

# Some Methods of Assessing the Conversational Ability of Non-native English Speakers

Randall Cotten

Abstract

Evaluation of the oral communication ability of non-native speakers (NNSs) needs to be done in a way that is directly connected to what is taught in the classroom, as well as being clearly designed so that learners will know exactly what they will be tested on and how it will be done. Research suggests that well-designed exams can actually foster learning. The current report examines two methods in which the speaking ability of students may be tested in university EFL classes in Japan. The advantages and disadvantages of each method are also discussed.

Keywords: oral communication evaluation, EFL teaching, teacher action research

## Introduction

Evaluation of EFL learners is a necessary part of every course, particularly at the university level, yet it is coupled with a great deal of stress both for teachers and students. Well-made tests that accurately assess learners not only take time to make, but implementing and grading them is time consuming as well. For language students in Japan, testing of any kind can give them a great deal of pressure due to their experience with exams that have traditionally required a great deal of memorization of minute details about grammar and vocabulary. The danger in these types of exams is that they can create in students a fear of being evaluated; that poor scores are the root of a lack of confidence not only in their language ability, but in themselves as young adults.

Yet, if tests are designed well, they can actually foster in learners a desire to learn and improve. Talandis (2017), in citing Brown (1995), points out that, if teachers want to encourage students to learn and practice oral communication, then they need to create and implement tests which directly test for that ability. Furthermore, well-designed tests will also inform teachers about the effectiveness of their teaching. Are students actually learning what the teacher is aiming to teach? In *What the Best College Teachers Do*, Bain (2004: 151) reports that “outstanding teachers” endeavored to not just attach numbers to their students’ abilities. Rather, they use tests

in order to help students learn. It is possible in the EFL classroom in Japan as well to also carry out speaking tests which not only encourage students to review and prepare, but also give clear feedback to the teacher, which can then be used to to assess the learners’ abilities and help understand how better to teach oral communication skills.

## Need for clear evaluation criteria

When evaluating the conversational ability of NNSs, one of the most important first steps is having a concrete set of criteria. To make a more accurate evaluation of students, the instructor must first know what to test the students on. Rather than just noting that certain students are ‘good’ or ‘bad,’ the instructor should be able to describe the students in terms of ‘can’ or ‘cannot,’ or perhaps more precisely, as either having or not having certain abilities at using the language.

Similarly, by knowing exactly what they are expected to do during a speaking test, students are able to thoroughly review and prepare for the evaluation. This is especially important for conversation tests, which tend to sound very vague. As long as students are aware of the features of the language that they will be tested on (i.e., strategies or techniques for carrying out natural conversations rather than grammar and vocabulary items), then will know what to aim for when speaking during the test.

Possibly the most powerful effect that a clear set of evaluation criteria has is that of informing the course syllabus. Known as the *washback effect*, by planning from the start what will be tested, the instructor then teaches *for* the test, making sure that students have learned, practiced, and reviewed the skills that they will be evaluated on. Both Mendelsohn (1989) and Weir (2005) point out that being clear from the outset about what will be tested makes for a more accurate evaluation of students. It is altogether too tempting (even for English conversation teachers) to give their students a written test which is based on grammar and vocabulary items. Such a test is fairly easy to produce, implement, and score; however, it is not an accurate method of evaluating natural spoken ability. This is not to say that grammar and vocabulary are not important items to access. However, in order to more comprehensively evaluate the spoken ability of students, they actually need to be tested on that ability.

### Course design

The method and criteria used for evaluating NNSs described within this report were implemented in an oral English communication course based on a textbook called *Conversations in Class* (3rd ed.) written by Talandis & Vannieu (2015). The course materials are designed to teach learners various strategies which will enable them to carry out successful conversations in English. The text also includes a wide array of explanations which help students to understand differences between how Japanese speakers and English speakers communicate, then shows through examples ways which students can improve their communication ability and have more natural conversations.

Within *Conversations in Class*, there are review sections at the end of every two units. Each review section has a clear description of what type of spoken language features students need to incorporate in a conversation which reviews what they have learned and practiced up to that point. These instructions also serve as the basis for the criteria on which students are evaluated. As there are eight units in the text, this works out to four review sessions/speaking tests in one year. Talandis & Vannieu (2015) include some guidelines for various speaking tests in the instructor's manual for *Conversations*

*in Class*. A more in-depth exploration of giving evaluating oral communication in Japan, as well as more detailed descriptions of different types of speaking tests are given in another publication by Talandis (2017). The speaking tests described in this report were closely based on two of the methods outlined in these two sources. One of the methods incorporated some differences used on speaking tests carried out previously by the author (c.f., Cotten 2016). A discussion about the facilitation and effectiveness of these methods follows below.

### Method 1: Record & transcribe

In this speaking test, one 90-minute class period is used to carry out a short review session and the actual speaking test. First, students are given a sheet with instructions about how the test will be carried out. On it, they are given a choice of two topics to talk about. These are topics which they have practiced in the previous two units before the test day. Students are also given a list of language features to use in their conversation (with indications to specific places in the text to review). Below are some examples of the features students were told to include in their conversation for one speaking test:

- Use an opening & a closing.
- Use *English* thinking sounds.
- Answer “implicit questions.”
- Talk about yourself to give long replies.
- Agree or disagree politely.
- Make reactions *in English*.
- Use transition expressions to change topics.
- When you don't understand, tell your partner.
- Use at least one “no questions” pattern:
  - A: <statement or opinion>
  - B: <reaction + comment>
  - A: <reaction + comment>
  - B: <reaction + comment>

As mentioned above, these are also the very points on which students will be evaluated on.

After the students are paired up, they decide which topic they will talk about, then begin planning their conversation. They are not allowed to write down their conversation, but they are allowed to make some brief notes which they can look at while talking during the test.

The notes can be in English or Japanese. During this time, the instructor circulates and helps students as the need arises.

In this method, students use the recording function on their 'smart' phones to record their conversations. Experience has shown that there is no student who does not have one of these devices in this day and age and they make fairly decent recordings. Students are reminded to bring them in the previous week's class. However, occasionally students will sometimes forget their phones, so as long as one of the people in a pair has hers, the pair will be able to record their conversations. As a backup, it is a good idea to have a small voice recorder to lend to a pair if, for one reason or another, they are unable to use one of their phones.

Students are given three opportunities to record their conversations. Each time, they are to talk about the same topic (with the same person). Pairs record their conversations each time as well. The instructor signals to the students when to begin and, at the end of four minutes, signaled to stop. While talking, students are not allowed to use dictionaries or look in their books. Each time, after they finish talking, students are instructed to listen to their recording and think about how they might improve their conversation. During this time, they can use their dictionaries to look up words, or look back in their textbook to refresh their memories about some expressions they wish to use.

This cycle is repeated three times. After the last time, pairs are asked to choose one conversation that they think is the 'best' one. In choosing which one they think was the most successful, students are reminded of the list of language features which they were to have incorporated into their conversations.

The last thing for students to do in class is to begin work on a transcription of the conversation which they chose. Each student is given a large piece of paper on which to write their conversation transcription. Some examples of how to transcribe their conversation are given to the students, but basically, they just need to listen closely to the conversation and write everything that they hear – mistakes, Japanese words, laughter, silence, etc.

To enable the instructor to evaluate the conversations, students are asked to complete and hand in their

transcriptions by the following week. In addition, students are asked to upload the recording of their conversation (one per pair) to an Internet storage service, such as *Dropbox*. A folder for each class is set up in advance, and all students are given a link to it via email sent to them by the teacher. Once everything is handed in, the instructor can then evaluate each student's performance and, subsequently, give them feedback and a score.

### **Method 2: Write & perform**

In the second method of evaluating oral communication skills discussed here, an almost reversed procedure is used. In this speaking test, students are given an instruction sheet (similar to the one described in the first method) which outlines concrete language features to incorporate into their conversations along with specific pages to refer to for review. This is where the similarity ends. In Method 2, the test is explained and students are paired up in the class prior to the test. Each pair is given one large sheet of lined paper on which to write a 3-minute conversation in English. During the week before the test, pairs are to work together to write their conversation then memorize it. They are told that they will not be able to look at their written conversation during the test; however, they will be allowed to use brief notes in Japanese.

In the following class, the speaking test is held. The order in which pairs will perform their conversations is announced, then students are given a brief time to warm-up. When the test begins, pairs are called, one by one, either to a secluded corner of the classroom or an adjacent empty classroom. The instructor collects each pair's written conversation and checks any notes to make sure that they are not too lengthy and contain only Japanese.

Rather than using smart phones to record the conversation, a video camera is set up and used to record the conversations. This method has previously been used with success by the the author (c.f., Cotten 2016). Before performing in front of the camera, pairs can practice their conversation. Then after they are finished, they spend the rest of the class period completing a worksheet.

To evaluate the students, the instructor can use both the written conversation and the video for each pair. As

pairs have extensive time to develop and create their conversation, scores are divided between those for content of the prepared dialogue and those on individual performance. See Figure 1 for an example of a form which can be used both for scoring and giving feedback to students.

### **Pros and cons**

Below is a discussion of the advantages and disadvantages of each method of carrying out a speaking test. This will be framed in terms of ease of implementing the test and evaluating the students, as well as the effectiveness of each method for evaluating natural oral communication ability.

With regard to ease of carrying out the test, Method 1 is simplest due to the fact that all of the students are speaking and recording their conversations at the same time. A downside to this is that it does get quite noisy when all students are talking at the same time. If the size of the classroom allows it, having pairs distance themselves somewhat from each other can cut down some of the outside sounds when recording. Having students hold the microphone close to their mouth also helps, albeit it makes it a bit awkward to have a conversation. Furthermore, when students listen to their recordings after each 'take,' it is almost impossible for them to hear their conversations unless they use ear buds or headphones. This is one more thing to remind the students to bring if they have them.

Another technical issue was uploading the sound files. Either due to incompatibility of devices or lack of computer know-how, some pairs found it quite hard or impossible to upload the recording of their conversation. Often this could be completed after consultation with individual students outside of class, but time constraints often led to missed deadlines and delayed feedback and scoring. In other words, sometimes file sharing was quite a hassle.

Method 1 enables students to have three chances to improve their conversations. Most often, students pick the third and final conversation to transcribe because it was the one that they finally 'got right.' While there is a review period at the beginning of class, it often is not enough time to completely review the language features

that they are asked to use when speaking. This means that often after listening to the first or second take, students realized that they forgot to use one of the required features. However, this is a good practice for getting learners to self monitor and become more aware of the skills they should use in order to have successful conversations.

Certainly, Method 1 allows students to have much more natural conversations – even after talking with the same person about the same topic three times. Yet, even with the cycle of speaking, listening, reflection and review, some pairs still are not able to incorporate all of the required features in their conversations. Being a speaking test, and not just practice, this lack of meeting the requirements is reflected in their score. Audio recordings are very unforgiving for students who make few audible responses. Head nods just do not come through on an audio recording. On a similar note, because not a lot of planning goes into the conversations ahead of time, Method 1 makes it much easier to evaluate students individually. Because managing a conversation successfully is the individual responsibility for each person in a pair, any strengths or weakness in either one of the participants is easy to spot.

Method 2 is almost the complete opposite to the above method. Students have one whole week to create, then memorize their conversations. As such, this method relies heavily on memorization (something Japanese are famed for) and they have only one shot to get it right. Due to the ability for pairs to work closely together to prepare a dialogue, it is more difficult to evaluate students on their individual ability to use the required language features. Even though all of the required items might be in the conversation, it cannot be determined who thought to put them there. As a result, evaluation has to be done both for the pair, as a team, and on an individual basis where possible (see Figure 1).

Because pairs are video recorded one by one, it necessarily takes more time to carry out the Method 2 test. In large classes, this is a potential problem. However, in classes where this testing method was used, as many as 16 or 17 pairs (32 to 34 students) could be recorded in a 90-minute class, with a little time left over. One benefit of recording pairs either partially or completely isolated from other students is that it is much easier to hear what they are

saying. If an external microphone is attached to the camera, quite high sound quality can be captured. One more bonus of using a camera is that, during the evaluation process, the teacher is able to check for eye contact and other gestures that aid or interfere with communication.

As can be expected, when using Method 2, the conversations generally lack spontaneity, however, there is a much higher probability that students will include all of the required language features into their dialogues. Experience has shown that, when students are required to use a concrete set of language features then given time to prepare a dialogue in which those features can be incorporated, they actually use them. While Method 1 does allow for more free, ‘off-the-cuff’ conversation, students frequently do not use language features which they know they are required to use. In such a test, it could be quite difficult to remember exactly what one is supposed to do if the communication skills and expressions have not been thoroughly acquired yet. For example, when students are told to take longer turns, they can do it by talking more about themselves without waiting for questions. However, a problem arises when the listener merely sits and listens without giving a sign that she understands or is interested in what the speaker is saying. Yet, when students are able to write out their dialogues in advance, they can see that response are necessary. If a large enough number of students are neglecting to use this kind of language feature, lessons can be developed to give more direct instruction in them (Barraja-Rohan, 2011).

The use of notes written in Japanese did not seem to detract from or interfere with students’ ability to say what they planned to say. At times, the actual words spoken differed from what was written in the ‘script,’ but the general meaning of the utterance remained unchanged. It is quite interesting to notice how students can look at the Japanese in their notes for support and instantly say the words in English. Since students were reading something which they had prepared by themselves, not too much can be read into this, yet it is one instance of being able to say something in the target language that was originally conceived of in their native language. This is something that all language learners strive to do.

## Conclusion

The advantage of having a clear set of criteria when doing oral communication assessments is that it is easier to evaluate students on their ability to use strategies which enable them to have successful conversations. Likewise, learners can know exactly what they are expected to do when having conversations in the target language. Of the two methods for evaluating conversation ability described here, both have their strong points and drawbacks. One method allows for more natural exchanges with a chance of less target features being used, while the other offers a chance for students to incorporate a large number of required features at the cost of spontaneity. Each method also varies in the degree to which teachers can effectively assess individual abilities. There are, of course, many other methods for assessing the oral communication ability of students. Over the course of one year, it is possible for the teacher to perhaps use a more controlled test, such as Method 2 at the beginning of the course to familiarize students with the means of evaluation, then use a more open speaking test in the latter half of the course to give more individualized evaluation of students of their ability to have more natural conversations.

## References

- Bain, K. (2004). *What the Best College Teachers Do*. Cambridge: Harvard University Press.
- Barraja-Rohan, A. (2011). Using conversational analysis in the second language classroom to teach interactional competence. *Language Teaching Research* (15)4, 479-507.
- Brown, J. D. (1995). Differences between norm-referred and criterion-referenced tests. In J. D. Brown & S. Okada Yamashita (Eds.), *Language Testing in Japan* (pp. 12-19). Tokyo: JALT.
- Cotten, R. (2016). Conversation evaluations for increasing student motivation: Rationalizations and methods. 岐阜市立女子短期大学研究紀要第 66 輯, 15-22.
- Mendelsohn, D. (1989). Testing should reflect teaching. *TESL Canada Journal* (7)1, 95-108.
- Talandis, Jr., J. (2017). *How to test speaking skills in Japan: A quick-start guide*. Kyoto: Alma Publishing.
- Talandis, Jr., J. & Vannieu, B. (2015). *Conversations in class* (3rd ed.). Kyoto: Alma Publishing.

Weir, C. (2005). *Language testing and validation: An evidence-based approach*. New York: Palgrave MacMillan.

Appendix

Figure 1: One type of scoring and feedback form

		(A)					(B)					
NAMES / STUDENT #												
CONVERSATION STRATEGIES	<ul style="list-style-type: none"> <li>▪ Echo questions</li> <li>▪ Expressing uncertainty</li> <li>▪ Reacting</li> </ul>	<ul style="list-style-type: none"> <li>▪ Asking for repetition</li> <li>▪ Saying you don't understand</li> <li>▪ Getting more details</li> </ul>	1	2	3	4	5					
SOUNDING NATURAL	<ul style="list-style-type: none"> <li>▪ Used appropriate expressions</li> </ul>	<ul style="list-style-type: none"> <li>▪ Used model sentences &amp; questions from text</li> </ul>	1	2	3	4	5					
GOLDEN RULES	<ul style="list-style-type: none"> <li>▪ Answers with extra information</li> </ul>	<ul style="list-style-type: none"> <li>▪ 6-turn "no-questions" sequence</li> </ul>	1	2	3	4	5					
		(A)					(B)					
INDIVIDUAL PERFORMANCE	Pronunciation	1	2	3	4	5	Pronunciation	1	2	3	4	5
	Intonation	1	2	3	4	5	Intonation	1	2	3	4	5
	Eye contact	1	2	3	4	5	Eye contact	1	2	3	4	5
	Few pauses	1	2	3	4	5	Few pauses	1	2	3	4	5
GRADE												

Comments:

(提出日 平成 30 年 1 月 9 日)

Some Methods of Assessing Conversational Ability