameras, and I want to strongly advocate
3 for the implementation of body cameras for police officers. …
4 For the public, it allows them to really see evidence and clear up that sort of nebulous space
5 that allows all these protests litigation cases to emerge,
6 when we're not really sure which side committed what crime.
7 For police officers who are often accused of these cases, it allows them
8 to have solid evidence as back up, justifying their actions for doing what they did at the time.
9 I think everyone can be in agreement that a lot of these cases where we accuse
10 police of corruption or major issues is definitely the result of only a few bad actors.
11 And I don't think anybody is accusing the police force in general of being a negative body.
12 But I think that, despite the fact that these cameras
13 may be expensive and impact city budgets, the savings that it would generate
14 by reducing litigation and internal affairs investigations would be hugely beneficial. …
15 Thank you.

Formal Register   Well-articulated Thesis   Clear Organization   Detailed Nuance   Insightful Counterargument

Tier Four Response

1 Fellow citymen, fellow councilmen, law enforcement, and everyone else here in attendance,
2 I think this is a very terrible idea to establish body cameras
3 on our local law enforcement officers.
4 Due to the climate of the past eight years in our country,
5 there seems to have been a greater target put on our law enforcement officers' backs. …
6 By asking them to do something by which inherently betrays their trust,
7 you're further dividing our country, dividing our community. …
8 Trust is ultimately a two-way street. …
9 Perhaps there could be an after-school program
10 by which law enforcement officers can walk with your sons and daughters home from school,
11 free of body cameras, to show that they are an integral member of the community.
12 Let's not ask them to further alienate themselves.
13 Let's instead ask them to show us how much – ask – let's instead – excuse me.
14 Let's instead show them how much trust we have in them,
15 so that this in turn can be reciprocated.

Formal Register	Informal Register	Well-articulated Thesis	Divisive Tone	Problematic Solution	Fluency Breakdown	Unaddressed Nuance

- Clarity and rhetorical effectiveness can manifest in a variety of ways:
    - How well the examinee organizes their overall argument
    - How articulate or "fluent" they are
        - Are their ideas easy to follow?
    - How well they convey intended tone and register
        - Is their language sufficiently formal?
        - Is the established tone conducive to their end goal?
    - How well they meet rhetorical expectations
        - Do they address counterarguments?
        - Do they consider important nuances?

5. Does the examinee summarize their main point(s)?

"Manner" by Tier

Percentage of Responses that Accomplish the Sociocultural Function				
	Tier One	Tier Two	Tier Three	Tier Four
Does the examinee summarize their main point(s)?	54%	37%	28%	10%

Tier One Response

Question: If the U.S. were to retain or redirect our foreign aid to domestic projects, **how would that alter the foreign perception of Americans?**

7. … I think were we to withdraw or redirect foreign aid, …
13. I think it would be one of the most destabilizing factors **THESIS**
14. that I could possibly imagine.
16. … For instance, Palestine. …
18. We are providing aid to Palestine. How does that work? **JUSTIFICATION**
19. Well, that works because we are still trying to have a connection with them.
20. We still want a dialogue.
21. You remove the aid, and potentially you remove the impetus for that dialogue.
27. … But to redirect that foreign aid, I think, would be ill-advised,
28. simply because of the role that we have
29. in helping get countries talking with one another. **SUMMARY**

| Question's language is used	Position is repeated	Justification is restated

"Manner" in High-level Speakers

- Summaries are not required but tend to happen more frequently among higher-level examinees

- Along with restating the thesis and main justification, summaries often repeat language from the question

Conclusions

- Linguistic forms alone may not effectively distinguish between Level 4 and Level 5 speakers

- Sociocultural functions may help differentiate between these higher- and lower-level speakers

- Sociocultural functions that significantly distinguish Level 5 from other levels include the following:
  - Ability to deliver a clear thesis
  - Ability to stay on topic and answer the original question
 - Ability to explore novel but related ideas

Sociocultural Functions
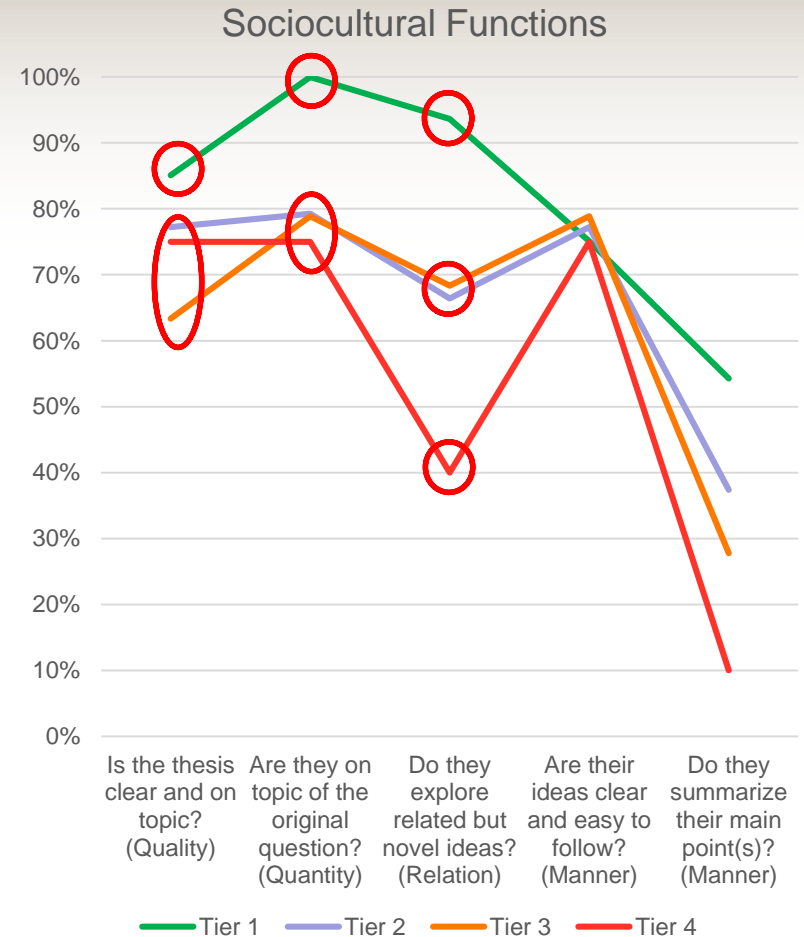
FBI Speaking Applications

- ILR Level 5

- Can set the tone of interpersonal official, semi-official and non-professional verbal exchanges for varied audiences, purposes, tasks, and settings. Can effectively weave together different concepts to create novel ideas and perspectives.

- Employs a wide range of rhetorical devices exceptionally well to achieve a desired goal or effect.

- Presents information to support ideas clearly, thoroughly, and concisely.

- Can use intonation, tone, stress patterns, humor, and other features to add emphasis, certainty, uncertainty, authority, or other stances.

Linguistic Toolbox
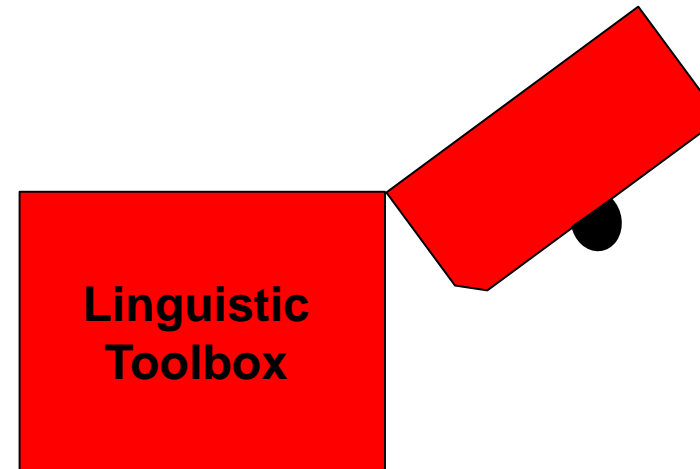
THANK YOU!

Rachel Brooks
rlbrooks@fbi.gov

Tanner Call
tscall@fbi.gov

Appendix

Question

Is their thesis clear and succinct?

Are they on topic of the original question?

Are their ideas clear and easy to follow?

Do they explore related but novel ideas?

Do they summarize their main point(s)?

Interactive

Do they appropriately greet the rater and establish rapport?

Do they bring up the issue in a non-confrontational way?

Do they proactively address and/or empathize with the concerns the rater mentions or implies?

Do they offer reasonable solutions?

Do they offer to help resolve the problem?

Do they use the appropriate register?

Do they summarize the plan near the end?

Do they close the conversation on a friendly note to maintain a positive relationship?

Non-Interactive

Do they have an appropriate greeting and acknowledge the audience?

Is their thesis clear and succinct?

Do they explain why this topic is personally important to them?

Do they support their thesis with logical arguments?

Are their arguments organized and easy to follow?

Do they address possible counterarguments?

Do they address the needs of the audience?

Do they summarize their overall argument?

Do they close their argument appropriately?

Do they use the appropriate register?