

## DOCUMENT RESUME

ED 453 674

FL 026 734

AUTHOR Yoshida, Yuichi  
TITLE Authentic Progress Assessment of Oral Language: Oral Portfolios.  
PUB DATE 2001-02-00  
NOTE 34p.  
PUB TYPE Information Analyses (070)  
EDRS PRICE MF01/PC02 Plus Postage.  
DESCRIPTORS Class Activities; \*English (Second Language); Foreign Countries; High School Students; High Schools; Instructional Materials; \*Oral Language; \*Portfolio Assessment; Second Language Instruction; Second Language Learning; Secondary Education; Student Evaluation; Teaching Methods; Test Format; Test Validity  
IDENTIFIERS \*Japan

## ABSTRACT

Student testing in Japan is not effectively used. In many cases test results are used only for ranking and sorting students into whatever the designated purposes of the tests are. They are not focused on monitoring individual student progress. This is an especially inadequate approach for teaching oral communications courses. This paper proposes a different assessment tool which emphasizes monitoring individual student progress and achievement on a daily basis: the oral portfolio. Part one of this paper explains the present situation in oral language instruction and assessment in Japanese high schools. Part two defines authentic assessment as well as key factors and technical qualities. Part three focuses on key concepts in portfolios that are essential to authentic assessment, and how these processes could be implemented in Japanese high school classrooms. The final part demonstrates how authentic oral portfolio assessment can be incorporated into classroom instruction and provides sample activities. Oral portfolios also help learners become more self-directed and self-starting by encouraging them to take more responsibility for their own learning. Appended are three forms appropriate for handouts: an analytic scoring rubric for formal speaking; formal speaking peer-assessment sheet; and sheet for Self-Assessment of Process. (Contains 14 references.) (KFT)

# Authentic Progress Assessment of Oral Language: Oral Portfolios

by

Yuichi Yoshida

PERMISSION TO REPRODUCE AND  
DISSEMINATE THIS MATERIAL HAS  
BEEN GRANTED BY

*Yuichi Yoshida*

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

This document has been reproduced as  
received from the person or organization  
originating it.

Minor changes have been made to  
improve reproduction quality.

• Points of view or opinions stated in this  
document do not necessarily represent  
official OERI position or policy.

Kumamoto Senior High School

February, 2001

FD026734

## Table of Contents

<b>Introduction</b> .....	<b>1</b>
<b>I. The Present Situation in Japan</b> .....	<b>2</b>
<b>II. What is Authentic Assessment?</b> .....	<b>3</b>
<b>A. Definition of Authentic Assessment</b> .....	<b>3</b>
<b>B. Key Factors in Authentic Assessment</b> .....	<b>4</b>
<b>C. Validity and Reliability of Authentic Assessment</b> .....	<b>6</b>
<b>III. Key Concepts of Portfolios</b> .....	<b>7</b>
<b>A. Systematic and Reflective Collection</b> .....	<b>7</b>
<b>B. Self-assessment</b> .....	<b>8</b>
<b>C. Collaboration</b> .....	<b>10</b>
<b>IV. Oral Portfolios</b> .....	<b>11</b>
<b>A. What to Include</b> .....	<b>12</b>
<b>B. How often to assess</b> .....	<b>13</b>
<b>C. How to assess</b> .....	<b>15</b>
<b>D. Management</b> .....	<b>17</b>
<b>V. Simulation of Formal Speaking Portfolio</b> .....	<b>20</b>
<b>Conclusion</b> .....	<b>24</b>
<b>References</b> .....	<b>26</b>
<b>Appendices</b> .....	<b>28</b>

## Introduction

An average Japanese senior high school student would take a written test of English almost every month, but what are the purposes of those many tests? Most of them are achievement tests, a few are proficiency tests, and one, if any, a diagnostic test. The question is whether both the teachers and students really know their purposes, which seems doubtful. Every time students get their tests back, all they are interested in is their entire test scores, T-scores, and rankings among their group. They don't care very much about what skills they have to improve and how they should improve them. On the other hand, many of the teachers look only at the average test scores of their classes and seldom care about how and why each student has improved or not improved since the last test or for a certain period of time.

What has made these situations? In many cases test results are used only for ranking or sorting the students whatever the designated purposes of the tests are. Of course, the ranking function of tests is necessary, but there should be tests or assessments which are done with another purpose: to monitor students' progress and help them work actively toward their own goals. This is especially important in Oral Communication classes where traditional assessments are not realistic.

This paper will propose a more useful assessment tool which could be used in Japanese classrooms in order to monitor students' progress in speaking on a daily basis: the Oral Portfolio. Chapter 1 will explain the present situation in oral language instruction and assessment in Japanese high schools. Next, the definition of authentic assessment as well as its key factors and technical qualities will be discussed in Chapter 2. Chapter 3 will focus on key concepts of portfolios, which are essential in authentic assessment. This will be followed by a chapter describing the details about how oral portfolios could be implemented in Japanese classrooms. Finally, the last

chapter will demonstrate how an oral portfolio assessment is incorporated into classroom instruction with a complete example of activities for formal speaking portfolios.

### **I. The Present Situation in Japan**

In Japan more and more oral activities are done in all English classes because communication has been very much valued in English education for the past decade. Since "Oral-Aural Communication A/B/C" was introduced by the Ministry of Education, a great many efforts have been undertaken to encourage students to speak/listen to English for communication, and the students get involved in more and more activities in pair and group work. Thanks to help from Assistant Language Teachers, who are native speakers of English, the language spoken there is more "authentic." The students seem to enjoy those speaking activities and to be taking more risks to speak out in the classroom.

However, many teachers don't evaluate students' oral language on an ongoing basis. If the teachers ever evaluate oral language, it is through interviews and oral presentations which are done very few times a year without any systematic plans. Why? First, many teachers seem to be just satisfied with the fact that their students are enjoying talking. Second, college entrance examinations, which greatly influence high school education, give a few listening tests and no speaking tests at all. They may give test questions on paper which they argue are measuring "speaking ability." Third, many teachers believe it is next to impossible to secure high reliability of speaking tests, which demand from them more subjective judgement than other tests. The large class size makes things more difficult. But, most importantly, Japanese teachers of English are simply not informed of ongoing progress achievement assessments carried out

based on the concept that assessment is done to help students improve. Therefore, the main purpose of this paper is to provide Japanese teachers of English with information on "authentic assessment," which I believe will improve the instruction in Japanese classrooms.

## **II. What is Authentic Assessment?**

Before discussing details of oral portfolio assessment, this chapter will explain characteristics of authentic assessment, which is the basis of oral portfolio assessment.

### **A. Definition of Authentic Assessment**

Since the 1990's the concept of assessment has been changing in the United States. In place of multiple-choice tests, or "machine-generated scores," new student-centered assessments are appearing (Katz 2000, p.138). These assessments are often called "alternative assessments" compared to "traditional assessments." Daring-Hammond (1994) gave these alternatives a name: "authentic" assessments (p.5), because students are asked to do "real-world" tasks and are evaluated based on criteria corresponding to the importance in "actual performance." She used the term performance-based assessment interchangeably with the term authentic assessment. But Meyer (1992) claimed that there is a difference between authentic and performance assessment. According to her, "Performance assessment refers to the kind of student response to be examined; authentic assessment refers to the context in which the response is performed." (p.40). In other words, all authentic assessments are performance assessments whereas all performance assessments are not necessarily authentic assessments. Does it mean that assessment of oral proficiency has to be done in real-life situations in which a student communicates with

native speakers of English, which is almost impossible in Japanese classrooms? The answer is "No." Then what makes an assessment authentic?

O'Malley and Valdez Pierce, the authors of "Authentic Assessment for English Language Learners," used the term "authentic assessment" in their book to describe "the multiple forms of assessment that are consistent with classroom goals, curricula, and instruction." (p.4). They argue that multiple-choice tests are not authentic because their focus is on "discrete skills" and not on "representations of classroom activities." (p.2). By their definition, in order for an assessment to be authentic, it has to be authentic in the students' world of learning the language, that is, their classroom. This agrees with the statement by Paris and Ayres (1994), that "what is authentic in one school or in one state may not be authentic in other schools or states" for the reason that "authentic assessment is defined by the situational appropriateness of teaching and learning practices." (p.7). Meyer (1992) introduced another view of authenticity, saying that "real life may be in terms of the student." She illustrated this by saying, "The significant criterion for the authenticity of a writing assessment might be that the locus of control rests with the student; that is, the student determines the topic, the time allocated, the pacing, and the conditions under which the writing sample is generated." (p.40).

Based on the discussions above, in this paper I would like to define authentic assessment as a set of various kinds of performance assessments which are real in terms of the classroom and in which students as well as the teacher have active control.

## **B. Key Factors in Authentic Assessment**

O'Malley and Valdez Pierce (1996) mentioned eight types of authentic assessment:

oral-interviews, story or text retelling, writing samples, projects and exhibitions, experiments or demonstrations, constructed-response items, teacher observations and portfolios. As they stated, many teachers have employed these kinds of assessments. They pointed out, however, that these assessments have been done informally, so that information acquired from them are not "systematic." (p.11). For example, teacher observations are not assessments unless the teacher keeps systematic records of students' progress in observing their performance. In this respect, a portfolio, "a purposeful collection of student work" plays an important role thanks to its best asset: "collecting information purposefully and systematically over time to reflect learning with regard to instructional objectives." (p.14). (I will discuss more about portfolios in the following chapter.)

O'Malley and Valdez Pierce (1996) also suggested that self-assessment should be used in all of these assessments except in teacher observations. Paris and Ayers (1994) emphasized the importance of student self-assessment in making students self-regulated learners who get actively involved in their own learning. They argue that self-regulated learners "select goals to pursue and work on a variety of tasks," "adjust the challenge of the tasks they choose," "know how to use the resources available to them in a classroom," "collaborate" with their classmates and teachers, "construct meaning," "evaluate and interpret their behavior in ways that promote further effort," and "monitor their own performance and evaluate their progress." (p.26-30). This is exactly what I, as a teacher, want my students to be able to do. Portfolios will play another important role besides the systematic collection of evidence: in self-regulated learning, because students "select samples of their own work" in making them (O'Malley and Valdez Pierce 1996, p.5).



### **C. Validity and Reliability of Authentic Assessment**

Before going on to further and concrete discussion on portfolios, two vital qualities of assessment have to be considered in this authentic assessment. They are its validity and reliability.

As for validity, O'Malley and Valdez Pierce (1996) and Darling-Hammond (1994) showed concern about content validity and consequential validity. Content validity is "its ability to describe the nature of performance that results from learning." (Darling-Hammond, p.11). To heighten content validity, or curriculum validity, consideration should be given to establish "correspondence between local curriculum objectives and the content of the assessment," and to make "the assessments represent thinking skills in your local curriculum," stated O'Malley and Valdez Pierce (1996, p.25). Consequential validity is "the extent to which an assessment tool and the ways in which it is used produce positive consequences both for the teaching and learning process and for students." (Darling-Hammond, p.11). O'Malley and Valdez Pierce (1996) explained that assessments are consequentially valid if "they lead teachers to focus on classroom activities which support student learning and responsive to individual student needs." (p.26). They further argued that if a teacher is doing on-going authentic classroom assessment, he gets information on whether his instruction is beneficial to his students, and therefore he should review his instruction to make it more effective.

Reliability, or the consistency of the assessment can be threatened by "the person or among the people collecting information," "the person about whom information is being collected," and "the procedures used for collecting information." (Genesee and Upshur, 1996, p.58). As for authentic assessment, the greatest concern is paid to the first factor, or raters, because teachers make judgements in scoring. In order to

enhance inter-rater reliability, use of scoring rubrics and rater training are suggested. O'Malley and Valdez Pierce (1996) illustrated rater training with six steps: "Orientation to the assessment tasks," "Clarification of the scoring rubric," "Practice scoring," "Record the scores," "Check reliability," and "Follow-up." (p.21-24). In Japan, several teachers are in charge of the students in one grade; therefore, establishing inter-rater reliability is crucial. As far as the last two factors are concerned, reliability can be improved by "assessing on several occasions." (Genesee and Upshur, 1996, p.59).

### **III. Key Concepts of Portfolios**

I'd like to begin this chapter on portfolios with a quote from O'Malley and Valdez Pierce (1992).

Performance assessment and portfolios are complementary approaches for reviewing student language development and academic progress. Together they represent authentic assessment, continuous assessment of student progress, possibilities for integrating assessment with instruction, assessment of learning and higher-order thinking skills, and a collaborative approach to assessment that enables teachers and students to interact in the teaching/learning process. (p.2)

As shown in this and as I briefly commented in the previous chapter, in authentic assessment portfolios play important roles; therefore, in this chapter I'd like to explain some key concepts of portfolios.

#### **A. Systematic and Reflective Collection**

A portfolio is traditionally a collection of best works by hopeful artists in order to show their talents. Because of this origin, the word 'portfolio' tends to remind many people of "showcase portfolios" (O'Malley and Valdez Pierce, 1996, p.36), which show only the final products and don't show the process to reach the products; however, there are other types of portfolios. According to O'Malley and Valdez Pierce (1996),

they are "collection portfolios" and "assessment portfolios." The former consists of every work of a student, from first draft to final product, achieved throughout the year, while the latter is a "systematic collection of student work, student self-assessment, and teacher assessment." (O'Malley and Valdez Pierce, 1996, p.37). Even though collection portfolios may show process as well as product, it cannot be said that they are carefully planned with a specific assessment purpose in mind. As stated in the previous chapter, the key factors of authentic assessments are being systematic and self-assessment. Thus, the type of portfolios I describe in this paper is assessment portfolios, which "afford students the opportunity to reflect on their development in a certain area." (Jensen and Harris, 1999, p.212).

## **B. Self-assessment**

O'Malley and Valdez Pierce (1996) expressed the importance of self-assessment as follows: "Without self-assessment and reflection on the part of the student, a portfolio is not a portfolio." (p.35). But some teachers doubt that a student can assess himself accurately. A great deal of research has been done on "response effects" in self-assessment, or "tendencies for certain people to respond to factors other than question content." (Heilenman 1990, p.175). In her research, Heilenman found evidence of two response effects: acquiescence effects ("a tendency to respond positively") and overestimation effects. (p.175).

LeBlanc and Panchaud (1985) reported, on the other hand, their successful experiment of using "self-assessment as a second language placement instrument." (p.673). What, then, makes self-assessment reliable? Heilenman, admitting response effects to be "natural consequences of asking people to give judgement about themselves under less than ideal circumstances" (p.191), suggested that "making

criteria known and giving learners access to records (i.e. video tapes) of their performance will improve accuracy." (p.195). In other words, students need some assistance from their teachers. As O'Malley and Valdez Pierce (1996) put it, "self-assessment is a process through which students must be led." (p.39).

From the discussion above, you can tell it is important that criteria are clearly stated in order for students to assess themselves accurately. But my experience tells me that it is not enough. I have tried self- and peer-assessment of students' oral presentation with an assessment sheet on which criteria are clearly written, only to find students failing to make fair and accurate assessment of themselves and their peers. Why? It is partly because my Japanese students are not familiar with this kind of assessment, especially unfamiliar with evaluating their classmates honestly or severely. (I will discuss this problem further in Chapter 4) It's partly because the criteria are external; that is, it was I who set up the criteria, not my students. Therefore, they do not know why the characteristics stated in the criteria are important, nor do they know the standards, for example, they don't know how slow or fast good speech should be.

Sperling (1993) made a suggestion on this problem, illustrating how a 4th grade teacher had set standards together with her students. The teacher, first having reached a pleasing set of criteria by herself, had her students experience the same procedure. Her students were given four writing samples "each with the attributes of one of the four grades she would give" (p.74) and they were told to grade the samples in pairs. Then they were asked to explain the reasons for giving the papers respective grades. In this way, she led her students to play an active role in setting their own standards. This illustration matches with the proposals by O'Malley and Valdez Pierce (1996) on setting criteria: "working with students to specify the criteria," "providing samples of exemplary work (called benchmarks)," "asking students to identify the

characteristics of exemplary work," "having students work in cooperative learning groups to examine samples," and "applying the criteria as a group to actual work samples." (p. 39-40).

In portfolio assessment it can be said that self-assessment means students assess "themselves" based on criteria set by "themselves."

### **C. Collaboration**

It is not only in setting initial criteria that a student cooperates with his teacher and classmates; in fact, teacher and students collaborate in every stage of making portfolios. For example, even after they have set initial criteria, as students go on using the criteria, they may find it necessary to modify the criteria. If this happens, they will discuss again how they can improve the criteria with their classmates and teacher.

Sperling (1993) called the activities like the one illustrated in the previous section "collaborative assessment." (p.73). She wrote, "Sometimes a student's most probing thinking occurs during an assessment session with a peer." (p.75). As an example, she mentioned a student asking her partner for advice on how to improve her own writing. The student would learn something from her partner to improve hers, but at the same time the partner, giving some suggestions for improvement, reflects her own writing and may improve it, too. "Collaborative assessment" surely encourages collaborative learning because the students have clear criteria and a picture of good final product, or benchmark, in their minds.

A student also works with his teacher. Jensen and Harris (1999) illustrated how students achieve "active reflection" through their "Public Speaking Portfolio." (p.214). First, by having two kinds of "snapshots" of students' performances: "static snapshots" and "dynamic snapshots," students can "identify past/current experiences."(p.214).

According to Jensen and Harris, the static snapshot, which is made up of the videotape and evaluations by peers and teacher, illustrates "a picture of speech performance at one point in time," whereas the dynamic snapshot, made up of "the Speech Log," portrays how the student has reached a certain performance, which helps them to find necessary "changes for the future."(p.215). Next, "by comparing past and future actions", they continued, "students recognize what future behavior is needed to bridge the gap" between "what they did and what they want to do in the next speech." (p.215). In the final step a student takes a specific approach to bridge the gap. However, they argued, this is not easy for the student because he may not know what approach he should take, or, even if he does, he may not be eager enough to employ the approach or he may not employ it long enough. Now is the time when the teacher assists the student, or they work together. In a dialogue journal like "Speech Process Log" or in a conference, the teacher extends a necessary helping hand to the student, or encourages him to continue his effort, and through doing this the teacher reflects on his own instructions and takes some action to improve it. Therefore, in portfolio assessment, it is very important to maintain opportunities to communicate between the student and teacher, through which they will develop a good relationship to promote collaborative learning.

#### **IV. Oral Portfolios**

As stated in the first chapter, many Japanese teachers have two problems: lack of systematic assessment of oral language and lack of assessment which evaluates not only final products but process to them. Oral portfolios can be a solution to both of these problems. This chapter will discuss details about how oral portfolios could be implemented in Japanese classrooms.

## A. What to Include

In order to decide the contents of portfolio, it is necessary to have a clear picture of your assessment purpose and it requires examination of your local curriculum objectives. Once you have specified your assessment purpose, the next question is what kinds of entries will meet the purpose. (O'Malley and Valdez Pierce, 1996). O'Malley and Valdez Pierce proposed "having two types of entries for all students: required and optional entries." (p.47). Required entries are the "basis for assessment" and consist of "student self-assessment, samples of student work, and some type of teacher assessment," while optional entries supply complementing additional information, or "evidence of not only what they have produced but also the process." (p.47).

Describing the process mentioned in a real Japanese situation may be helpful. The course in which I am planning to employ portfolio assessment is Oral Communication B, whose officially set main focus is to improve students' listening ability. In my school, however, the focus is not only on listening but also on speaking because in everyday life situations listening takes place more combined with speaking than by itself. Since my oral portfolios are geared for assessing speaking ability, here my focus is on speaking. The course has two main objectives for speaking. One is to make students able to communicate orally in English in every day life situations. The other is to make students able to give a formal presentation in English. The former requires students to engage in more improvisation, and the recording of their performance will be done in the classroom without much preparation on the part of students. In this case, the entries are basically "required entries" as defined by O'Malley and Valdez Pierce (1996, p.47). On the other hand, the latter is more formal and requires students' preparations at home and in the classroom and involves more



process, which means "optional entries," such as some kind of a dialogue journal, or "Speech Process Log" (Jensen and Harris, 1999, p.215), are included in addition to "required entries." In both cases, student's self-assessments and teacher assessments of the entire process in each cycle, term, and year will be included. Table 1 is the list of entries which will be collected throughout the year.

Table 1: Contents of Portfolios

Kind of Portfolio	Contents
Formal speaking portfolio	4 audio tapes: first and second practice are recorded on each tape 1 video tape: all four final presentations are recorded 4 self-assessment sheets: 1 for each cycle 4 peer-assessment sheets: 1 for each cycle 12 teacher-assessment sheets: 4 for first practice, 4 for second, 4 for final 4 process self-assessment sheets: 1 for each cycle 4 process teacher-assessment sheets: 1 for each cycle dialogue journals (number varies depending on students) 1 student year-end final summary sheet 1 teacher year-end final assessment sheet
Informal speaking portfolio	3 audio tapes: one tape per term, each contains four recordings 12 self-assessment sheets: 1 for each recording 12 peer-assessment sheets: 1 for each recording 12 teacher-assessment sheets: 1 for each recording 3 term-end self-assessment sheets: 1 for each term 3 term-end teacher-assessment sheets: 1 for each term 1 student year-end final summary sheet 1 teacher year-end final assessment sheet

## B. How often to assess

Based on the discussion in the previous section, oral portfolios have two subdivisions: *formal speaking portfolios* and *informal speaking portfolios*.

To finish one cycle of making formal speaking portfolios would require students to record their oral work two to three times and the students and teacher to assess each recording. One complete cycle includes five steps. Activities involved in the cycle are basically as in Table 2. The process is exactly like writing an essay and takes a certain amount of time, at least three weeks, and repeating this cycle on different topics



several times completes the entire formal speaking portfolio. Each cycle gives a picture of what a student has done to improve from his first sample to the final presentation over the three-week period of time, and the whole portfolio illustrates the student's growth over one year. Considering the time and energy that Japanese teachers and students can afford, a teacher could have two formal speaking portfolios each term, that is, five to six for a year.

Table 2: Process of Formal Speaking Portfolio

		Student	Peers	Teacher
I	1	Set criteria.	Set criteria.	Set criteria.
II	1	Record the first sample. Listen and self-assess.		
	2		Listen and peer-assess.	
	3			Listen and assess.
III	1	Re-record revised sample based on assessment by himself, peers and teacher.		
	2		Listen and peer-assess.	
	3			Listen and assess.
IV	1	Revise for presentation based on assessment by himself, peers and teacher.		
	2	Make final presentation in class.	Listen and peer-assess.	Videotape presentation. Watch and assess.
	3	Watch the video and self- assess.		
V	1	Self-assess whole process.		
	2			Assess whole process.

Informal speaking portfolios, on the other hand, consist of several single recordings of a student's oral performance interacting with his peer in classroom; completing each entry is much simpler and takes shorter time. The activities involved in each entry would be almost the same as Step I and II in Table 2. Thus, a teacher could have informal speaking portfolios every other week when the formal speaking portfolios are not in progress, which means four entries each term and twelve entries for a year.

The possible time line for the whole oral portfolio would be such as the one outlined in Table 3.

Table 3: Time line for Oral Portfolio

Term	I				II				III		
Month	4	5	6	7	9	10	11	12	1	2	3
Formal Speaking Portfolios	No.1			No.2	No.3			No.4			No.5
Possible Topic	Self-introduction		My dream		My summer vacation		My favorite book		My reflection of the year		
Informal Speaking Portfolios		No.1 & 2	No.3 & 4			No.5 & 6	No.7 & 8		No.9 & 10	No.11 & 12	
Possible Setting	Greetings, Talking over the phone		At a restaurant, Giving directions		On an airplane, At the post office		Interview, A guessing game		Discussion		

### C. How to assess

Each entry is assessed by the student, his peer, and his teacher based on the criteria which they set together. Since one major purpose of portfolio assessment is to monitor students' progress, it is advisable to use one fixed assessment form throughout the year, or at least the term, which illustrates clearly what has improved and what hasn't. As suggested in Chapter II, scoring rubrics would be used for teacher assessment in order to increase reliability. Analytic scoring rubrics rather than holistic scoring rubrics would be suitable for attaining another purpose of portfolio assessment:

promoting students' growth through reflective learning. (See Appendix 1.)

In portfolio assessment a teacher needs to decide whether to assess portfolio itself, and if he so, how. O'Malley and Valdez Pierce (1996) argued, "Whereas the portfolio itself does not receive a grade or a rating, the different entries may be weighted to reflect an overall level of student achievement." (p.37). On the other hand, Case (1994) introduced her successful experience with her 7th grade students in setting "how their portfolios should be graded." (p.46). Their criteria are: "how well the student compares pieces and explains why they were included," "the actual content (Is there enough? Is there too much?)," "neatness and organization," "the clarity and completeness of the cover letter." (p.46). Jensen and Harris (1999), admitting the challenge of evaluating portfolios, explained four basic elements to evaluate: "Table of contents," "Positive appearance and organization," "Completion of all assignments," "Assessment of reflective thought." (p.220). For the last element, the students were asked a question, "Do the portfolio contents show integration of course theory and concepts into the student's own life?" and asked to mention "Specific instances where this was illustrated." (p.227).

Oral portfolios described here are assessment portfolios and are devices to make learning activities and assessments systematically organized, and are not the objects to evaluate themselves. As for formal speaking portfolios, as shown in Table 2 in the previous section, assessment of the whole process is done after every final presentation is made. As for informal speaking portfolios, assessment of students' progress is done on the completion of every four entries. These assessments may be described as assessing portfolios at a shorter period of time. However, considering the previous examples of evaluating portfolios, portfolios could be evaluated at the end of the school year on the following two respects. One is if the portfolio includes all the

assigned entries. The other is how well and clearly a student describes his process and growth over the year in his final summary of his portfolio, which he is assigned to write at the end of the school year.

#### **D. Management**

My review of the portfolio literature revealed how rewarding portfolios are and, at the same time, how challenging it is to employ portfolios. This section explains what the challenges are and how they could be overcome.

##### **(1) Time**

In Japan, there are 40 students in a class and a teacher is usually in charge of three classes, which means 120 audio and video tapes to listen to/watch and assess. If a sample is five minutes long, just listening to 120 of them takes 600 minutes, or ten hours. One way is to stagger the cycle. Instead of three classes of students recording samples on the same schedule, staggering the cycle decreases the number of tapes to assess at a time. Another way is to limit the number of points to assess at a time. For example, in assessing the first sample of a presentation, only the organization of the speech could be assessed, and in assessing the second, accuracy could be assessed. Use of symbols would also save time. In using symbols, the meanings of the symbols should be known to students; otherwise, students would not get the message and that is a fatal loss. (M. Kama, personal interview, January 23, 2001).

Not only teachers but also students could make some contributions for saving time. First, students have to respect the time limit of their presentation. Second, rewinding the tape to where he started the speech and telling his name and the date at the beginning of the recording will be a great help to the teacher. (M. Higginbotham,

personal interview, January 31, 2001).

Another option would be asking Assistant Language Teachers (ALT) for help. With clear criteria and some training, you can establish high inter-rater reliability. When the students have the final presentation, a teacher can divide the class into two and take care of one half while the ALT supervises the other. This enables 40 students to finish their five-minute presentations in two periods of class time. In addition to that, this reduces the stress of the student speaking in front of the entire class, and at the same time it could prevent the other students from getting bored by listening to 40 presentations.

Time is also precious for Japanese students who have to study other subjects as well and are busy with club activities after school. A teacher should make the most of the class time and limit the number of assignments students have to do at home.

## **(2) Self- and peer-evaluation**

As stated in the previous chapter, this type of assessment is new to Japanese students. This can be said of the students other than those in Japan as well. One of the challenges of portfolios Jensen and Harris (1999) mentioned was "a paradigm shift" for students and teachers from solely evaluating the final product to also evaluating the process." (p.223). In order to help students through this change, they argue, a teacher has to clearly explain the purpose and give his expectations clearly. They emphasize the importance of reminding the students of "the value of active reflection," or that "reflection in the process of speech creation can ultimately lead to a more satisfying, well thought-out product." (p.224).

Reflection can be done not only in self-assessment but in peer-evaluation, where a student reflects on his own work though assessing the work of his peer. This could be

an answer for an often-asked question about pairing: "What does a good student get from his poorly-skilled peer?" But, still, a teacher has to pair the students carefully because pairing two poor students is not effective. In formal speaking portfolios, pairing stays the same during one cycle of completing the presentation. It should be arranged so that during the five cycles throughout the year a student will have a peer who has different level of skills in each cycle. In informal speaking portfolio, changing pairing each time will also increase its reliability because in pair work the personality and proficiency of his partner could influence a student performance.

Last, but not the least important, an aspect to consider in peer-assessing is the classroom atmosphere. A warm and supporting atmosphere among the students cannot be nurtured in a short time and only through portfolio assessment. A teacher should use many group activities throughout his instruction. It is advisable to ask students to make only positive comments in the first few peer-evaluations and to introduce more severe comments as they build a good relationship and warm supporting classroom atmosphere.

### **(3) Equipment**

Having a language laboratory (LL) at school would make oral portfolio assessment easier to carry out, but unfortunately few Japanese schools have LLs. In a formal speaking portfolio a student records his first and second samples at home, which can be done without any trouble, whereas in informal speaking portfolio the recording has to be done in the classroom. Even in formal speaking portfolio peer-evaluation has to be done in the classroom. Here a happy situation may solve the problem: almost all Japanese high school students have their own Walkman, with which they enjoy listening to music on the way to and from school. With two Walkman in each pair,

peer-evaluation can be easily done. It is not sure how clearly conversation in pair work can be recorded in the classroom with Walkman, but using microphones would increase the quality of recordings.

#### **(4) Organizing contents**

It is necessary to organize the contents of portfolios so that students can reflect on their own past work and the teacher can monitor their progress. Obviously every entry has to be dated. To accommodate the large number of entries shown in Table 1 in the previous section of "What to Include" in an orderly manner, each set of entries including an audio tape should be in an envelope labeled such as "Presentation No. 1." At the end of the year, there should be four envelopes of formal speaking tasks and three of informal speaking tasks. Those envelopes can be bound together in a binder. O'Malley and Valdez Pierce (1996) recommended that a table of contents should be listed "on the left side of the inside cover" and the scoring rubrics should be listed on the "right side of the inside cover." (p.51). Regular evaluation of portfolios, Jensen and Harris (1999) reported, would help students organize their portfolio regularly as well as save time.

#### **V. Simulation of Formal Speaking Portfolio**

This chapter, through illustration of day by day steps in making a self-introduction speech for a formal speaking portfolio, will give a clear picture of how portfolio assessment is incorporated into classroom instruction.

##### Week 1, Class 1: Setting Criteria

A teacher gives several excerpts of self-introductions as examples, focusing on one

or two aspects such as loudness and speed of speech. Students, after listening to those examples, discuss and figure out what the characteristics of a good speech are. The teacher, according to the students' findings, will make and give to the students scoring rubrics, self-assessments and peer-evaluation sheets the following day so that they can use those sheets as they record their first practice at home. (Since this class could be one of the first classes of the year, it is important for the teacher to make a good start for a desirable relationship with the students by using his personal information as an example.)

After setting criteria, the students could spend the rest of class time brainstorming on what can be included in self-introduction speeches.

#### Week 1, Class 2: Peer-assessment (I)

Each student comes back to the next class with an audio tape recording of his first practice, its self-assessment, dialogue journal entry and a Walkman. The teacher divides the class into pairs, and the students listen to their peer's audio tape with a Walkman and each assesses his peer's speech using a peer-assessment sheet. (See Appendix 2). Then the students exchange the sheets and have a brief conference. The entire process would take less than fifteen minutes. A teacher can use the rest of time for any activity necessary. In the early stages, it is recommended that students share their feelings and impression toward taping their practice, self-assessment, and peer-assessment with their classmates. By sharing several constructive comments from peers, the students learn how to assess their peer's work. As stated in Chapter 3, a teacher should also give his students opportunities to practice peer-assessing by having them assess a couple of samples and discuss among themselves.

At the end of this class, each student hands in his tape, self-assessment sheet,



peer-assessment sheet, and dialogue journal. The teacher, after class, listens to the tape, assesses, and makes some comments and suggestions in the dialogue journal. He should try his best to finish this as soon as possible so that the student can reflect on his process based on these accumulated sources of information and set a new goal for his next recording.

### Week 2, Class 3: Peer-assessment (II)

At home before the class, each student reflects on his first practice based on the feedback from his peer and teacher and records his second practice on the tape, which he self-assesses. He comes to the class with his portfolio, which includes the tape recording his first and second practice, self-assessment sheets of those recordings and dialogue journal entry. In class, the students peer-assess their same partner's second practice in the same way as Class 2. But before the students start peer-assessment, the teacher should give them some feedback based on his findings through reading self-assessment and peer-assessment sheets, giving some good examples. This would be helpful in improving students' assessment skills, and, if not more importantly, this could be a message to the students that the teacher is checking their assessment and encouraging the students to get involved in assessment with more responsibility. In peer-assessing, before listening to the tape, each student tells his partner what changes he has made to improve his first practice. The student listens to his partner's tape and peer-assess it with these changes in mind. Sharing with the students some examples of good ways of revising would be useful for the final revision. Since the next class is students' presentations, it is advisable that, after peer-assessment, students discuss what the audience's responsibility is, i.e. to listen actively, to show interest. (N. Overman, personal interview, February 6, 2001).

After the class, the teacher takes the same procedure as the last class and has the students prepare for the final presentation.

### Week 3, Class 4&5: Presentation

In the next two periods the students make the final presentations, class being divided into two groups as suggested in the previous chapter. The teacher and an ALT videotape their performance, and assess it as they listen to it at the same time. After the class, each of them watches the videotapes from the other group and assesses their performance. In order to keep inter-rater reliability, as discussed in Chapter 2, the two as well as other teachers taking care of other classes have to achieve agreement on the standards before beginning the process. On the part of the students, they assess not only their partner's performance but their other classmates' performance. By doing this, the presenter can get more feedback and other students, at the same time, would actively get involved in listening and not become bored. If you have some time to spare, having Q&A sessions after each presentation or having quiz sessions on the content of speeches in small groups after every two or three speeches would make the classroom more active. (N. Overman, personal interview, February 6, 2001).

After the class, each student watches his performance on the video at home and self-assesses his performance and the entire process through which he came to the final performance. He hands in the self-assessment and process assessment sheet to the teacher, who assesses them. (See Appendix 3).

### Week 4, Class 6: Reflection

The teacher and students reflect on their entire process and discuss what worked

well and what didn't among them. They also check their criteria and modify them if necessary.

As in shown in Table 3 in the previous chapter, the teacher and students carry out the same process four more times throughout the year; however, unless a new format of presentation such as a Power Point presentation is introduced, the first process of criteria setting could be eliminated. Of course, it is recommended that those criteria be reviewed in class at the beginning of every cycle. It is also advisable that the teacher should see to it that each student has a clear goal for the new presentation based on his reflection on his past performances. In setting a goal, it is desirable that certain aspects of the criteria be more weighted than the others so that the students clearly know what to focus on when preparing and performing their presentations. The weighting can be done as an entire class or individually. Especially for poorly skilled students, the weighting is very helpful because, otherwise, they would be at a loss as to what to improve among the many aspects needing improvement.

## **Conclusion**

A teacher, I believe, has to be like a soccer coach. In soccer, once a game has started, the coach cannot take timeouts; therefore, the coach has to nurture players who are able to think by themselves and cooperate well with other players on the field: self-regulated players. For a teacher, the game does not mean a college entrance examination but it means each student's' entire life, in which he never stops learning.

Oral portfolio assessment can be a tool for assessing speaking, which is not being done satisfactorily in the present Japanese classrooms. But, at the same time, it can

be a tool for raising self-regulated learners. It requires students to reflect actively, self-evaluate process as well as product, and collaborate with other students and the teacher, which will enhance self-regulated learning. It is true that carrying out this assessment takes a great amount of time and has challenges, but my review of the literature confirms that it is more rewarding. I hope the ideas introduced here stimulate other teachers' creativity and can be adapted to their instruction, not only language classrooms but other subjects as well.

## References

- Blanche, P., & Merino, B. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. Language Learning, *39*, 313-340.
- Case, S. (1994). Will mandating portfolios undermine their value? Educational Leadership, *52* (2), 46-47.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. Harvard Educational Review, *64* (1), 5-30.
- Genesee, F., & Upshur, J. (1996). Classroom-based evaluation in second language education. Cambridge, UK: Cambridge University Press.
- Hargett, R. (1998). Assessment in ESL & bilingual education. NWREL's comprehensive Center, Region X. Retrieved November 7, 2000 from World Wide Web: <http://www.nwrac.org/pub/hot/assessment.html>
- Heilenman, K. (1990). Self-assessment of second language ability: The role of response effects. Language Testing, *7* (2), 174-201.
- Jensen, K., & Harris, V. (1999). The public speaking portfolio. Communication Education, *48* (3), 211-227.
- Katz, A. (2000). Changing paradigms for assessment. In Snow, A. (Ed.), Implementing the ESL standards for pre-K-12 students through teacher education (pp. 137-166). Alexandria, VA: TESOL.
- LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. TESOL Quarterly, *19*, 673-687.
- Meyer, C. (1992). What's the difference between "authentic" and "performance" assessment? Educational Leadership, *49* (8), 39-40.
- O'Malley, M., & Valdez Pierce, L. (1996). Authentic assessment for English language learners: Practical approaches for teachers. Reading, MA: Addison-Wesley.

Paris, G., & Ayers, R. (1994). Becoming reflective students and teachers with portfolios and authentic assessment. Washington, DC: American Psychological Association.

Sperling, D. (1993). What's worth an "A"? Setting standards together. Educational Leadership, 50 (5), 73-75.

Valdez Pierce, L., & O'Malley, M. (1992). Performance and portfolio assessment for language minority students. NCBE Program Information Guide Series, 9. Retrieved November 7, 2000 from World Wide Web:  
<http://www.ncbe.gwu.edu/ncbepubs/pigs/pig9.htm>

Appendix 1: Analytic Scoring Rubric for Formal Speaking

	1	2	3	4	5
Comprehensibility	Cannot be understood at all.	Can be understood with great difficulty.	Most can be understood.	Almost everything can be understood.	Everything can be understood.
Fluency	Stops for too long a time and too many times.	Stops for a long time several times.	Stops for a short time several times.	Speaks fluently with a few stops.	Speaks fluently.
Speed	Speaks too fast or too slowly.	Speaks very fast or very slowly.	Speaks sometimes fast or slow.	Speaks almost always at the right speed.	Speaks always at the right speed.
Clarity of pronunciation	Pronounces words too ambiguously to be heard.	Pronounces many words very ambiguously.	Pronounces some words ambiguously.	Pronounces almost all words clearly.	Pronounces all words very clearly.
Loudness	Cannot be heard at all.	Can barely be heard.	Sometimes difficult to be heard.	Easy to be heard most of the time.	Always enough to be heard.
Vocabulary	Too many misuses of words and very limited vocabulary.	Many misuses of words and limited vocabulary.	Several misuses of words and insufficient vocabulary.	A few misuses of words and rich vocabulary.	Use of vocabulary and idioms is always correct.
Grammar	Too many errors in grammar and word order.	Frequent grammar and word order errors.	Several errors in grammar and word order.	A few errors in grammar and word order.	Grammar and word order are very accurate.
Eye-contact	Never looks up at the audience.	Looks up at the audience only a few times.	Often looks up at the audience.	Almost always looks up but not enough eye-contacts.	Always looks up and makes many eye-contacts.
Gestures	No gestures at all and always stand still.	Tries to use a few gestures but very awkwardly.	Mixture of good gestures and awkward ones.	Uses many gestures but some are awkward.	Very natural use of gestures.
Facial Expressions	No facial expressions and looks very nervous.	A few expressions but very nervously.	Sometimes smiles and looks not so nervous.	Often smiles and looks relaxed.	Full of facial expressions and looks confident.

\*Thanks to Hargett (1996) for some of the criteria used in this chart.

**Appendix 2: Formal Speaking Peer-assessment Sheet**

Speaker's Name:	Date:
Your Name:	
Part 1: Complete the following sentences	
<p>1. What I liked most about his/her speech is: (First Practice)</p> <p>(Second Practice)</p> <p>(Final Presentation)</p> <p>2. His/Her speech has improved compared with the last one in that: (Second Practice)</p> <p>(Final Presentation)</p> <p>3. His/Her speech would improve if: (First Practice)</p> <p>(Second Practice)</p> <p>(Final Presentation)</p>	

**Part 2: Circle the word Yes, Some, or No to tell how you feel about the speech.**

	First Practice	Second Practice	Final Presentation
4. I understood what the speaker was talking about.	Yes Some No	Yes Some No	Yes Some No
5. The speaker spoke fluently.	Yes Some No	Yes Some No	Yes Some No
6. The speaker spoke at the right speed.	Yes Some No	Yes Some No	Yes Some No
7. The speaker pronounced words clearly.	Yes Some No	Yes Some No	Yes Some No
8. The speaker spoke loudly.	Yes Some No	Yes Some No	Yes Some No
9. The speaker used easy words.	Yes Some No	Yes Some No	Yes Some No
10. The speaker looked up at the audience.	N/A	N/A	Yes Some No
11. The speaker used gestures.	N/A	N/A	Yes Some No
12. The speaker smiled often and looked confident.	N/A	N/A	Yes Some No

\*Thanks to O'Malley and Valdez Pierce(1996) for some of the criteria used in this chart.





**Appendix 3: Sheet for Self-assessment of Process**

Name: _____		Date: _____
Topic of your speech: _____		
Part 1: Circle the word Yes, Some, or No to tell how you feel about your speech and add comments in each blank.		
1. I listened to the tape often.	Yes Some No	
2. I practiced a lot before recording the practices and final presentation.	Yes Some No	
3. I self-assessed very carefully.	Yes Some No	
4. I assessed my partner's performance carefully.	Yes Some No	
5. I asked my teacher and friends for advice.	Yes Some No	
6. I had a clear goal in my mind.	Yes Some No	
Part 2: Write your answers of the questions and directions below.		
7. What do you evaluate most highly in your process?		
8. What were your strengths and weaknesses in your presentation?		
9. What are you going to do to overcome your most serious weakness?		
10. Set a clear new goal for the next presentation.		
11. Write any comments and suggestions for your teacher and friends.		

\*Thanks to O'Malley and Valdez Pierce(1996) for some of the criteria used in this chart.

FL 026 734



U.S. Department of Education  
Office of Educational Research and Improvement (OERI)  
Educational Resources Information Center (ERIC)



REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

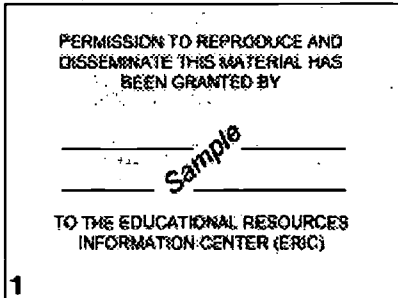
Title: <i>Authentic Progress Assessment of Oral Language</i>	
Author(s): <i>Yuichi Yoshida</i>	
Corporate Source:	Publication Date: <i>3/1/01</i>

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in monthly abstract journal of the ERIC system, *Resources in Education (RIE)*, are usually made available to users in microfiche, reproduced paper c and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following three options and sign at the bottom of the p

The sample sticker shown below will be affixed to all Level 1 documents

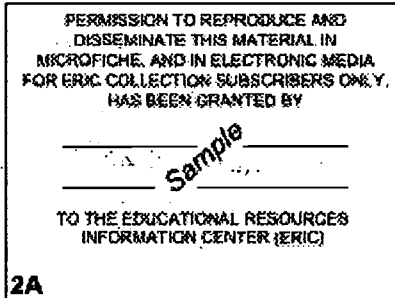


1

Level 1

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

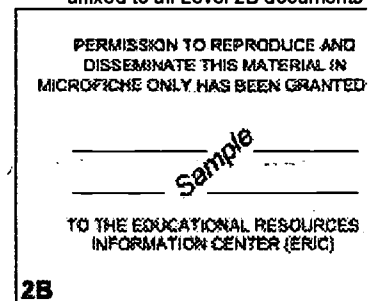


2A

Level 2A

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

The sample sticker shown below will be affixed to all Level 2B documents



2B

Level 2B

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contra requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to s information needs of educators in response to discrete inquiries.

Sign here please

Signature: <i>Yuichi Yoshida</i>	Printed Name/Position/Title: <i>Yuichi Yoshida</i>	
Organization/Address: <i>5-38-23 Obiyama Kunamoto 862-0924 JAPAN</i>	Telephone: <i>096-382-6977</i>	FAX: <i>096-382-6977</i>
	E-Mail Address: <i>yhyoshida@aol.com</i>	Date: <i>5/19/01</i>

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependal source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot t made available through EDRS).

