

The Efficacy of Placement Interviews for English Language Classes at a National Japanese University Based on a CEFR-J Model

Christopher Robert Hennessy, University of Fukui, Japan
Nicolangelo Becce, University of Fukui, Japan

The Asian Conference on Education 2015
Official Conference Proceedings

Abstract

From January 2015 to March 2015, the authors, in conjunction with other faculty members at a Japanese national university, created a system of five-minute English language oral interview protocol system based on the Common European Framework of Reference for Languages (CEFR) and its adapted version, the Common European Framework—Japan (CEFR-J), and used these protocols to conduct 817 English placement interviews in April 2015 at the above mentioned Japanese national university. The first part of this paper will focus on the development of the interview protocols, including use of Can-Do lists, instructor collaboration, and interview norming. The second part is an analysis of quantitative data obtained through in-interview data collection, including: accuracy of interview results through comparison with TOEIC results, breakdown of student body by CEFR level, and interview protocol accuracy. The analysis suggests the efficacy of such an interview system for accurately placing and assessing students according to their spoken English level. In the discussion, the paper offers an overview of issues surrounding the development of English interview protocols - including question creation and norming issues - and future research planned by the authors on the oral interview protocol system.

Keywords: EFL, CEFR, CEFR-J, oral placement interview, Japan, tertiary education, national university.

iafor

The International Academic Forum

www.iafor.org

Introduction

In 2013, the Japanese national university where this research was conducted reformed its English language curriculum after having become one of the funding recipients of the MEXT Global 30+ program. This program, supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT), aims at promoting the internationalization of the academic environment of Japanese universities¹.

The new language program developed through MEXT Global 30+ currently involves 16 full-time instructors and around 1,600 first- and second-year students of the faculties of Engineering, Medical Sciences, and Education and Regional Studies. The students are divided into 67 English language classes of 24 students each that meet twice, on average, a week for 90 minutes. The curriculum starts with a focus on personal communication and gradually shifts towards English for professional communication and TOEIC (Test of English for International Communication) Listening and Reading test preparation.

Since, in compliance with the MEXT requirements, this recently implemented English language program has specific goals in terms of average TOEIC scores among the student population, the TOEIC test was also used during the academic years 2013-2014 for placement purposes. But, at the same time, given that the first part of the language curriculum starts with a focus on communication-based activities, the authors developed a new placement system that could place students according to their spoken English level, as well as an activity capable of producing a positive affective outcome in the interviewees. In order to develop a system that could be valid on an international level, but that at the same time could be flexible enough to be tailored to the specific needs of our student population, we decided to follow the Common European Framework of Reference for Languages (CEFR) and its Japanese counterpart, the CEFR-J.

The Common European Framework of Reference for Languages (CEFR) is a language framework developed by the Council of Europe as a method of learning and assessing language use in Europe. The CEFR divides learners into 6 levels (A1, A2, B1, B2, C1, C2) according to their reading, listening, speaking and writing abilities. A language-independent framework by definition, the CEFR has been adapted to the specific needs of the English language teaching contexts in Japan by Tono and

¹ For more information, see the website on *Global 30 Project -Establishing University Network for Internationalization-*: <http://www.mext.go.jp/english/highered/1326725.htm>.

Negishi (2013), who developed a new framework, the CEFR-J. This new language framework refines the “can do” descriptors developed by the Council of Europe, and further divides the original framework into 12 levels (Pre-A1; A1.1, A1.2, A1.3; A2.1, A2.2; B1.1, B1.2; B2.1, B2.2; C1; C2). Moreover, the CEFR-J introduces a wordlist of 5,639 words, covering the levels from Pre-A1 to B2. This wordlist is based on the analysis of major English textbooks used in Asian regions (Tono and Negishi, 2012).

Both the CEFR and CEFR-J "can do" descriptors for speaking were used as a starting point for the creation of a set of new descriptors for this interview protocol. Given the time constraints, in terms of both general preparation and norming process for the interviewers, as well as in terms of available time to concretely interview the students, the authors decided to develop the following simplified "can do" descriptor list:

A1 – I can ask and answer simple questions, initiate and respond to simple statements in areas of immediate need or on very familiar topics.

A2 – I can give simple descriptions of things and make comparisons. I can describe past activities and personal experiences.

B1 – I can explain and give reasons for my plans, intentions and actions.

B2 – I can develop an argument well enough to be followed without difficulty most of the time. I can speculate about causes, consequences and hypothetical situations.

The A1 level is preceded by a Pre-A1 level that does not currently have any "can do" description, given that we decided to define it as the result of the interview result in case the interviewee performance was not sufficient enough to be considered at the A1 level.

Apart from the Pre-A1 level, the four "can do" descriptors here reported follow the original CEFR scale model, except that they cover only the first four original levels (A1-B2). This choice has been made for two main reasons: the first one is that we expected our student population to be composed, in the vast majority of cases, of A- and B-level students, with only a limited number of C-level students based on previous TOEIC scores by students as well as our understanding of the CEFR level system and the level of text our students use. Consequently, we assumed that, for our expected student population, regrouping the levels B2, C1 and C2 into a "B2+" would not have affected in a determinant way the grouping of students in different classes. The second reason is directly related to time constraint: as explained in the following pages, since the interview protocol workflow always starts from the lowest levels (from Pre-A1 to B2+), the choice to avoid the assessment of the interviewee

performance beyond the B2 level would have helped the interviewers to stay within the time limit of five minutes per interview.

Interview Logistics

Placement interviews took place from April 7 to April 9, 2015 (the first week of classes) during students' regular English class times. A total of 817 first-year students from three different departments (Engineering, Education, and Medical) were interviewed by 14 different instructors over six class periods during this three-day period, requiring a total of 69 periods of actual instructor labor time. There were 33 distinct English classes interviewed. Interviews were conducted in a number of places on the university campus including classrooms, instructor offices, and small multipurpose rooms. In general, one class (around 24 students for Engineering and Education; around 36 for Medical) required two instructors (three instructors for Medical) per one 90-minute class period in order to conduct each interview within the five-minute time limit.

Development of the Interview Protocol

The 4 Stages

The authors break up the development of the devised Interview Protocol (IP) into four distinct stages as described below

Stage 1

The original idea for the IP incorporated CEFR-J spoken production and interaction descriptors described above in order to devise a points-based system that would fit onto one piece of paper, making it easier for the interviewer to organize information and determine the interviewee's level (See *Figure 1*). All questions are on one sheet and the interviewer works progressively up from Pre-A1 to B2 level questions as he/she determines the accuracy of the interviewee's answers. Interviewers assigned 0 ('Communication does not happen'), 1 ('Communication needs help in order to happen') or 2 ('Communication happens') points depending on the answer given by the interviewee. Point explanations and CEFR-J descriptor guidelines are also provided on the sheet to aid interviewers in conducting interviews efficiently. Issues arose with this initial IP, though, as it contained a lot of text, making it difficult to follow in a timely manner. Also, the three-point system proved too arbitrary to allow interviewers to determine expected CEFR levels with consistency.

Name:		ID:		Interviews 2015			CEFR - Self-Assessment Grid	
Questions	Score	Instructions	2 Points Communication happens	1 Point Communication needs help in order to happen	0 Points Communication does not happen	Spoken Interaction	Spoken Production	
Warm-up questions:								
What's your name?	2 1 0	Always go on	Acceptable answer without hesitation					
Where are you from?	2 1 0							
How old are you?	2 1 0							
Set 1:								
What do you like about your town?	2 1 0	3 or more points – go on	Acceptable answer without hesitation	a) General hesitation (for example Japanese answer first, then English equivalent) b) The interviewer needs to repeat the question c) Unclear answer d) Inconclusive response		Silence or totally unrelated answer	A1: I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.	A1: I can use simple phrases and sentences to describe where I live and people I know.
Where do you buy your clothes?	2 1 0							
What do you do in your free time?	2 1 0							
Set 2:								
What did you do last weekend?	2 1 0	5 or more points – go on	Acceptable answer without hesitation (one complete sentence or thought at least)				A2: I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.	A2: I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job.
How important is music in your life?	2 1 0							
How often do you go out with your friends?	2 1 0							
Set 3:								
Is there a place you would really like to visit?	2 1 0	4 points – you're B1	Acceptable and complete answer				B1: I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversations on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).	B1: I can connect phrases in a simple way in order to describe experiences and events, my accents, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.
(Follow up) Why?	2 1 0							
Set 4:								
Are there any new skills you'd like to learn in the future?	2 1 0	4 points – you're B2	Acceptable and complete answer				B2: I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.	B2: I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
(Follow up) Why?	2 1 0							
Exit question:								
Welcome to the campus? Do you like this campus?	2 1 0	Always ask this question						
Comments:								

Figure 1: Stage 1 Interview Protocol.

Stage 2

Several changes were made to develop the second iteration of the IP system. The main progressive level-up style was maintained, but a visual binary tree similar to a gameboard was developed in order to allow the interviewer more ease in following the question process (See *Figure 2*). This change reduces the time needed for interviews. Also, the three-point system was amended to a binary two-point system ('Communication occurs' or 'Communication does not occur') in order to reduce ambiguity and create more consistency among interviewers' expected CEFR-level determinations. Finally, a 'soft-landing' exit question was kept to ensure interviewees did not end the interview on 'wrong' answers. However, the binary nature was determined to still be too confusing to follow, resulting in longer interview times.

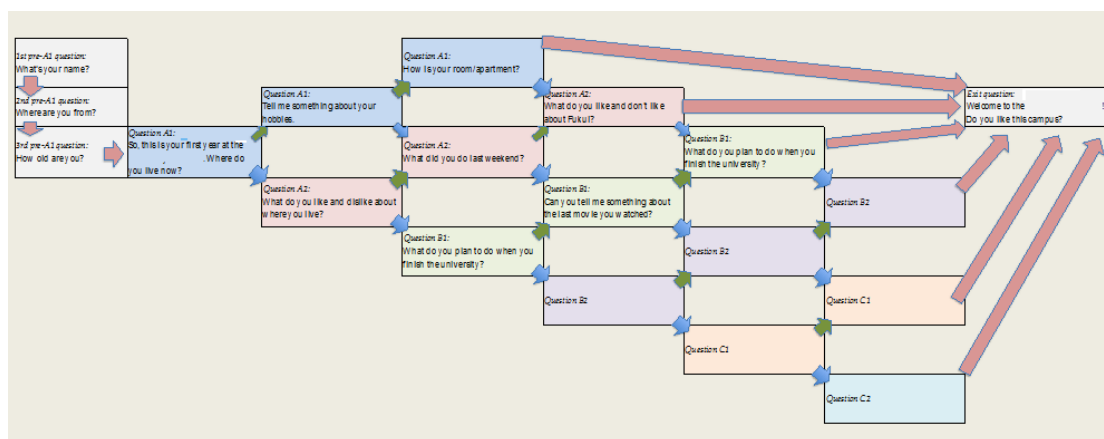


Figure 2: Stage 2 Interview Protocol.

Stage 3

In Stage 3 of the IP, most of the characteristics of Stage 2 IP are maintained; however, the binary tree system of Stage 2 was further developed to a visually represented level-up system (See *Figure 3*). Interviewees would have to accurately answer two of three questions at the A1 and A2 level to move up to the next level, and the question at the B1 level in order to move up to the B2+-level question. Levels are clearly written on the right side of the page for quick level-reporting by the interviewer. In addition, base conditions for each level are also written to aid interviewers in determining the acceptability of an answer. Also, acceptable follow-statements for interviewers were added in order to elicit more responses for the interviewee if the need were to arise. A comments section was added as well to allow interviewers to record any additional information deemed necessary. Finally, a check box was added to each question for recording the occurrence of a question being posed to an interviewee.

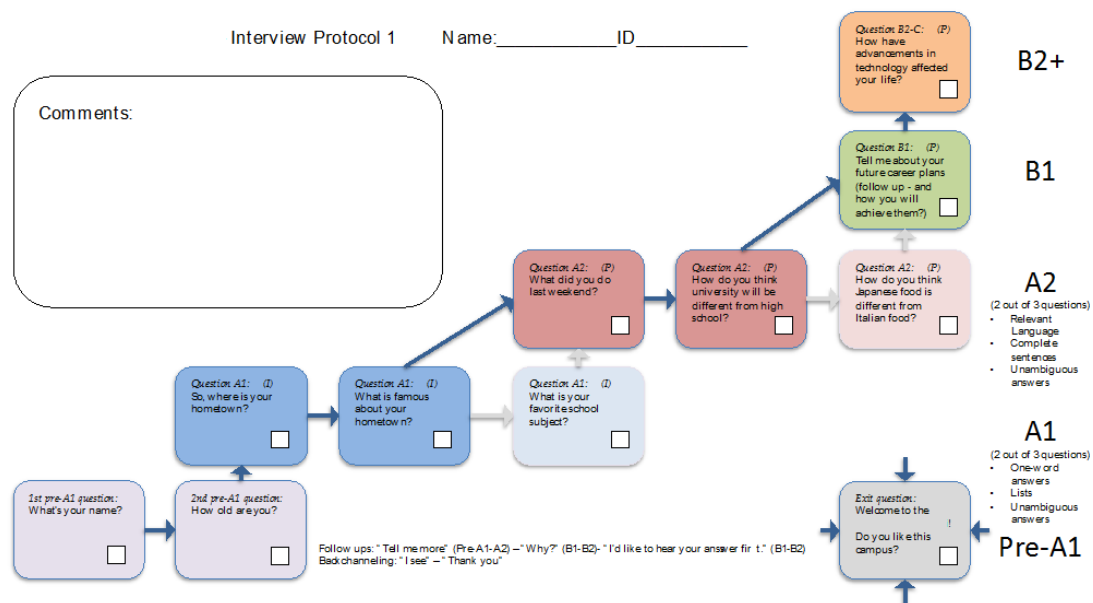


Figure 3: Stage 3 Interview Protocol.

Stage 4 (Final Version)

In Stage 4 of the IP, most of the characteristics from Stage 3 were maintained. An additional B1 question was added and minor cosmetic alterations were made in layout to further ease visual reference (See *Figure 4*). Also, another check box was added ('2x') for recording when a question was asked two times as well as a place to record time length of the interview. Using this Stage 4 Interview Protocol model, a total of three IPs with different (but CEFR-equivalent) content were developed to be used for interviews.

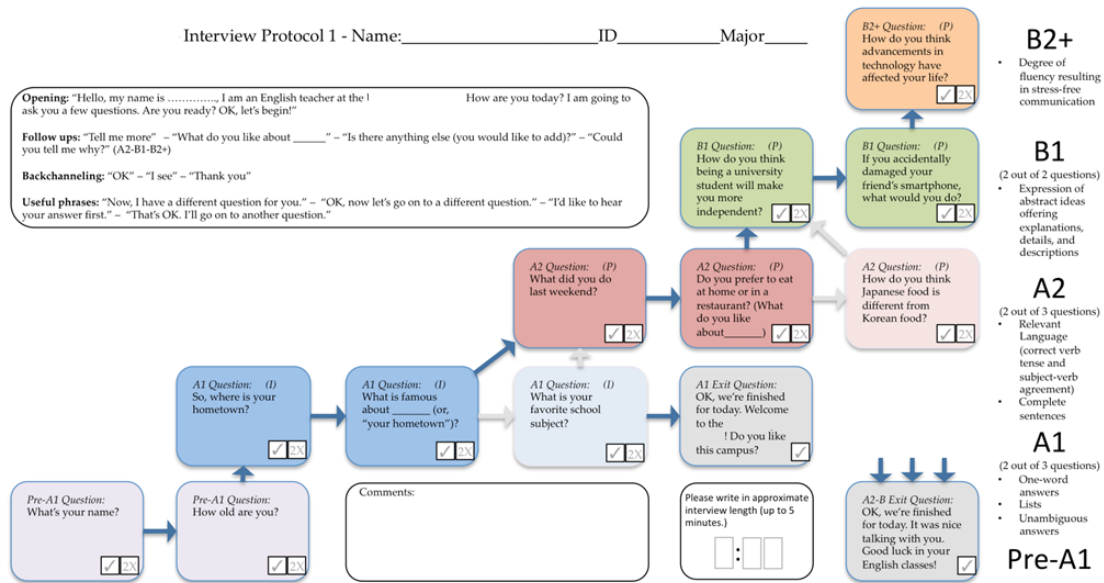


Figure 4: Stage 4 Interview Protocol.

Norming Process

Norming for interviewers was conducted over two three-hour sessions in March 2015, approximately one week before interviews were to be conducted. This section will explain the general content of each session.

Session 1

Session 1 began with a general explanation of, and discussion on, the process for conducting an interview and the logistics on the days the interviews were to be conducted. Instructors were given a walkthrough on the actual interview process and explanation of the IP. Also, instructors were given their interview dates, times, places, and interviewee list. This was followed by training in conducting interviews through blind assessment and discussion of a number of CEFR interviews – both video-taped mock interviews conducted with students of a national university and CEFR interview-training resource videos found online. Finally, the session ended with discussion and refinement of IP questions.

Session 2

Session 2 began with final logistics discussion and further discussion on the interview process. Instructors also trained by conducting mock interviews with students from a national university. Each student was interviewed at least two times by a different instructor each time. This activity was followed by discussion to allow instructors to compare and analyze results with each other. The end of this session featured a discussion section on the IP to finalize all details of each of the three IPs.

Findings

An analysis of the results obtained from the interviews shows a general consistency with our expectations in terms of number of interviewees that could be assessed at a hypothetical C level. More specifically, the results were the following: 28 students at the Pre-A1 level (3.4%); 384 students at the A1 level (47%); 288 students at the A2 level (35.3%); 99 students at the B1 level (12.1%); 18 students at the B2+ level (2.2%). Moreover, these percentages are very similar to those found in the surveys conducted by Tono and Negishi (2012), who stated that: "the surveys on Japanese ordinary people's English proficiency reveal that more than 80% of Japanese EFL learners are Non/Basic Users (A1 or A2), with less than 20% in B levels (Independent Users) and almost nil in C levels (Proficient Users)."

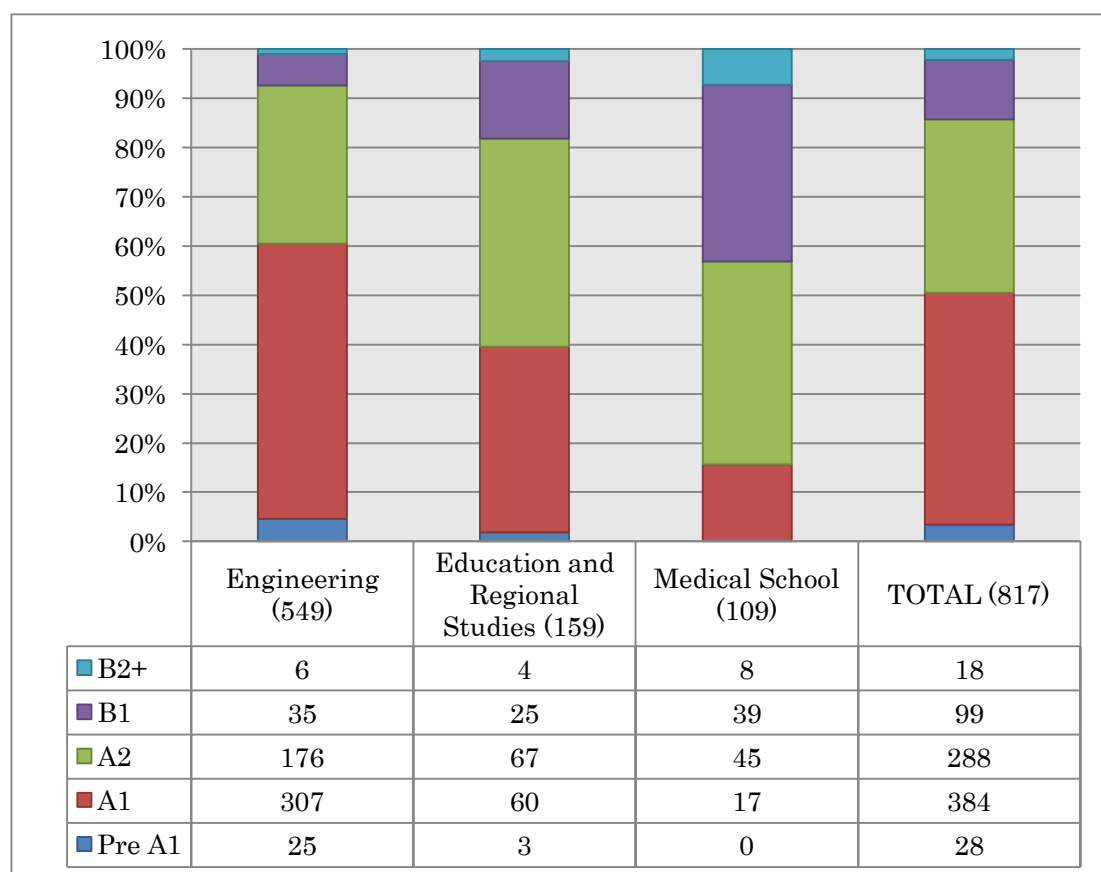


Figure 5: Interview results.

As already mentioned, the oral interview system conducted in April 2015 substituted for the TOEIC test as a placement instrument, and consequently the students took their first TOEIC test of their university career only in July 2015, which means after roughly 30 hours of communication-based language instruction. Except for a mock

test taken to develop some familiarity with the TOEIC test format during the very last week before the real test, the students did not receive any explicit TOEIC test preparation during the 30 hours of language instruction subsequent to the placement interviews.

Even though the CEFR-based oral interview system and the TOEIC Listening and Reading test assess totally different aspects of language proficiency, we tried to compare the results of the CEFR-based oral interview system with the standardized TOEIC test developed and administered by ETS, in order to see how different the results would have been had we used the TOEIC Listening and Reading test, instead of the new interview system, as a placement test for the current academic year.

The following charts describe the results obtained.

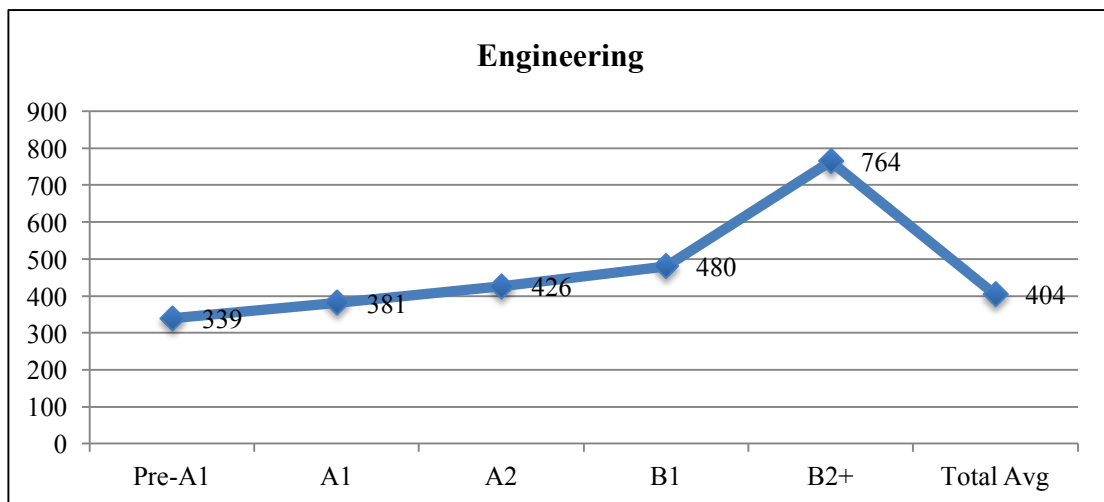


Figure 6: Comparison between TOEIC and interview results - Faculty of Engineering.

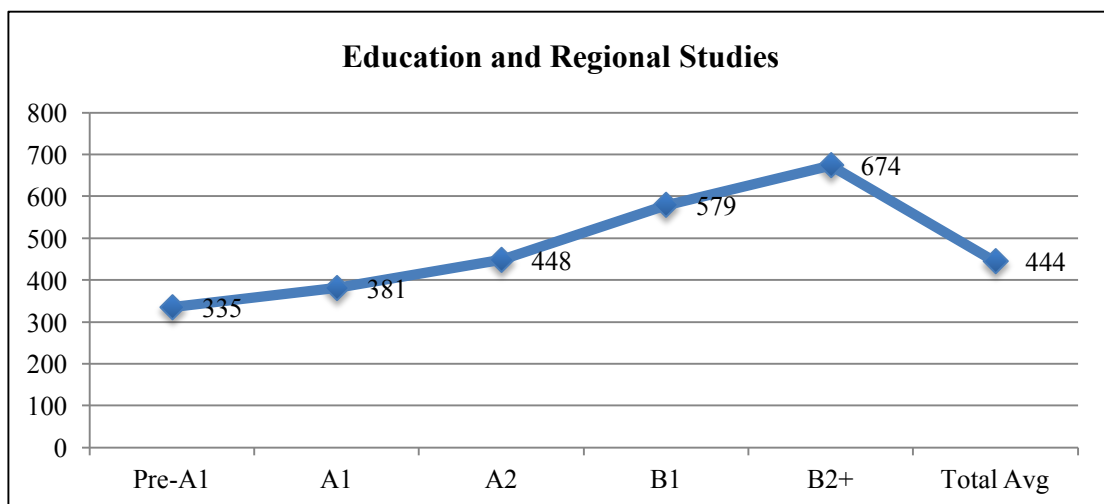


Figure 7: Comparison between TOEIC and interview results - Faculty of Education and Regional Studies.

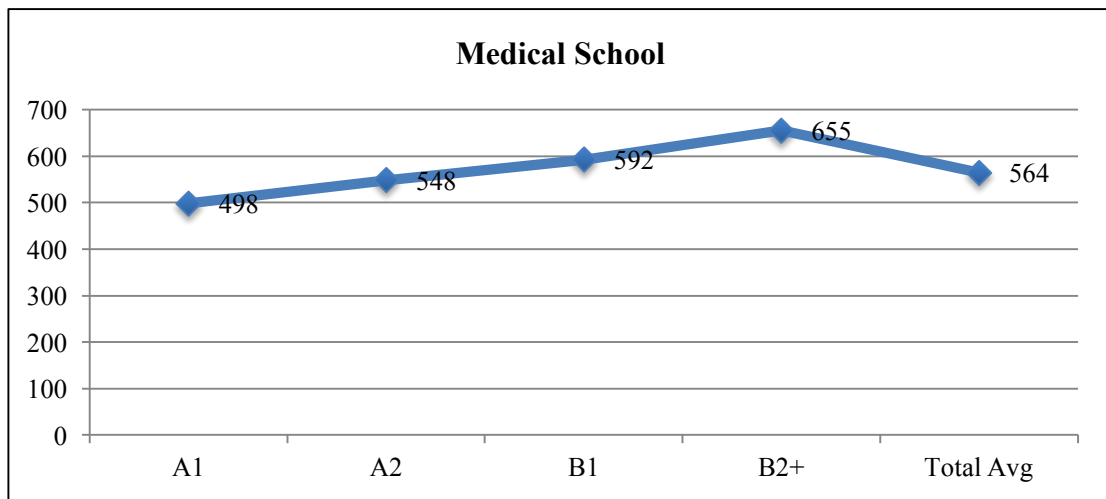


Figure 8: Comparison between TOEIC and interview results - Medical School.

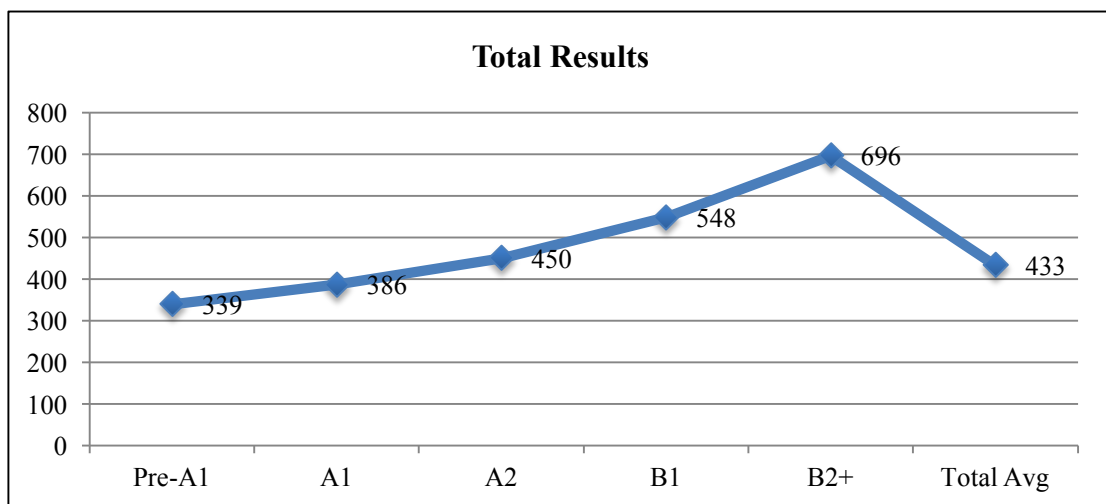


Figure 9: Comparison between TOEIC and interview results - Total results.

To compare the results of the interviews of April 2015 with those of the TOEIC test of July 2015, we divided the students into groups according to their majors and their interview results, and we then calculated the average TOEIC score of each of these groups (the minimum score on the TOEIC test is 0 points, while the maximum score is 990 points²). The results obtained, as the charts above represent, show a marked

² The TOEIC (Test of English for International Communication) Listening and Reading test is a standardized English-language proficiency test which, according to the Educational Testing Service (ETS), measures the everyday English skills of people working in an international environment. ETS reports that more than 2.3 million people in some 150 different countries took the TOEIC Listening and

difference in terms of average TOEIC scores for the groups of students who shared the same interview results. Moreover, considering that, according to ETS, the standard error of measurement (SEM) of the TOEIC Listening and Reading test is 50 points (ETS, 2013), the average TOEIC scores for the groups of students who shared the same interview results shows that the average proficiency in reading and listening between the groups of students with the same interview results was, in most cases, objectively different. This appears to validate the assumption that the interview system developed by the authors was an effective way to discriminate language proficiency according to different levels (Pre-A1, A1, A2, B1, B2).

Given the differences in terms of interview results according to majors, we also tried to compare the results according to interviewers in order to try to understand the efficacy of the norming process. Unfortunately, this analysis is problematic for a number of reasons. First of all, as can be noticed from the bar chart that shows the total results of the interviews, there are important differences in terms of results according to different majors. At the same time, for mere logistical reasons, each interviewer was randomly involved in interviewing only a different number of students from just a few (and not all) majors. Consequently, each interviewer conducted, in total, a different number of interviews (ranging from 12 to over 60), and sometimes the number of interviews conducted by each interviewer was simply too small to determine if, in those specific cases, their performances could be considered as the confirmation of a general trend, nothing more than a casual occurrence, or as a real norming-related issue. Nonetheless, conscious of the pivotal role of the norming process in this kind of language assessment, the authors will further analyze this issue through follow-up interviews scheduled for the end of the current academic year.

Another useful perspective to understand the protocol system is through a comparison of the results among the three different interview protocols, and whether these differences affected the interviews in any way. Anecdotally, the interviewers observed during the interviews that the third protocol seemed to be more challenging than the other two.

Reading test in 2013. It has two sections with 100 multiple-choice questions in each section, and lasts approximately 2 hours.

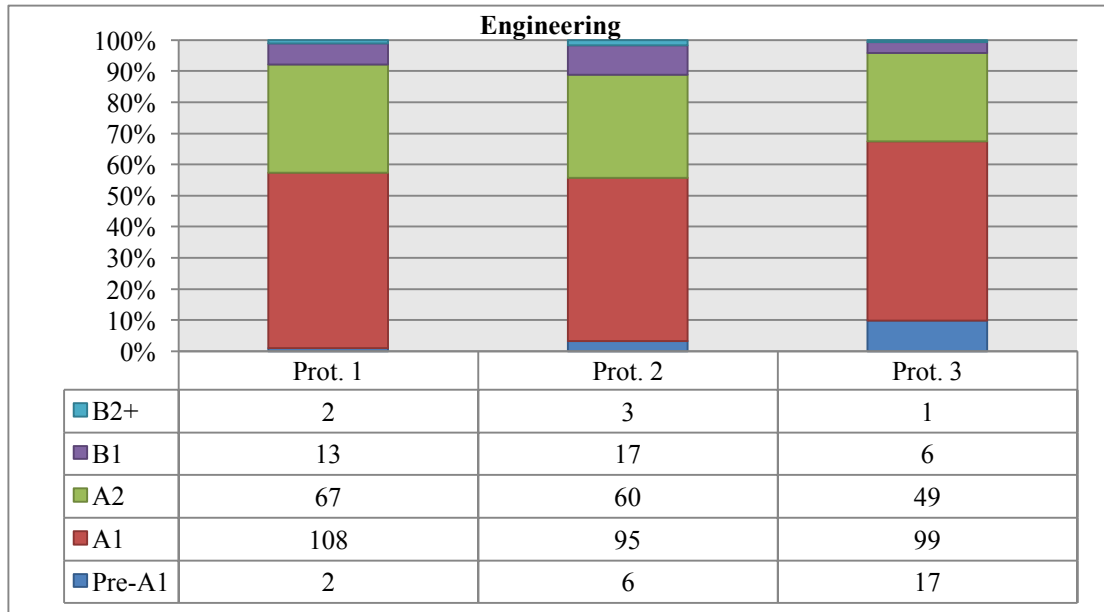


Figure 10: Comparison of results among interview protocols - Faculty of Engineering.

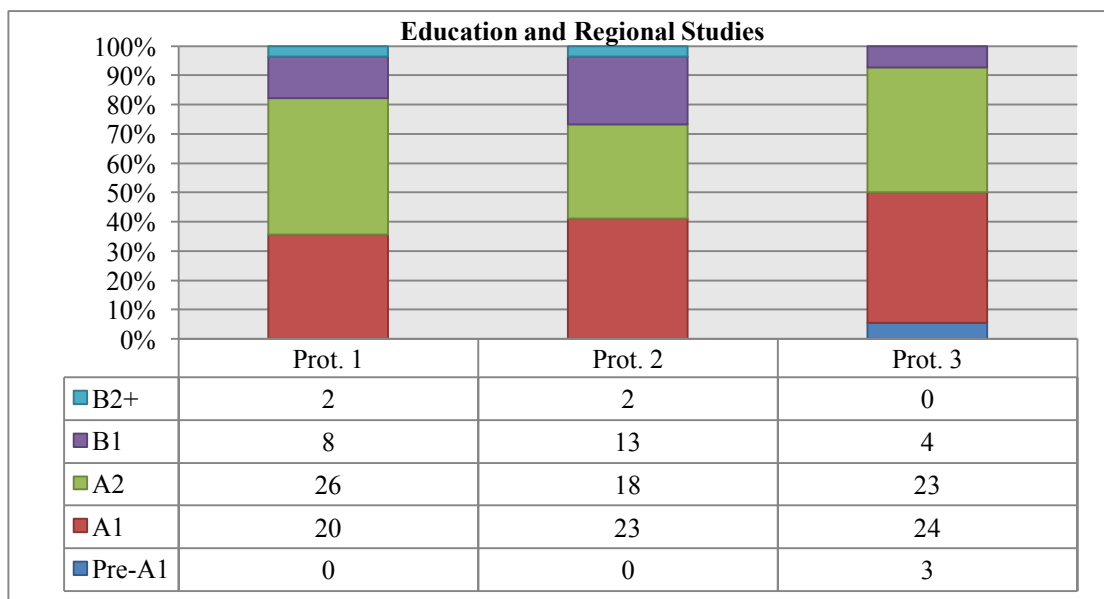


Figure 11: Comparison of results among interview protocols - Faculty of Education and Regional Studies.

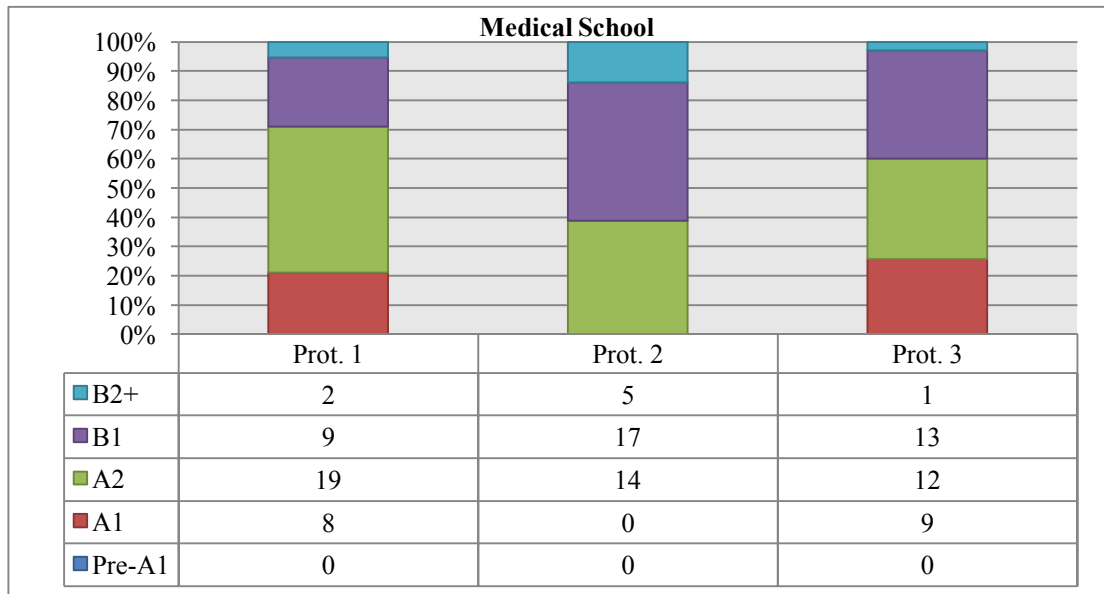


Figure 12: Comparison of results among interview protocols - Medical School.

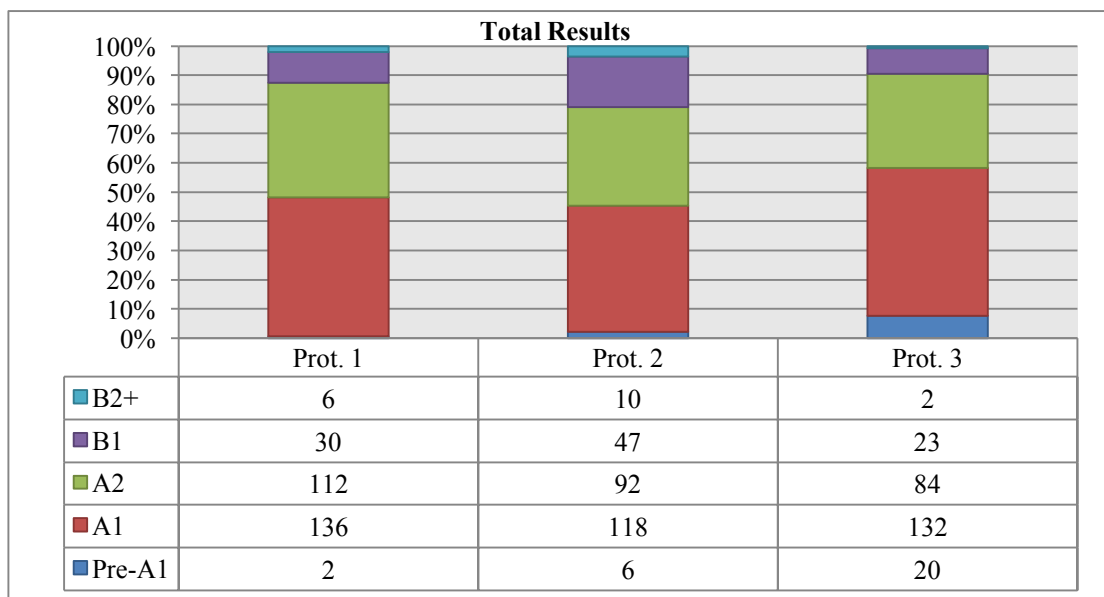


Figure 13: Comparison of results among interview protocols - Total results.

These four charts seem to confirm this trend, and demonstrate the need to rework the protocols and all the single questions, in order to improve the balance among the different protocols while keeping a stress-free conversational style at the same time.

Discussion

In conclusion, efficacy of the oral interview protocol system developed by the authors at a national university is suggested objectively when comparing students' projected

CEFR levels from the oral interview system against the students' TOEIC results. The total average TOEIC scores increase with each successively advanced CEFR level. However, further comparison with future TOEIC tests taken by students in this study as well as comparison against other oral interview systems is necessary to further show validity of the authors' system.

In addition, another issue to be addressed in the future includes refining the norming process for all interviewers through additional training sessions, professional development seminars, and experience through mock interviews. Furthermore, interview protocols need to be refined more to ensure accuracy in all interviews conducted. As for the research presented in the paper, the authors' believe it is necessary to analyze the data even further by breaking down the Education and Engineering department results by specialty. This additional analysis may yield interesting results for the further refinement of this oral interview protocol system.

The authors plan to continue and complement this research in a number of ways. First, the authors conducted qualitative research through questionnaires approximately two weeks after the actual interviews presented in this paper on all student participants and will use this data to develop research results which give insight into the qualitative aspect of this interview system from the participant's perspective. Also, follow-up interviews on randomly selected students will be conducted in January 2016 in order to (1) test new interview protocols developed and (2) attempt to track English progress compared to the first oral interview conducted approximately 10 months earlier. Finally, this oral interview protocol system along with any revisions made to the system will be utilized again in April 2016 on first-year students entering a national university and data will be taken again to further assess the validity of this system.

References

ALTE (2011) *Manual for Language Test Development and Examining*. Council of Europe. http://www.coe.int/t/dg4/linguistic/ManualLanguageTest-Alte2011_EN.pdf (last retrieved November 18, 2015).

Council of Europe (2001). *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Council of Europe (2011). *Manual for relating Language Examinations to the Common European Framework of Reference for Languages (CEFR)*. http://www.coe.int/t/dg4/linguistic/Manuel1_EN.asp (last retrieved November 18, 2015).

Educational Testing Service (2013). *TOEIC User Guide-Listening and Reading*. Princeton, NJ: Educational Testing Service.

MEXT (2014). *Report on the Future Improvement and Enhancement of English Education (Outline): Five Recommendations on the English Education Reform Plan Responding to the Rapid Globalization*. <http://www.mext.go.jp/english/topics/1356541.htm> (last retrieved November 18, 2015).

Shimizu (2001). Survey research on the use of placement tests at four-year universities in Japan. *Ritsumeikan Studies in Language and Culture Studies*, 14, pp. 231-243.

Tono, Y. & Negishi, M. (2012). The CEFR-J: Adapting the CEFR for English language teaching in Japan. *JALT Framework & Language Portfolio SIG Newsletter No. 8*, pp. 5-12.

投野 由紀夫／編 (2013) 『英語到達度指標CEFR-Jガイドブック CAN-DOリスト作成・活用』. 東京: 大修館書店.

Contact email: hennessy@u-fukui.ac.jp
nbecce@u-fukui.ac.jp