Integrated speaking tasks used for teaching and assessment

Rie KOIZUMI
rie-koizumi@mwa.biglobe.ne.jp
Juntendo University
Presented at Ibaraki JALT on February 17
Integrated tasks

• Need to use two or more skills to complete the task
  – Reading + Speaking    Listening + Speaking
  – Reading + Listening + Speaking
  – Listening + Reading + Speaking

• vs. Independent tasks
  – e.g., Explain what you did during summer vacation.

• The new Course of Study: Junior and senior high school teachers should use integrated tasks to enhance balanced English proficiency. This also applies to university contexts.
Teaching example (based on Fukazawa, 2011)

• Read a text and check comprehension. Practice the pronunciation of important words. Read the text aloud.
• Receive a handout that lists keywords or shows a summary with fill-in-the-blanks.
• Form pairs or groups.
• Retell the text. Each individual in the pair/group will retell a different section.
• Add thoughts and opinions about the text.
• Some pairs/groups make a presentation in class. Teacher and peer feedback is given.
Text and handout example

• Text: Prof. Shinya Yamanaka was awarded the Nobel Prize in Physiology or Medicine in a ceremony held Monday at the Stockholm Concert Hall. Yamanaka is the first Japanese to win the physiology or medicine prize in 25 years, following Susumu Tonegawa. (From Daily Yomiuri Online, Dec. 12, 2012)

• Handout: Please fill in the blanks and retell the story. Add your opinion on the story at the end.

• Yamanaka got the (                                        ) last (                 ). He is the (                               ) to win the prize in 25 years. I think this story is (                              ).
Example answer

• Yamanaka got the (the Nobel Prize in Physiology or Medicine) last (Monday). He is the (first Japanese person) to win the prize in 25 years. I think this story is (excellent and I am proud of his success).

• Presentation structure
• Examples of tasks
• Advantages and weaknesses
• Scoring of integrated tasks
• Issues
Examples: TOEIC(R) Speaking Test

• Respond to questions using information provided (no planning, 15 sec speaking; Reading + Speaking)
  – e.g., Could you give me information about the activities in the morning, before lunchtime?

• Propose a solution (30 sec planning; 60 sec speaking; Listening + Speaking)
  – e.g., In your response, be sure to show that you recognize the problem, and propose a way to dealing with the problem. (ETS, 2012)
• **Task features**
  • Information transfer (0 to 3 for each question)
  • Summary + add your opinion (0 to 5)

• **Scoring points:**
  • Pronunciation   Intonation and stress
  • Grammar
  • Vocabulary
  • Cohesion
  • Relevance and completeness of content
TOEFL iBT(R) (ETS, 2004)

• Reading + Listening + Speaking

• Respond to questions based on the passage and the conversation (campus-related & academic; 30 sec planning, 60 sec speaking)

• The ability to speak clearly and coherently and to (integrate and) accurately convey (key) information provided
TOEFL iBT(R)

• Reading (OR Listening) + Speaking
• e.g., Respond to questions based on the passage (campus-related & academic; 20 sec planning, 60 sec speaking)

• The ability to briefly summarize the problem, state which solution you prefer, and explain reasons for your preference.
• OR The ability to speak clearly and coherently, and to accurately convey information provided
• e.g., The man expresses his opinion about the proposal described in the article. Briefly summarize the proposal. Then state his opinion about the proposal and explain the reasons he gives for holding that opinion.

• Explain how the examples from the professor’s lecture illustrate the relationship between verbal and nonverbal communication.
R (OR L) + S

• e.g., Briefly summarize the problem the speakers are discussing. Then state which of the two solutions from the conversation you would recommend. Explain the reasons for your recommendation.

• Using points and examples from the lecture, explain the importance of visual elements in painting.
• **Task features**

• Summary (+ add your opinion) (0 to 4)

• **Scoring points:**

• Delivery

• Language use

• Topic development
  – Accuracy and completeness of the content provided
Story Retelling Speaking Test (SRST)

- Reading + Speaking (Koizumi & Hirai, 2012)
- Read a story silently (2 min)
- Answer three comprehension questions orally (30 sec)
- Retell the story and include opinions by looking at only the keywords (2.5 min)
• Last year, Bob and his sister Jean went to Florida for their summer vacation with their parents. They visited Florida for one week. The weather was very nice. So everyone was really happy. ...

• Q1: Where did Bob and his family go for their vacation?

• Retelling + opinion
• Keywords: Bob, Jean, Florida, beach
• Rating scale: EBB (empirically derived, binary-choice, boundary-definition) scales (Upshur & Turner, 1995)
1. Communicative Efficiency (伝達能力)

   With some fluency
   (流暢さはややある)
   No  Yes

1. Coherent story retell with no long awkward pauses
   （話に一貫性があり、長く不自然なポーズがない）
   No  Yes

2. Elaborations of the story with sufficient opinions
   （話の詳細を含み、意見を十分に述べている）
   No  Yes

3. With few hesitations and self-corrections
   （言いよどみや言い直しがほとんどない）
   No
   4  Yes  5
2. Grammar & Vocabulary (文法と語彙)

A variety of sentence patterns with almost no grammatical or lexical errors
(様々な文構造を使い、文法や語彙の誤りがほとんどない)

No  Yes

With some verbs marked for incorrect tense and aspect
(いくつかの動詞の時制やアスペクトが正しく使えていない)

Yes  No

With frequent grammatical and lexical errors
(文法や語彙の間違いが頻繁にある
または発話が少ない)

Yes  No

1  With some prominent grammatical and lexical errors or
lack of use of pronouns and prepositional phrases
(文法や語彙の誤りが目立つ。あるいは代名詞や前置詞句をあまり使用していない)

Yes  No

2  3
3. Pronunciation (発音)

Accurate pronunciation with correct stress and natural intonation
（正確な発音でかつ強勢位置が正しく、イントネーションも自然である）

With almost no prominent prosodic errors
（目立った韻律上の誤りがほとんどない）

Yes

With frequent prosodic errors
（韻律上の誤りが頻繁にある）

No

Yes

With a strong accent
（なりりが強い）

Yes

No

1

2

3

4

5
Advantages

• Pedagogical or test tasks become similar to real-life tasks.

• The skills developed or tested are likely to be useful outside the classroom.

• Input texts are likely to be used in the output.

• This enhances the acquisition of English grammar, vocabulary, speaking, and other skills.
Weaknesses

- Integrated tasks are more difficult than others because they require more skills. The text/instruction should be adjusted to students’ levels.
- When students do not perform well, it is hard to identify why they failed.
  - Because they could not comprehend the text well
  - Because they could not express themselves well
  - Mix of the two
- To understand why students failed, teachers can use independent tasks OR select easy texts for input.
Case for/against Writing + Speaking

• Writing --> Speaking
  • Students are likely to read their written texts aloud.
  • If the development of spontaneous speaking skills is the focus, writing in detail before speaking should be discouraged.
    – Writing only keywords or phrases is acceptable.
• Speaking --> Writing works better.
  – They can focus on meaning first and then on form.
Assessment tasks

• Could be the same as teaching tasks or similar but different from them.
  – Using both the same (old) and similar tasks gives students the chance to review and to try something new using the skills developed in class.

• Before the test, teachers can give students a task list and ask them to review the tasks they performed in class.

• During the test, students select one task card from a box containing old task cards and another from a box containing new task cards.
Scoring of integrated test tasks:
Example from a medical university

• Reading + Speaking
• Read a text. Tell the test-takers that retelling comes next.
• Answer comprehension questions while reading the text.
• (1) What did Yamanaka and Gurdon win?
• (2) Why were they awarded the Nobel Prize?
  – A reason common for both
  – A unique reason for Yamanaka/Gurdon
Instruction

• Your listener does not know this story. **Retell** as much of the story as you can in English for 30 seconds. You can use the keywords in the retelling. At the end of your retelling, be sure to include your **opinions** on the story. You will get more points if you paraphrase.

• Keywords: Yamanaka   Gurdon   Physiology because mature cells reprogram
Order the five students

• The highest: Student ?
• 2nd:
• 3rd:
• 4th:
• 5th:

• Reasons?
Order the five students

• The highest: Student C
• 2nd: Student B (grammar error)
• 3rd: Student A (content error but paraphrase)
• 4th: Student E (content error; no opinion)
• 5th: Student D (no summary)
Scoring: Points to consider

• Language use: Vocabulary, grammar, and pronunciation
  – Is the speech comprehensible, accurate, and appropriate?
  – Does it use paraphrased words?

• Content
  – Does the speech content reflect the prompt content?
  – Does it use the keywords provided?

• Delivery
  – Is the speech fluent?
  – Is it logical? Does it contain discourse markers?
Possible scale: Holistic (based on TOEIC S)

- **3:** The response is generally full, relevant, and appropriate to the task (summary and opinion). Information from the text is accurate. Errors are limited, and they do not obscure meaning. It contains paraphrased words. (Student C)
- **2:** The response is partially effective, but it is not complete, fully appropriate, or fully accurate. The use of vocabulary or grammar does not cause misunderstanding, but minor errors are noticeable. (Students B and A)
- **1:** The response does not complete the task effectively. Relevant information is not conveyed successfully. Some vocabulary or grammar choices cause misunderstanding. (Students E and D)
- **0:** No response OR the response is unrelated to the test.
Possible scale: Analytic

• **Language use: Vocabulary, Grammar, Pronunciation:**

• 2: The language is comprehensible, accurate, and appropriate to the question. (Students A, C, D, and E)

• 1: The language is comprehensible but sometimes inaccurate or inappropriate. (Student B)

• 0: The language is largely incomprehensible, inaccurate, or inappropriate.
Possible scale: Analytic

• **Content** 2: Information from the prompt is accurate. (Students B and C)

• 1: Information from the prompt is mostly accurate. (Students E and D)

• 0: Information from the prompt is largely inaccurate. (Student A)

• **Paraphrasing**: If some words are paraphrased, give one extra point. (Student A)
Possible scale: EBB

• Communicative efficiency
• Grammar and vocabulary
• Pronunciation (not used this time)
• Total (0 to 6)

• Can be simplified or elaborated based on the learning objectives.
1. Communicative efficiency (伝達能力)
   With a summary and an opinion (要約と意見を両方述べている)

   No  Yes

   0 → Coherent story retell with no long awkward pauses or repairs
       （話に一貫性があり、長く不自然なポーズや修正がない）

       No  Yes

       1 → Elaborations of the story using all the keywords
           with sufficient opinions
           （キーワードを全部使い、話の詳細を含み、意見を十分に述べている）

           No  Yes

           2  3
2. Grammar and vocabulary (文法と語彙)

With almost no grammatical or lexical errors
(文法や語彙の誤りがほとんどない)

No | Yes
--- | ---
3

With frequent minor errors but no major errors
(マイナーな間違いが頻繁にあるが理解は妨げられない)

No | Yes
--- | ---
2

With some major errors or with few sentences (理解を妨げる誤りがある、または発話が少ない)

Yes | No
--- | ---
1

0
Selecting rating scales

• Efficiency: Holistic
• Feedback to students: Analytic or EBB
  – Holistic scale + comments (verbal or written)
• Example: Student E’s score: 1
• Comments: Check who discovered mature specialized cells. Add your opinions.
Issues in developing tasks

• Whether to ask test-takers to add their opinions (or other types of creative tasks)
  – Important to include this, because Japanese learners of English are poor at expressing their opinions (National Institute for Educational Policy Research of Japan, 2007).

• How many tasks are used
  – Better to use multiple integrated tasks along with multiple independent tasks to develop and assess a wide range of skills
Issues in scoring tasks 1
• What type of rating scale to use
• How many levels to use in the scale (e.g., 0 to 3)
• How the content of the retelling and opinion is evaluated
  – E.g., How many blanks students can fill appropriately OR how many points they can retell successfully.
  – Criterion of content is necessary for integrated speaking tasks.
  – Which to value more, the content or language
Issues of scoring tasks 2

- Whether the use of the same words in speaking is penalized or encouraged
  - Depends on learners’ proficiency
- Whether to evaluate fluency
- What level descriptors to use
  - Think of teaching objectives. Write down the performances you expect from most students (e.g., 2 on a scale of 0 to 3), top students, and low-achieving students.
After administrating tasks

• Have students reflect on their own performance
  – Write down what you said. Write good and weak points of your test performance.
  – Write expressions you wanted to use during the test. Look them up in the dictionary for memory and future use.

• Give students scores and feedback early.
After scoring tasks

• Write down unexpected performances that you had difficulty scoring.
  – This improves the rating scale.

• Compare each student’s scores with his or her other performances or scores. Ask a second rater to rate or rate again by yourself after an interval.
  – This is to check the consistency and divergence.
  – We expect the overall tendency to be the same, but there are some fluctuations in scores. This step checks the reliability and validity of the tasks and scoring.
References 1


References 2


