



筑紫女学園大学リポジット

Testing English Speaking in Japan

メタデータ	言語: eng 出版者: 公開日: 2019-02-14 キーワード (Ja): キーワード (En): 作成者: WOOD, David John, WOOD, David John メールアドレス: 所属:
URL	https://chikushi-u.repo.nii.ac.jp/records/977

Testing English Speaking in Japan

David John WOOD

筑紫女学園大学研究紀要 第14号別刷

2019年 1 月

福岡県太宰府市石坂

Reprinted from *Journal of Chikushi Jogakuen University*

No. 14, pp. 13 – 28, January 2019

Ishizaka, Dazaifu-shi,

Fukuoka-ken, Japan

Testing English Speaking in Japan

David John WOOD

Abstract

This paper analyzes problems facing spoken English teaching and testing in the context of why MEXT (the Japanese education ministry's acronym) has not achieved government pledges to develop students' English ability, increase study abroad or make English a functioning language in society (PM's Commission, 2000). We examine MEXT's latest bid, forcing students applying to college to prove their ability to communicate by passing costly spoken tests produced, conducted and graded by private for-profit companies. This won't enable Japanese students to speak in English as it sidesteps the basic needs of professional teaching development and effective methodology, which aren't taken into account in the education ministry's recent schemes.

A fundamental dilemma is, testing and teaching are not the same thing, though that seems to be the wishful thinking behind MEXT's strategy. Delegating standards to commercial testing companies while ignoring the virtual absence of successful spoken English teaching in regular education is simply turning a blind eye to the crucial issue of communication completely. Thus, we propose 3 major steps vital to resolving this fatuous state of failure:

- | |
|---|
| 1. <i>Professionalization</i> ; 2. <i>Decentralization</i> ; and, 3. <i>Recurricularization</i> . |
|---|

These aims are developed throughout and evaluated in the later sections. One goal of our discussion is better teacher training via the re-education of administrations' and companies' conception of what constitutes a teacher. The default definition seems to be anyone with native-like spoken ability. Clearly, vastly improved opportunities and requirements for those charged with teaching Japanese students how to speak in English are urgently needed. After examining the form that spoken testing has taken over recent decades, including recent examples of tests produced by those now charged with providing them for college entry, we pinpoint fundamental gaps between what is possible in most testing and what spoken communication actually is. We present an alternative form of assessment to improve communication. The aims and results of this diametrically opposed approach are highlighted by measuring the effectiveness of the latter in terms of interactional fluency (Wood, April 2015). In a wider context, multi-cultural communication needs motivating and activating sincerely to let English function more socially.

1. Situational analysis

MEXT (2002) was tasked with establishing effective communicative English ability in all students to “support English conversation activities (to) hold normal conversations on everyday topics at *eiken* (the government’s testing monopolist) level 2 (by) enriching training for English teachers and promoting the utilization of native speakers”. (Parenthetical additions throughout are the present writer’s.) The 3 aims are far from being achieved. MEXT’s newest strategy, costly spoken testing, risks commercial overkill (Wood, July 2018). Such testing can’t be substituted for learning as stereotypical “latitudinal” one-off testing has patently terminal limitations. Spoken pedagogy and teaching professionalism need prioritizing instead. Substituting testing for teaching means learning to fail, not to communicate. Any prospect for skilled English teaching has been indefinitely shelved. Appropriate methodology bows out to powerhouse test or publishing profits.

Eiken has a weak spoken segment (**3. Neglect of teaching quality and nature of spoken testing**) but MEXT will include it and 7 more (like ETS’s TOEIC and TOEFL) to take up to 2 from (or 1 twice) for college entry yet even TOEIC LR (Listening and Reading) seems more apt (Wood, 2010) as major firms use it in the workplace. TOEFL is meant for high level study at US colleges, beyond most students’ purses and prowess. One of Japan’s top universities, Tokyo University (*Today*) adamantly rejects the idea. Long-term longitudinal spoken teaching and testing promise more interactional success. Solutions to Japan’s fundamental problems in English teaching include:

- | |
|--|
| 1. <i>Professionalization</i> : Pedagogical policymaking needs discussing with and delegating to the most able teachers of spoken English communication. |
| 2. <i>Decentralization</i> : To progress, decision-making for English teaching needs wresting from the hierarchy unilaterally deciding all key policy. |
| 3. <i>Recurricularization</i> : Regular external audits of methods and aims of valid TESOL systems can create more effective class contents. |

Proficient teacher training is essential. Regular Japanese teacher training is inappropriate, defaulting to prescriptive styles, while real TESOL requires communicative interaction. The word “teacher” of spoken English is a euphemism for someone with native speaker (NS) ability equivalency, but no linguistic or teaching skill. Companies seek the veneer of respectability so accept token qualifications like CELTA, a month’s course for high school graduates but a TESOL Master’s is over a year full-time for college graduates. The solution is more professionalism and less bureaucracy using businesses to seek quick fixes for windfall profits. Many graduate schools in Japan include “English” in syllabi but few offer the necessary TESOL training. This must change in order to employ innovative methodology and pioneering projects in authentic learning and interaction, incorporating and utilizing the social media revolution currently underpinning real-world communication.

2. Spoken testing validity versus paradox

Formal spoken tests are matters of personal and professional concern. The first such test I took was in 1968, a timely reminder to assess a half century's progress. There were oral and aural tests to acquire French qualifications in UK high schools. They taught me the gap between what is instilled in class and how that knowledge can help in real-life situations, when communication, even basic survival in the outside world is on the line. What is needed is experience of the language beyond what is only for tests. The best text to learn from as it were is actual face-to-face experience, trying and seeing what actually works outside the lecture theater by one's self. Testing in Japan is a commercially conspired practice of habit. It shows less of individuals' speaking ability than of the inability to challenge educational administrations' substitution of superficial looks for language-learning skills. The teaching and testing of spoken foreign languages face paradoxes like:

1. *Students are treated as passive vessels for teacher and text rote input only.*
2. *The only permissible answer to a question is that of the teacher or the text.*
3. *Tests are written though speaking is the essence of most communication.*

Conservative hierarchies obstruct communicative learning. The reliance on texts and testing over teaching stymies speech. Specific problems like insularity are as commonplace today as decades ago. Curricula contain only a minimal number of classes nominally related to real L2 acquisition. Part-timers are preferred by administrators. The crushing result is "teaching" badly in the wrong language only items of questionable or no active worth.

To test speech is conceptually irreconcilable with conventional curricula: "Compromising testing are the epistemological foundations underlying the methods we apply to engage in tests and our communicative/social/literacy practices in real life. These are so different and so incredibly hard to overcome." (De Groot, 2018). A conspicuous incongruity is the term *kokugo* or "the nation's language" as opposed to *gaikokugo* or any outside language, including English. Any non-Japanese language is thus okay as students' learning requirement, but the vast majority defaults to English in the yearly NCTUA *Center-shiken* or central college entry exam for 500,000 students at 700 sites around all of Japan. Without spoken Japanese tests in high school, can students with no proven L1 oral mastery suddenly be expected to speak the world's L2 fluently by taking tests conducted by for-profit companies? The grammar testing tradition has been markedly ineffective, accounting in large part for Japanese students' reluctance to view English positively and MEXT's long-standing inability to realize even minimal spoken standards. It's over half a century since English was no longer prohibited after war time. Interaction, the missing link, is a prerequisite to achieving communication. The stipulated tests fail to even consider this. Speaking has no limits yet text-bound "spoken" testing abounds. Companies spend heavily on blanket research to bolster face validity. Powers (2013) is typical of ETS's crusade. Quoting celebrity researcher Jack Richards, Powers justifies spoken tests using non-sequitur logic: "*If listening isn't tested, teachers won't teach it.* This

applies with equal force to the reading, speaking, and writing domains.” Not everything that needs teaching is testable. As Heaton (1975) showed oral tests are just reading, pronunciation and contrived “situations”. He wrote 5% on speaking mainly using pictures for storytelling like *eiken*. Brown (1995) bemoaned the lack of English oral testing in Japan. Fulcher (2006), has only a page about speaking out of over 400. Spoken tests’ limitations include:

1. *They are uniform, but speaking is spontaneous so harder to standardize.*
2. *In writing there is rarely any present audience affecting what’s written.*
3. *Examiners react subjectively to paralinguistic features, skewing scores.*

There are no national high school L1 spoken exams in Japan, bringing L2 spoken testing itself into existential question. Would there be any written or grammatical English exams in Japan if students didn’t have some kind of equivalent expertise in their own language? And why does writing ultimately and inevitably have more prestige in TEFL and TESOL when speaking is more fundamental to communication, more crucial in linguistic acquisition and more widespread globally? As a UK national spoken English examiner from 40 years ago, the need seems clear to me. Those exams have continued there for decades, and were extended up to college entry with success.

These latter nurture communication and self-expression, two qualities in short supply in Japan’s system, especially at secondary and tertiary levels. UK topics cover individual interests plus higher level discussion and debate. While Japanese college English curricula boast these latter two disciplines, most students are unable to discuss or debate proficiently even in their L1. How could they in English without any L1 conceptual skills? Their English study is L2 to L1 grammar translation, leaving them sentenced to silence as grammar creates aversion to speaking. To study English anywhere in the world begins with a spoken interview. Japanese studying abroad often land in the lowest classes shocking teachers because of the former’s hybrid knowledge of grammar but zero ability to vocalize it in English. News stories of Japanese in trouble abroad due to communication inability point to this.

3. Neglect of teaching quality and nature of spoken testing

Japan’s education budget is lower than the OECD average, but wastes money on inefficient aspects of English education. MEXT opts for quantity over quality, favoring unqualified “teachers” at unjustifiable costs. Neither the appropriate qualifications nor the proper experience are required. This situation allows unethical for-profit companies to control transient employees by easily hiring then firing them. The Internet is flooded with predatory “teaching” job ads, luring takers with no teaching qualifications. Routine profiteering makes the process twice as expensive as it needs be, yet only half as valid as it should. MEXT prefers a middle-man system and testing to effective teaching methods and personnel. Companies secure profits as long as students don’t acquire English,

so the only hope of mastering it is costly outside-school self-funding. The JET Program employs native speakers with any degree for irregular high school visits. Language schools and mercenary social media companies have neither standards nor accountability, like Nova which collapsed in 2007, leaving foreign teachers homeless. The aim to enforce spoken tests nationwide via such companies is a regurgitation of a recipe that renders students incapable of communication.

English has been introduced as a standard subject at elementary school, but similar problems occur as regular Japanese teachers with no English proficiency or aptitude for teaching languages are recycled to teach as a cost-cutting exercise. Without English ability or TESOL training, it's a face-saving gesture that falls flat on its face. TESOL developed in reaction to globalizing war and refugee streams (De Groot, 2018). Historically it has long been enacted as missionaries imposed the use of their language. This is one reason for aging administrations' deeply rooted opposition to language learning. Japan's IOC chief (a former prime minister) stereotypically referred to English as "the language of the enemy" (Reuters, 2014). Such fears are stoked by present-day religious groups offering "free" English. Europe became a leader in TEFL (Van Ek, 1975) establishing language learning as an instrument of trust, respect and cooperation within the EU. Based on that reinvigorated interest in formal foreign language learning, many programs were established boosting TEFL-related Master's degrees. MEXT will superimpose Europe's CEFR spoken scale here, but it can't be applied without teaching expertise or genuine administrative cooperation.

TESOL is for speakers. Grammar translation is outdated and unrealistic. Proposed latitudinal testing by profit-seeking companies is inappropriate. As grammar's modern-day 'founder' stressed when quoting himself from 1909: "Grammar must be flexible, not a set of stiff dogmatic precepts, according to which some things are correct and others absolutely wrong, but something living and developing, founded on the past (it) prepares the way for the future, something that is not always consistent or perfect, but progressing and perfectible - in one word, human" (Jespersen, 1935). His book has been used for decades at Japanese colleges without ever being properly understood.

The sole government sponsored spoken test of English since 1963, *eiken*, doesn't include spoken communication, just a cosmetic oral interview. Students tell the story of what they see in a picture series, with a constrained topical question. Conditions are regimented and restricted, and in no way resemble human interaction. Examiners can't smile, gesture or say anything outside a prescribed rubric, a sterile face-saving approach to avoid complaint. In a vain attempt to establish standards, each word or action is pre-scripted. TOEIC and TOEFL have in recent years developed spoken tests but they don't amount to anything like living communication, nor are they appealing. While TOEIC L&R attracted 2.5 million takers in Japan alone (similar to STEP's 2016 annual numbers) the S&W exam was taken by a mere 25,000. ETS's 5 to 10 minute spoken test includes reading a text, describing a picture, a highly constrained question session and a final problem to solve. TOEFL is designed for students going to study in USA, yet just 2,000 studied for a year or more there in 2016 (JASSO).

4. Spoken testing in Japan

Spoken testing of English ranges from a casual class quiz to pass-or-fail final exams. Real spoken teaching and assessment rarely occur in a text/test-bound milieu dominated by nation-wide standardized exams. The verbal varieties used would seem alien to most global English speakers. Dependent on L1 directions, arbitrary multiple choice questions and L1 translation, testing here makes communication a non-starter. The hierarchic follies of society are at their most evident in a scandal-prone education system. Its intrusive rules on what English is correct defy real world communication.

Teaching is heavily controlled by governmental administrative authorities. The latter increasingly see hiring private testing companies as the main answer to the problem of the chronic inability to communicate. They fail to see that testing itself is the cause. Better options are: making communication the aim; improving instructor quality; and updating methodology. A national environment of predetermined answering means tests focus on hybrid grammatical content monopolized by administrators and testing companies with zero concept of communication. External tests can't contain a single error on pain of public media humiliation but communication is by nature imperfect. Spoken tests and conversation texts are written dialogs tested for grammar via translation in unreal language. Communication can't be limited or set in type. It is spontaneous, interactively created and impossible to write in advance. Yet MEXT will impose high-stakes spoken tests on a nation of non-communicators (McCrostie, 2017). MEXT will purchase the products of companies traditionally monopolizing exams for massive profit. *Eiken's* whole test (average cost \$50) has been imposed 100,000,000 times at 20,000 high school testing centers in a staggering 50-year, \$5 billion ineffectual solo reign. The results speak (or rather fail to) for themselves in a culture where coerced language learners remain tongue-tied, too busy to complain having to cram for the next test. Commercial spoken testing, incurring extra exam, accommodation and travel expenses, will determine the admissions fate of a million high school students a year. Takers pay an extra \$250, twice, for spoken tests of 5 to 10 minutes (available at few sites) or be denied 4 years' college admission. A 2-day long *Center Shiken* is \$150 for all subjects.

Some experts support testing as a motivational boost to communication, but the exams proposed don't offer solutions. The burden they impose risks the opposite effect on many candidates. It also conveniently allows the government and such educators to shirk responsibility as they assert that tests are the answer to improving communicative ability (McCrostie, 2017) when improved training, teaching and techniques are the only credible solutions. MEXT persists in opting for unqualified, exploitable foreign teachers, but creating a "vicious cycle of outsourcing teaching jobs ultimately leads to everyone except the companies that win government bids losing as students are taught by ever worsening staff as the companies bidding lower estimates" (Wood, Jul 2018). The only products are the damaged goods of study and teaching failure.

5. Commercial latitudinal versus in-class longitudinal spoken testing

5.1 Latitudinal commercial testing: is written based; is a few minutes long (the available ETS, 2018 sample referred to is about 5 minutes); and, has only a single non-communicative speaking section; (ETS research suggests up to 10 minutes, but content types are similar.) This recalls Heaton's study (1975) with almost identical components. Available sample contents break down as follows:

Section 1: **Read** a short paragraph advertising a hotel aloud in 45 seconds

Section 2: **Describe** a test photo (shoppers at an outdoor shop) in 45 seconds

Section 3: **Answer** 3 questions in 1 minute: *How often do you watch TV? What kind of programs do you usually watch? Describe your favorite.*

Section 4: 3 questions on a **written** conference announcement in 60 seconds: "What time . . . ?" "How much . . . ?" "Describe (one part of the text)."

Section 5: Problem solving based on a **reading** passage (1 minute): *Help someone who lost a card in an ATM.*

Sections 1, 4 and 5 are **written**. Section 2 is a random photo, no conversation.

Section 3 is a short, predetermined question and answer session.

Depending on the candidate answers could be: (i) *Never*. (ii) *None*. (iii) *How?*

Without asking if the candidate even watches TV, communication is absent.

Answers (i)–(iii) justify full marks as much as longer, affirmative answers.

Such commercial tests nearly always lack any basic communicative validity.

Japanese hierarchical turn-taking is reflected both in *eiken* and the above.

The exam falsely assumes its own infallibility, but is in fact fatally flawed.

5.2 Longitudinal class testing: is the cumulative result of a course of study; directly results from and is graded with reference to ongoing class activity; and, includes, meaningful testing, unlike any of its one-off counterparts.

The examples below demonstrate interactional rhythm, an essential quality in communication, but one that is absent from standardized spoken testing.

(**5.1** is coercive as candidates have to answer or fail non-validated items.)

There are 15×90-minute classes; 2 students answer others about their photos each class (80 minutes); and, review and develop the questions (10 minutes).

5.3 Class spoken summary example: Review event length: 140 seconds

Content: 36 utterances (18 questions and answers) 138 words

Interaction speed: 4 seconds per utterance; 1 word per second

<i>Where did you go?</i>	I went to Miyazaki.
<i>Why did you go?</i>	Driving license.
<i>When did you go?</i>	Last spring vacation.
<i>Was it difficult to get a license?</i>	Yes.
<i>Do you have a plan to drive?</i>	No.
<u><i>Where do you want to go with car?</i></u>	Outlet.
<i>Do you like roast meat?</i>	Yes.
<i>Was it good?</i>	Yes.
<u><i>Which is your favorite restaurant?</i></u>	Egg restaurant.
<i>Which is your favorite shopping?</i>	Aeon Mall.
<u><i>Who did you go with?</i></u>	My friends.
<i>Tell me more?</i>	High school friends.
<i>Who are they?</i>	Same driving school students.
<i>Where are they from?</i>	Some people are from Fukuoka other people Nagoya.
<i>What's your nickname - why?</i>	I like Yamada Yusuke.
<i>What else did you do in Miyazaki?</i>	I ate chicken.
<i>Do you have a car?</i>	No.
<i>Are you a good driver?</i>	So so. (<u>Underlining</u> = repetition.)

5.4 Course-end test interaction: (5 statements, 5 questions) 76 words

Length: 30 secs/Content: 15 utterances (=simple replies or verbal clauses)

Interaction Speed: 2 secs per utterance; 2.5 words per sec

<i>Tell me about you photo.</i>	I <u>went</u> to Miyazaki last spring vacation to get a driving license with my friend of university. Then I <u>met</u> an old friend & <u>got along with</u> her friends so she <u>suggested</u> going to a barbecue restaurant & we <u>took</u> this picture when we <u>went</u> there.
<i>Tell me more about driving –</i>	Yes.
<i>– do you enjoy it?</i>	Yes.
<i>Have you got a car?</i>	No.
<i>Whose car do you use?</i>	My boyfriend's.

5.5 Comparison of class and test samples:

The spoken production rate of the subject's test was more than twice as fast as the class interaction. This indicates that some development took place. The breakdown rate was lower and the level of complexity increased. This showed that students had acquired more speaking ability and expressive proficiency, suggesting that by using this method they become more confident with each lesson and interact successfully at greater speeds each time. The sample subject's progress is summarized in *Table 1* below:

Table 1 – Comparison of class and test spoken production rate increase:

Event	Length	Interactions	Per second	Words	Per second
Class	140 seconds	36	2.5	138	1.0
Test	30 seconds	10	3.0	76	2.5

A rough estimation of interaction fluency equals the total number of syllables divided by the time of the interaction, including hesitations. The resulting interactional fluency (or IF) speed echoed the increased spoken production rate increase per second in *Table 1*. While rudimentary and limited in the scope of data used, *Table 2* below suggests clear acceleration:

Table 2 – Comparison of interaction fluency (IF):

Event	Length	Syllables	IF Speed	Difference
Class	140 seconds	173	1.23	– 2.17
Test	30 seconds	102	3.40	+2.17

6. Alternative longitudinal spoken teaching assessment

The prevailing approach to spoken testing is random one-off latitudinal testing which fails competent communicators as much as favoring guessers. This tests candidates once, making results neither scientific nor objective. Testing writing is also a flawed process, but as the latter is in static paper or cyber form, it seems controllable. As writing is susceptible to grammatical inertia, the mechanism to test it is in place, but the very concept of testing speaking is as feasible as capturing the air and soundwaves in which it resides, physically possible perhaps, but irrelevant to speaking's purposes; communicating thoughts and feelings live to share and develop with others.

The *eiken* interview test is a perfect example of the very opposite and symptomatic of what "speaking" degenerates to in the stranglehold of commerce. What is needed is methodology that allows speaking to flourish as communication. Interactional rhythm (not in the narrow sense developed amongst others by Couper-Kuhlen, 2009) can be holistic and measurable.

With the focus on effective communication, functional fluency can be generated through basic

vocabulary and well-paced interchanges, stimulating and maintaining smooth and efficient speaking between multiple participants. Interaction can't be circumscribed by pronunciation or grammar, both of which are random and illogical to a high degree. In a borderless age, why confine communication? Speech has been defined as "the verbal use of language to communicate with others" (Fulcher, 2006). Any attempt to control it sidesteps communication altogether.

Ockey and Li (2015) further divide spoken communication into three:

1. *Interactional competence*;
2. *Appropriate fluency*, and;
3. *Appropriate use of phonology, vocabulary and grammar*.

They define interactional competence as "An individual's underlying ability to actively structure appropriate speech in response to incoming stimuli, such as information from another speaker, in real time" (p. 5). At best the Japanese attempt only focusses on the third point, but nowhere do Ockey and Li (or any other recognized authorities) include writing as a central active element.

There are various text-free methods for helping students develop their speaking ability. This writer has developed one extensively tested approach using photographs. The method has been outlined with examples in book-length detail (Wood, 2015). An outline of the method will be given here, followed by suggestions for maximizing its effectiveness. To overcome text dependency and the stifling influence of the teacher as the only person in a classroom who determines questions and answers, students use their photos.

As only the student is the authority on the content, class confidence grows instead of shrinking for fear of making what the teacher or text deems wrong in conventional classes. Next, the method puts the onus on the students themselves to elicit information from each other by the teacher's silence while they interact in English to find out as much as they can. The teacher pre-teaches frequent questions and vocabulary before delegating the development to students. This requires effort to control the urge to intervene.

This is the only way in which they can become independent L2 speakers. It doesn't mean that the teacher does nothing. To maximize the process, more effort than usual is essential. Students need to be encouraged and also guided especially in setting up the process as well as in using their interactions comprehensively. Cooperation and negotiation are central to this approach.

The basic style of questioning needs repeating carefully with repeated explanation and clarification, as well as flexibility on the teacher's part to experience and evaluate previously unknown strategies. To ensure students' cooperation requires a shift in our paradigm. We must assume a democratic role wherein we share the role of learner as we discover methods to communicate with students not common in traditional teaching approaches. Interactions are processed in various ways. Students summarize by asking questions fellow students already asked in a concluding exchange. The goal is to interact quickly, focusing on smooth communication over

grammatical correctness. Rehearsal achieves a successful interaction rhythm, the essence of real communication. This is recorded to re-listen to in audio MP 3 files on social media (like LINE and so on) to improve student listening ability and affective engagement in the responsibility for learning beyond class and texts.

Assessment is longitudinal. Recordings are compared to measure fluency. When measured via set criteria, they get feedback showing improvement, making testing meaningful and satisfying as this real-world assessment will indicate their actual ability far more accurately. Rather than a latitudinal test, ongoing testing each class is a truer indicator of actual ability as well as a motivating teaching process in itself. The final test is not a pass-or-fail test, but a confirmatory check. This can be carried out using the photo method with students presenting their photos in a set maximum time, for example, in a minute or so with 25-30 students in a class, the average size for this writer. To make the session interactive, students can rehearse in pairs, and ask each other: "Tell me more about (any point the listener wants to know)" and then continue with a few follow-up questions, before the same presentation and question style, including different questions arising spontaneously at speed.

Recording the interaction helps assess accuracy, fluency and complexity. Factors are compared to measure relative performances and add to the overall portfolio of interactions that take place every lesson to produce a final overall grade and also see how students' ability changes. This can be a highly efficient and self-supporting experience. Speaking test interviews need flexibility to promote long-term learning and testing for communication.

Traditional latitudinal tests fall short making most spoken testing a literal waste of breath when dependent on for-profit companies testing students for non-communicative purposes. Revising Japanese entrance exams to directly assess oral communication might help English education in Japan if testing is over a period of time and can truly accommodate communicative principles. But that must be preceded by the sincere upgrading and financing of high quality language teaching, requiring international TESOL degrees.

Japanese oracy also cannot be ignored. Oral communication assessment currently faces many challenges. The suffocating stranglehold of administrative overkill means they are unlikely to be faced head-on any time soon. With high-stakes, large-scale testing taking center stage, prospects of achieving goals will fail, causing classes to continue to suffer in silence.

Heaton (1975) concluded that the most reliable method is that which "involves the teacher (and) continuous assessment". This is significant as his study, despite the minimal 5% treatment of spoken testing and the limitations of spoken testing evident in nearly all spoken tests, was central to the study and practice of spoken testing for several decades at the end of the 20 th century. Speaking and testing therefore seem non-connectible in most circumstances, because:

1. *Spoken timing and situational aptness require extensive natural use.*

2. *Learners need culturally transparent communicative ability.*

3. *Communicative fluency is the aim. Testing is short-lived and inflexible.*

The photo technique can be used actively in the language classroom by setting time limits and developing higher goals in the learning process. Speaking is more transient than other language skills. Therefore, recording interactions can help learners perceive their speaking objectively. They can review the contents at their own pace. With this practice, language learners need others' feedback and reflection at each learning opportunity.

The test of a test is how those taking it respond both in and after classes. Twice yearly course feedback consistently indicates that students react positively to the personal photo approach. The latest feedback (August 2018) from the class sampled to illustrate the longitudinal approach above confirms this. Numerically, students in the same year group classes rated classes lower. The other 3 classes overall averaged 4.3/5.0 maximum compared to mine 4.6. Their entire responder rate was half with just 4 comments from 77 students, compared to this writer's 8 comments from 25 students, or 5% versus 30%.

Unlike their all negative responses, mine were almost all positive, e.g.:

"This was a happy class!"

"I was happy to acquire practical speaking skills I can use anywhere."

(Both these comments were echoed by other students.)

"I enjoyed learning what everyone thought by using photos to talk about."

"There were more opportunities to speak and listen than in any other class."

In addition, as junior students had already had 4 conversation classes before, and don't get any more in the 4th year, comments are more definitive than in most classes. As sophomores, they rated other conversation classes very low and were extremely critical. The results received above were similar in all the researcher's communication classes, regardless of subject, year or department. There are some negative reactions, but proportionately very few. Most striking was my response rate was the highest as a result of interaction.

7. Deepening problems arising from ignoring key issues

TESOL is for *Speakers*. Speaking is the aim, not testing or reading. Children learn to communicate successfully without reading or grammar. Japanese students are forced to run with L2-to-L1 translation before they can stand with listening, leading to failure, resistance or just passive apathy. Long-term in-course assessment is better than one-off latitudinal testing. Utterance fluency can help measure speaking ability over a valid period (see [5.5 Comparison of class and test samples](#)) by firstly deferring from regular texts and tests. Once-only tests are meaningless minuses to communicative development. Such testing examines language from a written/grammatical

perspective. Enforcing commercial spoken tests aids neither communication nor long-term learning (Wood, 2010). It only increases widespread teaching to the test. To help students achieve lasting communicative ability, greater numbers of better employed professional EFL teachers and methods are vital.

State policies persist in avoiding this for example by substituting teachers with robots! (Japan Times, August 2018) and with instant out-sourced “solutions” like Tokyo’s new English Village, a perfect symbol of putting the veneer of English over the substance of education. A conversation school cartel (ELEC, 2018) hires staff with “native-like English” only. The rationale, to compensate for the absence of spoken English education at regular schools, is used as an excuse to exploit the situation when MEXT should be establishing real spoken classes with proper full-time staff in schools instead.

Such villages (EVs) have existed for decades in poorer EC countries like Hungary, but only in Korea and Japan are bureaucracies interceding directly. The Japan model is to call any native speaker a “teacher” to maximize profits. The concept of native speakers (NS) is moot, however. Proficient speakers (PS) regardless of nationality are often the best linguists and teachers.

EVs were first intended as cheap, effective alternatives to study abroad. Tokyo’s is an expensive lame duck, confined to inauthentic communication. Fees touted as affordable will rise and content default to non-English as in Korea. More than company tests, texts and “teachers” the right techniques, training and teachers guarantee results. Most commercial testing narrowly examines language from a grammatical perspective so fails to promote communication. (Wood, 2017 & 2018). EV conceptual flaws include:

1. *Immersion learning is ineffective on a short term basis like this;*
2. *Hourly fees aren’t cheap so students seeking any result would have to spend more than if they went abroad to acquire even minimal skills; and,*
3. *Genuine education shouldn’t be shored up by for-profit companies in non-school contexts when what’s need are better school staff and curricula.*

8. Conditions required for fluency

Fluency requires conditions antithetical to the MEXT textbook. Nation (2009) lists the following three aims, which are all good guides for classes:

1. *Communication needs to be “real time” and meaning-focused.*
2. *Language items must be within previous experience.*
3. *Learners perform should be at a higher than normal level. (2009).*

These precepts are all well served by the method described previously. (See **Section 6. Alternative longitudinal spoken teaching assessment.**) Although the content there appears limited,

it is actually infinite as student questions can develop in any and all directions, as real communication does. At the same time, by using their own photos chosen by themselves, they learn responsibility and gain confidence as the foremost authority on the content. It goes without saying that most of the content will be within their own experience, but they are challenged to dig deeply into their memories and to adjust their viewpoint and raise their performance in response to the questions that they receive from other students. Thus performance levels rise naturally as they need to make every effort to explain themselves to others.

9. Conclusion

Returning to those areas in urgent need of reform (See **1. Introduction** 1. *Professionalization*, 3. *Decentralization*, and 2. *Recurricularization*) their difficulty and denial must be solved. This may be possible in phases. Individual effort and example to find better alternatives are starting points. As instructors we have enough control of our own classrooms to maneuver. We can affect administrations' decisions to make a meaningful difference. Instead of accepting the role of passive resource they impose on us, we must determine our goals and materials by modifying syllabi even in teaching time. We can reinforce this through both official feedback channels (e.g., the student evaluations quoted above) and in more depth by informal questionnaires of our own design. The findings there can be disseminated in domestic and international presentations, plus academic papers, to invite feedback, comments and active co-operation from peers both near and far, by inspiring other teachers to test out developing theories and beliefs. There are excellent resources in our connected world, often free and readily available through the Internet plus our application in locating them. Sharing between students is important, but it is even more vital to seek opportunities to share with colleagues at conferences, through publishing papers and active networking. Conferences are often the target of commerce, hampering content and communication so we must seek sufficient question and interaction time for follow-up networking as our professional duty.

As a profession, we need to be more active, assertive and innovative, instead of servitude to teacher's notes or administrative directives. It's easy to turn on and off again when the bell chimes the start and finish of a class. However, as a profession with students depending on the decisions we make, to grow and evolve, passive compliance with a status quo that doesn't yield the results that parents and taxes pay for is unacceptable. The number of properly qualified TESOL instructors is still too limited. Opportunities to acquire such qualifications in Japan are too few in proportion to the vast number of Japanese urgently requiring to be taught by skillful teachers, not just "near equivalent native speakers". Programs also need diversifying as they are limited to a few US schools like Temple University's, a private institution graduating just 38 Masters of Education a year (TUJ, 2017). Domestic teacher training programs, reliant on Japanese language in the form of L2 to L1

grammar translation, need radical recalibration to increase bilingual native Japanese teachers fluent in TESOL methods and theory, and capable of communicating in spoken English, the perennial stumbling block. Most Japanese graduate schools lag decades behind other advanced countries in teaching ESOL. They struggle to survive, with non-productive staff resting on laurels they never won. This needs changing from the inside out, with Japanese teachers learning how to learn and teach English in English communicatively, instead of just drilling grammar into passive students. Education budgets are almost completely for non-educational use as decided by administrations. Japanese teachers have little say, non-Japanese even less, but non-Japanese still have a duty to represent their cultures and dialects fairly and openly, both locally and further afield. While acquiring and respecting language and culture is essential for the overall development of Japanese and non-Japanese alike, it will also help understand the way Japanese think and behave. This can also enhance our own communication, as well as giving us important insights into how to teach languages better.

On a wider social level, we can communicate in various ways, for example, through social media and opinion pieces to national news outlets by studying Japanese language and society. Examples of opinion pieces (see **References**) often encourage others to express their own ideas, and have been referred to frequently throughout this present study. The intersections between school and society on the one hand, and study and self-development on the other, are all around us. We simply need to look around us and learn.

In multi-lingual communities including English (for example, Canada and many parts of Europe) genuine cross-cultural social interaction is a basic precept of both linguistic and personal progress, one so far eluding Japan. Bilingual kindergartens would lead to major intercultural and inter-linguistic evolution, culminating in more bilingual elementary and high schools, the keys to a more communicative future for Japan which it desperately needs.

While Japan has long prided itself on its ethnic homogeneity, it already has a significant population of bicultural families. Their present and future children would thrive in a more openly communicative society. Japan will be the winner in this equation, as many ground-breaking attainments are being made by bicultural residents here as well as by Japanese living abroad. Such achievements can increase many times over if Japanese education applies its resources and abilities to increasing opportunities to foster English education and enable a more multi-cultural and -lingual society to flourish. This means promoting communication, not tests or texts or “teachers” without training. Japan has more dictionaries (paper, electronic and cyber) than anyone per capita, yet the ability to communicate is one of the lowest. Successful speakers use their ears more than eyes to interact meaningfully. Open face-to-face talk is rare in Japan’s one-way classrooms, but the best spoken test is one teachers don’t know the answers to beforehand.

References

- Brown, J.D. (1995). English Language Entrance Examinations at Japanese Universities. *TLT* 19(11). JALT.
- Brumfit, C. (1984). *Communicative Methodology in LT*. CUP.
- Couper-Kuhlen, E. (2009). Prosody. *Pragmatics of Interaction*, 174-189.
- De Groot, F. (2018). E-mail correspondence, with author's permission.
- ELEC. (2018). <http://www.elec.or.jp/aboutelec/en/>.
- ETS. (2018). www.ets.org/toeicpdfspeaking-writing-sample-tests.pdf.
- Fulcher, G. (2006). *Language Testing and Assessment*. Routledge.
- Heaton, J. (1975). *Writing English Language Tests*. Longman.
- IMF. (2018). *Report for Selected Country Groups and Subjects*.
- Japan Times. (Aug 29, 2018). Japan turns to robots in the classroom to boost English skills.
- Jespersen, O. (1935 ed.) *Essentials of English Grammar*. Allen & Unwin.
- McCrostie, J. (Jul 5, 2017). Spoken English tests among entrance exam reforms Japan's students will face in 2020. *The Japan Times*.
- MEXT. (2002). *Action Plan to Cultivate Japanese with English Abilities*.
- Nation, P. (2009). *Teaching ESLEFL Listening and Speaking*. Routledge.
- Ockey, G & Li, Z. (2015). Assessing Oral Communication. *LV*, 7, 1. 1-21.
- O'Donnell, K. (2005). Japanese Secondary English Teachers. *Language, culture and curriculum*, 18, 300-315.
- Powers, D. (2013). Assessing English Proficiency in All Four Domains. *ETS TOEIC Compendium Study 2*.
- Prime Minister's Commission. (2000). *Japan's Goals in the 21st Century*. <http://www.kantei.go.jp/jp/21century>.
- Reuters. (Feb 24, 2014). Quotes of the Sochi Winter Games.
- Richards, J. (2008). *Teaching Listening and Speaking*. CUP.
- TUJ. (2017). <https://www.tuj.ac.jp/news/2017/06/09/tuj-2017-graduation>
- Van Ek. (1975). *Threshold Level*, Council of Europe ML Project.
- Wood. (2010). TOEIC Materials and Preparation Questions. *TLT* 34 (6).
- _____. (2015). *Using Photographs to Learn a Second Language*. Mellen.
- _____. (Apr 4, 2015). Fluency requires direct interaction. *The Japan Times*.
- _____. (Apr 16, 2017). Another bad idea for English education. *Ibid*.
- _____. (Jul 13, 2018). Teachers and students lose while companies rake in profits. *Ibid*.
- _____. (Sep 7, 2018). Stick to legitimate English techniques. *Ibid*.

(デイビッド ジョン ウッド：英語学科 教授)