HAL
open science

Developing a Japanese Language Test for a Multilingual Online Assessment System: Towards an Action-oriented Approach to Japanese Instruction in Europe

Tomoko Higashi, Chieko Shirota, Michiko Nagata

To cite this version:

Tomoko Higashi, Chieko Shirota, Michiko Nagata. Developing a Japanese Language Test for a Multilingual Online Assessment System: Towards an Action-oriented Approach to Japanese Instruction in Europe. Learning and Assessment: Making the Connections, ALTE (Association of Language Testers in Europe), May 2017, Bologna, Italy. hal-01908200

HAL Id: hal-01908200

https://hal.science/hal-01908200

Submitted on 30 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Developing a Japanese Language Test for a Multilingual Online Assessment System: Towards an Action-oriented Approach to Japanese Instruction in Europe

Tomoko Higashi, Grenoble Alpes University, Lidilem, France
Chieko Shirota, Grenoble Alpes University, France
Michiko Nagata, Grenoble Alpes University, France

Abstract: In order to respond to the need for transparent and common assessment criteria for European students learning the Japanese language, we have undertaken to develop a CEFR-based Japanese test in the multilingual assessment system "SELF" as a part of the Innvalangues project (Université Grenoble Alpes). In this paper, after an overview of SELF, we will present our approach and point out particular difficulties in the development of a test in a non-European language as a part of a multilingual common framework. We will highlight the risk of bias relative to sociocultural knowledge in a language test when the target language is socio-culturally distant from the learner's language and present our reflection with examples of our tasks and items.

1 Introduction

The number of learners of Japanese as a foreign language has been steadily increasing for the last decade in France (The Japan Foundation, 2003, 2017). In 2015, France is the first country in Europe regarding the number of learners of Japanese, with more than 20,000 learners (The Japan Foundation, 2003, 2017). Taking into account the increase in the number of learners and their diversity, as well as the development of international mobility, there is a clear need for transparent and common assessment criteria for European students learning Japanese. In these situations, undertaking to develop a CEFR-based Japanese test in the multilingual assessment system "SELF" would contribute to fill this gap. In this paper, after an overview of SELF, we will point out particular difficulties the development of a test in a non-European language as a part of a multilingual common framework. In this study, we will highlight the risk of bias relative to sociocultural knowledge in a language test when the target language is socio-culturally distant from the learner's language and present our reflection with examples of our tasks and items.

2 Overview of multilingual online test SELF

SELF means "Système d'Évaluation en Langues à visée Formative" (Assessment system of foreign languages with formative aim), which is a part of the "Innovalangues" project, winner of a National Research Grant, IDEFI, ("Initiative of Excellence for innovative formation") supported by Grenoble University. SELF can be used as placement test, but also as a proficiency test with a formative and diagnostic aim. It is an online-based and adaptive test, assessing three abilities: listening, reading and short writing. The general result and three separated results are shown, allowing each learner to become aware of their strong and weak points. It takes into account partial competence, suggested by the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001) (CEFR 6.1.3.4), and the general result advises on the optimal level course to attend.

SELF is a multilingual assessment system, available in Italian and English (for A1 to C1 level), and Chinese (A1 to B1). Furthermore, Japanese, Spanish and French as foreign language

tests are in the process of being developed. The Japanese test is scheduled to come into service in September 2017 (A1 to B1).

3 Developing the Japanese test in the multilingual framework

3.1 Process in the test development cycle

SELF is based on a common methodological approach for these languages. The test development is based on a model of qualitative and quantitative validation, represented by an iterative process of successive steps (see Figure 1).

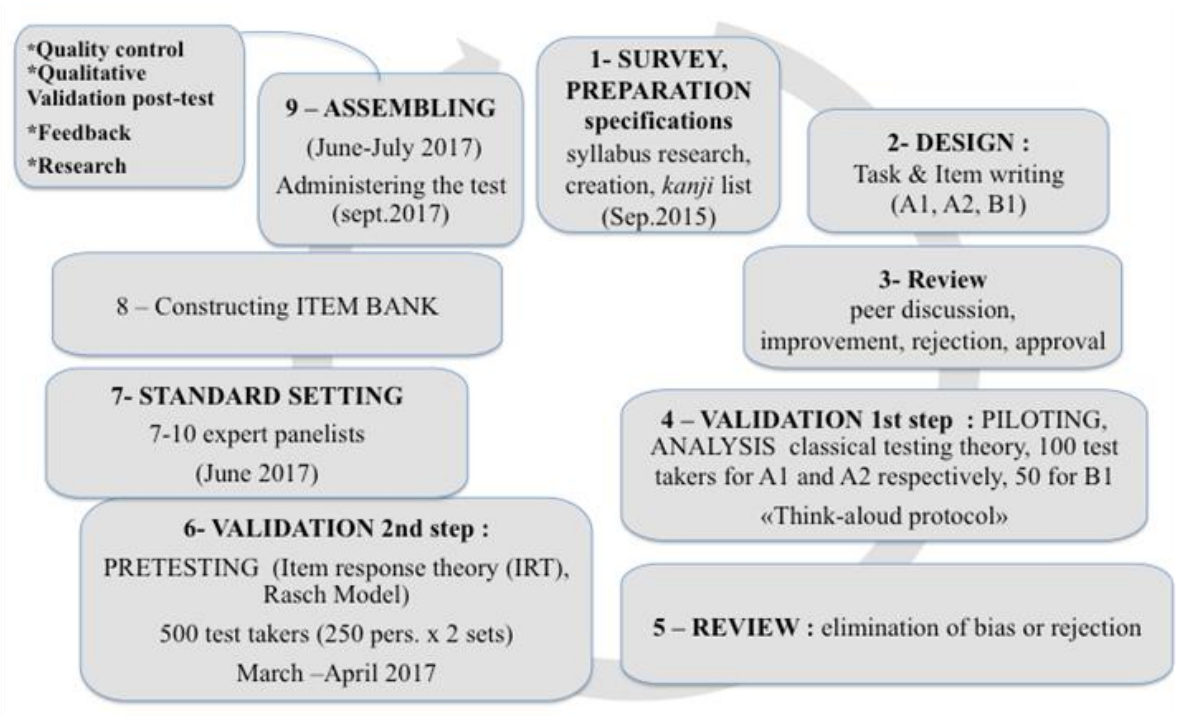


Figure 1. The test development cycle of SELF (ALTE, 2011; Cervini, 2016)

The first step of the cycle is to research the available syllabuses for the creation of tests. Unlike European languages, Japanese does not have a completed CEFR-based syllabus. So we have had to develop our own syllabus referring to the CEFR descriptors and a few CEFR-based Japanese syllabi, and we have constituted lists of Chinese characters, kanji (see sections 3.2 and 4). Step 2 concerns the task and item writing, referring to CEFR descriptors and Japanese language characteristics. Step 3 is to review these tasks designed by test developers, to improve, approve or reject by peer discussion. Only approved tasks and items will be tested in step 4, “piloting”. During this step, we also collected data by think aloud protocol with two learners by level, which allowed us to conduct qualitative analysis. At Step 5, in the light of the result of statistical analysis (classical testing theory), the items with inadequate value were rejected. Step 6 is the last validation through pretesting with almost 500 learners of Level A1 to B1. The result is analyzed with Item response theory. Then we held standard-setting meetings to

fix cut-off points and construct an item bank with validated items. The last step is to determine the algorithm and to assemble the test.

3.2 Authenticity of the tasks

Authenticity is the central notion of task conception in SELF (Cervini & Jouannaud, 2015). However, due to the Japanese graphic system, usage of authentic resources is difficult. It is important to know that the Japanese graphic system includes two systems of phonetic writing named hiragana and katakana, and another system of Chinese characters, named kanji (with the official kanji list containing about two thousands characters), and that these 3 writing systems are used conjointly even in a short sentence. Generally, about 50 kanji are taught at A1 level. So, for example, an A1 user can't understand a simple notice at a railway station, because a lot of kanji are used for the proper name of the station, for example. The type and genre of text described as A1 or A 2 level, like posters, city maps, restaurant menus, is not usable without modification.

So, the majority of our tasks are fabricated or rewritten, respecting situational and interactional authenticity (ALTE, 2011). We focused on situational authenticity, that is "tasks and items representing language activities in real life" and created a similar text type or text genre that learners of Japanese in Europe should face in daily life. The majority of learners have never been to Japan, but they often practice online-based language activities such as social networking, blog chat, etc. (Project on Language Activities and Competences of the CEFR B1 level, 2012). With regard to interactional authenticity, that is "naturalness of the interaction between test taker and task and the mental processes which accompany it", our tasks ensure the interaction (dialogue) is always between a native speaker and a non-native Japanese user. We also make sure that the test-takers can put themselves in the place of the non-native speaker. To create or rewrite the conversation, we referred to conversational analysis to ensure the naturalness of the scheme.

3.3 Reflection on sociocultural knowledge

From the viewpoint of the CEFR, the user/learner's competences are sub-divided into two parts: "general competences" including declarative knowledge (such as knowledge of the world, sociocultural knowledge and intercultural awareness) (Council of Europe, 2001, pp. 110-112) on the one hand, and on the other, "communicative language competences", which are further subdivided into three: Linguistics, Sociolinguistic and Pragmatic competences (ALTE, 2011.pp. 10-11). The SELF test assesses the communicative language competences of learners, but the task should not focus on declarative knowledge, which might distort the result on language competences. However, when we create tasks, with a view of authenticity and communicative approach, the sociocultural aspect is intrinsic in the text. Considering that Japan is culturally distant for European learners, some words or topics may cause sociocultural problems for the comprehension of text. We have chosen well known and ordinary words or notions such as manga, sushi, Kyoto, both stereotypic and explicit ones. But, sometimes, the learners don't understand the underlying role and functioning of a word or a notion in Japanese

society. In this case, the test doesn't assess learners' language competences. This problem is going to be discussed in the next section.

4 Task and item

Here are two examples of reading task related to cultural events in which people participate wearing a yukata, a kind of cotton kimono.

Figure 2 is the first example, a reading task of B1 level, "B1_CE_aquarium".

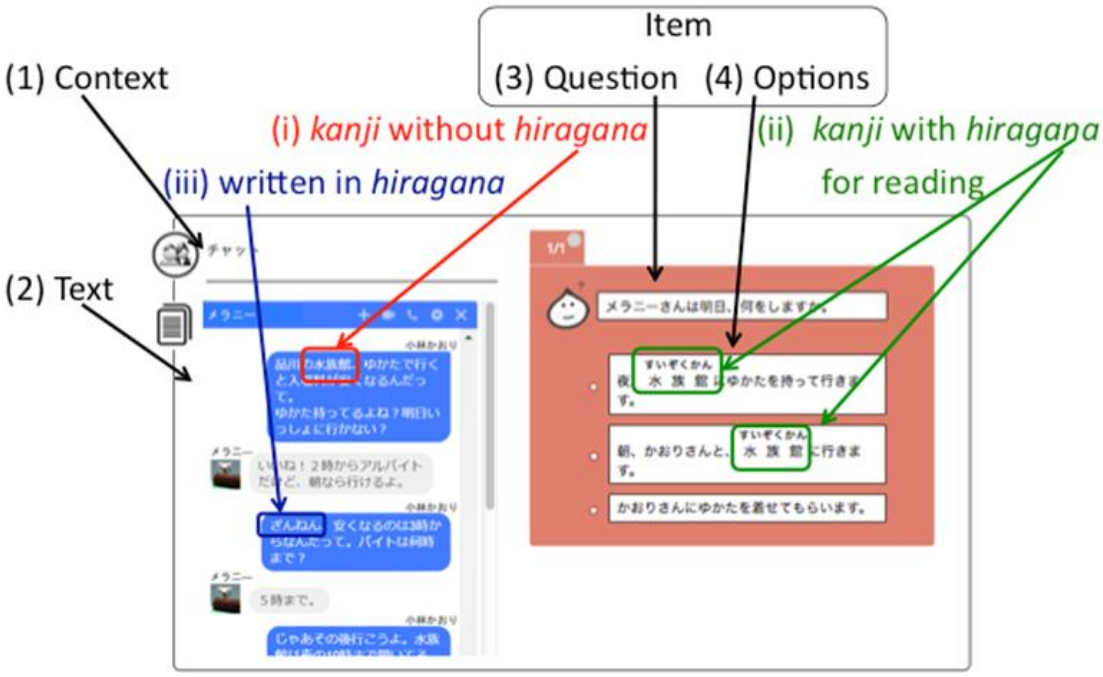


Figure 2. Reading task B1, "B1_CE_aquarium"

The tasks in SELF are composed of four elements: (1) context, (2) text, (3) question, (4) options, and the last two elements make up an item. Some tasks have more than two items. For Japanese written tasks, we decided how to write the words in kanji (Chinese characters) in the task. We made a list of kanji for each level, 57 characters for A1, 144 for A2 and 211 for B1, mainly based on the frequency of kanji in our original tasks. Basically, we use only the kanji in the lists to write the words in the task (i), and for other words, we write them in kanji with small hiragana added above the kanji to help the test takers to read these words, called "furigana" (ii), or in hiragana (iii).

The translation in English of each part is as follows (the key words to answer the question are in square brackets):

(1) Context: chat

(2) Text

Kaori: Hi, at the [aquarium] at Shinagawa, the entrance fee will be discounted if we go there [in yukata]. You've a yukata, don't you? Shall we go [tomorrow]?
Mélanie: Great! I've a job from 2:00, but I can go in the morning.
Kaori: Too bad. It's [from 3:00] they discount. Until what time, your job?
Mélanie: [Up to 5:00].

Kaori: Well, let's go [after that], right? It's open until 10:00 [at night].
Mélanie: OK! You'll [help me to put on yukata]?
Kaori: OK!

(3) Question: What will Mélanie do tomorrow?

(4) Options: A. She will [bring a yukata] to the [aquarium at night].
B. She will [go to the aquarium] with Kaori [in the morning].
C. She will [have Kaori put on her yukata]. (key)

The text type is a chat between two friends, Kaori, a Japanese student, and Mélanie, a French student living in Japan. Kaori begins this chat to propose Mélanie to go to an aquarium with a special discount entrance fee for the visitors wearing a yukata. The key is the third option, C, which means that Mélanie will be helped by Kaori to put on her yukata tomorrow. The other key words to eliminate distractors are related to the time and the verbs.

This task is based on a descriptor of CEFR, B1 level for written interaction: "Can write personal letters giving news and expressing thoughts about abstract or cultural topics such as music, films" (Council of Europe, 2001, p .83). We set the context of this language activity, as a leisure activity concerning a cultural event in which they participate in yukata.

This concept is inspired by the fact that many events in yukata for international students in Japan are organized by their university or local association. In addition, as this event at the aquarium was really organized in Tokyo, we consider this context is situationally authentic. However, according to the survey data on yukata, 60 to 70 percent of young Japanese women cannot put it on by themselves. So, they usually ask someone to help them to get it on, using an expression in causative-benefactive form of the verb "put on". The context in which this expression of the function "asking for help" is used is so natural to us that we focused on it for this task.

Contrary to our expectation, the results of the piloting test revealed that this item is too difficult for B1 level students, as the proportion of correct response is only 25 (see Table 1). We suppose that test takers believe that only children need to be helped to put on their clothes, at least in France where they do not wear a kimono or a yukata, except for a simple yukata as a nightdress when staying at a Japanese inn. This belief became a cultural bias that interfered with the test takers' comprehension.

The second example is a reading task of A2, "A2_CE_fete_d_ete", summer festival, of which the text is a festival poster (Figure 3). We also apply the same writing rules for words in kanji, written above, for the A level tasks. This task has 3 true-false type items, which the test takers can answer by clicking on the relevant numbers.

(i) *kanji* in the list (144) without *hiragana*: 大川, 夏, 来, 日時, 花火, etc.
(ii) *kanji* with *hiragana* for reading: 公園,留学生, 無料, etc.
(iii) written in *hiragana*: まつり(祭), きもの(着物), おどり(踊り), etc.

Figure 3. Reading task A2, "A2_CE_fete_d_ete"

The translation in English of each part is as follows (the key words are in square brackets):

(1) Context: Information of summer festival

(2) Text
Okawa Volunteer Group "Summer Festival"
Let's [dance wearing yukata]. [Then, do fireworks], too! Everyone, come to the festival!
Time and Date: Sunday, August 15
Dance: 6:00 p.m. – 8:00 p.m.
Fireworks: 8:00 p.m. – 9:00 p.m.
Place: Okawa Park
*[The international students can borrow a "yukata", a summer kimono for free].
*We will [rent a yukata to the Japanese people at 500 yen].

(3) Question: the keys are in parenthesis.
Item 1: You can [dance wearing summer kimono] in the festival. (True)
Item 2: You will [do fireworks before dancing]. (False)
Item 3: [The international students can borrow a yukata at 500 yen]. (False)

This task is designed based on a A2 level CEFR descriptor for reading activities: "Can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus, reference lists and timetables" (Council of Europe, 2001, p. 70). We imagined a festival poster organised by a local volunteer group as information of leisure activities concerning a cultural event in which they participate in yukata. As in the first example, this concept is inspired by the events in yukata organized in Japan, but to adapt to the language activities of A2 level learners, we made a bill for the international students. In this case, we can use some paraphrasing or additional explanations concerning traditional Japanese culture in the

text just as in authentic posters in Japanese universities or associations to aid international students to understand Japanese culture.

The results of the classical item analysis of piloting data of these examples are shown in Table 1. (ALTE, 2011)

	Difficulty (P-value)	Discrimination (Rir)	Options (A-value)
Ex.1 B1_CE_aquarium	25	25	59 16 25*
Ex. 2: item 1 A2_CE_fete_d_ete	80	31	80* 25
Ex. 2: item 2 A2_CE_fete_d_ete	80	49	20 80*
Ex. 2: item 3 A2_CE_fete_d_ete	76	14	24 76*

Table 1: Analysis of piloting data by Tiaplus (*key)/B1: 44 test takers, A2: 50 test takers

As we already mentioned, the proportion of correct response of the first example is 25, which indicates that this item is too difficult for B1 level, and the value of the discrimination index is inferior to 30. In contrast with the first example, the analysis data of the second example shows a high proportion of correct responses, 80 and 76, and the discrimination values of the first two items are superior to 30. These data mean that these items have a good validity as A2 level items.

Figures 4, 5, 6, and 7 show the graphic data of two examples. The main factor that affects the discrimination of the first example is that the strongest test takers could not answer correctly, which may also be an evidence of a cultural bias.

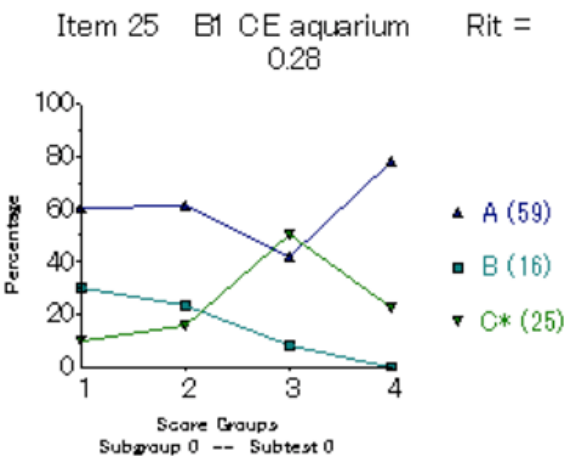
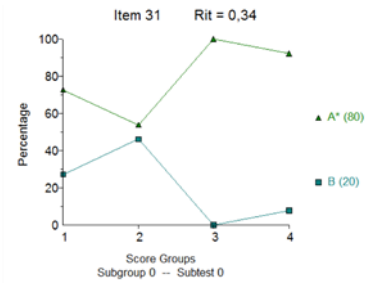
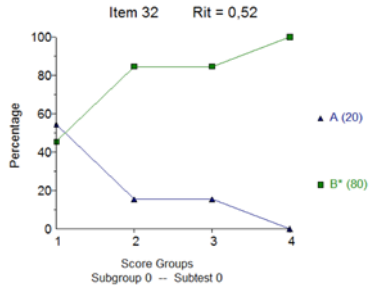


Figure 4. B1_CE_aquarium (*key)





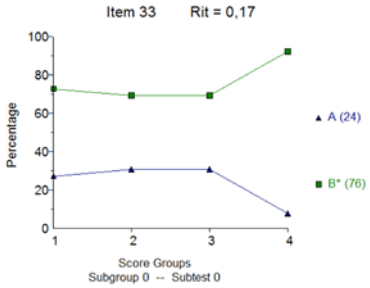


Figure 5. "A2_CE_fete_d_ete" item 1 (*key)

Figure 6. "A2_CE_fete_d_ete" item 2 (*key)

Figure 7. "A2_CE_fete_d_ete" item 3 (*key)

5 Conclusion

In this paper, we have demonstrated that the development of a Japanese test in a European environment and with the CEFR framework is possible but required some adjustments and consideration (Coste, 2007). SELF is a language test and assesses communicative language competence (with graphic, lexical, grammatical, sociolinguistic, pragmatic, and discourse components) but we showed that sociocultural knowledge plays a role in such communicative-type tasks. We have found that even an ordinary stereotypical sociocultural factor may distort an appropriate understanding of a situation if the learner didn't know the underlying functioning specific to the target culture. In other words, creating tasks with situational authenticity necessarily includes sociocultural factors. The higher the level, the more implicit and abstract the required sociocultural knowledge becomes. We therefore highlight the importance of "intercultural awareness", which would develop the sociocultural/intercultural competence, inseparable from communicative language competence (Byram, Zarate, & Neuner, 1997). This aspect should be taken into account in Japanese language education if it aims to use a real communicative/action -oriented approach.

References

ALTE (2011). Manual for Language Test Development and Examining. Council of Europe.

Bachman, L.F, Cohen,A (1998). Interfaces Between Second Language Acquisition and Language Testing Research, Cambridge University Press.

Byram, M., Zarate, G., Neuner, G. (1977) La compétence socioculturelle dans l'apprentissage et l'enseignement des langues, édition du conseil de l'Europe.

Cervini, C. (2016) Approcci integrati nel testing linguistico: esperienze di progettazione e validazione in prospettiva interlinguistica. In Interdisciplinarità e apprendimento linguistico nei nuovi contesti formativi. L'apprendente di lingue tra tradizione e innovazione. Bologna, Quaderni del CESLiC. pp.64-85

Cervini, C., Jouannaud, MP. (2015) Ouvertures et tensions liées à la conception d'un système d'évaluation en langues, numérique, multilingue et en ligne, dans une perspective communicative et actionnelle, Alsic vol 18, http://alsic.revues.org/2821;DOI : 10.4000/alsic.2821

Coste, D. (2007) Contextualiser les utilisations du Cadre européen commun de référence pour les langues, Gouiller, F. (ed).Le Cadre européen commun de référence pour les langues (CEFR) et l'élaboration de politiques linguistiques : défis et responsabilités, https://www.coe.int. pp.44-53.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press. Retrieved from http://www.coe.int/t/dg4/linguistic/source/framework_en.pdf

Project on Language Activities and Competences of the CEFR B1 level (2012), Activities' report of 2011.
The Japan Foundation. (2017) Survey report on Japanese Language Education Abroad 2015.
The Japan Foundation. (2003) Survey report on Japanese Language Education Abroad 2003.