

Teaching progressively ...for the *Senta Shiken*!

Mike Guest
Miyazaki University

Reference data:

Guest, M. (2006). Teaching progressively... for the Senta Shiken!.

In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.) *JALT2005 Conference Proceedings*. Tokyo: JALT.

University entrance exams have long served as a whipping boy for Japan's alleged English educational shortcomings. Among the popular and critical views of these entrance exams is that they focus inordinately upon narrow and arcane points of grammar (Brown, 1995; McVeigh, 2001), are largely limited to translation exercises, use de-contextualized sentence-level texts, present tasks and questions as discrete-items that foster only receptive learning (Brown & Yamashita, 1995), and that they do not address wide-ranging, holistic, practical skills (Murphey, 2001). The resultant claim is that the so-called "backwash effect" of these exams negatively influences high school English teaching pedagogy (Gorsuch, 1998) towards grammatical detail and translation drudgery. This paper intends to overturn these notions by showing that the foremost nationwide university placement exam, the *Daigaku Nyuushi Senta Shiken* (widely, and hereafter, referred to as the *Senta Shiken*) demands skills that can and do correspond to widely-held views of progressive or holistic pedagogies and thus should not be held responsible for any "grammar-translation" pedagogical backwash. In doing so, this paper will also attempt to demonstrate that the *Senta Shiken* does meet several standard criteria for test validity.

大学入試は長い間、効果の表れない日本の英語教育の原因として、非難の対象となってきた。最も一般的な批判的見解としては、入試が過度に細かく分かりにくい文法事項に焦点を当て、和訳問題に限定され、文脈のない文レベルのテキストを利用していることや、タスクや問題が個別項目であり受容形学習を促進し、幅広くホリスティックなスキル育成を着眼点に入れていないことが挙げられる。これらの試験のいわゆる否定的な「波及効果」として、高校での教授法が、詳細にわたる文法説明やつまらない和訳に徹していることが残念ながら指摘される。この論文では、主要な全国規模の大学入試、大学入試センター試験(広く「センター試験」と呼ばれ、ここでも以降そう呼ぶ)が、前進的、あるいはホリスティックな教授法について普及している見解に合致するもので、実際それに即しており、「文法訳読式」の教授法への波及効果を持つものではないことを示し、上記の概念を覆そうと試みる。また一方で、センター試験がテストの妥当性を満たすために、いくつかの基準を満たすべきであることを論証したい。

Features and functions of the *Senta Shiken*

In order to accurately address questions of validity and alleged backwash surrounding the *Senta Shiken*, it is first necessary that the general nature and structure of the *Senta Shiken* be outlined, and its role and function within the Japanese education system as well as some of the practical constraints surrounding the test be explained. After this, the 2004 *Senta Shiken* will be analyzed in detail, with those features that would seem to allow for a “progressive” backwash and, further, indicate a high degree of test validity, noted.

The *Senta Shiken* is the focal point of the Japanese university entrance exam system. The *Senta Shiken* is a standardized, nationally-applied test taken by the majority of students¹ who wish to enter a Japanese university. It is developed and administered by the Ministry of Education with input and help from various academic sources, including highly-qualified university professors, among them native speakers of English.

Each year, on a set weekend in February, over 500,000 examinees sit for these tests at several hundred strictly-controlled local examination centers throughout the country². Depending on the year, the *Senta Shiken* contains from 4 to 6 subjects, most commonly having a five-subject core. *Kokugo* (Japanese language), mathematics, and English are the subjects that appear on the *Senta Shiken* annually, and it is these subjects that carry the greatest weight in terms of final scores. Other test subjects (most often sociology and science) vary from year to year and carry less weight. Almost all academically-inclined high schools in Japan and thousands of *jukus* (cram schools) base their final year of

high school curricula largely toward the goal of achieving success on this exam, hence its alleged backwash effect on high school pedagogy.

Background of critical research

Some researchers have already questioned the extent and nature of this alleged backwash effect. Mulvey (2001) argues that the phenomenon is overstated—that only about half of all high-school graduates take the *Senta Shiken*³ and further, that the scores tend to really affect only those attempting to enter the most highly rated universities. Mulvey (2001) and Guest (2000) argue that there seems to be little correlation between the actual content of the university entrance exams and current high school pedagogy and curricula. Nonetheless, there certainly remains a widely held *perception* that the *Senta Shiken* score will severely affect university entry⁴, plus a further widely-held perception that there is a negative backwash, a perception fed by faulty notions of the nature of the university entrance exams⁵. Therefore, while the actual backwash may be an inaccurate reflection of the exams, the faulty perception alone does end up generating negative backwash regardless.

Both Stout (2003) and Mulvey (2001) correctly argue that the *Senta Shiken* has recently become less of a decisive factor in determining university entrance, largely due to demographic factors⁶ that have reduced the effect of backwash, although Stout is mistaken in his belief that this is particularly so for “public” universities because they administer their own exams. Stout appears to be under the impression that universities have a choice of employing either the *Senta Shiken* or their individual university exams

(widely known as *Niji Shiken*) as the singular means of entry, but this does not reflect the actual system.

In fact, the *Senta Shiken* does not contain a pass/fail criterion in and of itself. What is vitally important is the actual score (widely known as *souten*⁷) obtained on the *Senta Shiken*. It is this *Senta Shiken* score that will more-or-less determine which individual university entrance exam the candidate can sit for. Nor does the *Senta Shiken* confer university entry in and of itself. Rather, university entrance committees will set a basic standard for candidates sitting their *Niji Shiken* based on the candidate's *Senta Shiken* score. If an examinee has a sufficiently high *Senta Shiken* score, he or she will have a greater chance of passing a prestigious university's entrance exam. This is because all national universities, which are generally considered to be the most prestigious, will factor candidates' *Senta Shiken* results into the final score of their own examinations⁸.

Constraints upon the *Senta Shiken*

There are several constraints and conditions surrounding the administration of the *Senta Shiken* that have to be factored into any critique or analysis. First is the fact that results must be calculated quickly (in the case of the *Senta Shiken*) to allow examinees to make informed decisions about which *Niji Shiken* to sit for. This means that the *Senta Shiken* will have to contain many discrete answers—the kind that will be machine readable in the case of the *Senta Shiken*. Given the large number of candidates who sit for the *Senta Shiken* and the need for hasty results, there is no way around this constraint.

Given the competitiveness and seriousness with which the exams are held (entrance to a particular university may affect an examinee's lot in life), a high degree of objectivity is expected of both the *Senta* and *Niji Shiken*. Any subjectivity in the grading is believed to allow the prejudices of the evaluators to creep into the results, leading to criticism of imbalance or unfairness, hence another reason for adopting a mechanical approach to grading. This also means that the type of questions that demand a subjective evaluation on behalf of the grader do not appear on the *Senta Shiken*, and are limited on *Niji Shiken*. This is a particularly vexing problem for evaluating holistic language skills since—as the differences between grammatical competence and interactional qualities such as discourse, sociolinguistic, and strategic competencies clearly indicates—so-called objective features of language account for only a small part of overall communicative skill in any given language.

Such constraints mean that opportunities for testing productive skills, such as writing and/or speaking, will also be severely limited on the *Niji Shiken* and nonexistent in the case of the *Senta Shiken*. A subjective evaluation of the speaking skills of over 500,000 examinees cannot be achieved within the time frame required, even if there were some way to ensure scoring balance and fairness. Moreover, since the great majority of evaluators will be Japanese, they may not feel comfortable subjectively evaluating the dynamic, interactive aspects of what is for them a second language.

One should also keep in mind that English is only *one* of the subjects contained within the *Senta Shiken*. Therefore, the time set aside for taking the test is a further limiting

factor. The current allotted time for the English section of the *Senta Shiken* is 80 minutes. If the English section were to contain sections focusing upon expository writing or speech, the test would take up the better part of an entire day, reducing the amount of time available for other subjects to be tested and increasing what is already a considerable psychological burden for examinees⁹. However, the lower number of candidates taking individual *Niji Shiken* allows these examinations to contain more productive and subjectively graded tasks such as expository writing in English. Individual universities also occasionally demand complete written English essays over and above their actual *Niji Shiken*. On the other hand, some universities have no English section on their *Niji Shiken* at all.

Finally, it should be remembered that the ultimate function of the *Senta Shiken* is placement, and therefore the immediate goal is to rank and stratify examinees such that they can make reasonable choices about which *Niji Shiken* to sit for. As a result, it is unreasonable to criticize the test for being norm-referenced rather than criterion-referenced, as a criterion-referenced test would contradict the placement function that the tests have.

Unfortunately, in past critical evaluations of Japanese university entrance exams (see Brown, 1995; Brown & Yamashita, 1995; Murphey, 2001; Stout, 2003), some of the aforementioned factors have been ignored. Calls for extending the test to include all four skills seem not to recognize its practical unfeasibility. Brown and Yamashita's (1995) criticism of the *Senta Shiken* as being passive and receptive, as reflected in the high degree of multiple-choice questions, appears not to take into consideration the demands

for objectivity and fast results on a massive scale. Claiming that the *Senta* and *Niji Shiken* are not valid as tests of communicative English (Stapleton, 1996; Murphey, 2001) may be accurate but misses the point, since neither test is meant to be a test of communicative skill. And yet, while similar claims regarding the possible lack of validity and reliability of the tests are legion, no one has yet conclusively shown that they are unreliable or invalid.

Analysis of the 2004 Center English Shiken

Both Brown and Yamashita (1995) and Murphey (2001) argue that the validity of Japanese university entrance exams has not been established¹⁰. There are, of course, several means by which one can measure test validity. Paramount is the question as to whether the test measures what it presumes to measure and for the purpose it is designed. The *Senta Shiken* attempts to measure a basic academic aptitude for English with consideration given to the content as taught in the standardized high school curricula¹¹. In other words, the *Senta Shiken* does not measure high school English achievement per se but does employ high-school pedagogical content as a reference point when developing test texts and questions. The purpose of the test is to stratify examinees sufficiently so that they can choose appropriate *Niji Shiken* for which to sit. The *Senta Shiken* does not attempt to measure real-world skills needed for functioning in society. That is not its purpose. Therefore, criticisms that the exam does not reflect or enhance real-world needs and skills do not address the test's reliability or validity.

If the purpose is to measure the examinees' general academic aptitude in English, validity might best be

measured using the following criteria: Does the test use a variety of texts? Do the tasks and questions demand a variety of skills? Does the test demand holistic and integrated skills—as opposed to narrow, discrete-point shards of knowledge? These factors for determining content validity do not constitute a complete investigation of validity as a whole, but they are clearly factors in determining certain key aspects of test validity. They are a key element of the following analysis.

While Stout (2003) argues that a better-designed *Senta Shiken*¹² would and should lead to positive backwash effect, using the following analysis, it will be argued that the current *Senta Shiken* can and should already allow for a positive effect upon high-school pedagogy.

The 2004 *Senta Shiken*, using the version found in *Senta Shiken Kako Mondaisaku: Tanki Kansei Ban* (2004), contains six sections in total, most of these divided into a further two or three sub-sections. The test time allotted for the English section is 80 minutes, with the total score being 200 points.

Section 1:

...is worth 16 points (8%) and begins with two sets of questions on stress. The first set focuses upon syllable stress within individual words.

Sample:

The city made a lot of (a) progress in the area of (b) industrial development.

The test then provides four multiple-choice examples of possible stress patterns involving the two keywords above,

from which the examinees choose one.

A following subsection focuses on word stress in sentences taken from a casual dialogue of ten turns between two characters.

Sample:

Kei: What are you looking for?

Ted: (1) I can't find my wallet. Oh no, maybe (2) I left it somewhere.

Examinees are then asked to choose which word would be stressed in these sentences (and dialogue).

Section 2:

...is worth 38 points (19%) and is divided into three parts. This section involves selecting the correct word or phrase from a multiple-choice set to complete the sentence or short exchange. This demands lexico-grammatical rather than purely grammatical knowledge. Moreover, the choices are often practical signal words, phrasal verbs, examples involving pragmatic force, and phrases of social propriety—not random or obscure vocabulary.

Sample A:

The laundry won't dry quickly _____ it's sunny.

1) if 2) whether 3) unless 4) since

Sample B:

Osamu: It's very hot and humid today. Will it be any better tomorrow?

Betty: I heard it's going to be even worse!

Osamu: _____

1) *I'm sorry that it's going to be cooler.* 2) *That's too bad. I like humid weather.* 3) *It's a shame that it's so dry.* 4) *Oh, no. I'm planning to go hiking.*

The final part of Section 2 asks examinees to put jumbled vocabulary in the correct order in a sentence, with two items (marked with an X below) determining the correct position. This demands productive as opposed to receptive- skills.

Sample:

That gallery has many _____ X _____ X _____ miss.

1. that 2. not 3. art lovers 4. paintings 5. should

Section 3:

...is worth 34 points (17%) and begins by asking examinees to place the correct rhetorical signals (e.g., “in spite of this”) within a paragraph about the differing meanings of the word “hot” in English. In order to do this correctly, an examinee will need to be able to understand the rhetorical flow and discursive structure of the entire paragraph. The answers are once again provided in a multiple-choice format.

The subsection in Section 3 asks examinees to put in order a series of three sentences that have been removed from the middle of two short paragraphs (about Interpol and telegraph lines respectively). The third part involves a short essay about children and small animals in which examinees are instructed to correctly insert three sentences at appropriate places in the text. Six choices of placement are given. Correctly completing this task means that examinees will again have to display understanding of the holistic, rhetorical flow of the essay. This demands not only some

comprehensive reading skills but also the ability to re-order texts in a coherent, organized manner.

Section 4:

...is worth 35 points (18%) and involves a short but simple factual essay about holidays and stress, comparing various Western countries with Japan. Examinees are asked to abstract information in the essay and match that information to an adjacent chart as well as answer questions focusing upon recognizing and interpreting themes and summarization by extrapolating relevant information from the text.

Section 5:

...is worth 32 points (16%) and is based upon a casual dialogue between two newspaper editors. This dialogue is highly transactional—a lot of information is shared within the dialogue—and it is also quite natural stylistically. This section asks examinees to do three tasks: 1) be able to differentiate between different senses of the word “it”, 2) abstract written detail from the text in order to match adjacent drawings regarding the newspaper layout, and 3) sequence paraphrased ideas and events from the text.

Section 6:

...is worth 45 points (22%) and involves a lengthier, narrative essay. Again, the questions demand a variety of skills: holistic reading, sequencing, reading-between-the-lines, paraphrasing, interpreting, and making inferences. It is notable that none of the correct answers are found directly in the text.

Sample:

Why did Kate want to help Angela?

1) *She wanted Angela to win.* 2) *The coach insisted that she do so.* 3) *Angela was a newcomer.* 4) *She understood how Angela felt.*

Summary of 2004 English *Senta Shiken*

It is immediately notable that very few of the question items demand only discrete grammatical or other de-contextualized, sentence-level knowledge (54 points or 27%), and there are *no items* that focus specifically on arcane vocabulary or discrete grammar minutiae.

The texts are heavily weighted towards essays rather than limited-context discrete items (over 80% of the total text of the test is found in essays), gradually increasing in length and complexity as the test progresses. The tasks on these essays demand a variety of wide-ranging and comprehensive reading skills: making inferences, summarizing, recognizing themes, extrapolating information indirectly, knowledge of the functions of rhetorical signals and connectors, scanning for specific information, and socio-pragmatic understanding. In order to complete these varied tasks, comprehensive reading skills and a more integrated knowledge of English are required.

The topics and writing styles of the essays are varied (scientific essay, narrative, dialogue, opinion, etc.) as are the lengths. The *Senta Shiken* is also of a suitable length and contains sufficient variety in difficulty to adequately separate skillful and less skillful examinees for the purpose of placement¹³.

When all of the above factors are considered, it seems that they meet the basic criteria for validity. At the very least, an accurate understanding of the test content by high-school teachers should allow for the introduction of more holistic and progressive methodologies and practices in high school English classes.

How should these findings affect high school English pedagogy?

As the analysis has shown, the *Senta Shiken* does not meet the stereotype of a narrow, arcane, discrete, grammar and translation-based examination. Instead, a clear emphasis upon a variety of texts and genres and a variety of question types that demand a number of skills has been noted. While a very small portion of the *Senta Shiken* focuses upon the narrow and discrete properties of word stress, the great majority of the test is concerned with holistic and comprehensive understanding of a text's flow—it's cohesion, development and, most importantly, it's purpose or meaning. Examinees are asked to predict, summarize, expound, extrapolate, interpret, sequence, reconstruct, paraphrase, translate, skim, scan, read for specific information, and read between the lines using texts of a variety of genres (narratives, information transactions, casual dialogues, scientific essays, personal essays, etc). These skills are quite in conformity with what is usually considered to be a progressive approach to reading comprehension.

It is important here to analyze these results in light of Brown and Yamashita's (1995) seminal study. Brown and Yamashita criticize the content validity of Japanese English university entrance exams on three accounts in particular.

One is that knowledge of a particular topic might aid an examinee already versed in that topical area. However, it is almost impossible to escape this charge, no matter what text is chosen, since any topic will be familiar to at least some examinees¹⁴. The way to minimize this unreliability is to have several, varied topics which, as we have seen, appears to be the case on the 2004 *Senta Shiken*. Second, is that the *Senta Shiken* contains a largely multiple-choice format, which Brown and Yamashita characterize as discrete-point and passage-independent (p.25). Yet, a multiple-choice format can still demand integrative skills and engage the whole passage, as we have seen with this exam. Third, Brown and Yamashita claim that the variety of item types may be a factor in invalidating the exam. But an exam with very few item types—a narrower focus—could hardly be considered a holistic measure of an examinee's skill. Moreover, Brown and Yamashita actually imply the possible positive backwash of this variety of item types, noting that preparation for the exams should train for variation in item types and tasks. Finally, Brown and Yamashita criticize the widespread use of translation on the exams. While some *Niji Shiken* may still utilize translation questions, there are none on the 2004 *Senta Shiken*. Things have certainly changed since 1995.

Therefore, a pedagogy focusing upon a narrow bottom-up approach featuring the transformation of discrete grammatical features or memorization of de-contextualized word lists is not likely to lead to success on either the *Senta* or *Niji Shiken*. If learners must take these exams, it seems that exposing potential examinees to a similar variety of texts in varying genres and styles, and focusing more upon a top-down approach (moving from general meanings and

integrated, holistic frameworks to how these are realized grammatically or lexically within a text) would be more beneficial. If, in fact, most high school English pedagogy in Japan still takes a bottom-up, de-contextualized, discrete-item, structural approach, it clearly cannot be said to be the result of any alleged backwash from the construct of university entrance exams.

The Ministry of Education has reformed the *Senta Shiken* over the past several years (see Monbukagakusho, 2000). High School teachers should be aware of the possibility that they might wrongfully be imagining an older type of *Senta Shiken*¹⁵. Teaching in this style (narrow, de-contextualized, sentence-based syntax manipulations based upon discrete, arcane knowledge) will probably not help students succeed on the *Senta Shiken*.

Mike Guest is an English teacher in the Medical School of Miyazaki University. He also moonlights as a scout for a National Hockey League (NHL) team.

References

- Brown, J.D. (1995). English Language entrance examinations in Japan: Myths and facts. *The Language Teacher*, 19(10), 21-26.
- Brown, J.D. & Yamashita, S. (1995). English language entrance exams at Japanese universities: What do we know about them? *JALT Journal*, 17(1), 7-30.
- Gorsuch, G. (1998). Yakudoku EFL instruction in two Japanese high school classrooms: An exploratory study. *JALT Journal*, 20(1), 6-32.

- Guest, M. (2000). 'But I have to teach grammar!'. *The Language Teacher*, 24(11), 23-29.
- Guest, M. (in press). Teaching for those (expletive) university entrance exams. *ETJ Journal*.
- McVeigh, B. (2001). Higher education and post-meritocracy. *The Language Teacher*, 25(10), 29-32.
- Monbukagakusho (2000). *Daigaku nyuushi no kaizen ni tsuite* [How to improve the university entrance exams]. [Online] Available at:
www.mext.go.jp/b?Qmenu/shingi/12/daigaku/toushin/001102.html
- Mulvey, B. (2001). The role and influence of Japan's university entrance exams: A reassessment. *The Language Teacher*, 25(7), 11-15.
- Murphey, T. (2001). Nonmeritorious features of the entrance exam system in Japan. *The Language Teacher*, 25(10), 37-39.
- Senta Shiken Kako Mondaisaku: Tanki Kansei Ban* (2004). Kyougakusha, Tokyo.
- Stapleton, P. (1996). A reaction to J. D. Brown's recent inquiry on the English entrance exam. *The Language Teacher*, 20(3), 29-31.
- Stout, M. (2003). Not guilty as charged: Do the university entrance exams in Japan affect what is taught? *ETJ Journal*, 4(1), 1-7.

Endnotes

- ¹ Some students enter private universities directly via feeder high schools while some enter National Universities via special recommendations based on other criteria of merit, special skills, or circumstances. However, the candidate's *Senta Shiken* score is still a major factor in doling out such recommendations.
- ² These (and other) facts and figures can be found at <www.dnc.ac.jp/contact/toiawase.htm>, the homepage of the 大学入試センタ試験 (The National Center for University Entrance exams).
- ³ These numbers, though, do not appear to include so-called *ronin*—second and third-time test-takers. These *ronin* make up a considerable number of examinees.
- ⁴ Which remains so for most national universities.
- ⁵ See Mulvey (2001) for numerous citations.
- ⁶ Due to low birthrates, seats are available for almost any high school graduate who wishes to enter a university. However, competition at national universities can still be quite fierce.
- ⁷ Some researchers seem to have confused the function of *souten* with *hensachi*, a standard deviation score. The *hensachi* stratifies the results of mock exam takers, which are taken prior to the actual *Senta Shiken* as a predictor. *Hensachi* do not actually affect university entrance results.
- ⁸ In the author's university department for example, the final total is based 65% upon the *Senta Shiken* score and 35% upon the *Niji Shiken* results. This balance will vary by university and department, but the *Senta Shiken* weight is invariably over 50%.

⁹ A listening section was introduced into the *Senta Shiken* in 2006.

¹⁰ It should be noted that this claim is quite distinct from any claim that the exams are *invalid*.

¹¹ Interestingly, this suggests a type of reverse backwash effect, wherein university entrance exams are influenced by the content of high school pedagogy as well as vice-versa, a possibility which seems not to have been noted in the literature up to this point.

¹² For Stout, this means being criterion-referenced and based upon four skills. I have previously explained why these are not feasible for the *Senta Shiken*.

¹³ In fact, if the test did not serve this stratification function, it would be invalid as a placement test and the test would be re-designed. One of the reasons that changes in design do occur on the *Senta Shiken* is that, in the past, some tasks and questions have not aided in creating this stratification.

¹⁴ With the added fact that vocabulary not encountered in the high school curricula does not appear on the *Senta Shiken* and will be annotated on most *Niji Shiken*.

¹⁵ Guest (in press) contrasts the 2004 *Senta Shiken* with its 1981 counterpart, noting that the 1981 exam actually does conform to the negative stereotype and lacks several qualities associated with test validity.